



OPEN

Using metabarcoding and droplet digital PCR to investigate drivers of historical shifts in cyanobacteria from six contrasting lakes

Mailys Picard^{1,2,✉}, Xavier Pochon^{1,3}, Javier Atalah¹, John K. Pearman¹, Andrew Rees⁴, Jamie D. Howarth⁴, Christopher M. Moy⁵, Marcus J. Vandergoes⁶, Ian Hawes², Samiullah Khan⁵ & Susanna A. Wood¹

The frequency and intensity of cyanobacterial blooms is increasing worldwide. Multiple factors are implicated, most of which are anthropogenic. New Zealand provides a useful location to study the impacts of human settlement on lake ecosystems. The first humans (Polynesians) arrived about 750 years ago. Following their settlement, there were marked landscape modifications which intensified after European settlement about 150 years ago. The aims of this study were to reconstruct cyanobacterial communities in six lakes over the last 1000 years and explore key drivers of change. Cyanobacterial environmental DNA was extracted from sediment cores and analysed using metabarcoding and droplet digital PCR. Cyanobacteria, including potentially toxic or bloom forming species, were already present in these lakes prior to human arrival, however their overall abundance was low. Total cyanobacteria abundance and richness increased in all lakes after European settlement but was very pronounced in four lakes, where bloom-forming taxa became dominant. These shifts occurred concomitant with land-use change. The catchment of one deteriorated lake is only moderately modified, thus the introduction of non-native fish is posited as the key factor driving this change. The paleolimnological approach used in this study has enabled new insights into timing and potential causes of changes in cyanobacterial communities.

Freshwater ecosystems throughout the world are currently experiencing an increase in frequency and intensity of cyanobacterial blooms^{1–3}. Various causes have been identified, such as increased human disturbances including enhanced erosion caused by deforestation and land-use intensification, and major shifts in food-webs due to the introduction of non-native species (such as fish and macrophytes)^{4,5}. A further compounding factor is climate change, which is resulting in warmer water temperatures, longer periods of stratification, greater frequency of severe weather events such as storms or droughts, and reduced length of ice cover, all of which have also been implicated in the rise of cyanobacterial proliferations^{2,6–8}. Many bloom-forming species produce toxins collectively known as cyanotoxins, which can affect the reproduction and behaviour of aquatic organisms and have led to poisoning, or in extreme cases, death of humans and animals^{9,10}. Considerable effort has therefore been placed on developing lake management and restoration plans for water bodies experiencing cyanobacterial blooms^{11–13}. Unfortunately, historical records are very limited and monitoring programmes often start once a lake is already in an advanced state of degradation. There is uncertainty as to whether the recent increase in cyanobacterial blooms is entirely due to eutrophication and environmental change, or to a greater awareness and more monitoring programs contributing to these new bloom reports¹⁴. Increasing knowledge on how lake ecosystems, their catchments and regional climates have changed historically, and identifying links with cyanobacterial blooms, may help in the development of informed and successful restoration actions. New Zealand is an island nation in

¹Coastal and Freshwater Group, Cawthron Institute, 98 Halifax Street East, Nelson 7010, New Zealand. ²Department of Biological Sciences, School of Biological Sciences, University of Waikato, Hamilton 3216, New Zealand. ³Institute of Marine Science, University of Auckland, Private Bag 349, Warkworth 0941, New Zealand. ⁴School of Geography, Environment and Earth Sciences, Victoria University of Wellington, PO Box 600, Wellington, New Zealand. ⁵Department of Zoology, University of Otago, 340 Great King Street, North Dunedin, Dunedin 9016, New Zealand. ⁶GNS Science, 1 Fairway Drive, Avalon, Lower Hutt 5011, New Zealand. ✉email: Mailys.Picard@cawthron.org.nz

the South Pacific that has been isolated from all other landmasses for over 80 million years. The first humans to set foot on the mainland were from Polynesia (Māori), and only arrived around 1250–1350 AD^{15–17}. Following their arrival, there were marked landscape modifications in some regions of the country. For example, burning of native forests^{18–21} reduced overall coverage by about 50%^{20,21}. Land clearance intensified with the first major wave of European settlement beginning in the 1840s¹⁶, coupled with farming practices which used fertilisers leading to nutrient runoffs, urbanisation releasing contaminants in waterways, and the introduction of many non-native freshwater species affecting native food-webs^{22–24}.

Paleolimnology provides a suite of methods that allow inferences to be made about a lake's historical condition—its past biological communities, water quality, and vegetation changes in and around the lake. For example, shifts in biological communities, such as midges, diatoms, or vegetation, can be indicative of past disturbance events^{25–28}. Paleolimnological reconstructions have traditionally relied on microfossil indicators, however molecular methods targeting environmental DNA (eDNA) now allow a wider range of micro- or macro-organisms to be investigated in sediment cores, particularly soft-bodied organisms. While DNA can be extracted from the organism, it can also be recovered when it is bound to extracellular matrices in the sediment (exDNA)²⁹. Recent work has shown that it is possible to use DNA from sediment (sedDNA), and ancient DNA from sediment (sedaDNA) to reconstruct the historical dynamics of various lake organisms including mammals, fish, and plants^{30–33}. High-Throughput Sequencing (HTS) coupled with metabarcoding analyses allows for a large number of DNA sequences to be amplified directly from environmental samples and these can be used to characterise entire communities^{34–37}. A recent study used metabarcoding to reconstruct cyanobacterial diversity over the last 200 years in Lake Greifensee and Lake Zürich, Switzerland³⁸. Their results were congruent with an independent dataset obtained by microscopic identification of water samples for the last 40 years. Furthermore, quantitative PCR (qPCR) or droplet digital PCR (ddPCR) have also been successfully applied to paleolimnological studies for various targets including cyanobacteria and genes involved in cyanotoxin production in several countries^{39–42}.

New Zealand has 3820 natural lakes greater than a hectare⁴³, and due to its recent human settlement it provides a particularly useful location to study the ecological impacts of human settlement on lake ecosystem. It is possible to obtain sediment from pre-human times and reconstruct the past c. 1000 years with a relatively short sediment core (1–2 m) for most small lakes. Understanding the past of lakes is important because less than 200 of New Zealand's lakes have regular water quality monitoring programmes in place, most of which have been underway for less than two decades. It is estimated that 46% of all lakes are eutrophic⁴⁴ with many experiencing cyanobacterial blooms. The direct impact of anthropogenic activities is thought to be the main driver, but climate change may also be contributing. Over the last century, land temperatures have increased by 1 °C on average, droughts and floods have become more frequent and severe⁴⁴.

The aims of this study were to reconstruct cyanobacterial communities in six New Zealand lakes over a period spanning approximately 1000 years using sediment cores, ddPCR and metabarcoding, and to identify key drivers that may have led to increases in cyanobacteria abundance or shifts in community structure. We hypothesised that: (1) cyanobacterial abundance and composition would start to shift during the period following Māori settlement in lakes where the vegetation cover in the catchment had changed. However, the most pronounced change would occur post European settlement with bloom-forming species becoming more prevalent in lakes which had experienced land use change; and (2) in lakes with little modification in vegetation cover in their catchment, the increases in abundance and changes in composition would be related to the presence of non-native carnivorous fish having a top-down effect on the food chain. While we acknowledge that climate change is likely to have some impact on the study lakes, we do not directly explore this in this study.

Methods

Sampling. Sediment cores were collected from six lakes in New Zealand: Rotoehu, Pounui, Wairarapa, Paringa, Hayes, and Johnson (Fig. 1, Table 1). These lakes were chosen because they have a wide variety of catchment modifications, span a range of trophic levels, have varying histories of cyanobacterial blooms, and contain different species of non-native fish (Table 1). The lakes span a range of climatic regions (Table 1).

A sediment core was retrieved close to the deepest point of each lake (Table 1) using an Uwitech gravity corer with 2 m-long, 90-mm diameter polyvinyl chloride barrels. The exception was Lake Paringa, where the sediment core was collected using a Mackereth corer⁵⁵ and sealed within plastic barrels (50 mm internal diameter). All barrels were cleaned with 2% sodium hypochlorite (bleach) prior to coring. Core lengths ranged from 0.80 to 5.52 m (Table 1). After retrieval, the cores were sealed, stored at 4 °C and in darkness for up to 4 weeks until sub-sampling.

Sample processing. In a room separate from where all DNA work was undertaken, the cores were split in half using a manual saw and a guillotine. They were photographed and described in detail to relate sediment type and colour and the presence of organic material. To prevent cross-contamination caused by the splitting of the cores, the top 2–3 mm of one half-core per lake was carefully removed with a sterile spatula. Sub-samples (c. 0.5 g) were taken from the centre of the half-core using a sterile spatula at various depths down the core. In general, sub-samples were taken every 1–2 cm in recent sediments and every 4–5 cm in older sediments (Supplementary Table S1), resulting in 206 sediment subsamples. Sub-samples were kept frozen (–20 °C) and in the dark until DNA extraction. Sub-samples were also collected from a range of depths for pollen and charcoal analysis in all lakes as well as carbon-14 dating for Lakes Paringa and Pounui. These analyses are described elsewhere^{47,56–58} and a brief summary of methods is provided in Supplementary Table S2.

DNA extraction and amplification. Each step of molecular analyses (DNA extraction, PCR or ddPCR set-up, template addition, PCR/ddPCR analysis) was conducted in separate sterile laboratories dedicated to

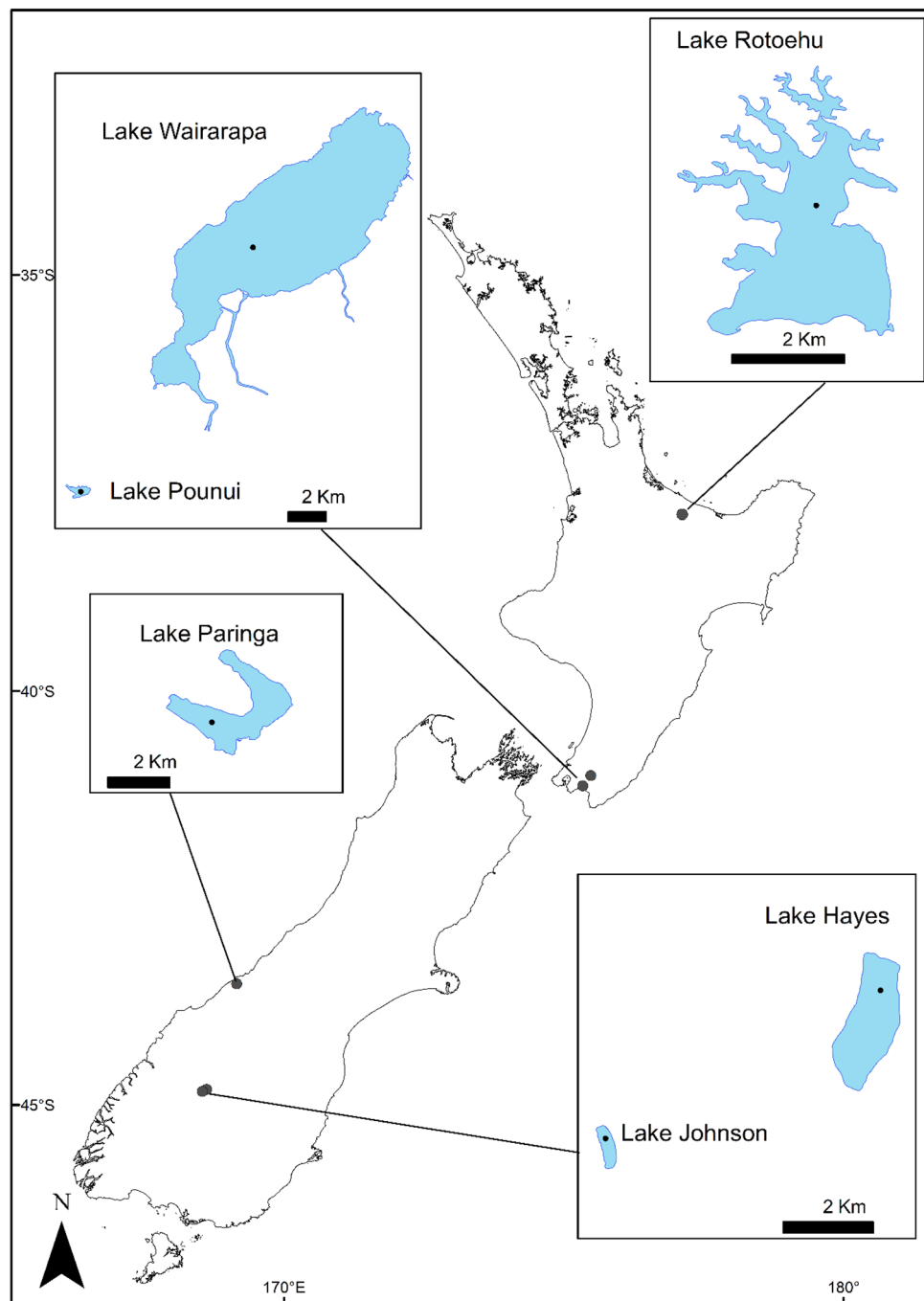


Figure 1. Location of the six study lakes, situated within the North and South islands of New Zealand and spanning 7° of latitudinal gradient. Black dots in each lake indicate coring sites. Map created using ArcGIS version 2.8 (<https://www.esri.com>) by Kati Doehring, Cawthron Institute.

these steps, with sequential workflow to ensure no cross-contamination. Rooms dedicated to DNA extraction, PCR set-up, or template addition were equipped with ultra-violet sterilisation which was switched on for at least 15 min before and after each use. The PCR/ddPCR set-up and template addition was always undertaken in laminar flow cabinets with HEPA filtration. Aerosol barrier tips (Axygen, USA for PCR or ep.T.I.P.S., Eppendorf, Hamburg, Germany for ddPCR) were used throughout.

Approximately 0.25 g of sediment from each sub-sample were weighed in the first tube provided in the DNeasy Power-Soil™ DNA Isolation Kit (QIAGEN, Germany) and exact weights recorded. Environmental DNA was extracted following the manufacturer's protocol. DNA extraction was performed in batches of eight to ten samples, including a negative control every two batches which contained all the reagents but no sediment. DNA concentrations and quality were measured using a spectrophotometer (Eppendorf AG, Hamburg, Germany).

| | Paringa | Pounui | Rotoehu | Hayes | Johnson | Wairarapa |
|---|--|---|--|---|---|---|
| FENZ ID | 46,725 | 229 | 40,188 | 54,190 | 53,707 | 1708 |
| Latitude (S) | 43°43'07.0" | 41°20'40.0" | 38°00'54.4" | 44°58'28.6" | 45°00'06.5" | 41°13'31.6" |
| Longitude (E) | 169°23'17.5" | 175°06'51.8" | 176°31'53.4" | 168°48'40.7" | 168°43'55.8" | 175°12'58.0" |
| Rainfall (mm/month) | 200–300 | 75–137 | 92–136 | 50–75 | 50–75 | 75–137 |
| Temperature (°C/month) | 7.4–15.6 | 8.9–17.7 | 7.8–17.9 | 3–15.8 | 3–15.8 | 8.9–17.7 |
| Coring date | December 2017 | August 2016 | April 2019 | November 2017 | November 2017 | June 2017 |
| Core total length (m) | 5.52 | 1.30 | 1.52 | 0.82 | 0.80 | 2.00 |
| Lake elevation (masl) | 16 ⁴⁵ | 14 ^{45,46} | 296.9 ⁴⁶ | 324.8 ⁴⁶ | 392.1 ⁴⁶ | 1.3 ⁴⁶ |
| Lake max depth (m) | 52* | 9.6 ⁴⁷ | 13.5 ⁴⁶ | 33 ⁴⁸ | 27 ^{46,49} | 2.5 ^{45,46} |
| Lake surface area (km ²) | 4.75 ⁴⁵ | 0.46 ^{45,46} | 7.90 ^{45,46} | 2.74 ^{45,49} | 0.25 ^{45,46} | 77.37 ^{45,46} |
| Lake total volume (km ³) | 0.29 ⁴⁶ | 0.002 ⁴⁶ | 0.035 ⁴⁶ | 0.31 ⁴⁵ | 0.002 ⁴⁶ | 0.64 ⁴⁶ |
| Lake type | Glacial ⁴⁶ | Tectonic ⁴⁶ | Volcanic ⁴⁶ | Glacial ⁴⁶ | Glacial ⁴⁶ | Riverine ⁴⁶ |
| Mixing type | No data | Polymictic ⁴⁷ | Polymictic ⁴⁹ | Monomictic ⁴⁸ | Monomictic ⁴⁸ | Isothermal ⁴⁹ |
| Current catchment area (km ²) | 79.09 ⁴⁵ | 6.27 ⁴⁵ | 73.85 ⁴⁵ | 44 ⁴⁵ | 1.90 ⁴⁵ | 654.25 ⁴⁵ |
| Current catchment composition | Native forest and shrubs ⁴⁶ | Native forest and a bit of pasture ⁴⁶ | Pasture, urban, native forest ⁴⁶ | Pasture and grasslands, urban, bit of native forest ⁴⁶ | Mainly pastoral land, bit of native forest and shrubs ⁴⁶ | Pasture, urban, native forest ⁴⁶ |
| Trophic status—TLI | Oligotrophic (2020)* | Eutrophic (2014) ⁵⁰ | Eutrophic (2018) ⁴⁹ | Eutrophic (2018) ⁴⁹ | Eutrophic (2018) ⁴⁹ | Supertrophic (2018) ⁴⁹ |
| Reported cyanobacterial blooms | Not monitored | Yes ⁴⁹ | Yes ⁴⁹ | Yes ^{48,49} | Yes ⁴⁸ | Yes ⁵⁰ |
| Non-native carnivorous fish species | Brown trout (<i>Salmo trutta</i>) in 1950s, Chinook/Quinnat salmon (<i>Oncorhynchus tshawytscha</i>) from 2011 | Rainbow trout (<i>Oncorhynchus mykiss</i>) from 1938 to 1958 (Adam Canning pers. comm.), European perch (<i>Perca fluviatilis</i>) in the 1960s (Rawhiri Smith pers. comm.) | Rainbow trout (<i>O. mykiss</i>) around 1900 ⁵¹ | Brown trout (<i>S. trutta</i>) and European perch (<i>P. fluviatilis</i>) from 1870s ^{52,53} | European perch (<i>P. fluviatilis</i>) from 1880s (Otago Fish & Game, unpublished data), rainbow trout (<i>O. mykiss</i>) from 1962 ⁴⁸ | European perch (<i>P. fluviatilis</i>) unknown date ⁴⁵ |

Table 1. Lake characteristics. Note that the FENZ ID is related to the Freshwater Ecosystems of NZ (FENZ) geo-database, and the Trophic Level Index (TLI) is based on total nitrogen, total phosphorous, water clarity, and chlorophyll-a. Rainfall and temperature values are the highest and lowest mean monthly values over one year, data from NIWA⁵⁴. Lakes are ordered from left to right by increasing lake trophic status. *Lakes380 unpublished data.

Quantifications of total cyanobacteria. Droplet digital Polymerase Chain Reaction (ddPCR) was used to quantify total cyanobacteria from each sample using the CYAN 108F and CYAN 377R primers targeting an approximate 270 base-pairs (bp) region of the 16S rRNA gene⁵⁹. All samples were diluted for quantification by ddPCR; dilutions ranged from 1/10 to 1/1000. The ddPCR was undertaken using a BioRad QX200 system. Each ddPCR reaction included 10 μ L of 2 \times BioRad QX200 ddPCR EvaGreen Supermix, 0.2 μ L of each primer at 10 μ M, 4 μ L of diluted template DNA, and 7.6 μ L of DNA/RNA-free water (Life Technologies) for a total reaction volume of 22 μ L per well, loaded on a semi-skirted ddPCR 96-well plate. The BioRad QX200 droplet generator partitioned each reaction mixture into tens of thousands of nanodroplets by combining 20 μ L of the mixture with 70 μ L of Evagreen droplet oil. After processing, this resulted in a total nanodroplet volume of 40 μ L, which was transferred to another semi-skirted ddPCR 96-well plate for amplification using specific thermo-cycling conditions (Supplementary Table S3). The plate was analysed on the QX200 droplet reader instrument. For each ddPCR plate run, at least one negative control (containing all reagents and DNA/RNA-free water instead of template DNA) and one positive control (genomic DNA extracted a cyanobacteria culture) were included. Droplet digital PCR droplet counts were fitted to a Poisson distribution by the QuantaSoft Analysis software (BioRad), resulting in target DNA concentration (gene copies/ μ L) which were then standardised to DNA gene copy numbers per gram of sediment using the following formula:

$$\text{Target concentration (gene copies/g)} = \frac{\text{nbcopies}}{\mu\text{L}} * \frac{\text{Total Mastermix volume (22}\mu\text{L)}}{\text{Template DNA volume (4}\mu\text{L)}} * \text{dilution factor (10 – 1000)} * \frac{\text{DNA extraction volume (100}\mu\text{L)}}{\text{sediment sample weight (g)} * (1 - \text{water content})}$$

Reconstructing cyanobacterial communities with metabarcoding. Polymerase Chain Reactions (PCRs) were performed using the cyanobacteria-specific primers CYB359-F and CYB784-R⁶⁰. These primers amplify an approximately 400 bp fragment of the V3-V4 regions of the cyanobacterial 16S rRNA gene. Sedimentary DNA samples were amplified in PCRs of 48 μ L volume each, containing AmpliTaq Gold[®] 360 Master Mix (Life Technologies), 360 GC enhancer (Life Technologies), Bovine Serum Albumin (Sigma), primers at 10 μ M,

DNA/RNA free water (Life Technologies), and template DNA (Supplementary Table S3). Batches contained 20 samples, including a negative and positive control, PCR cycling conditions are detailed in Supplementary Table S3.

Amplicon products, including all negative extraction and PCR controls, were visualised on a 1.5% agarose gel electrophoresis stained with Red Safe DNA Loading Dye (iNtRON Biotechnology Inc, Kyungki-Do, Korea), and UV illumination to ensure amplification of a single ~ 400 bp product. PCR products were purified (Agen-court AMPure XP Kit; Beckman Coulter, CA, USA), quantified (Qubit® 2.0 Fluorometer; Invitrogen, CA, USA), diluted to 5 ng μL^{-1} and submitted for sequencing to Auckland Genomics (University of Auckland). Sequencing adapters and sample-specific indices were added to each amplicon via a second short round of PCR using the Nextera™ Index kit (Illumina, CA, USA). Amplicons were pooled into a single library and paired-end sequences (2 × 250 bp) were generated on a MiSeq™ instrument using the TruSeq™ SBS kit (Illumina, CA, USA). Sequence data were automatically demultiplexed using a MiSeq Reporter (v2), and forward and reverse reads were assigned to samples.

Bioinformatic analyses. The R software⁶¹ and RStudio software⁶² were used for all bioinformatic and statistical analyses. The Tidyverse v1.3.0 and its associated packages⁶³ were used for data manipulation and exploration. All plots were drawn using the ggplot2 package⁶⁴ unless stated otherwise. Primer sequences were removed with Cutadapt⁶⁵ (Anaconda environment adapted in R) allowing 1 bp of mismatch. The DADA2 package⁶⁶ was used for general sequence quality assessment, quality profiles plots, and for the full amplicon workflow (quality filtering, merging of paired-end reads, dereplication, chimera identification, sample inference, and taxonomy assignment). Reads were truncated at 225 and 215 bp for forward and reverse reads respectively. The maximum number of “expected errors” (maxEE) per read was set at two for forward reads and four for reverse reads. Other parameters were set to default. The first 108 bp of the sequences were then used to calculate a parametric error matrix for forward and reverse reads each, which were checked for convergence. Sequences were dereplicated and sequence variants were inferred by pseudo-pooling based on their respective error matrix. Singletons were discarded and remaining paired-end reads were merged with a maximum mismatch of 1 bp and a required minimum overlap of 10 bp, producing a sequence table. Samples from all sequencing runs were then merged into one sequence table. Amplicon lengths were filtered so that only those within the range of 379 and 403 bp were retained and checked for chimeras using the consensus method, resulting in Amplicon Sequence Variants (ASVs). Taxonomy was assigned from Kingdom to Genus for each ASV using a training set from the SILVA database r138^{67–69}. If the prediction of the SILVA classifier was estimated to be correct at 85% or more (min bootstrap), then the taxonomy for that rank was assigned to the ASV.

The phyloseq package⁷⁰ was used to prepare the sequence table and associated information and undertake some analyses (merging, rarefaction with rarefy_even_depth, richness measures with estimate_richness). All non-bacterial reads (eukaryotes, mitochondria, chloroplasts) were removed from the dataset. Rarefaction curves were plotted using the vegan package⁷¹ to visualize sampling depth for each sample. Rarefaction of the entire dataset was undertaken at 10,400 reads per sample using phyloseq on all remaining bacterial reads through random sub-sampling with no replacement. Cyanobacterial-only ASVs were selected after rarefaction for further univariate analysis (richness). After rarefaction, some samples appeared to have no cyanobacteria ASVs because their very low numbers of cyanobacterial reads were removed during rarefaction. Multivariate analyses were therefore undertaken on unrarefied samples, where photosynthetic cyanobacterial reads counts were transformed to relative reads counts (%) per sample.

Defining phases and age models. Pollen and charcoal data were used to assign specific occupation periods which are referred to throughout this study as: Pre-Human (PH), Evidence of Māori Settlement (EMS, from approximately 1300 AD), and Post-European Settlement (PES, from approximately 1840 AD) phases (Supplementary Figs. S1 to S5). The main markers for the start of the EMS were peaks in charcoal and bracken fern spores, both of which are indicators of landscape disturbance^{15,20}. The primary indicator for the start of the PES phase was the appearance of exotic plant pollens such as the Monterey pine (*Pinus radiata*), sheep's sorrel (*Rumex acetosella*), and willow (*Salix* spp.)⁷². The catchment of Lake Paringa is relatively undisturbed and for this lake the palynology did not provide any useful information on the start or finish of the eras therefore the age model was used to assign these (see below).

Age models from ¹⁴C were developed in other studies for the sediment cores of Lake Pounui and Lake Paringa^{47,56} (presented in Supplementary Tables S4 and S5) and correlated and applied to the cores collected in this study. The cores from Lakes Rotoehu, Hayes, Johnson, and Wairarapa were not dated. In Lake Rotoehu, an approximately 5-cm tephra layer corresponded to the 1886 eruption of Mount Tarawera provides a defined point in time.

Statistical analyses. Cyanobacteria abundances (number of 16S rRNA gene copies per gram of sediment) from ddPCR data were reconstructed across lakes. The number of photosynthetic cyanobacteria (also now commonly called oxyphotobacteria) ASVs (species richness) from metabarcoding data were calculated and plotted as downcore profiles; the number of species within each Order and Class (photosynthetic or not) were plotted as barplots, and the relationship between cyanobacteria abundance and photosynthetic richness for individual lakes was assessed using scatterplots and Spearman's correlation.

Three binary categories of perturbations were assigned to each sample. The categories were: (1) land clearance as the start of large-scale vegetation disturbance—indicated by peaks of charcoal and/or bracken fern spores, which were indicative of burning, (2) land-use change – characterised by increases in pollen from non-native grasses and/or pine during the PES phase, and (3) presence of carnivorous non-native fish (trout or perch).

Catchment-related perturbations were observed in all lakes apart from Lake Paringa, indicated by increased in charcoal, bracken spores, and/or terrestrial non-native pollen (Supplementary Figs. S1 to S5). When the sediment cores were dated, the dates of non-native fish introduction in each lake were determined directly or indirectly from historical documentation or oral historical reports (Lake Paringa⁵⁶; Rawhiri Smith pers. comm. for Lake Pounui). Where no dating was available (Rotoehu, Hayes, Johnson, Wairarapa), non-native fish introduction was assumed to be associated with the start of the PES phase.

Generalised least squares (GLS) models were then used to investigate the drivers associated with changes in historical cyanobacteria total abundance and photosynthetic richness. GLS were chosen over other statistical methods to account for the temporal autocorrelation between proximate down-core samples⁷³. Five candidate predictor variables were initially considered (Supplementary Table S6): Phase (three levels: PH, EMS, and PES), Non-native Fish (binary), Native Vegetation Change (binary), Land-Use Intensification (binary) and Lake (six levels: Hayes, Johnson, Paringa, Pounui, Rotoehu, and Wairarapa). Because all predictors were highly correlated and collinear, in particular non-native fish presence and land-use intensification, their effects were summarised by the three phases: PH, EMS—encompassing the beginning of native vegetation clearance, and PES—encompassing non-native fish introduction and land-use intensification. Though not all lakes were affected by these pressures (Lakes Paringa and Pounui have not undergone significant changes in their catchment), this proved to be the best option for modelling purposes. Models were fitted using Phase and Lake as fixed effects and with a continuous first-order autoregressive process (corCAR1), with sub-sample core depth nested within Lake to correct for autocorrelation. Because cyanobacteria abundances consisted of strictly positive and right-skewed continuous data while richness was mildly right-skewed, they were log and square-root-transformed, respectively. Model assumptions of normality, homogeneity of variances and independence were checked by inspecting normalised residuals, including using autocorrelation function plots. Only photosynthetic cyanobacteria were selected for the multivariate analysis.

Shared ASVs across the dataset, across lakes, across phases, and across phases within lakes were investigated, and visualised using Venn diagrams⁷⁴. Changes in relative read abundance in the cyanobacterial community structure were plotted for each phase within each lake at genus level using stacked barplots. To further understand specific changes in cyanobacteria versus picocyanobacteria across time and lakes, all ASVs identified at Genus level were assigned to a cell diameter factor: either > 3 µm, or < 3 µm (picocyanobacteria), as well as a toxin category (none or potentially toxic, Supplementary Table S7). Community changes at genus level were visualised with barplots and aligned with dendrograms using the packages *rioja*⁷⁵ and *ggdendro*⁷⁶. Variations in ASV community structure were then visualised using principal coordinates analyses ordination plots (PCoA) based on Bray–Curtis dissimilarities using the *vegan* package⁷¹. Difference in community structure were explored using a distance-based permutational analysis PERMANOVA⁷⁷ based on Bray–Curtis dissimilarities of the relative abundance community data and 999 permutations. The experimental design consisted of Lake and Phase as fixed factors⁷⁸. Significant terms were then investigated using pair-wise comparisons with the PERMANOVA t-statistic and 999 permutations. Tests for homogeneity of multivariate dispersions based on Bray–Curtis distances⁷⁸ were used to compare distances to the multivariate centroids between Lake, and between Phase within individual lakes.

Results

General richness patterns. A total of 21,545 bacterial ASVs were recovered from the 208 samples sequenced. Of these, 1361 corresponded to all cyanobacteria and 1199 to photosynthetic cyanobacteria. After rarefaction (10,400 reads), a total of 1144 cyanobacterial ASVs remained, of which 928 were photosynthetic cyanobacteria. Several samples were discarded because they either did not have enough sequence reads to meet the rarefaction threshold (Supplementary Fig. S6, two samples), or they did not have any cyanobacterial reads following rarefaction (thirteen samples), or any photosynthetic cyanobacterial reads (further five samples).

Cyanobacterial ASVs were assigned to two different classes: Vampirovibrionia (previously Melainabacteria candidatus⁷⁹, non-photosynthetic cyanobacteria) and Cyanobacteria (photosynthetic cyanobacteria). Cyanobacteria ASVs were found in every lake and were amplified in 193 instances in the rarefied dataset (i.e., 93.7% of investigated [n = 206] samples). There were no common photosynthetic cyanobacteria ASVs across all samples, with only three ASVs belonging to the genus *Cyanobium* found in all lakes. When aggregating all samples by phase, there were 34 shared ASVs across the three phases (PH, EMS, PES), belonging to the genera *Cyanobium* (25 ASVs), *Dolichospermum* (2 ASVs), *Microcystis* (4 ASV), *Synechocystis* (1 ASV), and 3 ASVs unassigned at Genus level (1 unknown Nostocales, 1 unknown Cyanobacteriales and 1 unknown Nostocaceae). An assessment of individual lakes showed that in Lakes Hayes and Johnson a single ASV was found in all samples (a different one in each lake), taxonomically assigned to the genus *Cyanobium*. In all lakes the highest number of shared ASVs was between the PES and EMS phases (Supplementary Fig. S7). Lake Wairarapa was the only lake that did not have a core community across all phases (Supplementary Fig. S7). Shared taxa differed from one lake to the other, though they belonged mainly to the picocyanobacteria genus *Cyanobium*. Total Cyanobacteria, Vampirovibrionia and (photosynthetic) Cyanobacteria ASVs varied among lakes, with the highest numbers recorded in Lake Johnson (Table 2). A total of 47.8% cyanobacterial ASVs were unclassified at Genus level (493 out of 1031), including 297 ASVs that corresponded to photosynthetic cyanobacteria.

The majority of Vampirovibrionia ASVs could only be assigned at Order level (Caenarcaniphilales, Gastranaerophilales, Obscuribacterales, and Vampirovibrionales). Vampirovibrionia ASVs were found in every lake but only occurred in 104 out of the 193 samples and their richness differed greatly between lakes (>40 ASVs in Lake Hayes and Johnson and <40 ASVs in the other lakes, Supplementary Figs. S8 and S9). The Order Gastranaerophilales was the most common (7–72 ASVs per lake), followed by Obscuribacterales (4–11 ASVs per lake). The other two Orders (Caenarcaniphilales and Vampirovibrionales) were poorly represented and were

| | Paringa | Pounui | Rotoehu | Hayes | Johnson | Wairarapa | Total |
|---------------------|---------|--------|---------|-------|---------|-----------|-------|
| Number of samples | 29 | 36 | 29 | 31 | 37 | 30 | 187 |
| Total Cyanobacteria | 116 | 134 | 254 | 272 | 359 | 142 | 1034 |
| Cyanobacteria | 80 | 115 | 225 | 223 | 280 | 129 | 839 |
| Vampirovibrionia | 36 | 19 | 29 | 49 | 79 | 13 | 195 |

Table 2. Cyanobacterial amplicon sequence variants (ASVs) per lake. Results are given for total cyanobacteria and per Class.

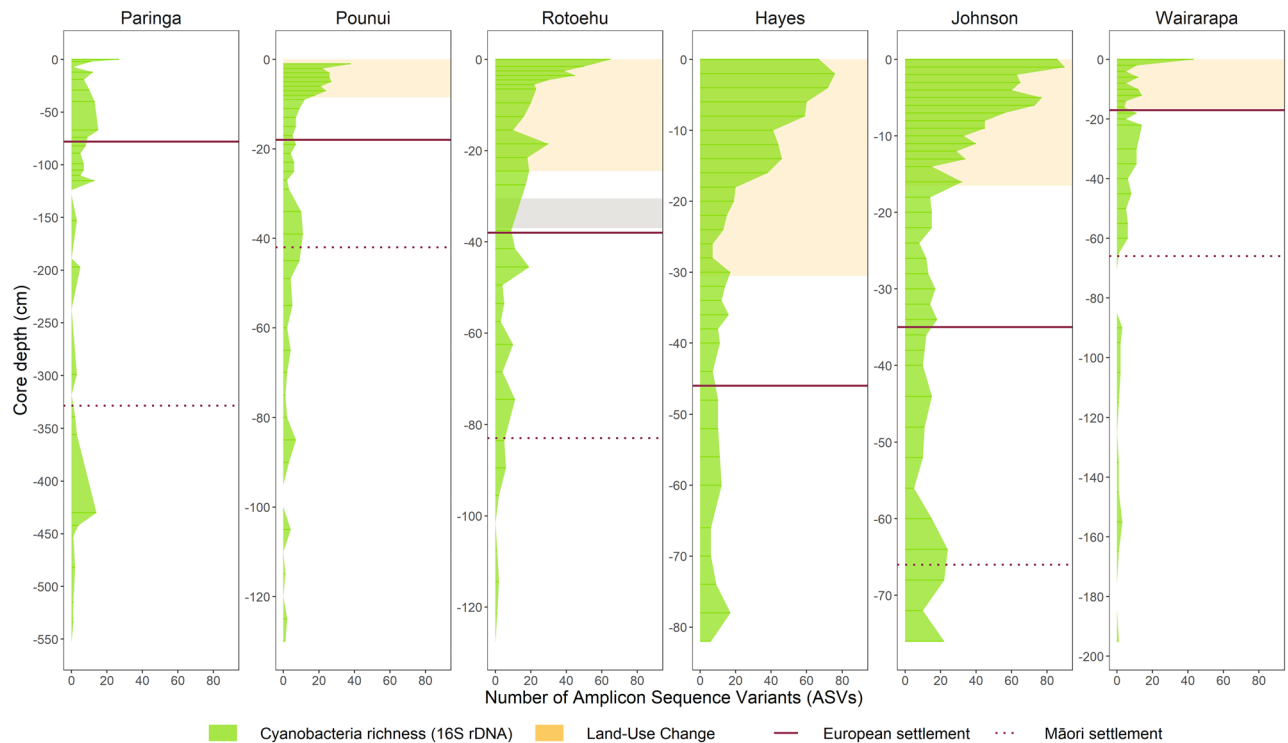


Figure 2. Richness of photosynthetic cyanobacterial Amplicon Sequence Variants (ASVs) from metabarcoding data, cyanobacterial 16S rRNA gene. The horizontal lines separate the three phases, pre-human (PH; bottom), evidence of Māori settlement (EMS; middle), and post European settlement (PES; top); note: the sediment core from Lake Hayes does not include any PH phase. Yellow shades indicate the depths associated with land-use intensification in relevant lakes and the grey shade indicates the 1886 Tarawera tephra layer in Lake Rotoehu.

not found in all lakes (Supplementary Fig. S9). Due to the lack of ecological knowledge on Vampirovibrionia, no further analysis was conducted.

The rarefied cyanobacterial photosynthetic dataset was composed of 8 orders, 21 families, and 51 genera across all lakes and all samples. The most common Cyanobacteria Order was Synechococcales, followed by Cyanobacteriales (formerly Nostocales; Supplementary Fig. S8). The number of Synechococcales ASVs was variable across lakes by an order of ten (55–153 ASVs per lake), and it was similar for Cyanobacteriales (14–120 ASVs per lake). All other photosynthetic cyanobacteria Orders were poorly represented (≤ 4 ASVs per lake) and were not detected in all lakes.

Cyanobacterial richness and abundance profiles. The richness (number of ASVs) of photosynthetic cyanobacteria increased significantly in all lakes during recent times compared to their historical levels (Fig. 2). Overall, Lakes Hayes and Johnson had the highest richness and Lake Wairarapa the lowest (Figs. 2 and 4). Shifts in richness after European arrival were particularly notable in lakes known to currently experience cyanobacterial blooms (Hayes, Johnson, Pounui, Rotoehu). This was particularly obvious in Lake Rotoehu (PH = 0–6 ASVs, PES = 9–65 ASVs), while Lakes Hayes and Johnson displayed a very high richness in recent times (80–90 ASVs at the top of the core). Despite having moderate ‘European land-use’ in its catchment, Lake Pounui also showed an increase in richness. Timing of shifts were different for Lakes Pounui (~1900s) and Paringa (~1700s). No specific dates could be approximated for Lakes Rotoehu, Hayes, Johnson, and Wairarapa due to the absence of age models, though the increase in richness coincided with land-use intensification in Lake Johnson. The richness profile of Lake Wairarapa was more consistent with the timing of phases, with a slight increase to 8 ASVs on average after EMS until near the top of the core, with a marked increase in the top sample (43 ASVs).

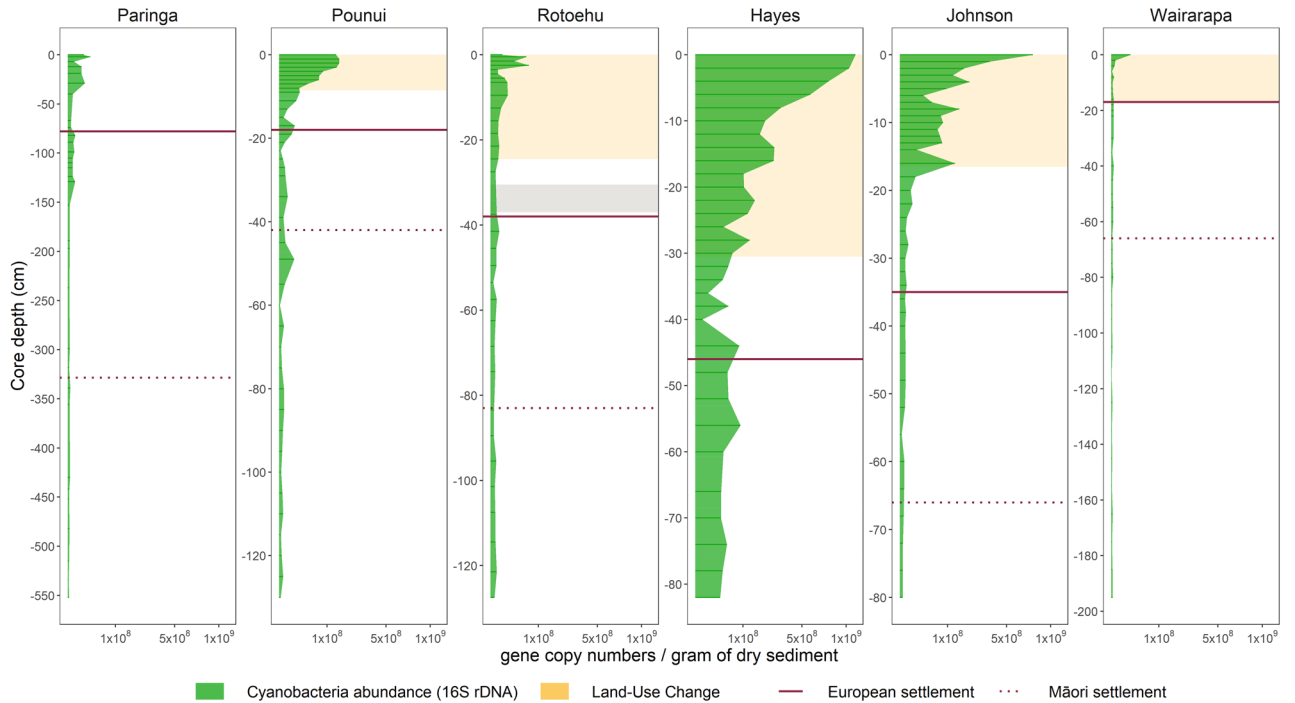


Figure 3. Cyanobacterial concentrations from droplet digital PCR data. Sediment depth profiles of cyanobacterial 16S rRNA gene copy numbers per gram of dry sediment, y-axis shown on square-root scale. The horizontal lines separate the three phases, pre-human (PH; bottom), evidence of Māori arrival (EMS; middle), and post European arrival (PES; top); note: the sediment core from Lake Hayes does not include any PH phase. Yellow shades indicate the depths associated with land-use intensification in relevant lakes, the grey shade indicates the 1866 Tarawera tephra layer in Lake Rotoehu.

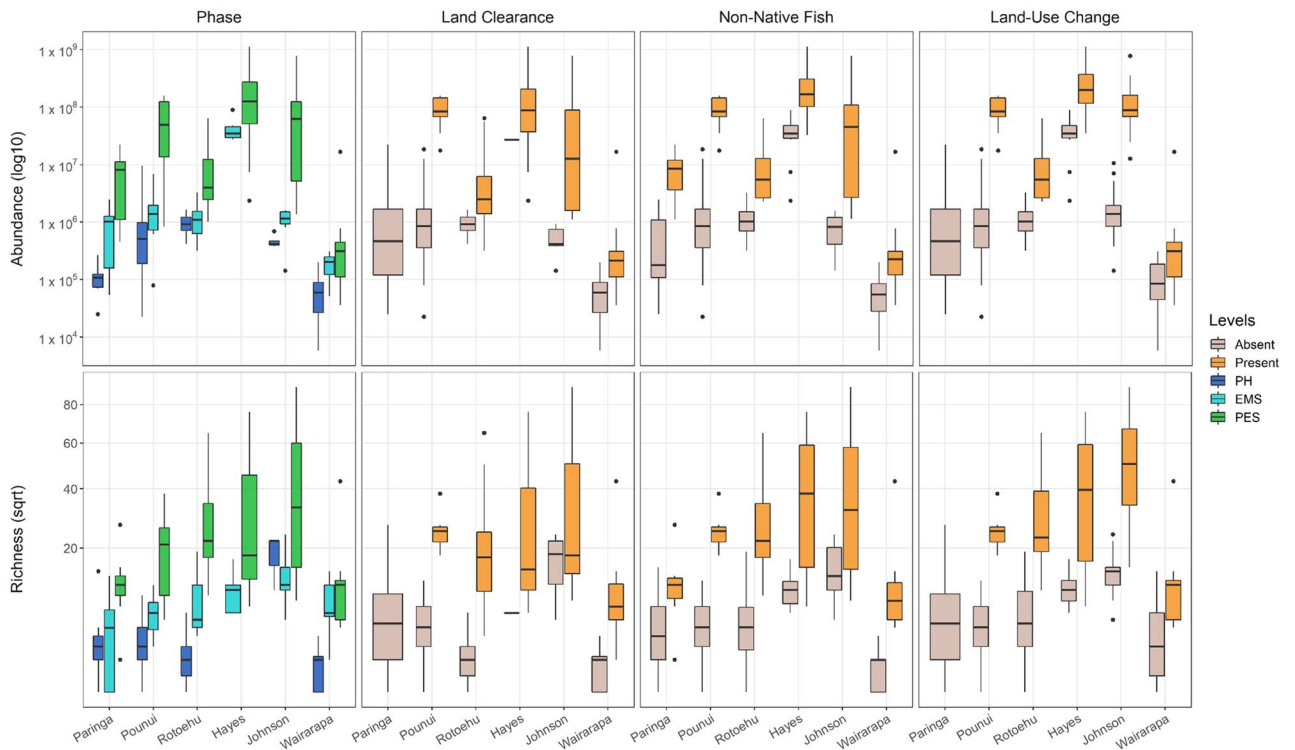


Figure 4. Differences in each response (cyanobacteria abundance or richness) depending on the lake and predictor variable. Cyanobacteria abundance was measured by droplet digital PCR (ddPCR–16S rRNA gene copy numbers/gram of dry sediment, log₁₀ y-axis scale) while cyanobacteria richness was calculated from metabarcoding data (number of photosynthetic cyanobacteria Amplicon Sequence Variants per sample, square-root y-axis scale). PH = Pre-Human, EMS = Evidence of Māori Settlement, PES = Post European Settlement.

| Predictors | Estimates | log(Abundance) CI | p | Estimates | sqrt(Richness) CI | p |
|----------------|-----------|----------------------|--------------|-----------|----------------------|--------------|
| (Intercept) | 12.91 | 12.32–13.49 | <0.001 | 1.98 | 1.29–2.67 | <0.001 |
| Phase PES | 1.64 | 1.04–2.24 | <0.001 | 1.04 | 0.39–1.69 | 0.002 |
| Phase PH | –1.13 | –1.75––0.51 | <0.001 | –1.05 | –1.77––0.32 | 0.005 |
| Lake Pounui | 4.28 | 3.32–5.25 | <0.001 | 0.62 | –0.49–1.73 | 0.27 |
| Lake Rotoehu | 1.62 | 0.68–2.57 | 0.001 | 0.90 | –0.21–2.01 | 0.11 |
| Lake Johnson | 1.44 | 0.60–2.28 | 0.001 | 2.57 | 1.27–3.87 | <0.001 |
| Lake Hayes | 1.13 | 0.29–1.97 | 0.009 | 1.76 | 0.47–3.05 | 0.008 |
| Lake Wairarapa | –1.08 | –1.86––0.30 | 0.007 | 0.21 | –0.79–1.20 | 0.68 |
| Observations | | 209 | | | 206 | |

Table 3. Results of the generalised least squares (GLS) models for cyanobacteria abundance (from droplet digital PCR data) and richness (from metabarcoding data). PH = pre-human, EMS = evidence of Māori settlement, PES post European settlement. PES and PH were compared to EMS since the Lake Hayes core does not have a PH phase, while all lakes were compared to Lake Paringa (most pristine lake). Significant p values ($p < 0.05$) are shown in bold.

Cyanobacterial abundance (16S rRNA gene copy numbers per gram of dry sediment) also increased in all lakes during recent times compared to historical levels (Fig. 3). These increases were particularly prominent in Lakes Johnson and Pounui (up to ~6,000 times their historical levels), medium in Lakes Hayes, Rotoehu, and Paringa (around 500 times their historical levels), and a slight increase in Lake Wairarapa, which was mostly in the surface sample. Timing of shifts in cyanobacterial abundance were different for all lakes: Lake Paringa suddenly shifted in the mid 1900's, Lake Pounui underwent a steady though steep increase post European arrival, while Lake Hayes and Johnson steadily increased with land-use intensification. Increase in abundance were more recent for Lakes Rotoehu and Wairarapa (top of the cores). Lastly, though the sediment core of Lake Hayes did not include the PH phase, its levels of cyanobacterial abundances during EMS were up to 1000 times higher than in the other lakes at the same time (Figs. 3 and 4).

When all lakes were combined, richness was positively correlated with 16S rRNA cyanobacterial gene copy numbers ($r = 0.74$, $p < 0.001$). Relationships were also significant when each lake was analysed individually (from $R = 0.49$ for Lake Wairarapa to $R = 0.82$ for Lake Pounui, all $p < 0.01$; Supplementary Fig. S10).

Drivers of shifts in cyanobacterial community. Both cyanobacteria abundances and photosynthetic richness tended to increase with land clearance, land-use change, and the introduction of non-native (carnivorous) fish (Fig. 4). Generalised least squares (GLS) models only incorporated Phase and Lake since the other predictors were highly correlated. These models showed a significant effect of Phase on cyanobacteria abundance and richness ($p < 0.01$); compared to EMS, richness and abundance were significantly higher during PES and lower during PH (Table 3, Fig. 4). When comparing all lakes to Lake Paringa (the most pristine lake), GLS models revealed that all lakes had a significantly higher abundance than Lake Paringa apart from Lake Wairarapa which had a significantly lower abundance (Table 3, Fig. 4). Only Lakes Johnson and Hayes had a significantly higher richness (Table 3). When comparing lakes to one another with pairwise comparisons, the models revealed that overall, Lakes Hayes, Johnson and Pounui had a significantly higher 16S rRNA gene copy abundance than Lakes Paringa and Wairarapa (Supplementary Table S8), while Lake Johnson had a significantly higher richness than Lakes Paringa and Wairarapa (Supplementary Table S9).

Historical and spatial variability in cyanobacteria community structure. Photosynthetic cyanobacterial reads were selected from the unrarefied dataset for the remainder of the analyses ($n = 190$ samples). *Cyanobium*, a picocyanobacterial genus, dominated the cyanobacterial community in almost all lakes and phases (Figs. 5 and 6). Other picocyanobacteria genera, such as *Synechocystis*, were present intermittently in low abundance in Lakes Pounui, Hayes, and Johnson (mainly PES). The main non-picoplanktonic genera were the potential toxigenic *Aphanizomenon*, *Microcystis*, and *Dolichospermum* and these became dominant in Lakes Johnson and Pounui in the PES phase (Figs. 5 and 6). These were found in at least one sample in most lakes. *Aphanizomenon* and *Microcystis* were found in all lakes and mostly in the PES phase; *Dolichospermum* was restricted to Lakes Pounui (PES), Rotoehu (PES), Johnson (PES) and Paringa (PH), although relative abundances differed greatly among lakes. Other relatively common pelagic genera were only found in specific lakes: *Geminocystis* in Lake Hayes (PES), Johnson (PES and PH), and Rotoehu (PES), *Tychonema* in all lakes except Lake Pounui. Some benthic taxa were also identified, such as *Phormidium* and *Geitlerinema* in Lake Wairarapa, *Calothrix* in one sample from Lake Rotoehu (EMS), *Nodosilinea* in Lakes Paringa (EMS + PH, one sample each), Johnson (PES), and Wairarapa (PES, one sample), and *Oscillatoria* in Lake Paringa (EMS). Some potentially interesting taxa (high enough relative abundance) could unfortunately not be identified down to genus level, especially in Lakes Pounui and Hayes (Unknown Cyanobacteriales, Microcystaceae, and Synechococcaceae, Fig. 5). Overall, toxic taxa were present in all lakes prior to human arrival, however their relative abundance increased in Lakes Hayes, Johnson, Pounui, and Rotoehu in the PES phase. They became continuously occurring from European arrival in Lakes Pounui and Johnson, and shortly after land-use intensification in Lakes Rotoehu and Hayes. In general, picocyanobacteria relative abundance remained constant, although slight decreases were observed

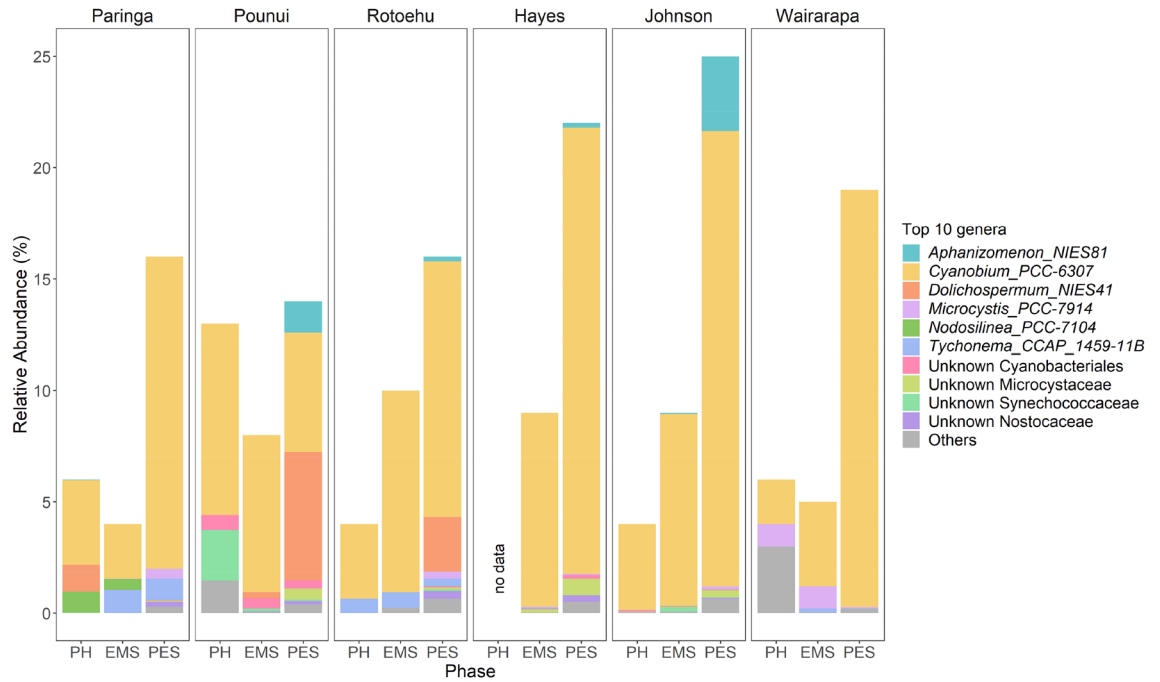


Figure 5. Relative abundance of the top 10 most abundant photosynthetic cyanobacterial genera across all lakes, less abundant genera are grouped as “Others”. All samples were grouped by Phase within each lake. PH = Pre-Human; EMS = Evidence of Māori Settlement; PES = Post-European Settlement. The sediment core from Lake Hayes does not include any PH phase.

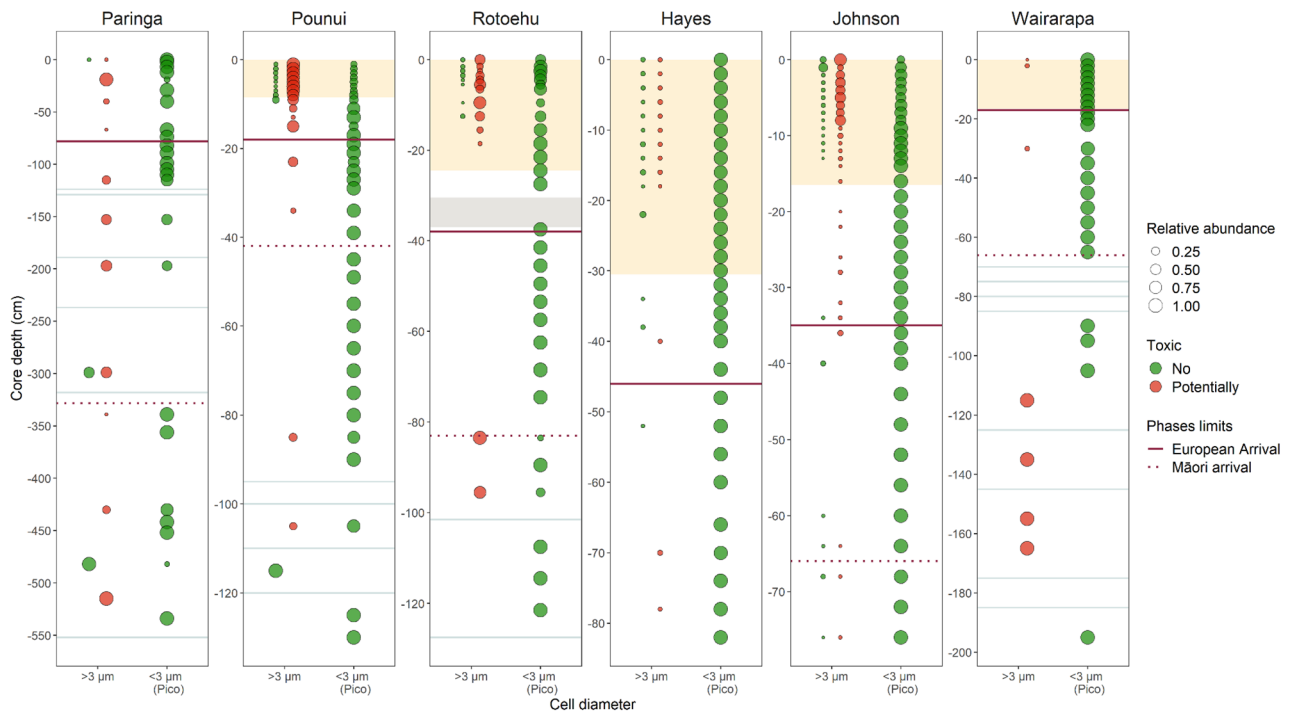


Figure 6. Relative abundance profiles of photosynthetic cyanobacterial genera classified by size (>3 μm or <3 μm—picocyanobacteria) and potential toxin production. Relative abundance is plotted as a continuous variable and the legend gives an indication of size for a set of abundances. Potential for toxin production of each genus was inferred from the literature. Phase delimitation is indicated by horizontal purple straight and purple dotted lines (see legend). Horizontal grey lines indicate samples with no cyanobacteria sequences after rarefaction. Yellow shades indicate land-use intensification in relevant lakes. Grey shade in the Lake Rotoehu core indicates the Tarawera tephra (1886).

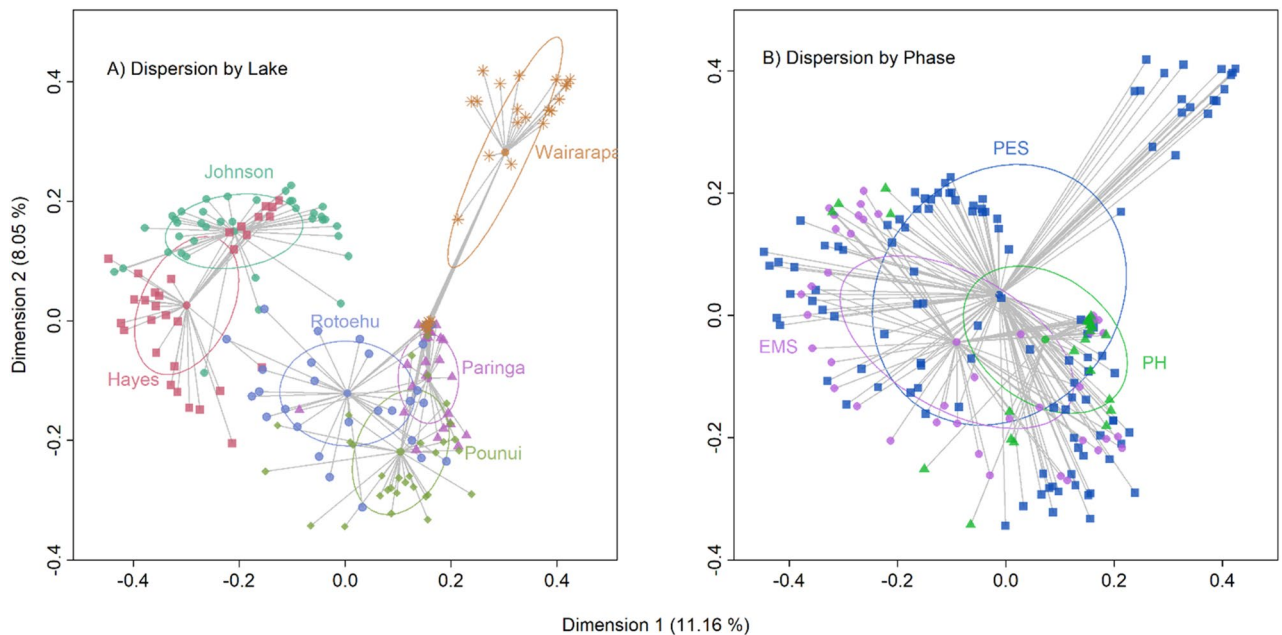


Figure 7. Principal coordinates analysis (PCoA) ordination plots of cyanobacterial communities based on Bray–Curtis dissimilarities for all six study lakes (A) and all phases (B). The ellipses indicate the mean distance from centroids, whereas the vectors show the distance to the centroid of each group.

in Lakes Johnson and Pounui as potentially toxic taxa became dominant (Fig. 6, Supplementary Fig. S11). The CONISS analysis showed that timings of shifts in the community composition varied depending on the lakes, though the main shift seemed to coincide with land-use intensification for Lakes Rotoehu, Hayes, and Johnson (Supplementary Fig. S11).

The PERMANOVA test showed that there were significant differences in community structure across lakes ($F_{189,5} = 11.726$, $p < 0.01$), and pairwise comparisons revealed that all lakes were significantly different from one another ($p < 0.01$). The test for homogeneity of multivariate dispersions showed that these differences could be due to differences in the mean distance from lake centroids among lakes ($F_{184,5} = 7.5613$, $p < 0.001$). The PCoA ordination revealed three distinct clusters: Hayes and Johnson; Pounui, Paringa and Rotoehu; and Wairarapa by itself (Fig. 7A). Only Lake Hayes's dispersion was significantly different from Lakes Rotoehu, Paringa, and Pounui (Tukey's post-hoc test on homogeneity of multivariate dispersion, $p < 0.05$).

When data from all lakes was combined, there was no clear evidence of a differentiation of cyanobacterial assemblages by phase (Fig. 7B). When each lake was analysed individually, phases tended to be quite distinct and significantly different from one another (Supplementary Fig. S12, Supplementary Tables S10 and S11). All phases were different from one another in Lakes Pounui, Rotoehu, Hayes and Johnson; the PH phase was not different from EMS only in Lakes Paringa and Wairarapa (Supplementary Table S10). While some of the results may be affected by heterogeneous dispersion (Supplementary Table S11), the data indicate that the cyanobacterial community changed Post-European settlement (PES vs. EMS) in all lakes except Lake Pounui and Lake Rotoehu (Supplementary Fig. S12, Supplementary Tables S10 and S11).

Discussion

General patterns. We inferred a large diversity of cyanobacteria present within our study lakes. Interestingly, many ASVs were identified as picocyanobacteria (small—usually less than 3 μm , unicellular cyanobacteria), and in particular *Cyanobium*, regardless of lake or phase. It is unclear whether their high numbers are due to; (1) primer biases (i.e. the primers preferentially bind to these taxa and/or these taxa have higher 16S rRNA copies in the genome compared to other cyanobacteria), (2) more variability in their 16S rRNA sequences (which resulted in our bioinformatics pipeline separating them as different ASVs), (3) better cell or DNA preservation, or (4) if they are actually present in high concentrations in the environment. It is also possible that some cyanobacteria sequences are wrongly assigned as *Cyanobium* in the SILVA database, leading to an over-representation of this genus in our results. Given their small size and difficulties with identifying them, it is also highly likely they are commonly overlooked in traditional light microscopy analysis of lake samples⁸⁰. A study of cyanobacterial diversity in 143 New Zealand lakes identified an equally high number of operational taxonomic units and associated sequence reads of picocyanobacteria and suggested that this group has been overlooked in most limnological studies in this country⁸¹. Other recent ecology studies^{82,83} have highlighted similarly diverse freshwater autotrophic picoplankton communities and suggested a need for further characterisation with particular regard to their response to environmental change. Although relative abundances of picocyanobacteria remained relatively constant across phases, there were differences in picocyanobacteria structure among lakes and between phases at the ASV level. Because so little is known about the ecology of these taxa, it is not feasible

to make ecological inferences based on these shifts. Enhancing knowledge on their ecology would greatly assist in understanding the observed changes and may allow these taxa to be used as early indicators of change in lake ecosystems.

Potentially toxic genera and/or well-known bloom-forming species (e.g., *Microcystis*, *Dolichospermum*, *Aphanizomenon*, or *Phormidium*) were present in all lakes including before human arrival (except Lake Hayes, no pre-human period reconstructed). This corroborated previous findings that potentially toxic cyanobacteria have been present historically in lakes, albeit in low abundance or sporadically^{42,84}. However, while bloom-forming species were occasionally common in some samples in pre-human times according to the metabarcoding data, this does not necessarily mean they were actually abundant in the lake. Rather, this could be an artifact due to few cyanobacteria sequences in the samples, since metabarcoding relies on relative abundance. This is reinforced by the ddPCR data which does not indicate high abundance of cyanobacteria in the pre-human period. Furthermore, shifts to dominance by potential bloom-forming species were only observed in recent times in lakes that are presently eutrophic (Pounui, Rotoehu, Hayes, Johnson), a trend which was reflected by an increase in total cyanobacterial abundance in the same lakes. Lake Wairarapa was the only exception, having general low levels of cyanobacteria despite its current supereutrophic status.

Some non-photosynthetic cyanobacteria (Vampirovibrionia Class) were identified by the metabarcoding data. They mostly occurred in the top half of the cores, and could generally only be identified at Order level. It is likely that they were also detected by the ddPCR assay, and they may account for some of the observed increases in abundance and richness. Further research is needed to enhance knowledge on their ecological role.

Lake-specific cyanobacteria histories. Total cyanobacterial abundance and richness significantly increased in all six study lakes over the 500–1500 years reconstructed (depending on the sediment core). However, it was most evident after the arrival of Europeans in Lakes Hayes, Johnson, Pounui, and Rotoehu. Although the GLS models could not specifically test the relationship between land-use and cyanobacteria response, the exploration of the data suggests that both abundance and richness increased significantly post European settlement, particularly for Lakes Hayes and Johnson. Palynology data from the cores used in this study shows that there was a marked decrease in native forests in the catchments of Lakes Hayes, Johnson, and Rotoehu, coinciding with the periods of increasing cyanobacterial abundance.

All lakes presented in this study have some historical environmental/social records, though the degree of knowledge varies across lakes (summarised in Supplementary Table S12). In Lake Hayes, cyanobacteria abundance has increased steadily since European settlement (1862⁸⁵) and especially since the early 1900s with the conversion of much of the lower catchment into sheep pasture. Furthermore, superphosphate fertiliser application began in the 1950s; by 1961 there had been significant drainage of the wetlands in the catchment, and by the 1970s Lake Hayes was considered eutrophic. This increase in trophic state aligned with the first reports of cyanobacteria blooms (1970–1972, *Dolichospermum flos-aquae*⁴⁸), and our own observations of potentially toxic cyanobacteria in the second half of the Post-European Settlement phase (PES, top 20 cm of the core). Similarly, the catchment of Lake Johnson, near Lake Hayes, was spared from urban development, but forest clearing from c. 1895 led to pasture development and application of superphosphate fertiliser from c. 1955⁴⁸. Subsequently, potentially toxic taxa appear continuously in the sediment core with European settlement (~45 cm), and in higher proportion in the top 8 cm. Cyanobacterial richness and abundance both increased with land-use intensification, which may be related to superphosphate application (1955). Lake Rotoehu, the northernmost lake of this study, receives phosphorus-rich waters from geothermal input(s), has numerous shallow arms and generally quite warm waters (min 8 °C, max 22 °C)⁸⁶. Its name in Te Reo Māori translates as “murky/cloudy (ehu) lake (roto)”, which could suggest that phytoplankton biomass might already have been high when Māori were living around the lake. However, the data in the present study indicate that if this was the case, it was likely not due to cyanobacteria, as historic levels were relatively low. The first cyanobacterial blooms were observed in the 1960s, coinciding with the first changes in land-use (forest and scrubs converted to pasture)^{86,87}. In the 1960–70s, Lake Rotoehu was considered mesotrophic but a marked decrease (4.2 m) in lake level in 1993 is thought to have increased nutrient levels and possibly enhanced cyanobacterial blooms, which have occurred on an annual basis since then⁸⁸. While we do not have any dates for the core of Lake Rotoehu apart from the tephra (1886, 30.5 to 37 cm), potential bloom-forming species also appear continuously from 18.5 cm to the top.

The results from these three lakes are in agreement with previous studies that have indicated a link between land use change, increased nutrient runoff, and cyanobacterial blooms and associated toxins. For example, in Baptiste Lake (Alberta, Canada) an increase in microcystin gene copies was correlated with nutrient enrichment⁴². Similarly, increases in agriculture in the surrounding landscape of Anderson Lake (Washington, USA) were linked with the first appearance of toxic *Dolichospermum* blooms⁸⁹, and saxitoxin gene copies increased in Laguna Blanca (Maldonado, Uruguay) after forestry developed in its catchment⁹⁰. Further work is required to fully identify the drivers of change as the definition of land-use intensification was categorical (presence/absence) rather than a quantitative measure of change in the catchment and was highly colinear with non-native fish introduction. The analysis would be enhanced by including more lakes and more variables that were not available for all the lakes studied.

In contrast to the previous three lakes, Lake Pounui provides an interesting anomaly to the proposed link between land-use change and the development of cyanobacterial blooms. Palynological data⁴⁷ and historic knowledge⁹¹ suggest that land clearance only occurred within a part of the catchment and mostly stopped after a few decades. Aerial photographs show that land clearing in the southern side of the catchment had already occurred in 1941⁹¹, but the native forest had mostly regenerated by 1961. A possible link to the severity of cyanobacteria blooms in this lake is the introduction of rainbow trout (*Oncorhynchus mykiss*) and the European perch

(*Perca fluviatilis*) into Lake Pounui and associated streams and lakes in the wider catchment, from 1938 and the 1960s, respectively (Fish and Game NZ, historic records; Rawhiri Smith, pers. comm.).

Non-native fish can disrupt native food webs^{92,93} and have a top-down impact on cyanobacteria⁹⁴. Pelagic food webs are quite simple in New Zealand, as they are lacking strongly piscivorous fish^{95,96}. Voracious zooplanktivorous fish such as juvenile perch are absent from the indigenous fauna, and when introduced, they can effectively shift zooplankton communities so that smaller species (e.g. rotifers) become dominant, thus reducing grazing pressure on phytoplankton⁹⁷. *Perca fluviatilis* has few effective predators in New Zealand, albeit they are cannibalistic, and, as such, their populations are often stunted and mostly composed of juveniles, therefore enhancing zooplanktivory and reducing grazing on phytoplankton⁹⁸. Brown trout (*Salmo trutta*) and perch were introduced to Lake Hayes in the 1870s, while for Lake Johnson, perch was introduced in the 1880s and rainbow trout have been stocked annually since 1962. Lake Rotoehu also contains rainbow trout, which were introduced between 1898 and 1903⁵¹. Only a handful of studies have explored the trophic effects of perch (*P. fluviatilis*) introductions in New Zealand^{99,100}, and more detailed food-web related research is required to fully understand the wider impacts of the introduction of these species on cyanobacterial blooms in these lakes.

Analysis of the metabarcoding data was required to further understand the recent increases in cyanobacterial abundance in Lakes Paringa and Wairarapa. The vegetation in the catchment of Lake Paringa is relatively unmodified⁵⁶, therefore we anticipated there would be no change in cyanobacterial abundance at this site. Conversely, the catchment of Lake Wairarapa has been highly modified^{50,57} but only very sporadic and localised cyanobacterial blooms have been reported (e.g. *Dolichospermum lemmermannii* in 2008⁵⁰). Lake Wairarapa is a very shallow lake with constant sediment resuspension; therefore any labile material could be thoroughly degraded before it settled down on the lakebed. This might have an impact on sedaDNA recovery and analysis. The metabarcoding data indicated that total increase in cyanobacteria copy numbers in these two lakes was due to picocyanobacteria. Picocyanobacteria are abundant in the oceans and often in oligotrophic lakes, but more recently they have also been reported in high numbers in eutrophic lakes, particularly coastal lakes and lagoons^{101,102}. Picocyanobacteria are sensitive to high light intensities¹⁰³, and the low-light conditions in both Lakes Paringa and Wairarapa (respectively brown waters due to tannins, and turbid waters from sediment resuspension) could potentially explain their abundance and dominance in these lakes. The increase in abundance in picocyanobacteria in Lake Paringa occurred in the 1960s. This coincides with the introduction of *S. trutta* (brown trout) in the 1950s, and the construction of a road in 1958 along the south-eastern side of the lake, thus enhancing human access to the lake. The dramatic increase in picocyanobacteria communities in the top 3 cm of the core in Wairarapa could suggest this is a recent change, possibly representing a further decline in the health of this lake.

Other potential drivers of change. In addition to the impact of land-use change, and non-native fish introduction, many other factors impacted these lakes. For example, Lake Wairarapa have been significantly affected by hydrological changes, such as changing the main inflow away from the lake in the 1960s. Earthquakes can also have significant impacts on lakes, potentially triggering landslides, especially in mountainous terrain^{104,105}. Landslides increase sediment yields directly in the lake, and tributaries flowing in to the lakes. The Alpine fault line near Lake Paringa has a magnitude 8 earthquake approximately every 300 years⁵⁶, with at least four earthquakes occurring during the time period covered by the sediment core. Based on the results in this study, these earthquakes do not always appear to align with major shifts in cyanobacterial abundance or composition in Lake Paringa. One exception seems to be the 1717 Alpine fault earthquake, which is followed by a slight increase in cyanobacteria abundance. In 1855, a major earthquake¹⁰⁶ (est. M 8.2) struck the Wairarapa region where Lake Wairarapa and Lake Pounui are situated, as well as smaller one in 1942¹⁰⁷ (max M 7.2 and 6.8). In Lake Pounui, these events were concomitant with the discrete appearance of *Synechococcus* after these two events, and post 1855 there was a shift to *Dolichospermum* which is now dominating the cyanobacteria community. However, Europeans also started more intensive use of the land in the catchment around the same time and it is therefore not possible to correlate these changes in community structure with earthquakes specifically.

The impact of climate change was not assessed in this study because its timing overlaps with the main anthropogenic stressors studied here, making it extremely challenging to tease apart the impacts of climate change from other stressors. Further proxies which provide more detailed indications of climate change are required. Furthermore, focusing on lakes not affected by land-use, such as high-altitude alpine lakes, would yield more significant results. It is likely that climate change is enhancing the blooms. In New Zealand, average air temperatures have risen by 1.13 °C (± 0.27) from 1909 to 2019 nationally and there are more extreme, storm/rainfall events and longer periods of droughts since the beginning of the twentieth century¹⁰⁸. It is extremely unlikely that this change in temperature alone is responsible for driving the changes in cyanobacterial abundance and composition observed in this study. If it was the case, blooms would be observed in Lake Paringa despite no major catchment changes, which is not the case.

Caveats when interpreting sedimentary DNA data. Although sedaDNA approach are now increasingly used in paleolimnology, there are a number of caveats that need to be considered when interpreting the data. Many of these are well acknowledged challenges with the use of molecular approaches for analysing eDNA from a variety of sample matrices including primer biases, appropriate selection of universal (or targeted) primers and incomplete reference databases^{33,109,110}.

One issue that requires further investigation in paleolimnological studies is DNA degradation. Degradation leads to DNA fragmentation; however it is currently unknown how quickly this occurs. Studies on contemporary samples indicate that amongst others, high amounts of light (UV irradiance), high temperatures, acidic and oxic conditions contribute to DNA decay^{111–116}. A recent study suggests that sedaDNA degradation mainly occurs over the first 250 years after sediment deposition¹¹⁷, however this research was undertaken in a tropical

swamp (Uganda) and thus the environment conditions (especially temperature and pH) are very different to the lakes studied here. It is likely that degradation rates would be much slower rates in deeper, temperate lake sediments. DNA degradation in lake sediment cores needs further exploration, using methodologies such as shotgun sequencing which shows the degree of fragmentation in DNA sequences.

A puzzling observation in the present study was the higher ASV richness in all lakes in the PES phase, albeit this was most pronounced in the lakes that experienced the greatest increase in cyanobacterial abundance. This is somewhat contrary to expectations, given that cyanobacterial blooms are generally dominated by one or few species. There are several plausible explanations for this observation. Sediments core samples integrate time, they capture activity across an entire year and over multiple years of within lake biological activity. Cyanobacterial blooms primarily occur in summer months¹¹⁸, yet the sediment samples capture an annual signal, thus it is feasible that richness escalates during non-bloom periods. A study which focused on collecting regular water samples for at least a year is required to determine whether seasonal differences in cyanobacterial ASV abundance drive the observed increase in ASV richness. Additionally, in this study we took 1-cm wide depth slices for most lakes, which we estimate captures a period of 2–5 years in Lakes Hayes and Johnson, approximately 10 years in Lake Pounui, and 0.5 cm slices in Lake Paringa which captures about one year of time. During these most recent periods there have been rapid shifts in many environmental variables not measured in this study such as nutrients, chemicals, and toxins. It is feasible that the high species richness may truly represent a community that is adapting and diversifying to respond to an increasing number of new pressures. A further possibility is that DNA degradation is accounting for the reduction of richness in deeper samples. In metabarcoding studies a large proportion of ASVs are rare and have a low number of reads¹¹⁹. In this study 18 photosynthetic cyanobacterial ASVs (out of 839) accounted for 50% of the reads after rarefaction. As DNA deeper in the core is subjected to degradation pressures for a longer period, there is a likelihood that the rare ASVs are more likely to be lost compared to more recent samples, thus decreasing richness.

Conclusion

Through the study of six different lakes over ~ 1000 years, this research has provided valuable new insights into changes in cyanobacterial communities in a range of lake types in New Zealand. Although cyanobacteria were present in these lakes prior to human arrival, including potentially toxic or bloom forming species, their overall abundance was low. Some changes in cyanobacterial abundance and structure occurred following Māori settlement in New Zealand, but the most pronounced shifts happened post European settlement, concomitant with land use change. The inclusion of Lakes Pounui and Paringa, where there has been only moderate (Pounui) or minimal (Paringa) land use change in the catchment, highlights that there are multiple drivers of shifts in cyanobacterial communities. While we speculate that the introduction of non-native carnivorous fish played a role in changing cyanobacterial communities, further work exploring multiple trophic levels is required. Our results demonstrate the high utility of combining ddPCR and metabarcoding for investigating shifts in cyanobacterial abundance and structure in lakes over many centuries. Paleolimnological studies utilising sedaDNA offer the potential to track the history of cyanobacteria, and when used in combination with other paleolimnological proxies the data can provide insights into the drivers of change.

Data availability

The metabarcoding dataset generated and analysed during the current study is available in the Sequence Read Archive (SRA) repository (BioProject PRJNA855252).

Received: 4 June 2021; Accepted: 2 June 2022

Published online: 27 July 2022

References

1. Paerl, H. W. & Huisman, J. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environ. Microbiol. Rep.* **1**, 27–37 (2009).
2. Paerl, H. W. & Paul, V. J. Climate change: links to global expansion of harmful cyanobacteria. *Water Res.* **46**, 1349–1363 (2012).
3. Huisman, J. *et al.* Cyanobacterial blooms. *Nat. Rev. Microbiol.* **16**, 471 (2018).
4. Reissig, M., Trochine, C., Queimaliños, C., Balseiro, E. & Modenutti, B. Impact of fish introduction on planktonic food webs in lakes of the Patagonian Plateau. *Biol. Conserv.* **132**, 437–447 (2006).
5. Britton, J. R., Davies, G. D. & Harrod, C. Trophic interactions and consequent impacts of the invasive fish *Pseudorasbora parva* in a native aquatic foodweb: a field investigation in the UK. *Biol. Invasions* **12**, 1533–1542 (2010).
6. Beaulieu, M., Pick, F. & Gregory-Eaves, I. Nutrients and water temperature are significant predictors of cyanobacterial biomass in a 1147 lakes data set. *Limnol. Oceanogr.* **58**, 1736–1746 (2013).
7. O’Neil, J. M., Davis, T. W., Burford, M. A. & Gobler, C. J. The rise of harmful cyanobacteria blooms: the potential roles of eutrophication and climate change. *Harmful Algae* **14**, 313–334 (2012).
8. Paerl, H. W. & Huisman, J. Blooms like it hot. *Science* **320**, 57–58 (2008).
9. *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring, and Management.* (E & FN Spon, 1999).
10. Sukenik, A., Quesada, A. & Salmaso, N. Global expansion of toxic and non-toxic cyanobacteria: effect on ecosystem functioning. *Biodivers. Conserv.* **24**, 889–908 (2015).
11. Ibelings, B. W., Bormans, M., Fastner, J. & Visser, P. M. CYANOCOST special issue on cyanobacterial blooms: synopsis—a critical review of the management options for their prevention, control and mitigation. *Aquat. Ecol.* **50**, 595–605 (2016).
12. Paerl, H. W. Mitigating harmful cyanobacterial blooms in a human- and climatically-impacted world. *Life* **4**, 988–1012 (2014).
13. Rastogi, R. P., Madamwar, D. & Incharoensakdi, A. Bloom dynamics of cyanobacteria and their toxins: environmental health impacts and mitigation strategies. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2015.01254> (2015).
14. Ewing, H. A. *et al.* “New” cyanobacterial blooms are not new: two centuries of lake production are related to ice cover and land use. *Ecosphere* **11**, e03170 (2020).
15. McGlone, M. S. & Wilmshurst, J. M. Dating initial Maori environmental impact in New Zealand. *Quat. Int.* **59**, 5–16 (1999).
16. Brooking, A. P. D. of H. T. & Brooking, T. *The History of New Zealand.* (Greenwood Publishing Group, 2004).

17. Wilmshurst, J. M., Anderson, A. J., Higham, T. F. G. & Worthy, T. H. Dating the late prehistoric dispersal of Polynesians to New Zealand using the commensal Pacific rat. *PNAS* **105**, 7676–7680 (2008).
18. McGlone, M. S. The Polynesian settlement of New Zealand in relation to environmental and biotic changes. *N. Z. J. Ecol.* **12**, 115–129 (1989).
19. McGlone, M. S. Polynesian deforestation of New Zealand: a preliminary synthesis. *Archaeol. Ocean.* **18**, 11–25 (1983).
20. McWethy, D. B. *et al.* Rapid landscape transformation in South Island, New Zealand, following initial Polynesian settlement. *PNAS* **107**, 21343–21348 (2010).
21. McWethy, D. B., Wilmshurst, J. M., Whitlock, C., Wood, J. R. & McGlone, M. S. A high-resolution chronology of rapid forest transitions following Polynesian arrival in New Zealand. *PLoS ONE* **9**, e111328 (2014).
22. Star, P. New Zealand environmental history: a question of attitudes. *Environ. Hist. Camb.* **9**, 463–475 (2003).
23. Clark, A. H. *The Invasion of New Zealand by Plants, People, and Animals* (Rutgers University Press, 1949).
24. Wilmshurst, J. M. Human effects on the environment: European impact. *Te Ara: The Encyclopedia of New Zealand* <https://teara.govt.nz/en/human-effects-on-the-environment/page-3> (2007).
25. Smol, J. P. The ratio of diatom frustules to chrysophycean statospores: a useful paleolimnological index. *Hydrobiologia* **123**, 199–208 (1985).
26. Rees, A. B. H., Cwynar, L. C. & Cranston, P. S. Midges (Chironomidae, Ceratopogonidae, Chaoboridae) as a temperature proxy: a training set from Tasmania, Australia. *J. Paleolimnol.* **40**, 1159–1178 (2008).
27. Epp, L. S., Stoof, K. R., Trauth, M. H. & Tiedemann, R. Historical changes in a sediment core from a Kenyan lake: intraspecific genotype turnover in a tropical rotifer is related to past environmental changes. *J. Paleolimnol.* **43**, 939–954 (2010).
28. Buchaca, T. *et al.* Rapid ecological shift following piscivorous fish introduction to increasingly eutrophic and warmer Lake Furnas (Azores Archipelago, Portugal): a paleoecological approach. *Ecosystems* **14**, 458–477 (2011).
29. Cristescu, M. E. & Hebert, P. D. N. Uses and misuses of environmental DNA in biodiversity science and conservation. *Annu. Rev. Ecol. Syst.* **49**, 209–230 (2018).
30. Giguet-Covex, C. *et al.* Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nat. Commun.* **5**, 3211 (2014).
31. Alsos, I. G. *et al.* Plant DNA metabarcoding of lake sediments: how does it represent the contemporary vegetation. *PLoS ONE* **13**, e0195403 (2018).
32. Nelson-Chorney, H. T. *et al.* Environmental DNA in lake sediment reveals biogeography of native genetic diversity. *Front. Ecol. Environ.* **17**, 313–318 (2019).
33. Capo, E. *et al.* Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: overview and recommendations. *Quaternary* **4**, 6 (2021).
34. Shokralla, S., Spall, J. L., Gibson, J. F. & Hajibabaei, M. Next-generation sequencing technologies for environmental DNA research. *Mol. Ecol.* **21**, 1794–1805 (2012).
35. Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* **21**, 2045–2050 (2012).
36. Thomsen, P. F. & Willerslev, E. Environmental DNA: an emerging tool in conservation for monitoring past and present biodiversity. *Biol. Conserv.* **183**, 4–18 (2015).
37. Keeley, N., Wood, S. A. & Pochon, X. Development and preliminary validation of a multi-trophic metabarcoding biotic index for monitoring benthic organic enrichment. *Ecol. Ind.* **85**, 1044–1057 (2018).
38. 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.* **82**, 6472–6482 (2016).
39. Pal, S., Gregory-Eaves, I. & Pick, F. R. Temporal trends in cyanobacteria revealed through DNA and pigment analyses of temperate lake sediment cores. *J. Paleolimnol.* **54**, 87–101 (2015).
40. Dodsworth, W. Temporal Trends in Cyanobacteria Through Paleo-Genetic Analyses. (Université d'Ottawa/University of Ottawa, 2020). <https://doi.org/10.20381/ruor-24401>.
41. Rinta-Kanto, J. M. *et al.* The diversity and distribution of toxigenic *Microcystis* spp. in present day and archived pelagic and sediment samples from Lake Erie. *Harmful Algae* **8**, 385–394 (2009).
42. Zastepa, A. *et al.* Reconstructing a long-term record of microcystins from the analysis of lake sediments. *Sci. Total Environ.* **579**, 893–901 (2017).
43. Schallenberg, M. *et al.* Ecosystem services of lakes. In *Ecosystem Services in New Zealand* (ed. Dymond, J.) 23 (Manaaki Whenua Press, 2013).
44. Ministry for the Environment & Ausseil, A.-G. *Our freshwater 2020*. www.mfe.govt.nz (2020).
45. Takiwa—Map Page. <https://lernz.takiwa.co/map>.
46. Leathwick, J. *et al.* Freshwater ecosystems of New Zealand (FENZ) geodatabase. Users guide. (2010).
47. Cochrane, L. Reconstructing Ecological Change, Catchment Disturbance, and Anthropogenic Impact over the last 3000 years at Lake Pounui, Wairarapa, New Zealand. (2017).
48. Burns, C. W. & Mitchell, S. F. Seasonal succession and vertical distribution of phytoplankton in Lake Hayes and Lake Johnson, South Island, New Zealand. *N. Z. J. Mar. Freshw. Res.* **8**, 167–209 (1974).
49. Lawa. *Land, Air, Water Aotearoa (LAWA)* <https://www.lawa.org.nz/> (2018).
50. Bunny, T., Perrie, A., Milne, J. & Keenan, L. *Lake water quality in the Ruamāhanga Whaitua*. 17 (2014).
51. McKinnon, M. Volcanic Plateau region: The lure of trout. *Te Ara—The Encyclopedia of New Zealand* <https://teara.govt.nz/en/volcanic-plateau-region/page-8> (2015).
52. Burns, C. W. & Mitchell, S. F. Seasonal succession and vertical distribution of zooplankton in Lake Hayes and Lake Johnson. *N. Z. J. Mar. Freshw. Res.* **14**, 189–204 (1980).
53. Schallenberg, M. & Schallenberg, L. *Lake Hayes restoration and monitoring plan*. 55 https://a234f952-dbf2-444e-983e-ef311d984ee7.filesusr.com/ugd/c1b10b_d2993ed023cd4bdbac7eef71a89c2de7.pdf (2017).
54. NIWA. NIWA <https://niwa.co.nz/>.
55. Mackereth, F. J. H. A portable core sampler for lake deposits. *Limnol. Oceanogr.* **3**, 181–191 (1958).
56. Howarth, J. D., Fitzsimons, S. J., Norris, R. J. & Jacobsen, G. E. Lake sediments record cycles of sediment flux driven by large earthquakes on the Alpine fault, New Zealand. *Geology* **40**, 1091–1094 (2012).
57. Trodahl, M. L., Rees, A. B. H., Newnham, R. M. & Vandergoes, M. J. Late Holocene geomorphic history of Lake Wairarapa, North Island, New Zealand. *N. Z. J. Geol. Geophys.* **59**, 330–340 (2016).
58. Khan, S., Puddick, J., Burns, C. W., Closs, G. & Schallenberg, M. Palaeolimnological evaluation of historical nutrient and food web contributions to the eutrophication of two monomictic lakes. *Submitted for Journal Publication* (2022).
59. Rinta-Kanto, J. M. *et al.* Quantification of toxic *Microcystis* spp. during the 2003 and 2004 blooms in Western Lake Erie using quantitative real-time PCR. *Environ. Sci. Technol.* **39**, 4198–4205 (2005).
60. Nübel, U., Garcia-Pichel, F. & Muyzer, G. PCR primers to amplify 16S rRNA genes from cyanobacteria. *Appl. Environ. Microbiol.* **63**, 3327–3332 (1997).
61. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2020).
62. RStudio Team. *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA. (2020).
63. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Sour. Softw.* **4**, 1686 (2019).

64. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2016).
65. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
66. Callahan, B. J. *et al.* DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581 (2016).
67. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
68. Yilmaz, P. *et al.* The SILVA and “all-species living tree project (LTP)” taxonomic frameworks. *Nucleic Acids Res.* **42**, D643–D648 (2013).
69. Glöckner, F. O. *et al.* 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J. Biotechnol.* **261**, 169–176 (2017).
70. McMurdie, P. J. & Holmes, S. phyloseq: an R Package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
71. Oksanen, J. *et al.* *vegan: Community Ecology Package. R package version 2.5–6.* 2019. (2019).
72. Williams, P. A. & Cameron, E. K. Creating gardens: the diversity and progression of European plant introductions. In *Biological Invasions in New Zealand* Vol. 186 (eds Allen, R. B. & Lee, W. G.) 33–47 (Springer-Verlag, 2006).
73. Simpson, G. L. Modelling palaeoecological time series using generalised additive models. *Front. Ecol. Evol.* <https://doi.org/10.3389/fevo.2018.00149> (2018).
74. Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinform.* **12**, 35 (2011).
75. Juggins, S. *rioja: analysis of quaternary science data.* (2020).
76. de Vries, A. & Ripley, B. D. *ggdendro: create dendrograms and tree diagrams using 'ggplot2'.* (2022).
77. Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32–46 (2001).
78. Anderson, M. J. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **62**, 245–253 (2006).
79. Soo, R. M. *et al.* An expanded genomic representation of the phylum Cyanobacteria. *Genome Biol. Evol.* **6**, 1031–1045 (2014).
80. MacKeigan, P. W. *et al.* Comparing microscopy and DNA metabarcoding techniques for identifying cyanobacteria assemblages across hundreds of lakes. *Harmful Algae* **113**, 102187 (2022).
81. Wood, S. A. *et al.* Trophic state and geographic gradients influence planktonic cyanobacterial diversity and distribution in New Zealand lakes. *FEMS Microbiol. Ecol.* <https://doi.org/10.1093/femsec/fw234> (2017).
82. Becker, S., Richl, P. & Ernst, A. Seasonal and habitat-related distribution pattern of *Synechococcus* genotypes in Lake Constance. *FEMS Microbiol. Ecol.* **62**, 64–77 (2007).
83. Sánchez-Baracaldo, P., Handley, B. A. & Hayes, P. K. Picocyanobacterial community structure of freshwater lakes and the Baltic Sea revealed by phylogenetic analyses and clade-specific quantitative PCR. *Microbiology (Reading)* **154**, 3347–3357 (2008).
84. Pilon, S. *et al.* Contrasting histories of microcystin-producing cyanobacteria in two temperate lakes as inferred from quantitative sediment DNA analyses. *Lake Reserv. Manag.* **35**, 102–117 (2019).
85. Queenstown's Pioneering Beginnings. <https://www.queenstownnz.co.nz/stories/post/queenstowns-pioneer-beginnings/> (2017).
86. Fish, G. R. A limnological study of four lakes near Rotorua. *N. Z. J. Mar. Freshw. Res.* **4**, 165–194 (1970).
87. Lake Rotoehu—Lakes Water Quality Society. <https://lakeswaterquality.co.nz/lake-rotoehu/>.
88. Bay of Plenty Regional Council, Rotorua District Council, & Te Arawa Lakes Trust. *Lake Rotoehu Action Plan.* 61 <http://www.rotorualakes.co.nz/vdb/document/76> (2007).
89. Hobbs, W. O. *et al.* Using a lake sediment record to infer the long-term history of cyanobacteria and the recent rise of an anatoxin producing *Dolichospermum* sp.. *Harmful Algae* **101**, 101971 (2021).
90. de la Escalera, G. M., Antoniadis, D., Bonilla, S. & Piccini, C. Application of ancient DNA to the reconstruction of past microbial assemblages and for the detection of toxic cyanobacteria in subtropical freshwater ecosystems. *Mol. Ecol.* **23**, 5791–5802 (2014).
91. Retrolens—Historical Imagery Resource. <https://retrolens.co.nz/>.
92. Strayer, D. L. Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. *Freshw. Biol.* **55**, 152–174 (2010).
93. Hall, S. R. & Mills, E. L. Exotic species in large lakes of the world. *Aquat. Ecosyst. Health Manag.* **3**, 105–135 (2000).
94. Gehrke, P. C. & Harris, J. H. The role of fish in cyanobacterial blooms in Australia. *Mar. Freshw. Res.* **45**, 905–915 (1994).
95. Burns, C. W. & Schallenberg, M. Impacts of nutrients and zooplankton on the microbial food web of an ultra-oligotrophic lake. *J. Plankton Res.* **20**, 1501–1525 (1998).
96. Rowe, D. K. & Schallenberg, M. Food webs in lakes. In *Freshwaters of New Zealand* (ed. Harding, J. S.) **23** (Wellington, N.Z.: New Zealand Hydrological Society, 2004).
97. Gliwicz, Z. M. & Pijanowska, J. The role of predation in zooplankton succession. In *Plankton Ecology: Succession in Plankton Communities* (ed. Sommer, U.) 253–296 (Springer, 1989).
98. Vanni, M. J. & Findlay, D. L. Trophic cascades and phytoplankton community structure. *Ecology* **71**, 921–937 (1990).
99. Smith, K. F. & Lester, P. J. Trophic interactions promote dominance by cyanobacteria (*Anabaena* spp.) in the pelagic zone of lower Karori reservoir, Wellington, New Zealand. *N. Z. J. Mar. Freshw. Res.* **41**, 143–155 (2007).
100. Smith, K. F. & Lester, P. J. Cyanobacterial blooms appear to be driven by top-down rather than bottom-up effects in the Lower Karori Reservoir (Wellington, New Zealand). *N. Z. J. Mar. Freshw. Res.* **40**, 53–63 (2006).
101. Caroppo, C. Ecology and biodiversity of picoplanktonic cyanobacteria in coastal and brackish environments. *Biodivers. Conserv.* **24**, 949–971 (2015).
102. Pulina, S. *et al.* Picophytoplankton seasonal dynamics and interactions with environmental variables in three Mediterranean coastal lagoons. *Estuaries Coasts* **40**, 469–478 (2017).
103. Callieri, C. Picophytoplankton in freshwater ecosystems: the importance of small-sized phototrophs. *Freshw. Rev.* **1**, 1–28 (2008).
104. Keefer, D. K. Investigating landslides caused by earthquakes: a historical review. *Surv. Geophys.* **23**, 473–510 (2002).
105. Fan, X. *et al.* Earthquake-induced chains of geologic hazards: patterns, mechanisms, and impacts. *Rev. Geophys.* **57**, 421–503 (2019).
106. Manighetti, I. *et al.* Repeated giant earthquakes on the Wairarapa fault, New Zealand, revealed by Lidar-based paleoseismology. *Sci. Rep.* **10**, 2124 (2020).
107. McSaveney, E. Historic earthquakes: the 1942 Wairarapa earthquakes. *Te Ara Encyclopedia of New Zealand* <https://teara.govt.nz/en/historic-earthquakes/page-9> (2006).
108. *New Zealand's environmental reporting series: our atmosphere and climate.* (Ministry for the Environment & Stats NZ, 2020).
109. Beng, K. C. & Corlett, R. T. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodivers. Conserv.* **29**, 2089–2121 (2020).
110. Freeland, J. R. The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA. *Genome* <https://doi.org/10.1139/gen-2016-0100> (2016).
111. Barnes, M. A. *et al.* Environmental conditions influence eDNA persistence in aquatic systems. *Environ. Sci. Technol.* **48**, 1819–1827 (2014).
112. Barnes, M. A. *et al.* Environmental conditions influence eDNA particle size distribution in aquatic systems. *Environmental DNA* <https://doi.org/10.1002/edn3.160> (2020).
113. 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.* **17**, 3939–3951 (2008).

114. Eichmiller, J. J., Best, S. E. & Sorensen, P. W. Effects of temperature and trophic state on degradation of environmental DNA in lake water. *Environ. Sci. Technol.* **50**, 1859–1867 (2016).
115. Strickler, K. M., Fremier, A. K. & Goldberg, C. S. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biol. Conserv.* **183**, 85–92 (2015).
116. Seymour, M. *et al.* Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. *Commun. Biol.* <https://doi.org/10.1038/s42003-017-0005-3> (2018).
117. Dommain, R. *et al.* The challenges of reconstructing tropical biodiversity with sedimentary ancient DNA: a 2200-year-long metagenomic record from Bwindi Impenetrable Forest, Uganda. *Front. Ecol. Evol.* <https://doi.org/10.3389/fevo.2020.00218> (2020).
118. Jöhnk, K. D. *et al.* Summer heatwaves promote blooms of harmful cyanobacteria. *Glob. Change Biol.* **14**, 495–512 (2008).
119. Sogin, M. L. *et al.* Microbial diversity in the deep sea and the underexplored “rare biosphere”. *PNAS* **103**, 12115–12120 (2006).

Acknowledgements

This research was funded by the New Zealand Ministry of Business, Innovation and Employment research program—Our lakes’ health; past, present, future (C05X1707). The authors thank Olivier Laroche (Cawthron Institute) and Frances Pick (University of Ottawa) for technical advice, and Marc Schallenberg (University of Otago) for advice on the manuscript. Ngāti Kahungunu ki Wairarapa, Rangitāne o Wairarapa, Te Arawa Lakes Trust, Ngāi Tahu, Hokonui Rūnanga, Kāti Huirapa ki Puketeraki, Te Rūnanga o Awarua, Te Rūnanga o Ōraka Aparima, Te Rūnanga o Ōtākou, Waihōpai Rūnanga, Te Rūnanga o Makaawhio, Otago Regional Council, Greater Wellington Regional Council, Environment Bay of Plenty, West Coast Regional Council, and the landowners are acknowledged for their support.

Author contributions

S.W. and M.P. conceived and designed the study. A.R., J.D.H., C.M.M., M.J.V., S.K. provided the lake sediment cores. M.P. processed the samples and M.P., J.K.P., and J.A. analysed the data. X.P., I.H., A.R., J.D.H., M.J.V., J.A., J.K.P. provided guidance during the study and assisted with draft manuscripts. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-14216-8>.

Correspondence and requests for materials should be addressed to M.P.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2022