Received: 6 July 2023 | Revised: 24 December 2023 | Accepted: 18 April 2024

DOI: 10.1111/gbi.12599

ORIGINAL ARTICLE

gebiology WILEY

A 1 Ma sedimentary ancient DNA (*sed*aDNA) record of catchment vegetation changes and the developmental history of tropical Lake Towuti (Sulawesi, Indonesia)

Md Akhtar-E Ekram[1] | Matthew Campbell[1] | Sureyya H. Kose[1] | Chloe Plet[1] | Rebecca Hamilton[2] | Satria Bijaksana[3] | Kliti Grice[1] | James Russell[4] | Janelle Stevenson[2] | Hendrik Vogel[5] | Marco J. L. Coolen[1]

[1]The Institute for Geoscience Research (TIGeR), Western Australia Organic and Isotope Geochemistry Centre (WAOIGC), School of Earth and Planetary Sciences (EPS), Curtin University, Bentley, Western Australia, Australia

[2]ARC Centre of Excellence for Australian Biodiversity and Heritage and Archaeology and Natural History, School of Culture, History, and Language, Australian National University, Canberra, Australian Capital Territory, Australia

[3]Faculty of Mining and Petroleum Engineering, Institut Teknologi Bandung, Bandung, Indonesia

[4]Department of Earth, Environmental, and Planetary Sciences (DEEPS), Brown University, Providence, Rhode Island, USA

[5]Institute of Geological Sciences & Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

Correspondence
Marco J. L. Coolen, The Institute for Geoscience Research (TIGeR), School of Earth and Planetary Sciences (EPS), Western Australia Organic and Isotope Geochemistry Centre (WAOIGC), Curtin University, Bentley, WA 6102, Australia.
Email: marco.coolen@curtin.edu.au

Funding information
Australian Research Council, Grant/Award Number: DP15102587; International Continental Scientific Drilling Program (ICDP); U.S. National Science Foundation (NSF); Swiss National Science Foundation, Grant/Award Number: 200021_153053; PT Vale Indonesia; Ministry of Research, Education, and Higher Technology of

Abstract

Studying past ecosystems from ancient environmental DNA preserved in lake sediments (*sed*aDNA) is a rapidly expanding field. This research has mainly involved Holocene sediments from lakes in cool climates, with little known about the suitability of *sed*aDNA to reconstruct substantially older ecosystems in the warm tropics. Here, we report the successful recovery of chloroplast *trn*L (UAA) sequences (*trn*L-P6 loop) from the sedimentary record of Lake Towuti (Sulawesi, Indonesia) to elucidate changes in regional tropical vegetation assemblages during the lake's Late Quaternary paleodepositional history. After the stringent removal of contaminants and sequence artifacts, taxonomic assignment of the remaining genuine *trn*L-P6 reads showed that native nitrogen-fixing legumes, C_3 grasses, and shallow wetland vegetation (*Alocasia*) were most strongly associated with >1-million-year-old (>1 Ma) peats and silts (114–98.8 m composite depth; mcd), which were deposited in a landscape of active river channels, shallow lakes, and peat-swamps. A statistically significant shift toward partly submerged shoreline vegetation that was likely rooted in anoxic muddy soils (i.e., peatland forest trees and wetland C_3 grasses (Oryzaceae) and nutrient-demanding aquatic herbs (presumably *Oenanthe javanica*)) occurred at 76 mcd (~0.8 Ma), ~0.2 Ma after the transition into a permanent lake. This wetland vegetation was most strongly associated with diatom ooze (46–37 mcd), thought to be deposited during maximum nutrient availability and primary productivity. Herbs (Brassicaceae), trees/shrubs (Fabaceae and Theaceae), and C_3 grasses correlated with inorganic parameters, indicating increased drainage of ultramafic sediments and laterite soils from the lakes' catchment, particularly at times of inferred drying. Downcore variability in *trn*L-P6 from tropical forest trees (*Toona*), shady ground cover herbs (Zingiberaceae), and tree orchids (*Luisia*) most strongly correlated with sediments of a predominantly felsic signature considered to be originating from the catchment of the Loeha River draining into Lake Towuti during wetter climate conditions. However, the co-correlation with dry climate-adapted trees

Indonesia (RISTEK); Brown University; The Institute for Geoscience Research (TIGeR); Research Office at Curtin University (ROC)

(i.e., *Castanopsis* or *Lithocarpus*) plus C_4 grasses suggests that increased precipitation seasonality also contributed to the increased drainage of felsic Loeha River sediments. This multiproxy approach shows that despite elevated in situ temperatures, tropical lake sediments potentially comprise long-term archives of ancient environmental DNA for reconstructing ecosystems, which warrants further exploration.

KEYWORDS
metabarcoding, paleodepositional environment, Quaternary, sedimentary ancient DNA, tropical paleovegetation record

1 | INTRODUCTION

Fossil pollen and spores provide a valuable record of the evolutionary history of vegetation and their response to past climate change (e.g., Anshari et al., 2004; Hamilton, Stevenson, et al., 2019; Hope, 2001; Stevenson, 2018), but overlapping morphological features can complicate the identification of various plant families at lower taxonomic levels. Notably, similar fossil pollen morphologies between genera of true grasses (Poaceae) make it impossible to distinguish between wet climate C_3 versus dry climate C_4 grasses and grasses that can switch between C_3 and C_4 carbon fixation pathways (Mander & Punyasena, 2014 and references therein). The same problem stands for dominant tropical tree families, which cannot be separated into broad habitat types without time-intensive morphological research (e.g., Hamilton, Hall, et al., 2019). Moreover, landscape reconstruction from pollen analysis is complicated by the variable production and transportation of pollen from different genera. For instance, the origin of pollen from wind-pollinated plants can be obscured by long distances, sometimes even cross-continental transportation (Gregory, 1978). Additionally, and problematically for tropical landscape reconstruction, rainforest angiosperms have low pollen production rates as they typically rely on animal vectors for pollination and successful reproduction (e.g., Ollerton et al., 2011). Such vegetation is, therefore, likely to be poorly represented in the fossil pollen record.

Several studies have shown that lacustrine sedimentary records also represent important archives of ancient vegetation DNA (Capo et al., 2021 and references therein) that can expand on and resolve critical issues from pollen-based reconstructions. Extraction of ancient sedimentary DNA (*sed*aDNA) and subsequent sequencing analysis of preserved vegetation metabarcoding genes, notably the short 10–134 base pair (bp) P6 loop of the chloroplast *trn*L (UAA) intron (Taberlet et al., 2007), can complement fossil pollen in reconstructing local vegetation communities and their responses to paleoenvironmental and more recent anthropogenic perturbations (e.g., Alsos et al., 2021; Courtin et al., 2021; Epp et al., 2015; Li et al., 2021; Niemeyer et al., 2017; Parducci et al., 2013, 2015; Pedersen et al., 2013; Zimmermann et al., 2017). Namely, since chloroplasts are organelles found predominantly in plant leaves where most photosynthesis takes place, *sed*aDNA records targeting chloroplast-derived metabarcoding genes are thought to mainly represent ancient local catchment vegetation that entered lake basins as decaying plant litter via riverine or terrestrial runoff (e.g., Alsos et al., 2018; Liu et al., 2020; Parducci et al., 2017; Paus et al., 2015; Voldstad et al., 2020). Stratigraphic analysis of *sed*aDNA thus has the potential to be an excellent archive of local vegetation change.

DNA from past phytoplankton communities has been extracted and sequenced from up to 2-Ma-old deep-sea sediments (Armbrecht et al., 2022; Kirkpatrick et al., 2016) and from up to 270-Kyr-old lake sediments (i.e., Lake Van, Turkey) (Randlett et al., 2014). However, most terrestrial *sed*aDNA records have been generated from Holocene lakes located in high-latitude regions where relatively cold conditions can promote the long-term preservation of *sed*aDNA (Capo et al., 2021 and references therein). A few recent studies have used *sed*aDNA to reconstruct tropical catchment vegetation change (Boessenkool et al., 2014; Bremond et al., 2017; Dommain et al., 2020). However, these records only date back several millennia and comprise tropical lakes in cooler, high-altitude locations. Therefore, it remains unknown how far back in time *sed*aDNA can be recovered and sequenced from tropical lakes after long-term exposure to less favorable, high in situ temperatures.

During the 2015 Towuti Drilling Project (Russell et al., 2016), sediment cores spanning more than 1 Ma of deposition (Russell et al., 2020) were obtained from Lake Towuti (Sulawesi, Indonesia; Figure 1) to reconstruct aquatic and terrestrial ecosystem responses to late Quaternary environmental and climate variability in the tropical western Pacific. The 205-m-deep Lake Towuti (2.75° S, 121.5° E) is the largest (500 km^2 surface area) member of the tectonic Malili Lake complex (which includes the smaller Matano and Mahalona lakes; Figure 1) and is positioned in a floristically biodiverse catchment area (~1144 km^2; Cannon et al., 2009; Whitten et al., 1988). The lakes' surface and bottom water temperatures were ~31 and 28.5°C, respectively. The uninterrupted continuous record from Lake Towuti, which extends well beyond the Holocene (Russell et al., 2020), offers a unique opportunity to study the long-term preservation potential of *sed*aDNA in tropical lakes and its suitability to reconstruct past vegetation change.

According to detailed geochronological, lithostratigraphic, mineralogical, and geochemical analyses spanning the entire sedimentary

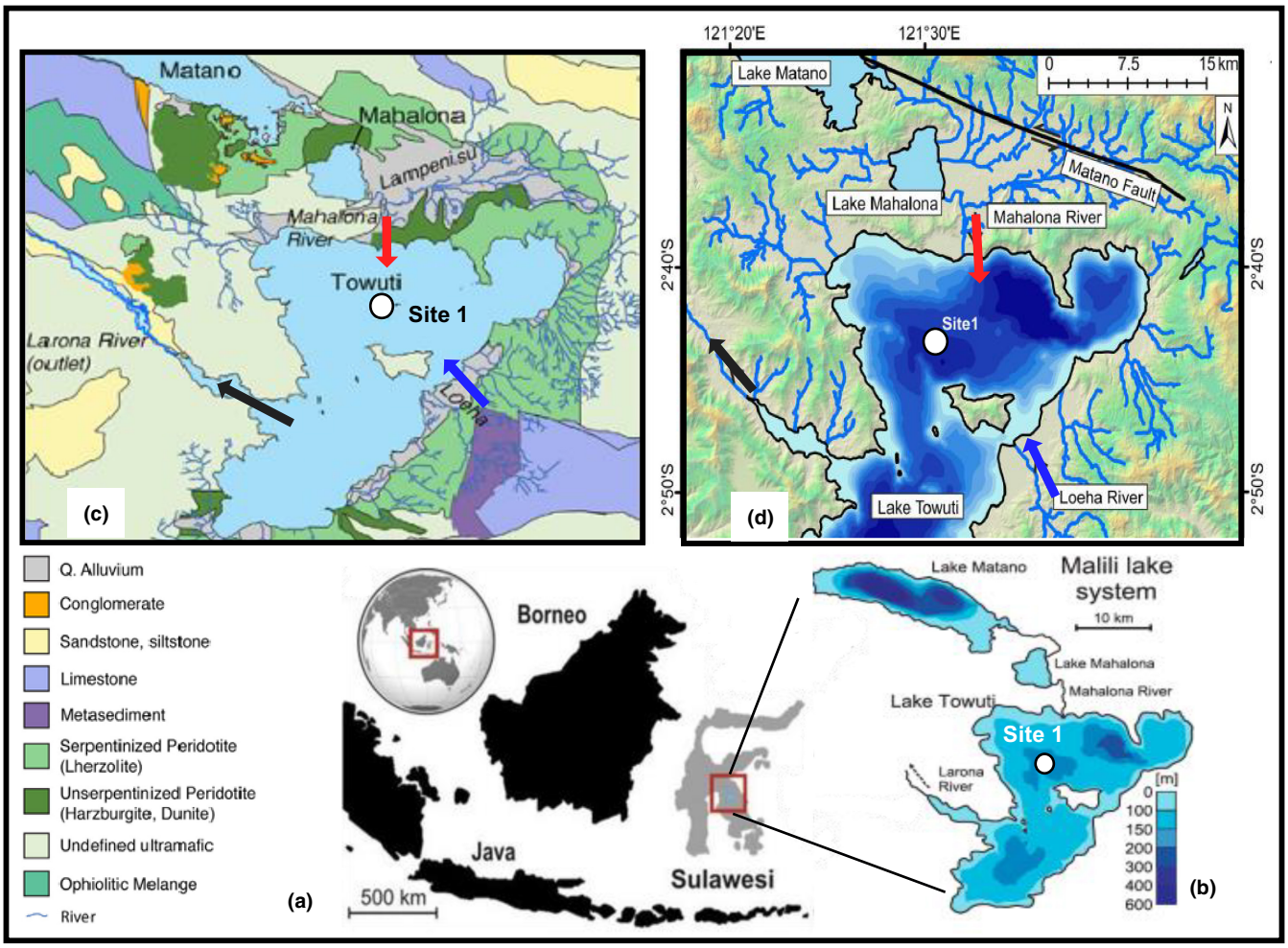

FIGURE 1 General overview of the sampling location. (a) Location of Lake Towuti on the Indonesian Island of Sulawesi as part of the Malili Lake system (modified after Friese et al., 2021). (b) Bathymetry of Lake Mahalona, Lake Mahalona, and Lake Towuti (changed from Russell et al., 2016). (c) Map of the local bedrock geology with legend below (modified from Costa et al., 2015). (d) A detailed map of the main rivers draining into Lake Towuti (modified from Morlock et al., 2021). Note the locations of ultramafic rocks, the source of Mg-rich eroded catchment material that nowadays drains into the lake via the Mahalona River (Northeast of Lake Towuti) (red arrow in c and d) during periods of reduced precipitation, resulting in low lake water stands versus limestone and metasedimentary bedrock as the source of K-rich eroded catchment material, which drains into the lake via the Loeha River (southeast of Lake Towuti) (blue arrow in c and d) during wetter conditions and high-lake-level stands (Russell et al., 2020). The direction of lake surface water outflow via the Larona River is indicated with a black arrow in c and d. The coring location 1 is marked with a white circle in (b-d).

record, Lake Towuti emerged during accelerated tectonic subsidence from a landscape initially characterized by active river channels, shallow lakes, and swamps and became a permanent lake at ~1 Ma (Russell et al., 2020). Drilling recovered cores from holes up to ~150 m deep, with sediment subdivided into two main stratigraphic units. The deeper sediment sequence, Unit 2, consists of alternating fluvial silty clays and peat layers (Figure 2; Russell et al., 2020). A thick peat interval between 101 and 98.8 mcd marks the transition into a permanent lake (Unit 1 (U1)). The U2/U1 transition at 98.8 mcd was estimated 1 Ma ago by extrapolating an accurately $^{40}Ar/^{39}Ar$ dated tephra layer at 72.95 mcd (i.e., 797.3 ± 1.6 Ka; Russell et al., 2020). The lacustrine U1 sediments contain alternating red sideritic and green organic-rich clay intervals, likely to reflect orbital-scale alterations between cooler, drier climates that promoted lower lake levels, lake mixing, and ultraoligotrophic conditions due to phosphate trapping by sedimentary iron oxyhydroxides (red sideritic clays) versus warmer, wetter climate stages that promoted a more productive lake through the release of sedimentary P and Fe^{2+} under seasonally stratified and anoxic conditions (green clays; Russell et al., 2020).

At much longer timescales, the lacustrine record reveals three major shifts in paleodepositional and paleohydrological conditions. Unit 1c (~98–76 mcd) shows frequent oscillations between more felsic (K-rich) and ultramafic (Mg-rich) sediments. The oscillations in sediment sources were likely caused by variable fault motion and a continued influence of tectonically driven changes in basin morphology and catchment hydrology (Russell et al., 2020). Unit 1b (76–30 mcd) has low %Mg and undetectable amounts of serpentine. A combination of a higher contribution of green clays and elevated %K in U1b suggests increased discharge of felsic sediments from the Loeha River to the east of

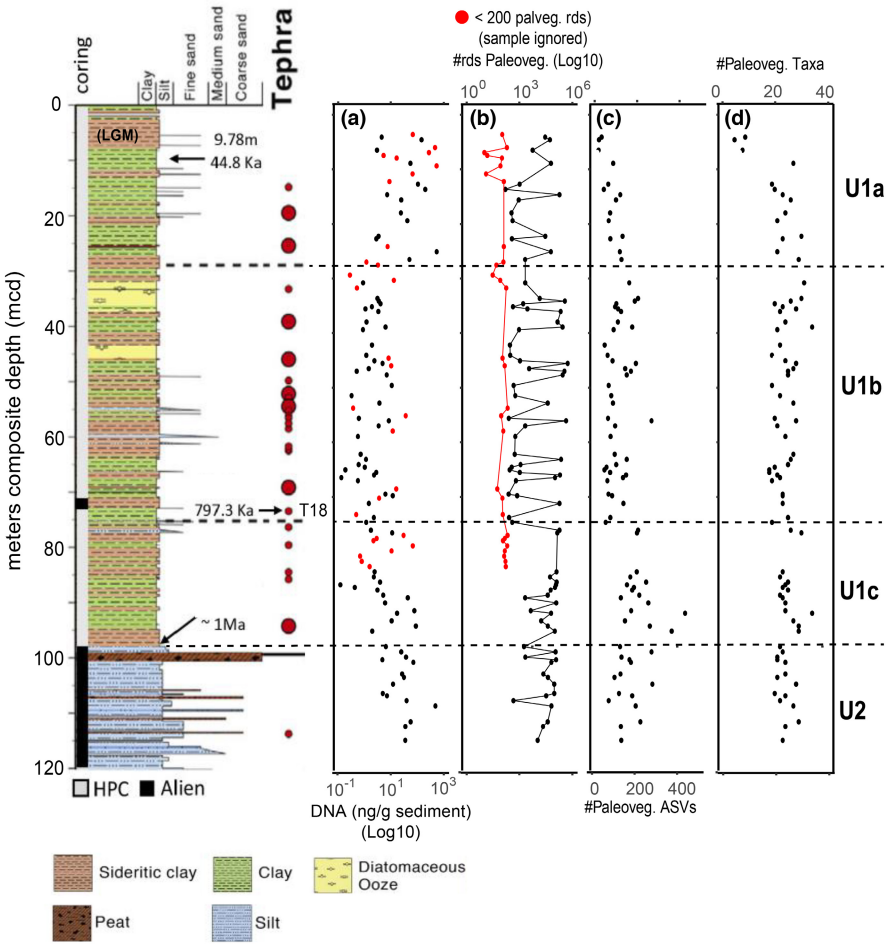

FIGURE 2 Total environmental DNA (eDNA) concentration and general overview of the recovered Paleogenic sequence data: (a) Amount of extracted total eDNA (ng/g sediment). (b) The total number of reads from paleovegetation. Samples containing fewer than 200 *trn*L-P6 reads of paleovegetation, which were excluded from further analysis are indicated in red. (c) Total number of ASVs that could be assigned to paleovegetation at family, clade, subfamily, tribe, or genus levels. (d) The number of taxa inferred from the sum of ASVs assigned to the same taxonomic ranks in the 82 remaining analyzed sediment intervals. See Figures S1–S3 for an overview of the contaminants and unassigned ASVs or sequencing artifacts removed from the dataset before processing. The lithology graph left of panel (a) shows the alternating deposition of silty clay and peat intervals during the pre-lake stage (Unit 2), followed by permanent lacustrine conditions above the U2/U1 transition at 98 mcd. See Russell et al. (2020) and results and discussion in the main text for a detailed description of the main lithologies (lacustrine green clays, red sideritic clays, and diatom ooze intervals vs. pre-lake fluvial silts and peats) and for the paleohydrological/paleodepositional conditions that were used to define the major pre-lake U2 and lacustrine U1a-, U1b, and U1c stages and their transitions. The lithology graph further shows composite sediment depth (mcd), coring method (hydraulic piston vs. alien), the position of tephra deposits T1–T23, and sediment ages. Radiocarbon dating on the bulk organic matter at 9.79 mcd revealed a sediment age of ~44.7 Ka, whereas $^{40}Ar/^{39}Ar$ dating of the tephra T18 layer at 72.95 mcd revealed a sediment age of 797.3 ± 1.6 Ka. The U2/U1c transition at ~98.8 mcd was estimated to have occurred 1 Ma ago through extrapolation (see Russell et al., 2020, for details).

Lake Towuti during warmer, wetter climates that promoted a more productive stratified lake. Long-term accumulation and weathering of tephra-bound P in the catchment may also have contributed to the development of mesotrophic conditions and maximum productivity, which ultimately resulted in the deposition of two several-meters-thick diatomaceous oozes between 32–37 and 43–46 mcd (Russell et al., 2020). The overlying U1a (top 30 m) shows a substantial increase in %Mg, indicating a connection of the Lampenisu and Mahalona Rivers between Lakes Mahalona and Towuti (Russell et al., 2020).

Here, we used sedimentary *trn*L-P6 amplicon sequencing (Taberlet et al., 2007) to generate a record of Quaternary tropical catchment and aquatic vegetation changes associated with the well-described transitions in limnological and hydrological changes throughout the lakes' more than 1 Ma history. Furthermore, the potential origin of the sedimentary chloroplast DNA was inferred from Pearson correlations between downcore relative changes in the *trn*L-P6-inferred paleovegetation assemblages and the previously analyzed (in)organic geochemical paleohydrological and paleoenvironmental parameters.

2 | MATERIALS AND METHODS

2.1 | Sampling

Drilling commenced at Site 1 (Figure 1) on May 23, 2015, using the ICDP Deep Lakes Drilling System (DLDS) operated by DOSECC Exploration Services. This priority coring site was chosen for highly resolved paleoclimate and paleoecology reconstructions spanning the entire history of Lake Towuti, as it is relatively free of thick turbidites and located in the proximity of core IDLE-TOW10-9B, which yielded high-quality paleoclimate data for the last 60 ka (Russell et al., 2014, 2020). Core 1F used for this study was from site IDLE-TOW10, obtained at a water depth of 156 m (Figure 1) using a PQ (122.6 mm hole, 66 mm core) diameter drill string with a hydraulic piston corer (HPC) for soft lacustrine sediments of Unit 1. Alien rotating coring, which required the addition of drilling fluid to lubricate the drill bit, was used to recover the more resistant lithologies of pre-lake U2 and a 2-m-thick interval of lacustrine red clays between 70.5 and 72.5 (Figure 1). Upon completion of the coring expedition in June 2015, the 3-m-long whole-round core sections were shipped inside standard capped butyrate liners via air freight to LacCORE (the National Lacustrine Core Facility) at the University of Minnesota, MN, USA, for processing, description, scanning, and subsampling (Russell et al., 2016). Sediment samples for DNA analysis ($n = 146$) were obtained at ~10 ka resolution from freshly split core sections between 2.95 and 114.57 mcd after Coolen et al. (2013) and transported inside sterile Whirl Pac bags to the 5.1. quarantined trace DNA lab facility (Australian Department of Agriculture Approved Arrangement # W3032), located within the Western Australia Organic and Isotope Geochemistry Centre (WA-OIGC) at Curtin University in Perth, where the samples were stored at −80°C until DNA extraction.

2.2 | Sedimentary DNA extraction

Depending on availability, both intracellular and mineral-adsorbed extracellular DNA was extracted from 2 to 8 g of sediment inside a bleach- and UV-sterilized HEPA-filtered horizontal laminar flow bench within lab W3032 using the DN*easy* Powermax Soil DNA extraction kit (Qiagen) with modifications after Direito et al. (2012) to efficiently release mineral-adsorbed extracellular DNA (Orsi et al., 2017). The concentration of extracted DNA was quantified fluorometrically using the Quant-iT™ Picogreen dsDNA Assay Kit (Thermo Fisher Scientific, Scoresby, VIC, Australia). Co-extracted PCR-inhibiting substances such as humic acids were removed using the OneStep PCR Inhibitor Removal Kit (Zymo Research, Irvine, CA, USA). Reagent mixtures without sediment were extracted and purified in parallel and served as controls for contamination during the lab procedures (extraction blanks; $n = 4$).

2.3 | Illumina MiSeq amplicon sequencing of sedimentary *trnL*-P6

In a general molecular biology lab that is physically separated from the trace DNA clean lab, PCR reactions were prepared using the SYBR Premix Ex Taq kit (Takara, Bio Inc.). One microliter of extracted and purified DNA (~2–25 ng) was added to a total volume of 20 µL reaction mixture and amplified (in triplicate) using a Realplex Quantitative PCR System (Eppendorf, Hauppauge, NY). The PCR reactions were stopped in the exponential phase or after a maximum of 33 cycles to prevent overamplification and minimize the production of PCR artifacts such as primer dimers. Each PCR cycle included a denaturation step (5 s at 95°C), primer annealing (30 s at 51°C), and primer extension plus imaging (15 s at 72°C) of newly formed SYBR®green-labeled double-stranded DNA. The short-chloroplast ~10–134 bp *trn*LP6 loop (Taberlet et al., 2007) was amplified using modified eco-PCR tested primers after Riaz et al. (2011): *trn*L-G2 (5′-GGG CAA TCCT GAG CCA A-3′) and *trn*L-H2 (5′-TTG AGT CTC TGC ACC TAT C-3′) since this primer combination yielded the least amount of primer dimers. Each initial round of PCR included triplicate reactions with 1 µL of each extraction blank and two reactions without template DNA present (background blanks).

The quality of the amplicons was evaluated by agarose gel electrophoresis. A total of 131 of 146 samples between 3.64 and 114.57 mcd revealed the expected average amplicon size of ~90 bp. Unfortunately, the expected amplicon size of ~90 bp could not be obtained from the two shallowest available samples (2.95 and 3.12 mcd). One microliter of the 131 *trn*L-P6 amplicons and the amplified background and extraction blanks served as a template for a nested PCR using the same primers carrying the required Illumina adapters. In addition, the reverse primers in the nested PCR carried 12 bp Golay barcodes after Caporaso et al. (2012) to support the pooling of samples and contamination controls for subsequent MiSeq Illumina sequencing. Identical barcodes were used for triplicate reactions of each sample and background versus extraction blanks. The nested PCR reactions were also followed in real-time and stopped after reaching the end of the exponential phase (12 cycles) to prevent overamplification and minimize the formation of PCR artifacts (More et al., 2018). The concentration of each barcoded amplicon was quantified fluorometrically (Quanti-iT™ PicoGreen TM ds DNA reagent; Thermo Scientific, Scoresby, VIC, Australia), and equimolar amounts were pooled and concentrated using an Amicon ultra centrifugal system (3 KDa). The concentrated barcoded library was run on a sterile 2% agarose gel at 120 V for 60 min, and the expected fragment was excised and gel purified using the Monarch Gel Extraction Kit (New England Biolabs). The DNA concentration of the concentrated and gel-purified library was measured again fluorometrically and sent for Illumina MiSeq paired-end sequencing (2 × 75 bp/150 cycles) at the Australian Genome Research Facility (AGRF) in Perth.

2.4 | Taxonomic assignments and removal of unassigned reads plus contaminants

Using ObiTools (Boyer et al., 2016), raw sequences were paired with an alignment score threshold of 20, assigned to samples, and dereplicated. Low-count sequences (<10) and short reads (<10 nucleotides) were removed using obigrep, and PCR errors were removed using obiclean. The remaining reads were assigned to amplicon sequence variants (ASV) using ecotag. The *trn*L reference database for taxonomic assignment of the ASVs was created using ecoPCR from release 143 of the European Nucleotide Archive. Before downstream analysis, unassigned and contaminant reads that occurred in both the samples and in extraction and background amplification blanks were removed, and the remaining error- and contaminant-free datasets were used for downstream bioinformatic and biostatistical analysis.

2.5 | Bioinformatics and biostatistics of the clean *trnL*-P6 datasets

Relative abundance and presence/absence matrices of ASVs that could be identified at the same lowest reliable taxonomic levels (i.e., family, clade, subfamily, tribe, genus, or species level) formed the template for downstream bioinformatics and biostatistical analysis. Partial least-square discriminant analysis (PLSDA) (Sisk-Hackworth & Kelley, 2020) was performed in the R package MixOmics (Rohart et al., 2017) as a sensitive ordination tool to visualize the level of dissimilarities in paleovegetation communities between samples and sample groups (i.e., paleodepositional units and main lithologies). Prior to this analysis, zero entries in the species abundance matrix were offset by 1 to allow a centered log-ratio (CLR) transformation (Lee et al., 2017). Global and pairwise similarity analysis (ANOSIM) was used to determine if *trn*L-P6 paleovegetation communities differed significantly between paleodepositional units and main lithologies. This analysis was performed using Bray–Curtis dissimilarities of square-root-transformed relative abundance data and Sorensen dissimilarities of presence/absence data in PRIMER-e v7 (Clarke & Gorley, 2015). Indicator species analysis (ISA) was performed using the R package IndicSpecies (De Caceres & Legendre, 2009) to reveal the significant association of *trn*L-P6 vegetation (classified at the lowest reliable taxonomic levels) with paleodepositional units and sediment lithologies. In addition, a similarity of percentage (SIMPER) analysis was performed in PRIMER-e v7 to identify the taxa that contributed most to group similarities or dissimilarities between groups that contained significantly different (ANOSIM $p < .05$) paleovegetation communities.

A Clustered Image Map (CIM) of Pearson correlations between *trn*L-P6-inferred paleovegetation and previously analyzed (in)organic geochemical parameters for changes in Lake Towuti's paleohydrology and paleoenvironment (Morlock et al., 2021; Russell et al., 2020; Sheppard et al., 2021) was performed in MixOmics to infer the origins of the different types of vegetation that drained into Lake Towuti in the form of chloroplast-rich plant litter. These parameters included:

(i) Cr and Ni (as oxides)—potentially phototoxic.
(ii) Fe and Mn (as oxides)—redox sensitive and enriched in the Mahalona and Lampenisu Rivers draining ultramafic bedrock catchments and laterite soils, respectively.
(iii) Mg—discharged into Lake Towuti as Mg-rich serpentines via the Mahalona River (Morlock et al., 2019; Russell et al., 2020).
(iv) K, Al, and Ti (as oxides), and increased Al/Mg ratio—indicative of increased erosion and sedimentation of laterite soils and drainage of felsic bedrock material from the Loeha River to the east (Morlock et al., 2019).
(v) TOC—elevated due to enhanced primary productivity, increased terrestrial OM influx, or increased SOM preservation during stratified and anoxic conditions.
(vi) $\delta^{13}C$ of bulk SOM—enriched in case of increased influx of OM from C_4 vegetation (Russell et al., 2020).
(vii) TLE/TOC ratio—elevated in the case of increased preservation of SOM under stratified and reducing anoxic conditions and signifying high-lake-level stands and a wetter climate (Russell et al., 2020).
(viii) $\%SiO_2$—Diatom ooze indicating mesotrophic conditions caused by long-term leaching of P from tephra, which contributed to the deposition of the several-meter-thick diatom oozes (Russell et al., 2020).

Bubble plots were prepared using the packages reshape2 and ggplot2 (Wickham, 2016) in R (https://www.R-project.org/) to visualize downcore changes in the relative abundance of the *trn*L-P6 inferred vegetation.

3 | RESULTS

3.1 | General overview of the downcore distribution of eDNA content and sequence data

The total amount of extracted environmental DNA (eDNA), which likely comprises a complex mixture of extant and past biological sources, notably bacteria, varied substantially between the analyzed intervals of the same lithologies (see Figure 2a, Table S1 for details). Sedimentary *trn*L-P6 was successfully amplified from 113 of 143 extracted intervals, and subsequent amplicon sequencing yielded 10^3 to 4×10^5 sequence reads per sample (Figure S1A). In total, 2770 ASVs were assigned from all samples combined. Aligning the recovered ASVs with available *trn*L-P6 sequences in the EMBL plant sequence database revealed that roughly one-third of these ASVs (962 of 2770) could only be identified at subphylum or higher taxonomic ranks and were not considered for further analysis. While this may be attributed to a lack of available closely related *trn*L-P6 sequences, similarity analysis using the Basic Local Alignment Search (BLAST) tool showed that 899 of the 962 poorly

assigned ASVs did not show any closely matching sequences available in NCBI's (National Center for Biotechnology Information) nucleotide collection (nr/nt) database (https://blast.ncbi.nlm.nih.gov). The remaining 63 ASVs showed the highest similarities with sequences of bacterial origin (Figure S1).

The parallel amplified and sequenced background and extraction blanks (BGB and EB) yielded a combined total of 227 ASVs that could be assigned to the common lab contaminants *Musa* (a genus of the family Musaceae that includes banana) and northern hemisphere temperate climate conifers (Pinaceae; *Pinus* and *Picea*) (Boessenkool et al., 2014; Pedersen et al., 2013), as well as unusual contaminants, notably unclassified Asparagaceae (subfamily Nolinoideae) (Figure S2). The remainder of relatively less abundant contaminant ASVs represented non-native taxa showing the highest sequence similarities (>96%) with *Allium* (a genus of herbs of the family Amaryllidaceae that comprises garlic and onion), *Arachis* (a genus of herbs of the legume family Fabaceae that includes peanuts), *Callitris* (conifers indigenous to Australia), and *Alnus* (deciduous trees of the family Betulaceae, native to temperate regions of the northern hemisphere) (Figure S2). In agreement with this result, Musaceae, Pinaceae, and Asparagaceae also represented the relatively most abundant assigned contaminants in many samples throughout the record, with minor contributions of *Allium*, *Arachis*, *Callitris*, and *Alnus* (Figure S3). In addition, 216 of the 962 unassigned ASVs were found to be relatively abundant contaminants in the blanks and many of the samples (Box B of Figure S3) and were removed from the dataset. The remaining 1375 ASVs, considered sourced from paleovegetation, could be assigned at taxonomic ranks low enough (i.e., family, clade, subfamily, tribe, or genus levels) for meaningful paleoenvironmental interpretations (Figure 3).

We decided to exclude 31 sediment intervals which, after the stringent removal of contaminants and unassigned reads, yielded less than 200 *trn*L-P6 sequences from taxonomically assigned paleovegetation in total (i.e., 18 red clays, 12 green clays, and 1 silty clay (i.e., the red-labeled intervals in Figure 2b)). Using the definition of depositional units after Russell et al. (2020), the remaining 82 intervals comprised 14 U2 intervals (5 peats and 9 silty clays), 17 U1c intervals (6 red clays and 11 green clays), 37 U1b intervals (1 silty clay, 10 red clays, 17 green clays, and 9 diatom oozes), and 14 intervals of U1a (5 red clays and 9 green clays) (Figure 3). The thick interval of sideritic red clays between ~3 and 7 mcd of Unit 1a span the Last Glacial Maximum (~15–30 cal. ka BP) (Russell et al., 2020). Eight red clay samples from the LGM (2.95, 3.12, 3.64, 4.12, 4.6, 6.03, 6.51, and 6.94 mcd) were initially available for DNA extraction. However, only three red clays (4.12, 4.6, and 6.51 mcd), which were deposited during the LGM, were found suitable for downstream analysis as they yielded more than 200 *trn*L-P6 reads from paleovegetation that could be identified at family or lower taxonomic ranks (Figures 2 and 3). Only 61 additional ASVs were lost after excluding 31 intervals with fewer than 200 paleovegetation reads.

After these measures, 1364 remaining ASVs were selected for downstream data analysis. The number of ASVs varied greatly

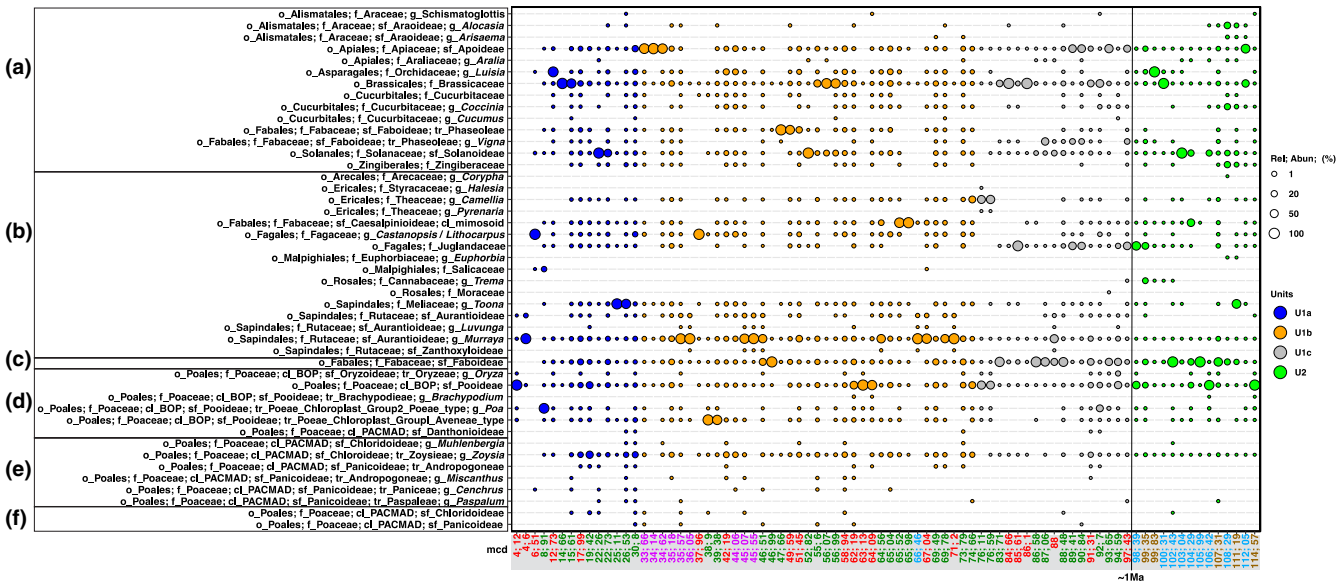

FIGURE 3 Bubble plot showing the downcore distribution of *trn*L-P6-inferred paleovegetation. Only samples are shown that contained more than 200 *trn*L-P6 reads from past taxa after removal of contaminants, poorly classified reads (i.e., subphylum or higher) in the EMBL plant database, or represent artifact reads that did not return any significant similarities after blasting against the NCBI's nr database. The taxonomic levels include order (o_), family (f_), subfamily (sf_), clade (cl_), tribe (tr_), and genus (g_). The size of the bubbles indicates the relative abundance at 1%, 20%, 50%, and 100% scale. The vegetation is ordered alphabetically within the following boxed categories: (a)—(herbs), (b)—(TRSH), (c)—(herbs or TRSH), (d)—(C_3 grasses), (e)—(C_4 grasses), and (f)—(C_3/C_4 grasses). See Russell et al. (2020) and the main text for details about the lithologies and depositional units. The *x*-axis denotes meters composite depths (mcd) of the analyzed sediment intervals. The samples have been color-coded according to the main lithology types. See the caption of Figure 2 for additional overlapping details.

throughout the core, being lowest in the red clays of the LGM (14–28 ASVs per interval) and highest in the older red clays of U1 (42–418 ASVs) and green clays (38–355 ASVs). See Table S1 for further details. The plant lists for Sulawesi were sourced from a digitized herbarium list of 28,000 specimens collected from Sulawesi and compiled by the Naturalis Biodiversity Center in Leiden, The Netherlands. Plant function types were determined by assessing the specimen notes in the herbarium list and published data (IUCN, 2019; Kessler et al., 2002; Whitten et al., 1988). The ASVs could be assigned to herbs ($n=12$), trees, or shrubs (TRSH; $n=18$, C_3 grasses ($n=6$), C_4 grasses ($n=6$), and C_3 or C_4 grasses ($n=2$; Figure 3)) at taxonomic ranks (family or lower) native to Sulawesi. In addition, unclassified Fabaceae (Faboideae) formed a separate category since they may represent herbs and TRSH (Figures 3 and 4). Roughly four times fewer paleovegetation taxa were identified from the LGM section of U1a ($n=6.3\pm2.1$) than in the older red clays of U1 ($n=22.5\pm3.5$; Table S1; 82 intervals). A comparable average number of taxa were identified from green clays (23.4 ± 3.9), diatom oozes (23.4 ± 3.8), clayey silts (22.9 ± 2.5), and peat intervals (22.2 ± 3.3) (Figure 2, Figure S1, Table S1). Herbs and TRSH showed the highest relative abundances throughout the record, followed by C_3 grasses. Obligate C_4 grasses comprised less than 10% throughout the core, reaching the highest levels in the green clays, notably in section U1b (Figure 4).

3.2 | Classification of samples using partial least-square discriminant analysis (PLSDA)

PLSDA enables selecting of the most predictive or discriminative features in data to help classify samples (Sisk-Hackworth & Kelley, 2020). Due to the dissimilarities in the composition and relative abundance of paleovegetation ($n=45$ taxa in 82 intervals), PLSDA analysis resulted in the formation of spatially separated clusters for each of the four previously defined major paleodepositional categories (i.e., pre-lake U2, and lacustrine U1c, 1b, and 1a) (Figure S4). Subsequent analysis of similarities (ANOSIM) using Bray–Curtis dissimilarity of standardized and square-root-transformed relative abundance data and using Euclidean dissimilarity of presence–absence (PA) data showed that the taxa identified through sedimentary *trn*L-P6 profiling differed significantly between these

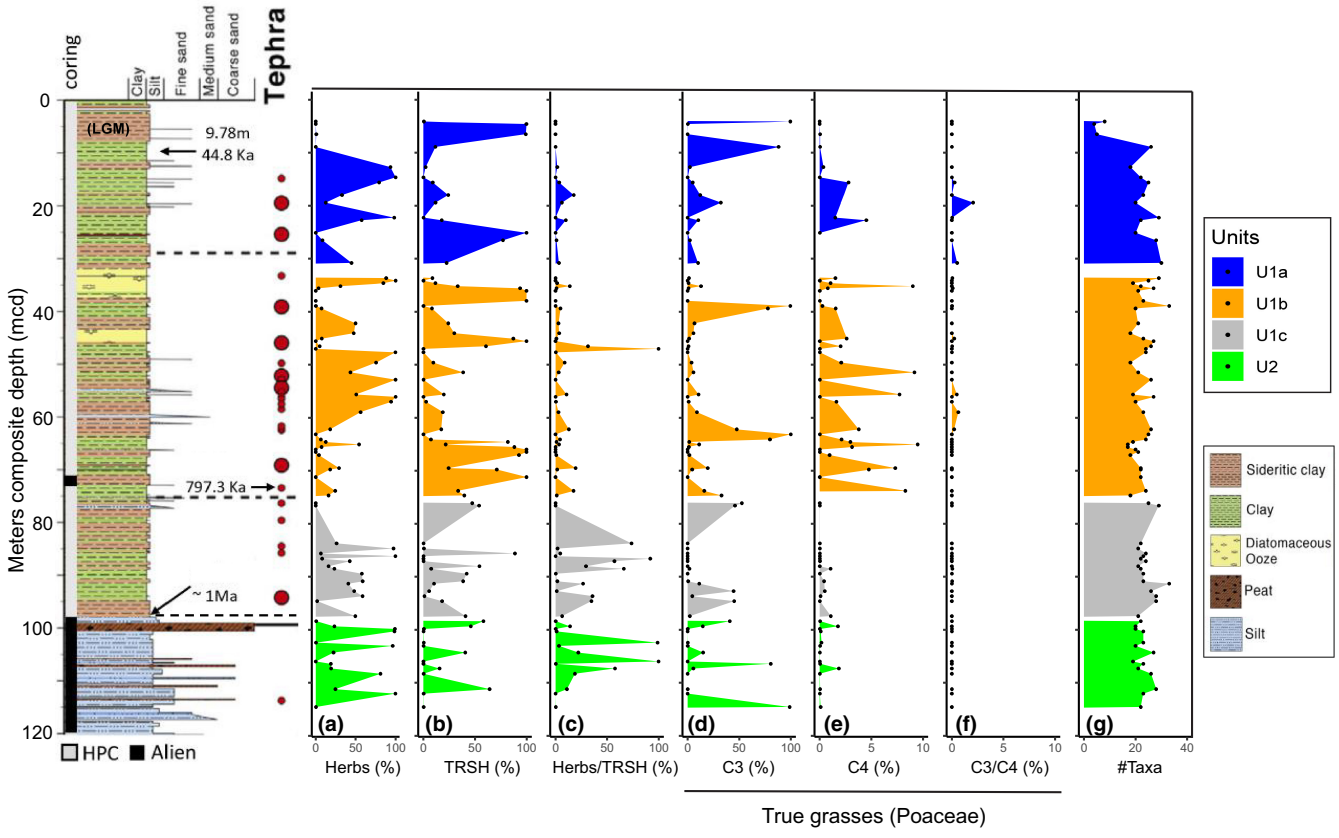

FIGURE 4 Downcore variability in *trn*L-P6-inferred paleovegetation. Relative abundance of *trn*L-P6 from (a) herbs; (b) evergreen trees and shrubs (TRSH); (c) Faboideae that could be herbs and/or TRSH; (d) true grasses (Poaceae) that use the C_3 carbon fixation pathway (BOP clade and the subfamily Danthonioideae of the PACMAD clade); (e) genera of true grasses within the PACMAD subfamilies Chloridoideae and Panicoideaea that only use the C_4 carbon fixation pathway as an adaption to drier conditions; (f) PACMAD subfamilies that could not be assigned at genus level and may have been using C_3 as well as C_4 carbon fixation pathways; (g) The number of *trn*L-P6 paleovegetation taxa identified at the lowest reliable taxonomic ranks. The paleodepositional units were defined by Russell et al. (2020). See the caption of Figure 2 for overlapping details.

paleodepositional units (global test: $r = .148$; $p = .003$ vs. $r = .150$; $p = .001$; Table 1). No significant differences in paleovegetation taxa were observed using Bray–Curtis versus Euclidean dissimilarities (global ANOSIM test) between pre-lake U2 and the entire lacustrine U1 section (top row in Table 1). Pairwise ANOSIM using all four paleodepositional categories showed that the Bray–Curtis dissimilarities in community composition were highest between U1a and U1c ($r = .153$; $p = .002$), between U1b and U1c ($r = .186$; $p = .012$), and between U1b and U2 ($r = .196$, $p = .024$) (Table 1). Highly significant differences in paleovegetation were observed between U1a and U2 only based on Euclidian analysis of presence/absence data ($p = .001$) versus a marginally significant difference ($p = 0.024$)

TABLE 1 Global and pairwise analysis of similarities (ANOSIM) showing significant dissimilarities in paleovegetation between sample categories: (i) pre-lake Unit 2 × lacustrine Unit 1; (ii) paleodepositional units (four categories; U2, U1c, U1b, and U1a); (iii) paleodepositional units with the diatom ooze considered as a separate sample category of Unit 1b (U2, U1c, $U1b_{DO}$, $U1b_{RC+GC}$, and U1a).

		SQRT + Bray–Curtis			PA + Euclidian		
Categories	**Global versus pairwise tests**	***R***	***p***	**Significance level**	***R***	***p***	**Significance level**
Stages (pre-lake × lake)	Global test	.063	.226	NS	.05	.268	NS
Paleodepositional units (U1a, U1b, U1c, and U2)	Global test	.148	.003	**	.15	.001	***
	U1a × U1b	.117	.094	NS	.204	.02	*
	U1a × U1c	.153	.002	**	.119	.009	**
	U1a × U2	.11	.036	*	.137	.003	**
	U1b × U1c	.186	.012	*	.101	.06	NS
	U1b × U2	.196	.024	*	.185	.01	*
	U1c × U2	.021	.256	NS	.072	.063	NS
Paleodepositional units (U1a, U1b(RC + GC), U1b(DO), U1c, and U2)	Global test	.193	.001	***	.123	.002	***
	U1a × U1b(DO)	.018	.364	NS	−.063	.8	NS
	U1a × U1b(RC + GC)	.074	.179	NS	.193	.01	*
	U1a × U1c	.153	.006	**	.119	.009	**
	U1a × U2	.08	.024	*	.137	.001	***
	U1b(DO) × U1b(RC + GC)	.04	.349	NS	.017	.399	NS
	U1b(DO) × U1c	.306	.001	***	.068	.184	NS
	U1b(DO) × U2	.374	.001	***	.151	.044	*
	U1b(RC + GC) × U1c	.161	.006	**	.114	.028	*
	U1b(RC + GC) × U2	.138	.041	*	.201	.004	**
	U1c × U2	.021	.28	NS	.072	.063	NS
Main lithologies	Global test	.044	.212	NS	.034	.265	NS
	RC × GC	.023	.281	NS	.056	.138	NS
	RC × DO	−.043	.628	NS	−.082	.784	NS
	RC × silt	.122	.099	NS	.004	.444	NS
	RC × peat	−.018	.519	NS	−.004	.582	NS
	GC × DO	.051	.319	NS	−.03	.582	NS
	GC × silt	.044	.330	NS	.085	.185	NS
	GC × peat	−.052	.595	NS	.069	.335	NS
	DO × silt	.487	.001	***	.12	.028	*
	DO × peat	.4	.002	**	.082	.236	NS
	Silt × peat	−.078	.718	NS	−.136	.897	NS

Note: ANOSIM results were compared using Bray–Curtis versus Euclidian dissimilarities of normalized and square root-transformed relative abundance versus presence–absence data of paleovegetation communities identified at the lowest reliable taxonomic levels. *R* shows the mean of ranked dissimilarities between categories to the mean of ranked dissimilarities within categories. $p < .05$ indicates significantly different global and pairwise ANOSIM tests. Significance levels: 0 "***" .001 "**" .01 "*" .05 and not significant (NS). Only samples with >200 total reads of genuine paleovegetation after removal of the artifact- and contaminant *trn*L-P6 reads were used for ANOSIM analysis (Figure 3). PLSDA analysis was performed to visualize the separation of samples belonging to the five developmental units (see Figure S4). See Table 2 listing the significant indicator species assigned to the developmental stages and Figure 5 for an overview of taxa with the highest percentage associated with the five depositional stages (i.e., SIMPER analysis).

using Bray–Curtis analysis of square-root-transformed quantitative sequence data. Paleovegetation assemblages between U1a and U1b and U1c and U2 did not differ significantly (Table 1). Treating diatom oozes and the clays from U1b as two separate sample categories for ANOSIM analysis resulted in higher global *r*- and lower global *p*-values mainly driven by the greater significant dissimilarities in vegetation assemblages between the diatom ooze and U1c ($r = .306$; $p = .001$) versus U2 ($r = .374$; $p = .001$). Paleovegetation assemblages did not differ significantly between diatom ooze and lacustrine clays, which is likely due to the imbalance in the availability of samples in favor of the clays. Treating diatom oozes and clays of U1b as two separate sample categories resulted in more significant dissimilarities in the presence/absence of paleovegetation, most notably between U1a and U2 (Table 1). Global and pairwise ANOSIM performed on the main lithologies revealed that paleovegetation assemblages only differed significantly between samples of the peat and clayey silt intervals of U2 and the diatom oozes of U1b (See Table 1 for further details).

3.3 | Identification of past vegetation through sedimentary *trnL*-P6 profiling

3.3.1 | Herbs currently native to Sulawesi

Wetland herbs of the order Allismatales (Arecaceae), notably *Alocasia* ("Elephant Ear") that grows from rhizomes and is native to tropical and subtropical Asia, including Sulawesi (http://www.powo.science.kew.org), were primarily identified in the peats and silts at the base of the studied interval (Figure 3). Indicator species analysis revealed that another genus *Arisaema* in this family of wetland herbs was significantly associated with this depositional stage ($r = .477$; $p = .008$; Table 2). Putative N-fixing legumes (Fabaceae) were relatively abundant throughout U2 and U1c. They also showed equally highest percentage associations (SIMPER; ~38%) with both depositional units (Figure 5) but represented only significant indicator taxa for U1c (Table 2). Unclassified Faboideae made the highest contribution among group members of U2 and U1c. Since this subfamily of Fabaceae could not be classified at lower taxonomic levels, they may also represent TRSH and were grouped as Herbs and TRSH (vegetation category C in Figures 2–4). ASVs of the cosmopolitan nightshade family Solanaceae, predominantly assigned to the native genus *Solanum* (http://www.powo.science.kew.org), were found to be associated with all subunits but did not show any association with diatom oozes. In contrast, the herb family Brassicaceae (Brassicales) showed the highest percentage of associations with U1c and U1a and the lowest association with the diatom ooze intervals (Figure 5). Moreover, ASVs assigned to herbs of the family Apiaceae (subfamily Apoideae) showed up to 100% sequence similarity with Java water dropwort (i.e., *Oenanthe javanica*). This native edible aquatic herb, which can be abundant in freshwater environments throughout Southeast Asia, showed the highest associations with U1b diatom oozes (SIMPER; 18%) (Figure 5) and represented a significant indicator species for this sample category (ANOSIM; $r = .657$; $p = .02$) (Table 2).

3.3.2 | Trees and shrubs (TRSH)

ASVs assigned to Juglandaceae (Fagales) were present in most pre-LGM intervals but were found to be a significant indicator taxon for U1c (Table 1) and revealed the highest SIMPER associations with this early lacustrine depositional unit (Figure 5). Although genus-level identification was not possible, it may represent *Engelhardia*, a native evergreen tree of Sulawesi often associated with dense primary forests on mountain slopes between 300 and 1600 m above sea level (Kessler et al., 2002), and fossil pollen type recorded within nearby lake sediment records (Hamilton, Stevenson, et al., 2019). ASVs with 99%–100% sequence similarities to *Toona ciliata*, a large native tropical forest tree of the Mahogany family (Meliaceae), were recovered from intervals throughout the record, except for the LGM red clays. *Toona* was most consistently identified from U1a (Table 2) and showed the highest SIMPER association with this depositional unit (Figure 5). ASVs related to Caesalpinioideae (Fabaceae), a subfamily of the Mimosoid clade, were most consistently identified from green clays (Figure 3). Although not a significant indicator taxon for any of the depositional units (Table 2), Caesalpinioideae were found to be only associated with clays of U1b in the case where these lithologies were treated as a separate group from diatom oozes (Figure 5). ASVs of this subfamily of Fabaceae showed ~98%–100% sequence similarity with the native genus *Parkia*, which can be found in various tropical habitats, including moist lowland forests and swamps (Hopkins, 1994). *Castanopsis* and/or *Lithocarpus* (Fagaceae) were most consistently present in U1a and U1b intervals, including the LGM (Figure 3), and revealed relatively low percentage associations (SIMPER: 5%–7%) with the clays of these subunits when diatom ooze was treated as a separate category (Figure 5). These drought-tolerant tropical evergreen trees now dominate at lower/mid-montane elevations in Sulawesi (Biagioni et al., 2015; Culmsee et al., 2010) and have been described from fossil pollen in the adjacent TOW-9 and satellite lake (Lantoa) records (Hamilton, Stevenson, et al., 2019; Stevenson, 2018). ASVs with up to 100% sequence similarity to native genera (*Murraya* and *Luvunga*) belonging to the Rutaceae subfamily Aurantioideae (http://www.powo.science.kew.org) were present in the majority of analyzed intervals, including LGM red clays. However, the relative abundance of Aurantioideae (notably *Murraya* that mainly inhabits peatland) was highest in the diatom oozes and showed very high association with this lithology (43%) and showed lower percentage associations with the clays of U1b and U1a but no substantial associations were observed with U2 and U1c (Figure 5). As expected, the Aurantioideae represented highly significant indicator taxa for the diatom oozes (Table 2). ASVs identified as *Trema* (Cannabaceae) were only detected in four consecutive intervals above the U2/U1c transition (Figure 3). In agreement with this result, *Trema* was a significant indicator taxon for U2 (Table 2) but was not substantially associated with this interval (Figure 5).

TABLE 2 Overview of taxa that differed significantly between paleodepositional units.

Indicator species	Category	A	B	*r* value	*p* value	Significance level
A						
U1a, U1b, U1c, and U2						
o_Poales; f_Poaceae; cl_PACMAD; sf_Panicoideae; tr_Paniceae; g_*Cenchrus*	U1a	0.5857	0.9	0.726	.048	*
o_Cucurbitales; f_Cucurbitaceae	U1b	0.4497	0.8108	0.604	.032	*
o_Fabales; f_Fabaceae; sf_Faboideae; tr_Phaseoleae	U1b	0.5997	0.973	0.764	.004	**
o_Poales; f_Poaceae; cl_BOP; sf_Pooideae; tr_Poeae_Chloroplast_GroupI_Aveneae_type	U1b	0.5168	0.973	0.709	.041	*
o_Sapindales; f_Rutaceae; sf_Aurantioideae	U1b	0.5103	0.8649	0.664	.007	**
o_Sapindales; f_Rutaceae; sf_Aurantioideae; g_*Murraya*	U1b	0.5248	1	0.724	.001	***
o_Fabales; f_Fabaceae; sf_Faboideae; tr_Phaseoleae; g_*Vigna*	U1c	0.6186	1	0.786	.001	***
o_Fagales; f_Juglandaceae	U1c	0.4891	1	0.699	.021	*
o_Alismatales; f_Araceae; sf_Araoideae; g_*Arisaema*	U2	0.7978	0.2857	0.477	.008	**
o_Malpighiales; f_Euphorbiaceae; g_*Euphorbia*	U2	0.9872	0.1429	0.376	.041	*
o_Rosales; f_Cannabaceae; g_*Trema*	U2	0.8742	0.4286	0.612	.01	**
B						
U1a, U1b(RC+GC), U1b(DO), U1c, and U2						
N/A	U1a	—	—	—	—	—
o_Cucurbitales; f_Cucurbitaceae	U1b (RC+GC)	0.3585	0.8571	0.554	.045	*
o_Fabales; f_Fabaceae; sf_Faboideae; tr_Phaseoleae	U1b (RC+GC)	0.5091	0.9643	0.701	.023	*
o_Apiales; f_Apiaceae; sf_Apoideae	U1b (DO)	0.4312	1	0.657	.002	**
o_Sapindales; f_Rutaceae; sf_Aurantioideae	U1b (DO)	0.4903	1	0.7	.001	***
o_Sapindales; f_Rutaceae; sf_Aurantioideae; g_*Luvunga*	U1b (DO)	0.5085	0.4444	0.475	.041	*
o_Sapindales; f_Rutaceae; sf_Aurantioideae; g_*Murraya*	U1b (DO)	0.4603	1	0.678	.001	***
o_Fabales; f_Fabaceae; sf_Faboideae	U1c	0.3388	1	0.582	.035	*
o_Fabales; f_Fabaceae; sf_Faboideae; tr_Phaseoleae; g_*Vigna*	U1c	0.5733	1	0.757	.002	**
o_Fagales; f_Juglandaceae	U1c	0.4407	1	0.664	.022	*
o_Alismatales; f_Araceae; sf_Araoideae; g_*Arisaema*	U2	0.7491	0.2857	0.463	.019	*
o_Rosales; f_Cannabaceae; g_*Trema*	U2	0.8707	0.4286	0.611	.025	*

Note: Indicator species analysis (ISA) was performed with (A) lacustrine subunits U1a, U1b (all lithologies), U1c, and pre-lake U2; (B) U1a, U1b (RC+GC separated from DO intervals), U1c, and U2. The best indicator species are those found only in samples/intervals belonging to one group/category ($A=1$) and are present in all samples within that group ($B=1$). Only taxa with p values <.05 are shown in the table. Significance levels (r): 0 "***" .001 "**" .01 "*" .05. Taxonomic levels: o_(order); f_(family); sf_(subfamily); tr_(tribe); cl_(clade); and g_(genus). Note that the indicator species in each sample category are ordered alphabetically.

3.3.3 | True grasses (Poaceae)

Sedimentary *trn*L-P6 profiling allowed us to distinguish between wet climate C_3 and dry climate C_4 grasses. Pooid grasses belonging to the BOP clade (Figure 3), which exclusively use the C_3 carbon fixation pathway (Chang et al., 2013), predominated over C_4 grasses in all analyzed intervals, including the LGM (Figure 4). These C_3 grasses included the Pooideae subtribes Poeae and Aveneae (type *Avena* L) found in cooler climate regions of the world as well as in tropical mountains (Lasut, 2009). According to SIMPER analysis, Pooideae revealed relatively high percentage associations with all depositional units except for the diatom ooze intervals. Native C_3 wetland grasses (*Oryza*) were sporadically found in green clays shortly after the U2/U1c transition and in green clays plus diatom ooze (Figure 3). A subfamily of PACMAD grasses that exclusively use the C_3 carbon fixation pathway (Danthanoideae) were also sporadically detected (Figure 3). *Zoysia* (subfamily Chloroideae) a genus of exclusively C_4 grasses within the PACMAD clade (Aliscioni et al., 2012) was detected throughout the record, except for the LGM red clays (Figure 3). Other C_4 grasses of the PACMAD clade (*Muhlenbergia*, *Miscanthus*, *Cenchrus*, and *Paspalum*) were only sporadically detected throughout the core, but most frequently in the green clay intervals of U1a and U1b, where they reached relative abundances of up to 10% of the total vegetation assemblages (Figure 4).

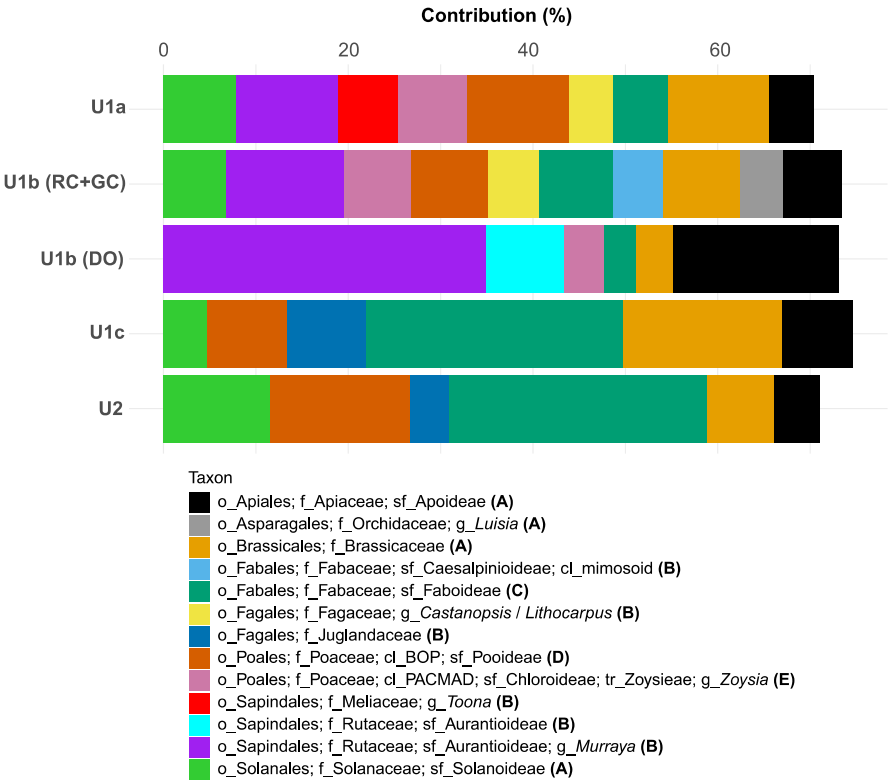

FIGURE 5 Similarity percentage (SIMPER) analysis showing the taxa that made the highest % contribution to the (Bray–Curtis) similarities observed within group members (cut-off = 70%): pre-lake U2 versus the lacustrine subunits U1a, U1b (RC and GC), U1b (DO), and U1c. (A) Herbs; (B) TRSH; (C) herbs and/or TRSH of Faboideae; (D) C_3 grasses; (E) C_4 grasses). See the caption of Figure 2 and the main text for a detailed description of the depositional stages.

3.4 | Pearson correlations between *trnL*-P6 vegetation and geochemical parameters

The cluster heatmap of Figure 6 shows Pearson correlations (Pearson's *r* values) between downcore changes in the relative abundance of taxa that comprised at least 1% of the *trn*L-P6 assemblage and previously analyzed (in)organic paleoenvironmental parameters. This analysis revealed three clusters of paleovegetation: Cluster A, which includes Juglandaceae and the putative N-fixing legumes, was predominantly identified from and associated with the pre-lake landscape and early lake development (U2 and U1c). This cluster of pioneer forest vegetation revealed the highest positive Pearson correlation with sedimentary calcium content, which, according to end-member (EM) modeling, indicates increased tectonic movement and a high-energy depositional environment at the coring site (Morlock et al., 2021). Cluster B includes taxa that showed highly significant SIMPER associations with the organic-rich diatom ooze intervals of U1b (Figure 5), deposited during elevated nutrient availability and primary productivity. Notable examples are *Murraya* (Rutaceae), which includes small native peatland forest trees that were most likely growing in organic-rich and anoxic muddy sediments along the lakes' shoreline (i.e., in a similar fashion as mangrove vegetation), and aquatic herbs within the Apiaceae subfamily Apoideae related to nutrient-demanding water dropwort (*Oenanthe javanica*). The concomitant highest negative Pearson correlations with $\delta^{13}C_{org}$ and highest positive Pearson correlations with %TOC, TLE/TOC, and %Si (Figure 6) suggest that chloroplast-rich litter from this local C_3 paleovegetation contributed substantially to the immature (well-preserved) and ^{13}C-depleted sedimentary OM in the diatom oozes.

In contrast, Cluster C consisted of herbs (Brassicaceae, Fabaceae), TRSH (Theaceae), and C_3 grasses (Poaceae) that most strongly (positively) correlated with sediments of a predominantly ultramafic signature (Fe, Cr, Ni, Mg, and Mn oxides) (Figure 5). Lastly, paleovegetation of cluster D comprised predominantly lower montane, large evergreen rainforest trees and ground cover herbs that spread via rhizomes and prefer moist and shady closed canopy conditions (Zingiberaceae), as well as tree orchids (*Luisia*). The vegetation of Cluster D is most strongly correlated with the Al/Mg ratio, indicating a more substantial contribution to the drainage of sediments from the Loeha River during periods of increased precipitation (Morlock et al., 2019). A possible explanation for why this cluster contains most of the assumed C_4 grasses of the PACMAD clade will be discussed below. See Figure 6 for a complete overview of the Pearson correlations between downcore changes in paleovegetation assemblages and inorganic and isotopic for past changes in paleohydrology, paleoecology, and paleodepositional development of late-Pleistocene Lake Towuti.

4 | DISCUSSION

4.1 | Origin of contaminants and proof of concept for the successful and strict removal of contaminants

In deep biosphere studies, drilling muds are often used during coring as a lubricate to prevent the drilling bit from overheating. Microbial cell counts in these muds often exceed those of the low

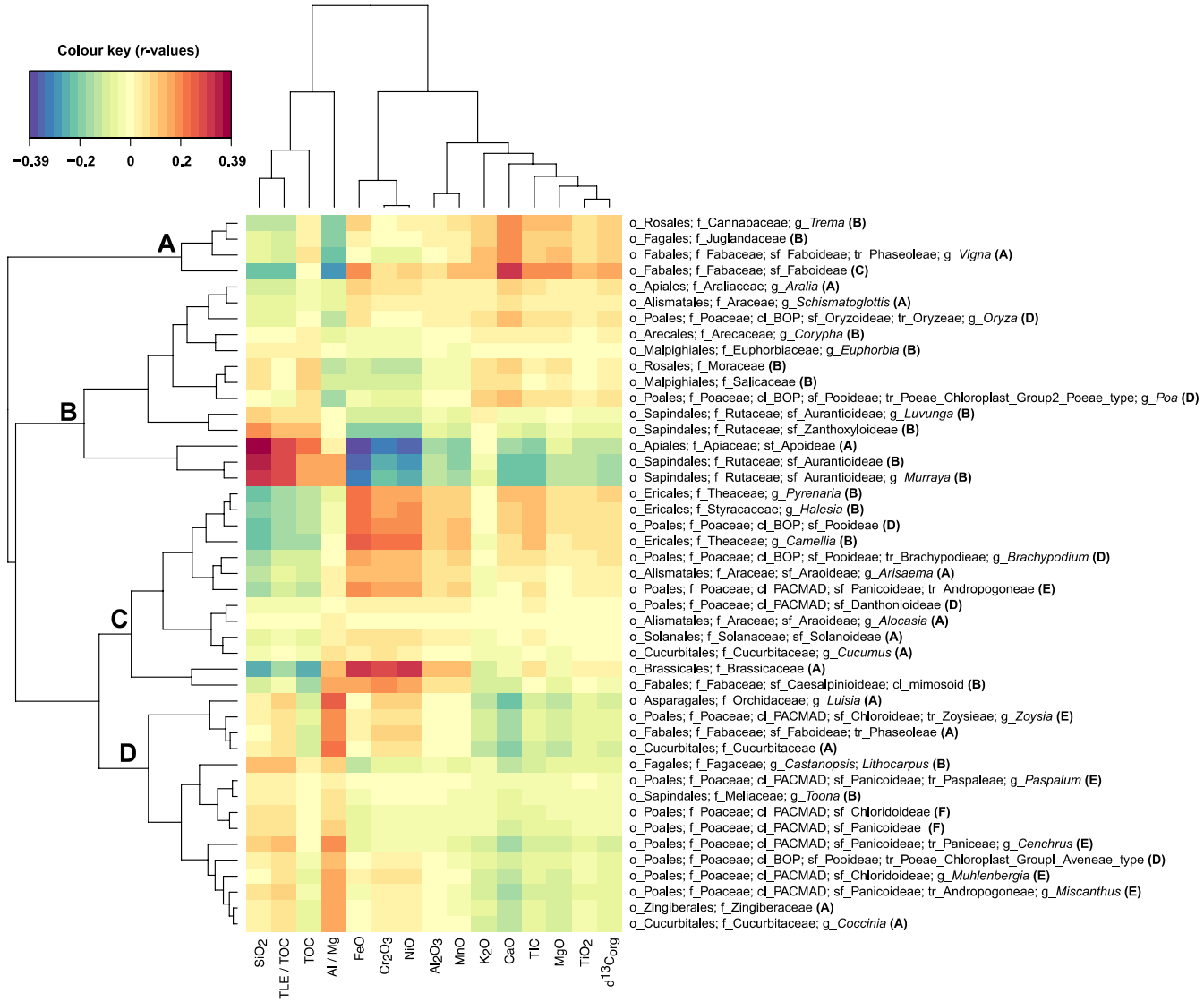

FIGURE 6 Clustered Image Map (CIM) showing Pearson correlations (*r* values in color key) between relative changes in *trn*L-P6-inferred paleovegetation communities and previously analyzed geochemical proxy data. Cluster A correlated most strongly with %Ca and comprised pioneering N-fixing trees of the legume family Fabaceae and N-fixing *Trema* (Cannabaceae) that prevailed during the pre-lake landscape characterized by active river channels, shallow lakes, and swamps. The upper part of Cluster B includes partially submerged vegetation that was likely rooted in muddy anoxic shoreline soils (e.g., *Oryza*, Alismatales) during early Lake Towuti. The lower part of Cluster B comprises peat swamp catchment vegetation (Rutaceae, notably *Murraya*) and nutrient-demanding aquatics (Apiaceae). This local vegetation prevailed during high nutrient availability that resulted in the subsequent deposition of diatom oozes (i.e., strong Pearson correlation with %Si) and when frequent stratified conditions and bottom water anoxia resulted in excellent preservation of SOM (high positive Pearson correlation with the TLE/TOC ratio, and %TOC). Cluster C comprises herbs (e.g., Brassicaceae), TRSH (e.g., Theaceae and Fabaceae), and mainly C_3 grasses (Poaceae), which most likely drained into Lake Towuti in the form of chloroplast-rich plant litter as the source of sedimentary *trn*L-P6 along with eroded ultramafic bedrock (Mg) and topsoil rich in Fe, Ni, and Cr during periods of reduced precipitation and low lake stands. The paleovegetation forming cluster D showed the highest positive Pearson correlations with the AL/Mg ratio indicative of predominant drainage from the Loeha River (Morlock et al., 2019) during elevated precipitation or periods of increased seasonality. See the main text for a more detailed discussion about the potential origin of the identified paleovegetation and inferred changes in the paleodepositional environment, plus the caption of Figure 2 for overlapping details and abbreviations.

biomass deep subsurface environment by orders of magnitude and, therefore, considered a major potential source of contamination with intact cells from surface-dwelling microbial communities (e.g., Cockell et al., 2021; Friese et al., 2017; Inagaki et al., 2015). Fluorescent particles (tracers) that mimic the shape and size of bacterial cells are usually mixed with drilling fluids so that after core retrieval, sediment slices can be analyzed microscopically to determine the extent and how deep these fluorescent tracers and possibly contaminant bacteria (could) have penetrated the core material (e.g., Friese et al., 2017). Drilling fluids were only required to lubricate the drill bit during Alien rotating coring to recover the more resistant lithologies of Unit 2 plus the 2-m-thick

red clay interval between 70.5 and 72.5 mcd. Therefore, contamination of the lacustrine sediments of Unit 1 with drilling fluids can be excluded and would not have impacted the observed Pearson correlations between relative changes in the *trn*L-P6 paleovegetation community composition and the geochemical parameters (Figure 6), which informed about concomitant changes in the paleodepositional environment of Towuti Lake.

While fluorescent tracers have not been used during coring of 1F, microscopic analysis performed on U1 sediments from a parallel core from site 1 (core 1A) showed that bacterial cell-sized fluorescent tracer particles only penetrated the outer few mm of sediment (Friese et al., 2017). Pollen grains represent the smallest plant diagnostic morphological features as a potential source of contamination with chloroplast DNA during coring. Being substantially larger than bacterial cells renders it highly unlikely that modern pollen sourced from the water column or surface sediments would have penetrated and contaminated the center of the core, which was exclusively subsampled for DNA extraction in this study. In contrast, contamination with airborne pollen during subsampling or DNA extraction and sequence library preparation is a well-known concern in *sed*aDNA studies. Most notably, Gymnosperms (conifers) are a common source of contamination as this tree vegetation produces vast amounts of airborne pollen that can travel large distances (e.g., Boessenkool et al., 2014; Pedersen et al., 2013). Stevenson (2018) reported that native montane conifers (*Agathis*, *Dacrydium*, *Phyllocladus*, and *Podocarpus*) combined comprised up to ~30% of the total pollen count in the Glacial and Holocene sediments from the TOW-09 record previously obtained from the same coring location. However, none of the *trn*L-P6 reads in the samples and controls could be assigned to these genera or their respective families. Therefore, chloroplast-rich plant litter from local catchment vegetation was the predominant source of paleovegetation identified in this study, which rules out the possibility that locally or remotely produced modern pollen represented a source of contamination during coring. Instead, all identified reads that occurred in both samples as well as in the background and extraction blanks were assigned to taxa not native to Sulawesi and/or to the common lab contaminants *Musa* (Musaceae), Nolinoideae (Asparagaceae), and *Pinus*/*Picea* (Pinaceae) (e.g., Boessenkool et al., 2014; Pedersen et al., 2013).

Indirect evidence that the randomly distributed contaminants were accurately identified and stringently removed from the *trn*L-P6 data was provided by (a) the significant concomitant shifts in *trn*L-P6-inferred paleovegetation with transitions in the paleodepositional environment of Lake Towuti and (b) the identification of plausible significant indicator taxa associated with the paleodepositional categories. Further evidence for a non-random downcore distribution of *trn*L-P6 from native taxa was provided by (c) the moderately strong Pearson correlations between relative changes in paleovegetation assemblages and geochemical parameters, indicative of simultaneous changes in the paleohydrology and paleodepositional environment of Lake Towuti as discussed in more detail below.

4.2 | Factors contributing to the long-term preservation of vegetation *sed*aDNA in Lake Towuti

Bacteria, archaea, and eukaryotes contribute to lake sediments' total pool of extractable DNA. They include a mix of intracellular DNA associated with intact cells and extracellular DNA released into the environment after the lysis of dead cells (e.g., Parducci et al., 2017; Vuillemin et al., 2017; and references therein). Intracellular DNA from intact and physiologically active microbial communities dominated the total pool of extractable DNA in the surface sediments, while the relative abundance of extracellular sedimentary environmental DNA from dead bacterial biomass increased with sediment depth in Lake Towuti (Vuillemin et al., 2017, 2019). These findings showed that a substantial amount of *sed*aDNA escaped initial microbial degradation in the surface sediments of Lake Towuti.

Since we used a non-selective destructive DNA extraction approach, it was impossible to verify whether plant taxa were identified from the pool of mineral-adsorbed extracellular DNA or from DNA still associated with intact chloroplasts in plant litter still undergoing microbial decomposition. Upon burial in the sedimentary record, microbial communities would be expected to preferentially degrade more reactive plant litter (e.g., leaves) and its associated DNA favoring the preservation of more recalcitrant wood structures. However, bias caused by preferential preservation of *sed*aDNA from woody plants and trees is unlikely to have played a substantial role in the outcomes of this study. Namely, leaves were likely the most abundant form of drained and sequestered plant biomass for all identified plant categories (herbs, TRSH, and grasses) and the richest source of chloroplasts per volume of plant remains. Moreover, leafy herbs and TRSH were equally well presented throughout the record (Figures 3 and 4).

Mineral adsorption of the negatively charged phosphate group in DNA contributes to extracellular DNA's immediate- and long-term protection against degradation by microbial nucleases (e.g., Coolen & Overmann, 2007; Demaneche et al., 2001; Freeman et al., 2023; Pietramellara et al., 2009; Vuillemin et al., 2017). Mineral adsorption in Lake Towuti's sediment is mainly facilitated due to the high amount of ferruginous clays and clay-sized particles (Vuillemin et al., 2017), making up >60% of the sediment particle composition (Hasberg et al., 2019). Similarly, efficient adsorption of phosphate by sedimentary iron hydroxides, especially in a low-sulfate environment like Lake Towuti, is thought to have contributed to ultraoligotrophic conditions and concomitant low primary productivity in the oxygenated sunlit surface waters of Lake Towuti during the deposition of the sideritic red clays (Caraco et al., 1989; Crowe et al., 2008).

In addition, as has been reported from many other stratified and anoxic basins (e.g., Coolen et al., 2004, 2013; Coolen & Overmann, 2007), bottom water anoxia in a stratified Lake Towuti would have resulted in lower microbial mineralization rates, and hence, an increase in the accumulation of sedimentary OM and the fraction of *sed*aDNA that escaped early degradation during

the deposition of the green clays as compared to red clays, which were deposited in a fully mixed and oxygenated water column (Russell et al., 2020).

Positively charged mineral surfaces generally exhibit a high affinity for DNA in any environment. In contrast, DNA binds to negatively charged surfaces, like non-clay silicates, via a cation bridge, and desorption will occur if the concentration or the ionic potential of the immediate environment is decreased (e.g., Pastré et al., 2006). Therefore, the adsorption strength of *sed*aDNA to the silicified diatom ooze material is likely lower than that of the clays. In that case, a low microbial mineralization rate under prevailing anoxic conditions was likely the main factor contributing to the accumulation of the immature/labile OM and the long-term survival of *sed*aDNA against microbial attack in the diatom ooze intervals.

We recovered *trn*L-P6 of wetland herbs such as *Alocasia* and Zingiberaceae in the oldest peat intervals of U2. This partially submerged vegetation was likely rooted in anoxic muddy soils in the peatland-dominated pre-lake landscape and could have survived waterlogged conditions through plant-mediated rhizosphere oxygenation (e.g., Koop-Jakobsen et al., 2021; Nijburg et al., 1997). Low microbial activities likely prevailed in the bulk of anoxic soils not associated with the oxygenated rhizosphere. Since the DNA adsorption capacity of sedimentary OM is deficient (Coolen & Overmann, 2007), reduced microbial mineralization rates under oxygen-depleted conditions likely also contributed most to the accumulation of organic matter and more labile biomolecules, including DNA in the peats of U2.

The strong mineral adsorption to the ferruginous clays most likely contributed uniquely to the long-term preservation of *sed*aDNA in tropical Lake Towuti (Friese et al., 2017). However, even in relatively productive and fully mixed oxygenated tropical lakes, oxygen as the highest-energy-yielding electron acceptor usually becomes depleted in the top few cm of sediments (e.g., Wang et al., 2022). Therefore, even if mineral adsorption did not play a substantial role in the long-term preservation of *sed*aDNA, reduced rates of microbial mineralization leading to sedimentary OM accumulation and enhanced preservation of *sed*aDNA in anaerobic subsurface sediments are likely to be relatively widespread in mesotrophic and eutrophic tropical lake environments and warrants further investigation.

4.3 | Insights into the developmental history of Lake Towuti from *sed*aDNA-inferred changes in catchment vegetation

4.3.1 | Taxa associated with the pre-lake landscape (U2) and early lacustrine development of Lake Towuti (U1C)

In agreement with previous mineralogical and geochemical findings (Russell et al., 2020), the paleovegetation identified from U2 indicates a pre-lake landscape characterized by shallow lakes and swamps before transitioning into a permanent Lake Towuti. For example, *Alocasia* was most consistently detected in U2. These wetland herbs prefer partial shade, spread through aerenchymous rhizomes, and can be invasive in swampy areas, but they need shallow waters and will not survive when completely submerged (Markesteijn et al., 2007). In addition, the presence of N-fixing pioneer vegetation indicates the presence of relatively disturbed and nutrient-depleted soils before developing a partly forested peat swamp. For example, *Trema* (Cannabaceae) was a source of sedimentary *trn*L-P6 in four consecutive intervals, marking the U2/U1c boundary and a significant indicator species for U2. ASVs derived from *Trema* cp DNA showed >99% similarity with *Trema orientale*, a fast-growing, N-fixing pioneer tree native to Sulawesi, which can tolerate poor soils and is often found in disturbed areas and tropical forest margins. It can provide shade and improve soil fertility to stimulate the germination of seedlings of secondary tropical hardwood trees (Haberle et al., 2006). *Trema* was accompanied by N-fixing pioneer herbs of the legume family Fabaceae (i.e., *Vigna*), which continued to be consistent throughout lacustrine U1c until ~76 mcd (i.e., ~0.8 Ma).

The timing of this shift in paleovegetation assemblages coincided with the U1c/U1b boundary (Russell et al., 2020). This pronounced change in the paleodepositional development of Lake Towuti seems to have had a more considerable impact on the local catchment vegetation that entered the lake in the form of chloroplast-rich litter than the transition into the permanent lake (i.e., U2/U1c boundary), which did not result in significant changes in *trn*L-P6-inferred paleovegetation assemblages. We infer that considerable tectonic movements and catchment morphological adjustments continued well past the U2/U1c boundary until ~0.8 Ma.

4.3.2 | Paleovegetation associated with increased drainage of (ultra)mafic substrates

Serpentine or ultramafic substrates are harsh environments for vegetation due to their low nutrient content (N, P, and K), the imbalance of Ca and Mg, and the presence of a high concentration of potentially phytotoxic trace elements, notably Ni and Cr (Garnier et al., 2021; Kazakou et al., 2008). Yet, ultramafic ecosystems are often biodiversity hotspots and host numerous endemic species (Myers et al., 2000). These include vegetation types that have developed strategies such as drastically lowering the impact of phytotoxic trace elements through hyperaccumulation of heavy metals (Prasad & Freitas, 2003) and the release of exudates in the rhizosphere to solubilize iron minerals to release Fe^{2+} needed for several processes including photosynthesis (Morrissey & Guerinot, 2009 and references therein). Brassicaceae, followed by Poaceae, the unclassified Faboideae (Fabaceae), and Theaceae had the strongest positive Pearson correlations with %Fe, %Ni, and %Cr derived from the drainage of ultramafic rocks from the lake catchment. These plant families are known to include many highly efficient taxa in extracting bioavailable Fe^{2+} from soils containing only biologically inert iron minerals (Morrissey & Guerinot, 2009 and references therein). Moreover, Brassicaceae

has the highest number of species capable of hyperaccumulating Ni and Cr (Prasad & Freitas, 2003). The strong Pearson correlations between downcore changes in the relative abundances of these taxa and %Ni, %Cr, and %Fe suggest that their chloroplast-rich plant litter drained into Lake Towuti mixed with eroded ultramafic sediment.

Ultramafic sediments rich in Ni, Cr, and Fe are deposited in Lake Towuti under drier conditions and lower lake levels (Russell et al., 2020). Drier conditions were pronounced during the LGM when the lake water level was 10–35 m lower than today (Vogel et al., 2015). The strong drying signal is derived primarily from the stable isotope measurements of terrestrial leaf waxes ($\delta^{13}C_{wax}$) (Russell et al., 2014). Vegetation utilizing the C_3 photosynthetic pathway (the bulk of herbaceous and woody plant life) has $\delta^{13}C$ values between −29‰ to −38‰, whereas plants utilizing the C_4 pathway (primarily tropical and warm-season grasses) have $\delta^{13}C$ values of −14‰ to −26‰ (Bi et al., 2005; Chikaraishi & Naraoka, 2003). Therefore, the observed enrichment of 15‰ over the LGM from approximately −40‰ to −25‰ was interpreted as an expansion of C_4 grasses in a savanna-like landscape with reduced tree cover (Russell et al., 2014).

In contrast, the pollen record from core TOW-09 revealed that tropical lowland tree forest vegetation was still abundantly present during the LGM, while Cyperaceae (sedges) outcompeted grasses as the most abundant C_4 vegetation (Stevenson, 2018). Therefore, instead of a widespread savanna-type open landscape, the pronounced lake-level low stands during the LGM most likely resulted in the expansion of the lakes' foreshore surface area covered by sedges as the predominant C_4 plants (Stevenson, 2018). Moreover, a local record of the deglacial period showed limited evidence of broad-scale grasslands in the lowlands (Hamilton, Stevenson, et al., 2019). Similarly, in this study, we do not find support for an expansion of C_4 grasses in the region during the LGM. Instead, the LGM red clays revealed the presence of sedimentary *trn*L-P6 assigned to wet-climate C_3 grasses. The inundation regime is a main driver shaping the vegetation community structure in wetlands (Baldwin et al., 1996). Most likely, the C_3 vegetation identified using *trn*L-P6 metabarcoding was adapted to or could have tolerated long-term exposure to flooding and inundation near the deeper coring location. Cyperaceae produce large amounts of wind-fertilized pollen that can travel long distances (Stevenson, 2018). The inability to detect the chloroplast marker gene assigned to Cyperaceae in our dataset implies that this vegetation was zoned further away from the coring location at more elevated, drier areas of the lakes' foreshore.

4.3.3 | Paleovegetation associated with increased drainage of lateritic soils and felsic substrates from the Loeha River

The paleovegetation forming cluster D (Figure 6) showed the strongest positive Pearson correlations with the Al/Mg ratio, indicating increased drainage of lateritic soils and felsic substrates from the Loeha River to the east of Lake Towuti. Previous studies (Morlock et al., 2019; Russell et al., 2020) inferred that these sediments were deposited during increased precipitation and lake-level high stands. The vegetation that forms cluster D indicates a warm, humid paleoenvironment with tropical evergreen trees (*Toona*) that created a closed canopy offering shady and moist conditions for tree orchids (*Luisia*) and ground cover herbs (Zingiberaceae), known to spread via rhizomes. The association between cluster D and the Al/Mg ratio indicates that this vegetation was widespread in the catchment area of the Loeha River. However, cluster D also included the PACMAD grasses identified as genera of the subfamilies Chloridoideae and Panicoideae that exclusively use the C_4 carbon fixation pathway (e.g., *Zoysia*, *Muhlenbergia*). In addition, cluster D included *trn*L-P6 assigned to *Castanopsis* and/or *Lithocarpus* that was most consistently present in U1a and U1b intervals, including the red clays from the LGM. As mentioned earlier, the presence of these drought-tolerant tropical evergreen trees that now dominate at lower/mid-montane elevations in Sulawesi (Biagioni et al., 2015; Culmsee et al., 2010) during the LGM was independently confirmed from adjacent fossil pollen records (Hamilton, Stevenson, et al., 2019; Stevenson, 2018). A mixture of *trn*L-P6 sourced from C_3 and C_4 vegetation in Cluster D implies that periods of increased seasonal rainfall prevailed and contributed to a predominance of felsic drainage.

4.4 | Nutrient-demanding aquatic herbs and partially submerged shoreline vegetation during stratification, mesotrophic conditions, and the deposition of diatom oozes

Rutaceae (subfamily Aurantioideae) and Apiaceae (subfamily Apoideae) were notable representatives of Cluster B paleovegetation. ASVs assigned to Aurantioideae showed the highest-sequence similarities with *Murraya* and *Luvunga*, small SE Asian swamp woodland trees native to Sulawesi (Kessler et al., 2002), while ASVs assigned to Apoideae showed up to 100% sequence similarity to *trn*L-P6 of native Java water dropwort (*Oenanthe javanica*). This perennial aquatic herb is widespread in Southeast Asia (http://www.powo.science.kew.org), where it is found in grassy forest margins, marshes, wet meadows, lake shores, and muddy rivers or stream banks and can become somewhat invasive when rooting in nutrient-rich soils or sediments. The upper part of cluster B also included *Oryza*, a genus of wetland C_3 grasses native to Sulawesi. The positive Pearson correlations between relative changes in their *trn*L-P6 content with %TOC and the TLE/TOC ratio suggest that this vegetation was rooted in muddy, organic-rich shoreline sediments. Anoxic conditions in the muddy sediments likely contributed to preserving the labile sedimentary OM and releasing bioavailable ferrous iron to support plant growth (Russell et al., 2020). An increase in nutrient availability combined with the slow release of

tephra-bound phosphorus (Russell et al., 2020) could have contributed to the expansion of Java water dropwort, which coincided with the formation of diatom blooms, as evident from the correlation between changes in the relative abundance of sedimentary *trn*L-P6 of this aquatic herb with silica content. The fact that sedimentary *trn*L-P6 from this shoreline vegetation and not that of Cyperaceae could be identified from the sedimentary record implies that the chloroplast-rich litter from the vegetation of cluster D was more efficiently transported to the coring location because lake levels were higher than during the dry LGM.

5 | CONCLUSIONS

We successfully amplified and sequenced sedimentary *trn*L-P6 from up to 114-m-deep and over 1-Ma-old tropical Lake Towuti sediments. After stringent removal of contaminants and unassigned/poorly classified ASVs, the remaining ASVs could be assigned to at least 45 native taxa at family or lower taxonomic ranks. Downcore changes in the relative abundance of these paleovegetation assemblages reflected changes in the paleodepositional environment as inferred from Pearson correlations with previously analyzed organic and inorganic parameters during the lakes' more than 1 Myr history. Nitrogen-fixing pioneer vegetation and shallow wetland herbs were most strongly associated with >1-Ma-old peats and silts deposited in a tectonically active landscape of active river channels, shallow lakes, and peat swamps. A statistically significant shift in the paleovegetation was observed ~200 ka after the transition into a permanent lake (i.e., ~800 kyr ago), which coincided with a decrease in tectonic activity and catchment adjustments. Most notably, the newly emerged shoreline vegetation comprised putative peatland forest trees and partially submerged C_3 grasses (Oryzaceae). Positive Pearson correlations between relative changes in their *trn*L-P6 content with %TOC and the TLE/TOC ratio suggest that this vegetation grew and rooted in muddy organic-rich shoreline sediments. Stratified and anoxic conditions likely contributed to preserving the labile sedimentary OM, and bioavailable Fe(II) would have been readily available. An increase in nutrient availability from frequent turnover of the water column combined with the slow release of tephra-bound phosphorus could have contributed to the expansion of Java water dropwort, which coincided with the formation of diatom blooms, as evident from the correlation between changes in the relative abundance of sedimentary *trn*L-P6 of this aquatic herb with silica content. In contrast, herbs (Brassicaceae, Fabaceae), trees/shrubs (Theaceae), and C_3 grasses (Pooideae) showed the highest positive Pearson correlations with sediments of an ultramafic signature (e.g., rich in Fe, Ni, and Cr). These plant families include the highest number of taxa with highly efficient strategies to extract bioavailable Fe(II) from iron-rich rocks and are capable of hyperaccumulation of phytotoxic metals, including Ni and Cr. Therefore, this vegetation was likely adapted to grow on the ultramafic rocks within the lake's catchment, and their chloroplast-rich biomass drained into Lake Towuti mixed with eroded ultramafic substrates during drier periods. Rainforest trees (e.g., *Toona*), shady ground cover herbs (Zingiberaceae), and tree orchids (*Luisia*) showed the highest Pearson correlations with inorganic parameters indicating increased drainage of felsic substrates from the Loeha River to the east of Lake Towuti during periods of increased precipitation. However, the co-presence of sedimentary *trn*L-P6 from dry climate-adapted vegetation (i.e., C_4 grasses and *Castanopsis/Lithocarpus*) implies that more seasonal climates also resulted in a predominance of felsic drainage, not just wetter conditions.

While the strong mineral adsorption to the ferruginous clays most likely uniquely contributed to the long-term preservation of *sed*aDNA in Lake Towuti, enhanced preservation of *sed*aDNA in anaerobic- and organic-rich subsurface sediments is likely to be relatively widespread in mesotrophic and eutrophic tropical lake environments. Therefore, to what extent tropical lake *sed*aDNA records could complement existing proxies in reconstructing local terrestrial and aquatic ecosystem changes and their responses to paleoenvironmental perturbations warrants further investigation. The limited availability of reference sequences of metabarcoding genes from tropical forest vegetation in public databases, notably from this tropical biodiversity hotspot, could explain why roughly half of the ASVs did not have close similarities with any of the vast amount of DNA sequences available in the NCBI nr database. In addition, a subset of unassigned ASVs could also represent PCR artifacts. Future studies could use hybridization capture-targeting multiple metabarcoding genes (e.g., Foster et al., 2022) to eliminate PCR bias and to improve the overall taxonomic resolution of tropical paleovegetation.

ACKNOWLEDGMENTS

This work was primarily supported by the Australian Research Council (ARC), Discovery Grant # DP15102587. The Curtin Office provided additional financial support for Research and Development (ORF) for funding AE's PhD stipend. This research was furthermore carried out with partial support from the International Continental Scientific Drilling Program (ICDP), the U.S. National Science Foundation (NSF), the Swiss National Science Foundation (SNSF; grant no: 200021_153053), PT Vale Indonesia, the Ministry of Research, Education, and Higher Technology of Indonesia (RISTEK), Brown University, and The Institute for Geoscience Research (TIGeR) at Curtin University. We thank PT Vale Indonesia, the US Continental Scientific Drilling and Coordination Office, the US National Lacustrine Core Repository, and DOSECC Exploration Services for logistical support. The research was carried out with permission from RISTEK, the Ministry of Trade of the Republic of Indonesia, the Natural Resources Conservation Centre (BKSDA), and the Government of Luwu Timur of Sulawesi. We thank Dr Cornelia Wuchter (Curtin University) for helpful discussions. Open access publishing facilitated by Curtin University, as part of the Wiley - Curtin University agreement via the Council of Australian University Librarians.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the Short Reads Archive under Bioproject ID PRJNA1105239 at https://www.ncbi.nlm.nih.gov/sra.

ORCID

Md Akhtar-E Ekram https://orcid.org/0000-0002-0446-7789
Chloe Plet https://orcid.org/0000-0003-1959-6732
Marco J. L. Coolen https://orcid.org/0000-0002-0417-920X

REFERENCES

Aliscioni, S., Bell, H. L., Besnard, G., Christin, P. A., Columbus, J. T., Duvall, M. R., Edwards, E. J., Giussani, L., Hasenstab-Lehman, K., Hilu, K. W., Hodkinson, T. R., Ingram, A. L., Kellogg, E. A., Mashayekhi, S., Morrone, O., Osborne, C. P., Salamin, N., Schaefer, H., Spriggs, E., … Grass Phylogeny Working, G., II. (2012). New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytologist*, *193*(2), 304–312.

Alsos, I. G., Lammers, Y., Kjellman, S. E., Merkel, M. K. F., Bender, E. M., Rouillard, A., Erlendsson, E., Gudmundsdottir, E. R., Benediktsson, I. O., Farnsworth, W. R., Brynjolfsson, S., Gisladottir, G., Eddudottir, S. D., & Schomacker, A. (2021). Ancient sedimentary DNA shows rapid post-glacial colonization of Iceland followed by relatively stable vegetation until the Norse settlement (Landnam) AD 870. *Quaternary Science Reviews*, *259*, 106903.

Alsos, I. G., Lammers, Y., Yoccoz, N. G., Jorgensen, T., Sjogren, P., Gielly, L., & Edwards, M. E. (2018). Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLoS One*, *13*(4), e0195403.

Anshari, G., Kershaw, P., van der Kaars, S., & Jacobsen, G. (2004). Environmental change and peatland forest dynamics in the Lake Sentarum area, West Kalimantan, Indonesia. *Journal of Quaternary Science*, *19*, 637–655.

Armbrecht, L., Weber, M. E., Raymo, M. E., Peck, V. L., Williams, T., Warnock, J., Kato, Y., Hernández-Almeida, I., Hoem, F., Reilly, B., Hemming, S., Bailey, I., Martos, Y. M., Gutjahr, M., Percuoco, V., Allen, C., Brachfeld, S., Cardillo, F. G., Du, Z., … Zheng, X. (2022). Ancient marine sediment DNA reveals diatom transition in Antarctica. *Nature Communications*, *13*(1), 5787.

Baldwin, A. H., McKee, K. L., & Mendelssohn, I. A. (1996). The influence of vegetation, salinity, and inundation on seed banks of oligohaline coastal marshes. *American Journal of Botany*, *83*, 470–479.

Bi, X., Sheng, G., Liu, X., Li, C., & Fu, J. (2005). Molecular and carbon and hydrogen isotopic composition of n-alkanes in plant leaf waxes. *Organic Geochemistry*, *36*(10), 1405–1417.

Biagioni, S., Wundsch, M., Haberzettl, T., & Behling, H. (2015). Assessing resilience/sensitivity of tropical mountain rainforests towards climate variability of the last 1500 years: The long-term perspective at Lake Kalimpaa (Sulawesi, Indonesia). *Review of Palaeobotany and Palynology*, *213*, 42–53.

Boessenkool, S., McGlynn, G., Epp, L. S., Taylor, D., Pimentel, M., Gizaw, A., Nemomissa, S., Brochmann, C., & Popp, M. (2014). Use of ancient sedimentary DNA as a novel conservation tool for high-altitude tropical biodiversity. *Conservation Biology*, *28*(2), 446–455.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). OBITOOLS: A UNIX-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182.

Bremond, L., Favier, C., Ficetola, G. F., Tossou, M. G., Akouegninou, A., Gielly, L., Giguet-Covex, C., Oslisly, R., & Salzmann, U. (2017). Five thousand years of tropical lake sediment DNA records from Benin. *Quaternary Science Reviews*, *170*, 203–211.

Cannon, C. H., Morley, R. J., & Bush, A. B. (2009). The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(27), 11188–11193.

Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P. D., Vuillemin, A., Ariztegui, D., Arnaud, F., Belle, S., Bertilsson, S., Bigler, C., Bindler, R., Brown, A. G., Clarke, C. L., Crump, S. E., Debroas, D., Englund, G., Ficetola, G. F., Garner, R. E., … Parducci, L. (2021). Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: Overview and recommendations. *Quaternary*, *4*, 6. https://doi.org/10.3390/quat4010006

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., & Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal*, *6*(8), 1621–1624.

Caraco, N. F., Cole, J. J., & Likens, G. E. (1989). Evidence for sulfate-controlled phosphorus release from sediments of aquatic systems. *Nature*, *341*(6240), 316–318.

Chang, Y. M., Chang, C. L., Li, W. H., & Shih, A. C. C. (2013). Historical profiling of maize duplicate genes shed light on the evolution of C4 photosynthesis in grasses. *Molecular Phylogenetics and Evolution*, *66*(2), 453–462.

Chikaraishi, Y., & Naraoka, H. (2003). Compound-specific δD–δ13C analyses of n-alkanes extracted from terrestrial and aquatic plants. *Phytochemistry*, *63*(3), 361–371.

Clarke, K. R., & Gorley, R. N. (2015). *PRIMER v7: User manual/tutorial*. PRIMER-E Ltd.

Cockell, C. S., Schaefer, B., Wuchter, C., Coolen, M. J. L., Grice, K., Schnieders, L., Morgan, J. V., Gulick, S. P. S., Wittmann, A., Lofi, J., Christeson, G. L., Kring, D. A., Whalen, M. T., Bralower, T. J., Osinski, G. R., Claeys, P., Kaskes, P., de Graaff, S. J., Déhais, T., … IODP-ICDP Expedition 364 Scientists. (2021). Shaping of the present-day deep biosphere at chicxulub by the impact catastrophe that ended the cretaceous. *Frontiers in Microbiology*, *12*, 668240. https://doi.org/10.3389/fmicb.2021.668240

Coolen, M. J. L., Muyzer, G., Rijpstra, W. I. C., Schouten, S., Volkman, J. K., & Damste, J. S. S. (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth and Planetary Science Letters*, *223*(1–2), 225–239.

Coolen, M. J. L., Orsi, W. D., Balkema, C., Quince, C., Harris, K., Sylva, S. P., Filipova-Marinova, M., & Giosan, L. (2013). Evolution of the plankton paleome in the Black Sea from the Deglacial to Anthropocene. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(21), 8609–8614.

Coolen, M. J. L., & Overmann, J. (2007). 217 000-year-old DNA sequences of green sulfur bacteria in Mediterranean sapropels and their implications for the paleoenvironment reconstruction. *Environmental Microbiology*, *9*(1), 238–249.

Costa, K. M., Russell, J. M., Vogel, H., & Bijaksana, S. (2015). Hydrological connectivity and mixing of Lake Towuti, Indonesia in response to paleoclimatic changes over the last 60,000 years. *Palaeogeography Palaeoclimatology Palaeoecology*, *417*, 467–475.

Courtin, J., Andreev, A. A., Raschke, E., Bala, S., Biskaborn, B. K., Liu, S. S., Zimmermann, H., Diekmann, B., Stoof-Leichsenring, K. R., Pestryakova, L. A., & Herzschuh, U. (2021). Vegetation changes in southeastern Siberia during the late Pleistocene and the Holocene. *Frontiers in Ecology and Evolution*, *9*, 625096.

Crowe, S. A., O'Neill, A. H., Katsev, S., Hehanussa, P., Haffner, G. D., Sundby, B., Mucci, A., & Fowle, D. A. (2008). The biogeochemistry of tropical lakes: A case study from Lake Matano, Indonesia. *Limnology and Oceanography*, *53*(1), 319–331.

Culmsee, H., Leuschner, C., Moser, G., & Pitopang, R. (2010). Forest aboveground biomass along an elevational transect in Sulawesi,

Indonesia, and the role of Fagaceae in tropical montane rain forests. *Journal of Biogeography*, *37*(5), 960–974.

De Caceres, M., & Legendre, P. (2009). Associations between species and groups of sites: Indices and statistical inference. *Ecology*, *90*(12), 3566–3574.

Demaneche, S., Jocteur-Monrozier, L., Quiquampoix, H., & Simonet, P. (2001). Evaluation of biological and physical protection against nuclease degradation of clay-bound plasmid DNA. *Applied and Environmental Microbiology*, *67*(1), 293–299.

Direito, S. O. L., Marees, A., & Roling, W. F. M. (2012). Sensitive life-detection strategies for low-biomass environments: Optimizing the extraction of nucleic acids adsorbing to terrestrial and Mars analogue minerals. *FEMS Microbiology Ecology*, *81*(1), 111–123.

Dommain, R., Andama, M., McDonough, M. M., Prado, N. A., Goldhammer, T., Potts, R., Maldonado, J. E., Nkurunungi, J. B., & Campana, M. G. (2020). The challenges of reconstructing tropical biodiversity with sedimentary ancient DNA: A 2200-year-long metagenomic record from Bwindi impenetrable Forest, Uganda. *Frontiers in Ecology and Evolution*, *8*, Article 218. https://doi.org/10.3389/fevo.2020.00218

Epp, L. S., Gussarova, C., Boessenkool, S., Olsen, J., Haile, J., Schroder-Nielsen, A., Ludikova, A., Hassel, K., Stenoien, H. K., Funder, S., Willerslev, E., Kjaer, K., & Brochmann, C. (2015). Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes. *Quaternary Science Reviews*, *117*, 152–163.

Foster, N. R., van Dijk, K. J., Biffin, E., Young, J. M., Thomson, V. A., Gillanders, B. M., Jones, A. R., & Waycott, M. (2022). A targeted capture approach to generating reference sequence databases for chloroplast gene regions. *Ecology and Evolution*, *12(4)*, e8816.

Freeman, C. L., Dieudonné, L., Agbaje, O. B. A., Žure, M., Sanz, J. Q., Collins, M., & Sand, K. K. (2023). Survival of environmental DNA in sediments: Mineralogic control on DNA taphonomy. *Environmental DNA*, *5*, 1–15.

Friese, A., Bauer, K., Glombitza, C., Ordoñez, L., Ariztegui, D., Heuer, V. B., Vuillemin, A., Henny, C., Nomosatryo, S., Simister, R., Wagner, D., Bijaksana, S., Vogel, H., Melles, M., Russell, J. M., Crowe, S. A., & Kallmeyer, J. (2021). Organic matter mineralization in modern and ancient ferruginous sediments. *Nature Communications*, *12*(1), 2216. https://doi.org/10.1038/s41467-021-22453-0

Friese, A., Kallmeyer, J., Kitte, J. A., Martínez, I. M., Bijaksana, S., & Wagner, D. (2017). A simple and inexpensive technique for assessing contamination during drilling operations. *Limnology and Oceanography: Methods*, *15*(2), 200–211.

Garnier, J., Quantin, C., Raous, S., Guimaraes, E., & Becquer, T. (2021). Field availability and mobility of metals in Ferralsols developed on ultramafic rock of Niquelandia, Brazil. *Brazilian Journal of Geology*, *51*(1), e20200092. https://doi.org/10.1590/2317-4889202120200092

Gregory, P. H. (1978). Distribution of airborne pollen and spores and their long-distance transport. *Pure and Applied Geophysics*, *116*(2–3), 309–315.

Haberle, S. G., Tibby, J., Dimitriadis, S., & Heijnis, H. (2006). The impact of European occupation on terrestrial and aquatic ecosystem dynamics in an Australian tropical rain forest. *Journal of Ecology*, *94*(5), 987–1002.

Hamilton, R., Hall, T., Stevenson, J., & Penny, D. (2019). Distinguishing the pollen of Dipterocarpaceae from the seasonally dry and moist tropics of south-east Asia using light microscopy. *Review of Palaeobotany and Palynology*, *263*, 117–133.

Hamilton, R., Stevenson, J., Li, B., & Bijaksana, S. (2019). A 16,000-year record of climate, vegetation and fire from Wallacean lowland tropical forests. *Quaternary Science Reviews*, *224*, 105929.

Hasberg, A. K. M., Bijaksana, S., Held, P., Just, J., Melles, M., Morlock, M. A., Opitz, S., Russell, J. M., Vogel, H., & Wennrich, V. (2019). Modern sedimentation processes in Lake Towuti, Indonesia, revealed by the composition of surface sediments. *Sedimentology*, *66*(2), 675–698.

Hope, G. (2001). Environmental change in the late Pleistocene and later Holocene at Wanda site, Soroako, South Sulawesi, Indonesia. *Palaeogeography Palaeoclimatology Palaeoecology*, *171*(3–4), 129–145.

Hopkins, H. C. F. (1994). The indo-Pacific species of *Parkia* (Leguminosae: Mimosoideae). *Kew Bulletin*, *49*(2), 181–234.

Inagaki, F., Hinrichs, K.-U., Kubo, Y., Bowles, M. W., Heuer, V. B., Hong, W.-L., Hoshino, T., Ijiri, A., Imachi, H., Ito, M., Kaneko, M., Lever, M. A., Lin, Y.-S., Methé, B. A., Morita, S., Morono, Y., Tanikawa, W., Bihan, M., Bowden, S. A., … Yamada, Y. (2015). Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. *Science*, *349*(6246), 420–424.

IUCN. (2019). *International Union for Conservation of Nature annual report 2019*. https://portals.iucn.org/library/node/49096

Kazakou, E., Dimitrakopoulos, P. G., Baker, A. J. M., Reeves, R. D., & Troumbis, A. Y. (2008). Hypotheses, mechanisms, and trade-offs of tolerance and adaptation to serpentine soils: From species to ecosystem level. *Biological Reviews*, *83*(4), 495–508.

Kessler, P. J. A., Bos, M., Sierra+Daza, S. E. C., Kop, A., Willemse, L., Pitopang, R., & Gradstein, S. (2002). Checklist of woody plants of Sulawesi, Indonesia. *Blumea Journal of Plant Taxonomy and Plant Geography*, *14*, 1–160.

Kirkpatrick, J. B., Walsh, E. A., & D'Hondt, S. (2016). Fossil DNA persistence and decay in marine sediment over hundred-thousand-year to million-year time scales. *Geology*, *44*(8), 615–618.

Koop-Jakobsen, K., Meier, R. J., & Mueller, P. (2021). Plant-mediated rhizosphere oxygenation in the native invasive salt marsh grass *Elymus athericus*. *Frontiers in Plant Science*, *10*(12), 669751. https://doi.org/10.3389/fpls.2021.669751

Lasut, M. T. (2009). *The floristic study of herbaceous grasses in Sulawesi*. https://docplayer.net/84661212-The-floristic-study-of-herbaceous-grasses-in-sulawesi-marthen-theogives-lasut.html

Lee, C., Lee, S., & Park, T. (2017). Statistical methods for metagenomics data analysis. *International Journal of Data Mining and Bioinformatics*, *19*(4), 366–385.

Li, K., Stoof-Leichsenring, K. R., Liu, S. S., Jia, W. H., Liao, M. N., Liu, X. Q., Ni, J., & Herzschuh, U. (2021). Plant sedimentary DNA as a proxy for vegetation reconstruction in eastern and northern Asia. *Ecological Indicators*, *132*, 108303.

Liu, S. S., Stoof-Leichsenring, K. R., Kruse, S., Pestryakova, L. A., & Herzschuh, U. (2020). Holocene vegetation and plant diversity changes in the northeastern Siberian treeline region from pollen and sedimentary ancient DNA. *Frontiers in Ecology and Evolution*, *8*, 560243.

Mander, L., & Punyasena, S. W. (2014). On the taxonomic resolution of pollen and spore records of earth's vegetation. *International Journal of Plant Sciences*, *175*(8), 931–945.

Markesteijn, L., Poorter, L., & Bongers, F. (2007). Light-dependent leaf trait variation in 43 tropical dry forest tree species. *American Journal of Botany*, *94*(4), 515–525.

More, K. D., Orsi, W. D., Galy, V., Giosan, L., He, L. J., Grice, K., & Coolen, M. J. L. (2018). A 43 kyr record of protist communities and their response to oxygen minimum zone variability in the northeastern Arabian Sea. *Earth and Planetary Science Letters*, *496*, 248–256.

Morlock, M. A., Vogel, H., Nigg, V., Ordonez, L., Hasberg, A. K. M., Melles, M., Russell, J. M., Bijaksana, S., & Team, T. D. P. S. (2019). Climatic and tectonic controls on source-to-sink processes in the tropical, ultramafic catchment of Lake Towuti, Indonesia. *Journal of Paleolimnology*, *61*(3), 279–295.

Morlock, M. A., Vogel, H., Russell, J. M., Anselmetti, F. S., & Bijaksana, S. (2021). Quaternary environmental changes in tropical Lake Towuti, Indonesia, inferred from end-member modelling of X-ray fluorescence core-scanning data. *Journal of Quaternary Science*, *36*, 1040–1051. https://doi.org/10.1002/jqs.3338

Morrissey, J., & Guerinot, M. L. (2009). Iron uptake and transport in plants: The good, the bad, and the ionome. *Chemical Reviews*, *109*(10), 4553–4567.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858.

Niemeyer, B., Epp, L. S., Stoof-Leichsenring, K. R., Pestryakova, L. A., & Herzschuh, U. (2017). A comparison of sedimentary DNA and pollen from lake sediments in recording vegetation composition at the Siberian treeline. *Molecular Ecology Resources*, *17*(6), e46–e62.

Nijburg, J. W., Coolen, M. J. L., Gerards, S., Gunnewiek, P. J. A. K., & Laanbroek, H. J. (1997). Effects of nitrate availability and the presence of *Glyceria maxima* on the composition and activity of the dissimilatory nitrate-reducing bacterial community. *Applied and Environmental Microbiology*, *63*(3), 931–937.

Ollerton, J., Winfree, R., & Tarrant, S. (2011). How many flowering plants are pollinated by animals? *Oikos*, *120*(3), 321–326.

Orsi, W. D., Coolen, M. J. L., Wuchter, C., He, L. J., More, K. D., Irigoien, X., Chust, G., Johnson, C., Hemingway, J. D., Lee, M., Galy, V., & Giosan, L. (2017). Climate oscillations reflected within the microbiome of Arabian Sea sediments. *Scientific Reports*, *7*, 6040. https://doi.org/10.1038/s41598-017-05590-9

Parducci, L., Bennett, K. D., Ficetola, G. F., Alsos, I. G., Suyama, Y., Wood, J. R., & Pedersen, M. W. (2017). Ancient plant DNA in lake sediments. *New Phytologist*, *214*(3), 924–942.

Parducci, L., Matetovici, I., Fontana, S. L., Bennett, K. D., Suyama, Y., Haile, J., Kjaer, K. H., Larsen, N. K., Drouzas, A. D., & Willerslev, E. (2013). Molecular- and pollen-based vegetation analysis in lake sediments from central Scandinavia. *Molecular Ecology*, *22*(13), 3511–3524.

Parducci, L., Valiranta, M., Salonen, J. S., Ronkainen, T., Matetovici, I., Fontana, S. L., Eskola, T., Sarala, P., & Suyama, Y. (2015). Proxy comparison in ancient peat sediments: Pollen, macrofossil, and plant DNA. *Philosophical transactions of the Royal Society of London Series B: Biological Sciences*, *370*(1660), 20130382.

Pastré, D., Hamon, L., Landousy, F., Sorel, I., David, M.-O., Zozime, A., Le Cam, E., & Piétrement, O. (2006). Anionic polyelectrolyte adsorption on mica mediated by multivalent cations: A solution to DNA imaging by atomic force microscopy under high ionic strengths. *Langmuir*, *22*(15), 6651–6660.

Paus, A., Boessenkool, S., Brochmann, C., Epp, L. S., Fabel, D., Haflidason, H., & Linge, H. (2015). Lake store Finnsjoen - a key for understanding Lateglacial/early Holocene vegetation and ice sheet dynamics in the central Scandes Mountains. *Quaternary Science Reviews*, *121*, 36–51.

Pedersen, M. W., Ginolhac, A., Orlando, L., Olsen, J., Andersen, K., Holm, J., Funder, S., Willerslev, E., & Kjaeer, K. H. (2013). A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. *Quaternary Science Reviews*, *75*, 161–168.

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M. T., Guerri, G., & Nannipieri, P. (2009). Extracellular DNA in soil and sediment: Fate and ecological relevance. *Biology and Fertility of Soils*, *45*(3), 219–235.

Prasad, M. N. V., & Freitas, H. M. D. (2003). Metal hyperaccumulation in plants - biodiversity prospecting for phytoremediation technology. *Electronic Journal of Biotechnology*, *6*(3), 285–321.

Randlett, M. E., Coolen, M. J. L., Stockhecke, M., Pickarski, N., Litt, T., Balkema, C., Kwiecien, O., Tomonaga, Y., Wehrli, B., & Schubert, C. J. (2014). Alkenone distribution in Lake Van sediment over the last 270 ka: Influence of temperature and haptophyte species composition. *Quaternary Science Reviews*, *104*, 53–62.

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: Inference of new DNA barcode markers from whole-genome sequence analysis. *Nucleic Acids Research*, *39*(21), e145.

Rohart, F., Gautier, B., Singh, A., & Le Cao, K. A. (2017). mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Computational Biology*, *13*(11), e1005752.

Russell, J. M., Bijaksana, S., Vogel, H., Melles, M., Kallmeyer, J., Ariztegui, D., Crowe, S., Fajar, S., Hafidz, A., Haffner, D., Hasberg, A., Ivory, S., Kelly, C., King, J., Kirana, K., Morlock, M., Noren, A., O'Grady, R., Ordonez, L., … Tamuntuan, G. (2016). The Towuti drilling project: Paleoenvironments, biological evolution, and geomicrobiology of a tropical Pacific lake. *Scientific Drilling*, *21*, 29–40.

Russell, J. M., Vogel, H., Bijaksana, S., Melles, M., Deino, A., Hafidz, A., Haffner, D., Hasberg, A. K. M., Morlock, M., von Rintelen, T., Sheppard, R., Stelbrink, B., & Stevenson, J. (2020). The late quaternary tectonic, biogeochemical, and environmental evolution of ferruginous Lake Towuti, Indonesia. *Palaeogeography Palaeoclimatology Palaeoecology*, *556*, 109905.

Russell, J. M., Vogel, H., Konecky, B. L., Bijaksana, S., Huang, Y. S., Melles, M., Wattrus, N., Costa, K., & King, J. W. (2014). Glacial forcing of central Indonesian Hydroclimate since 60,000 y BP. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(14), 5100–5105.

Sheppard, R. Y., Milliken, R. E., Russell, J. M., Sklute, E. C., Dyar, M. D., Vogel, H., Melles, M., Bijaksana, S., Hasberg, A. K. M., & Morlock, M. A. (2021). Iron mineralogy and sediment color in a 100 m drill Core from Lake Towuti, Indonesia reflect catchment and diagenetic conditions. *Geochemistry, Geophysics, Geosystems*, *22*(8), e2020GC009582. https://doi.org/10.1029/2020GC009582

Sisk-Hackworth, L., & Kelley, S. T. (2020). An application of compositional data analysis to multiomic time-series data. *NAR Genomics and Bioinformatics*, *2*(4), lqaa079.

Stevenson, J. (2018). Vegetation and climate of the last glacial maximum in Sulawesi. *The Archaeology of Sulawesi: Current Research on the Pleistocene to the Historic Period*, *48*, 17–29.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., & Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*(3), e14. https://doi.org/10.1093/nar/gkl938

Vogel, H., Russell, J., Cahyarini, S., Bijaksana, S., Wattrus, N., Rethemeyer, J., & Melles, M. (2015). Depositional modes and lake-level variability at lake Towuti, Indonesia, during the past similar to 29 kyr BP. *Journal of Paleolimnology*, *54*(4), 359–377.

Voldstad, L. H., Alsos, I. G., Farnsworth, W. R., Heintzman, P. D., Hakansson, L., Kjellman, S. E., Rouillard, A., Schomacker, A., & Eidesen, P. B. (2020). A complete Holocene lake sediment ancient DNA record reveals long-standing high Arctic plant diversity hotspot in northern Svalbard. *Quaternary Science Reviews*, *234*, 106207.

Vuillemin, A., Horn, F., Alawi, M., Henny, C., Wagner, D., Crowe, S. A., & Kallmeyer, J. (2017). Preservation and significance of extracellular DNA in ferruginous sediments from Lake Towuti, Indonesia. *Frontiers in Microbiology*, *8*, 1440.

Vuillemin, A., Wirth, R., Kemnitz, H., Schleicher, A. M., Friese, A., Bauer, K. W., Simister, R., Nomosatryo, S., Ordonez, L., Ariztegui, D., Henny, C., Crowe, S. A., Benning, L. G., Kallmeyer, J., Russell, J. M., Bijaksana, S., Vogel, H., & Towuti Drilling Project Sci, T. (2019). Formation of diagenetic siderite in modern ferruginous sediments. *Geology*, *47*(6), 540–544.

Wang, Z., Wang, C., & Liu, H. (2022). Higher dissolved oxygen levels promote downward migration of phosphorus in the sediment profile: Implications for lake restoration. *Chemosphere*, *301*, 134705.

Whitten, A. J., Mustafa, M., & Henderson, G. S. (1988). *The ecology of Sulawesi*. Gadjah Mada University Press.

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis* (p. 260). Springer International Publishing, Springer-Verlag. ISBN 978-3-319-24277-4.

Zimmermann, H. H., Raschke, E., Epp, L. S., Stoof-Leichsenring, K. R., Schwamborn, G., Schirrmeister, L., Overduin, P. P., & Herzschuh, U. (2017). Sedimentary ancient DNA and pollen reveal the composition of plant organic matter in late quaternary permafrost sediments of the Buor Khaya Peninsula (northeastern Siberia). *Biogeosciences*, *14*(3), 575–596.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ekram, M.-E., Campbell, M., Kose, S. H., Plet, C., Hamilton, R., Bijaksana, S., Grice, K., Russell, J., Stevenson, J., Vogel, H., & Coolen, M. J. L. (2024). A 1 Ma sedimentary ancient DNA (*seda*DNA) record of catchment vegetation changes and the developmental history of tropical Lake Towuti (Sulawesi, Indonesia). *Geobiology*, *22*, e12599. https://doi.org/10.1111/gbi.12599