Supplementary information

Coastal ecosystem degradation driven by decades of unregulated terrestrial mining

Mathisse Meyneng, Hugues Lemonnier, Dominique Ansquer, Florence Antypas, Nicolas Briant, Delphine Dissard, Axel Ehrhold, Thomas Haize, Gwenaël Jouet, Farid Juillot, Clémence Lavignon, Vladimir Mikryukov, Arthur Monjot, Pascal Le Roy, Sabine Schmidt, Leho Tedersoo, Raffaele Siano.

Corresponding author: mathisse.meyneng@ifremer.fr (https://orcid.org/0009-0000-7442-4613)

Summary

Supplementary Methods: Long-reads amplicon sequencing for fungi identification

Since many fungal ASVs from short-read (Illumina) remain without precise affiliation beyond the kingdom level, a subsequent long-read amplicon sequencing procedure was performed. Long rRNA gene amplicons spanning from the SSU-V4 region to the LSU-D9 region were generated using the primers EUK575F [1] and 21Rngs [2] and sequenced on a PacBio Sequel II instrument (Pacific Biosciences, Menlo Park, CA, USA). PacBio reads were processed using the NextITS pipeline v1.0.0 [3]. Briefly, libraries were demultiplexed with LIMA v2.12.0 (Pacific Biosciences) using dual asymmetric indices, adapters and primers were trimmed with cutadapt v5.0 [4], and reads lacking both primer sites were discarded. Full-length ITS regions were extracted with ITSx v1.1.3 [5], and sequences <250 bp, with >0.6 expected errors per 100 bp, or containing homopolymers >25 nt were removed. Chimeras were filtered in two steps: *de novo* UCHIME screening in VSEARCH v2.29.4 [6], followed by reference-based verification against EUKARYOME v1.9.4 [7], removing any sequence flagged in either step. The remaining reads were clustered at 98% pairwise identity with VSEARCH to define OTUs based on the full-length ITS. For each OTU representative, the SSU and LSU segments of rRNA were retrieved, and all rRNA parts were taxonomically annotated against the EUKARYOME database.

All long reads comprising each OTU (i.e., noted as LR (quality-filtered PacBio CCS members)) were used for matching to the Illumina short read ASVs (noted as SR). In brief, the rDNA SSU was extracted from LR sequences using ITSx. The sequences were then trimmed at the 3' end with cutadapt v4.2 [4] using the parameters *-a TYRATCAAGAACGAAAGT";min_overlap=18" --action retain -e 0.2 --discard-untrimmed*. This trimming step was designed to detect the reverse primer used in the short read (SR) amplicon sequencing procedure within LR sequences with a maximum error rate of 0.2 (i.e., allowing mismatch and therefore preventing primer biases) and to remove bases downstream of the primer; sequences without the primer were discarded. Since both forward primers used for LR and SR amplicon sequencing procedures have the same location in the V4 region, the resulting trimmed LR sequences align with the region previously sequenced (i.e., SR dataset). Trimmed LR sequences and SR-derived ASVs were aligned with VSEARCH v2.18.0 [6] using the parameters *--allpairs_global --id 0.8 --query_cov 0.8 --target_cov 0.8* in order to produce a blast-format file summarising sequence pairs that were at least 80% identical over 80% of their length. The resulting file, along with taxonomic affiliations for SR derived ASVs and LR sequences, was processed in R with the igraph package v2.0.3 [8] to build a sequence similarity network composed of connected components (CCs) gathering sequences with an identity score of at least 97% over at least 200 bp. Only CCs clustering at least one SR-derived ASVs detected differentially abundant in the sedimentary Unit I and affiliated to Unassigned Fungi, and at least one LR sequences were studied.

Supplementary Discussion:

SedaDNA study in a tropical ecosystem

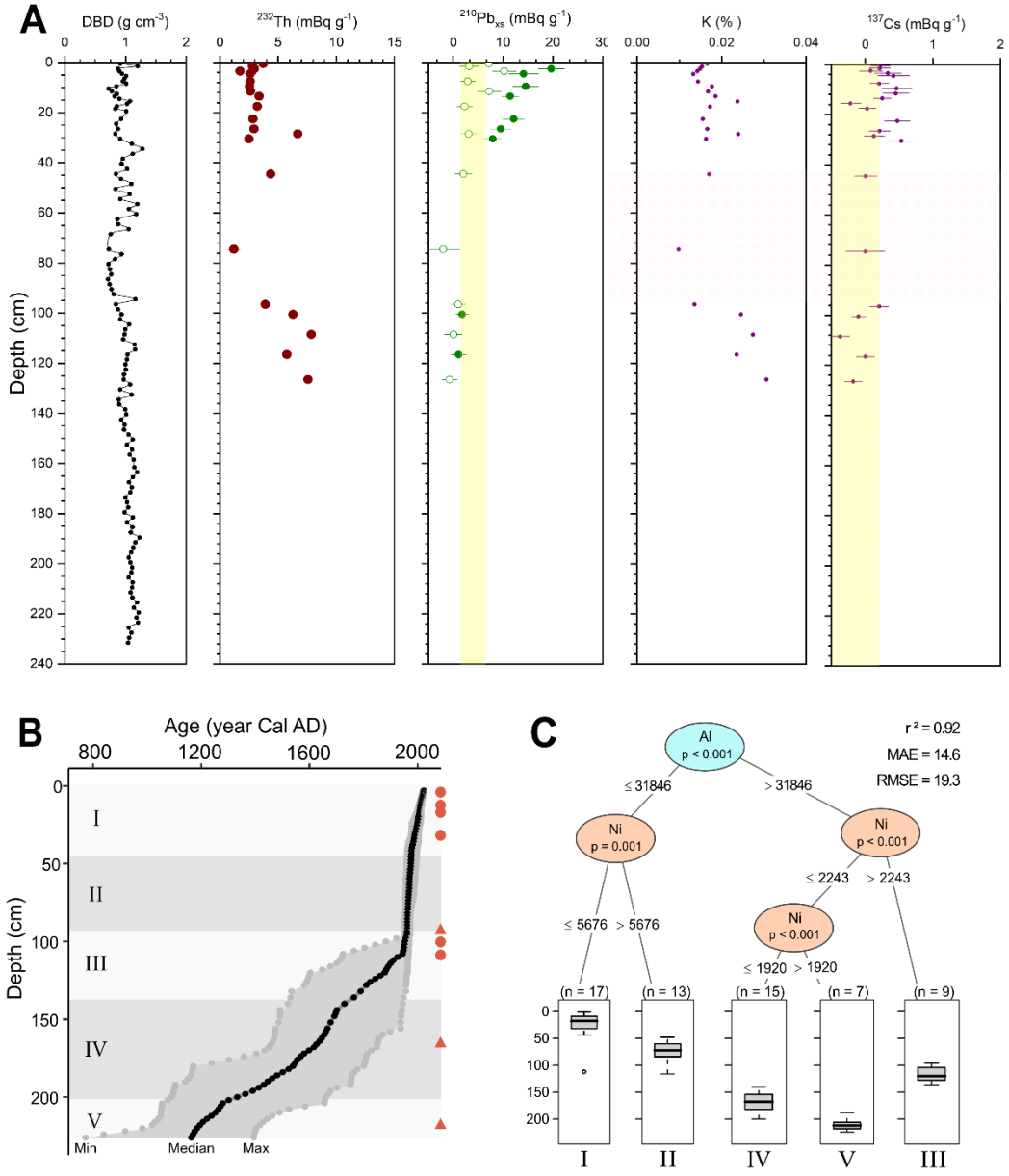
This study represents the first application of marine *sed*aDNA in New Caledonia and one of the earliest in a tropical/subtropical region. As expected, both DNA concentration and ASV richness declined with depth due to degradation **(Fig. 2-S7)**, primarily affecting extracellular DNA [9,10]. Nevertheless, ASVs were recovered across all sediment depths, likely facilitated by the fine-grained size of the sediments and the presence of resting stages, both of which enhance DNA preservation [11]. Because amplicon sequencing does not allow direct assessment of DNA degradation rates with depth, an advantage of shotgun sequencing [12], further research is needed to compare the benefits and biases of both approaches in tropical marine environments, where *sed*aDNA studies remain scarce. In this study, we selected amplicon sequencing to specifically investigate changes in the microeukaryotic community following mining impacts. The temporal scope of our analysis, extending back only to the onset of mining in the 1800s, limited the risk of severe degradation and allowed robust taxonomic annotation using a relatively long marker (18S-V4, ~400 bp). For longer timeframes, shotgun sequencing or shorter markers, commonly recommended in ancient DNA research, would be more appropriate [12,13].

Resting stage subcommunity

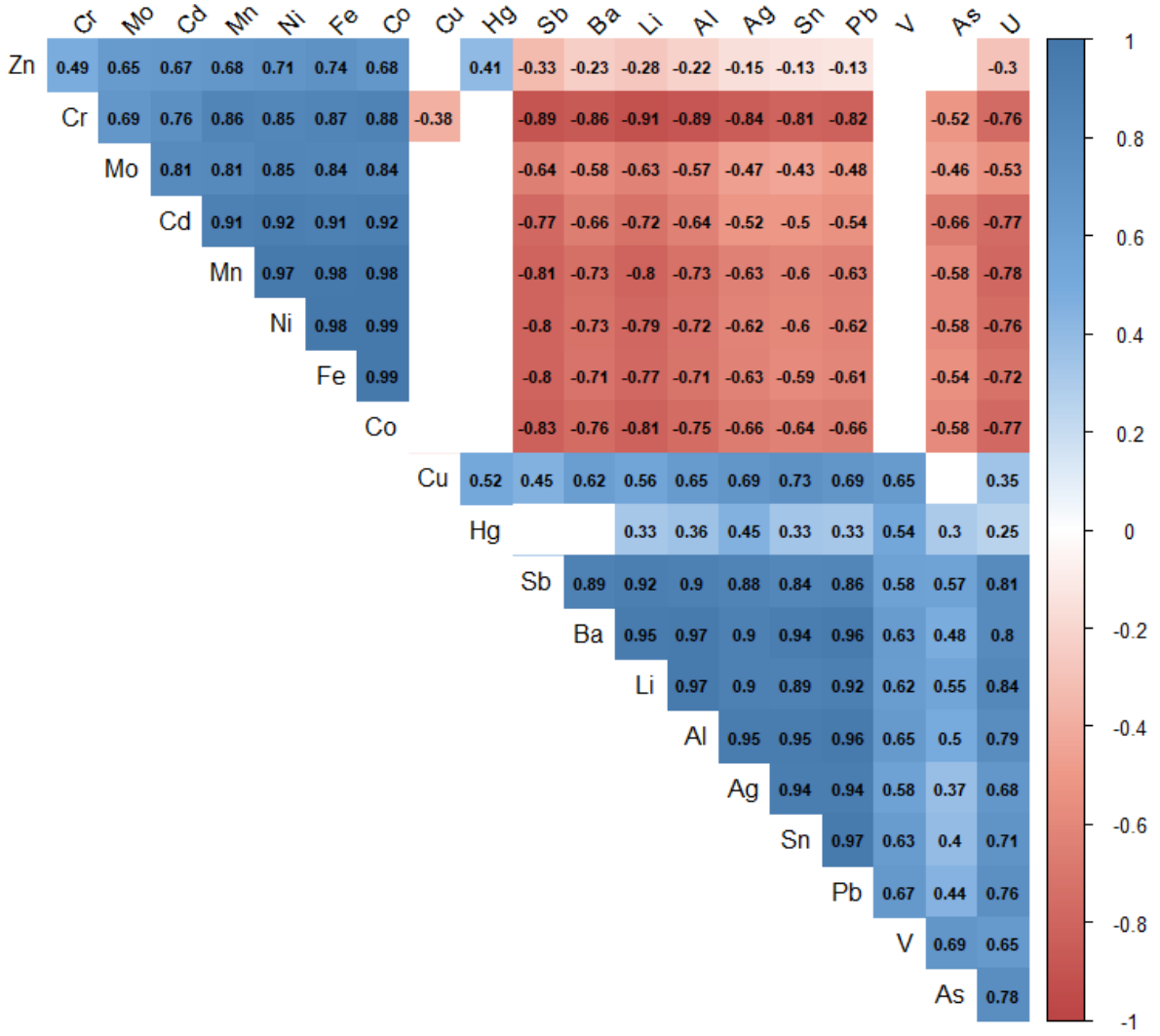
Because DNA degradation complicates the separation of taphonomic signals from ecological ones, we focused on taxa capable of forming resting stages, which accounted for 47% of sequences in the rarefied dataset **(Fig. 3)**. A key limitation, common in many eDNA studies (e.g., Ramond *et al.* 2019), was insufficient taxonomic resolution: on average, 26 ± 9.2% of ASVs per sample remained classified as unassigned Eukaryota. This lack of resolution is likely exacerbated by the limited representation of Pacific archipelago taxa in reference databases compared to regions such as the North Atlantic. Consequently, the proportion of resting-stage taxa in our dataset was probably underestimated, in contrast to studies in coastal sediments where they appear dominant [15,16].

The resting-stage community was dominated by Apicomplexa, followed by Fungi, Cercozoa, Chlorophyta, and Dinoflagellata **(Fig. 3D)**. Although the high relative abundance of Apicomplexa may be partly inflated, since all ASVs from this phylum with resting stages were annotated, this pattern is consistent with both marine [16] and terrestrial studies [17]. Their presence throughout the core, including deep sediments, supports their persistence over long timescales. Apicomplexa are widespread parasites infecting hosts from both marine and terrestrial ecosystems [17,18]. Dominant genera included *Rhodocystis* spp., parasites of marine annelids [19], as well as terrestrial parasites such as *Gregarina blattarum*, which infects cockroaches [20], and *Leidyana erratica*, which infects crickets [21]. While some misidentifications remain possible [22], the detection of terrestrial parasites distant from their hosts could plausibly result from their natural occurrence in the watershed and leaching of terrestrial material carrying their cysts.

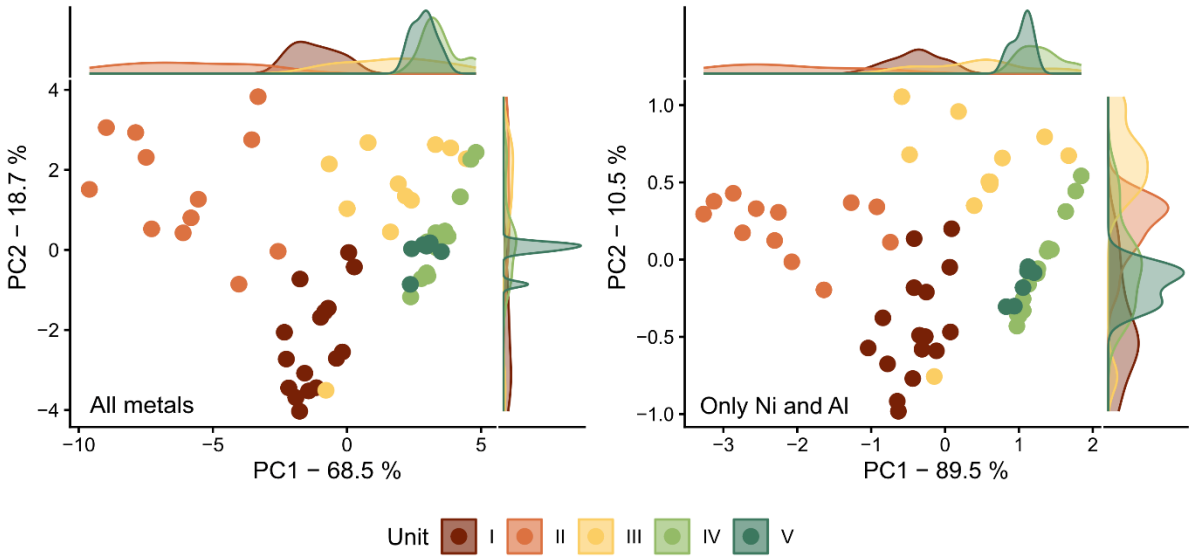
Supplementary Figure 1: Identification of sedimentary units associated with the period of the mine activity. (A) Profiles of dry bulk density (DBD), ^{232}Th, ^{210}Pb, K and ^{137}Cs with depth in the sediment core. ^{210}Pbxs, ^{232}Th and ^{137}Cs activities are given with their determination error based on statistical errors (number of counts in the peak of each measured radionuclide). The yellow area represents the estimated limit of detection. The $^{210}Pb_{xs}$ values plotted with solid circles were used in (B) the modelling of sediment dating. Red points correspond to the constraints applied to the model, based on the ages measured with ^{14}C (triangle) and ^{210}Pb (round) methods. The black dots represent the mean value, while grey dots represent the minimum and maximum estimated date, used to assess an error rate, and precise for each date along the manuscript. (C) Sedimentary unit identification was performed by clustering samples based on their concentrations (mg kg-1) of Ni and Al, separating the core into five units. The figure shows the threshold of concentration highlighted by the model and the repartition of the sample along the core which were used to establish the different period defined as sedimentary units.



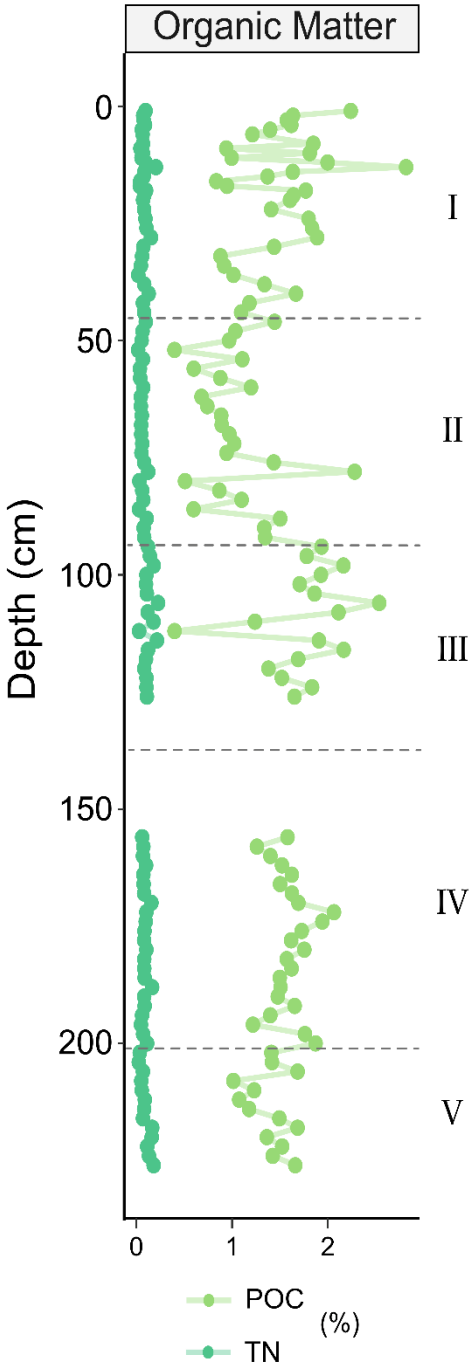
Supplementary Figure 2: Spearman correlation between 20 trace metal concentrations measured in the core sediment. Only correlations showing p-value < 0.05 are shown. Non-significant correlations are indicated by white boxes. Two main groups appear to support the choice of Ni and Al as representative metals, with many others. Al is positively correlated by Li, Ag, Sn, Pb, V, As, Ba, Sb, Hg, Cu, and U, with a mean correlation of 0.79 ± 0.22, min = 0.36 (Hg) and max = 0.97 (Li and Ba). Ni is positively correlated with Zn, Cr, Mo, Cd, Mn, Fe, and Co, with a mean correlation of 0.90 ± 0.10, min = 0.71 (Zn) and max = 0.99 (Co).



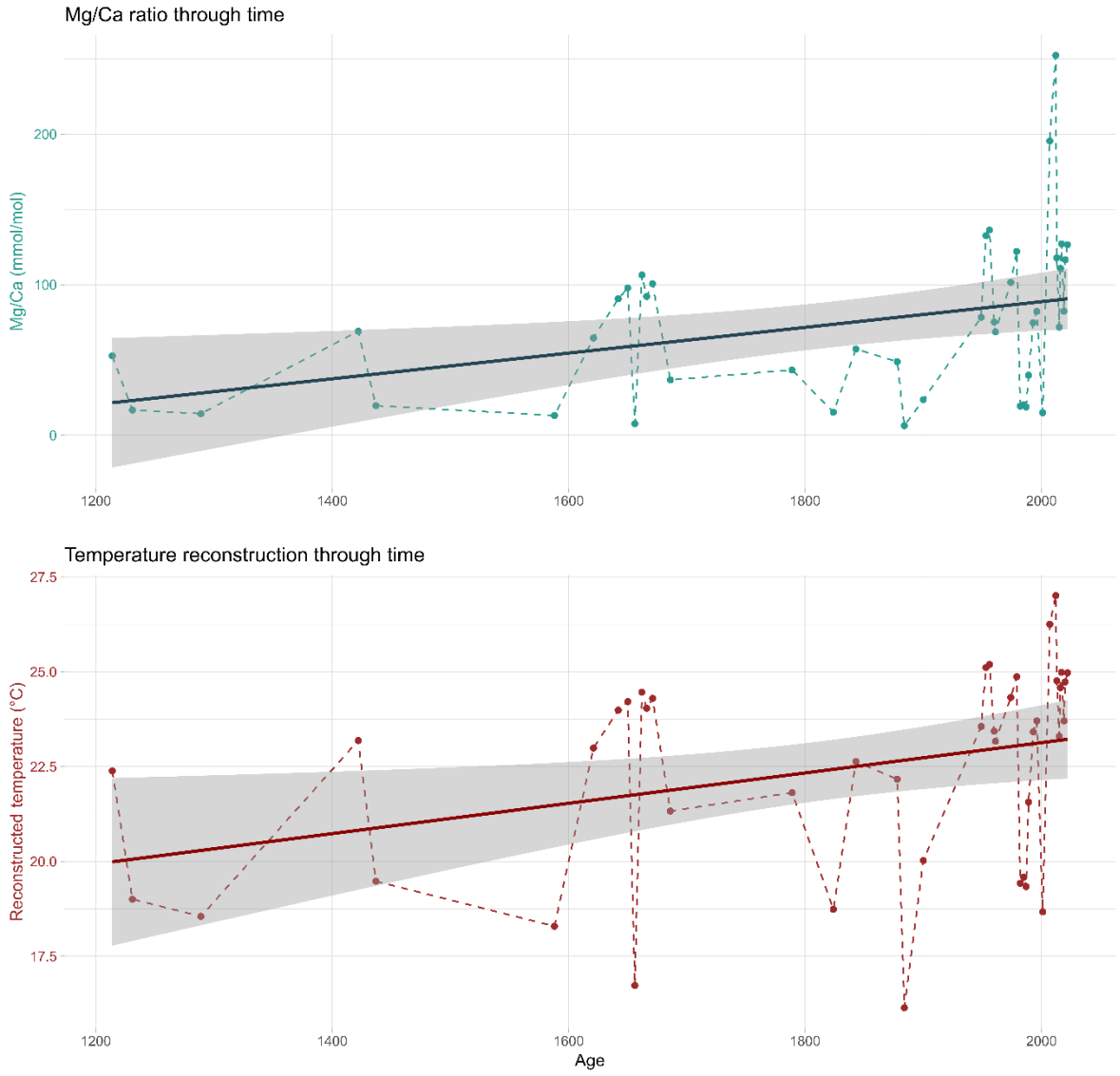
Supplementary Figure 3: Comparison of multivariate and bivariate PCA of sediment trace metals. Each point represents an individual sample. Colours indicate the sedimentary unit (Unit I–V), and the shaded areas along the axes represent the distribution of samples along each principal component (marginal densities). The first two components (PC1 and PC2) explain the percentages of variance indicated on the axes. Pairwise dissimilarity between samples was quantified using the Euclidean distance. These analyses illustrate that the bivariate Al–Ni PCA retains the major patterns of sample clustering and dispersion observed in the full PCA, supporting the selection of Al and Ni as key indicators of metallic composition in the sediment core.



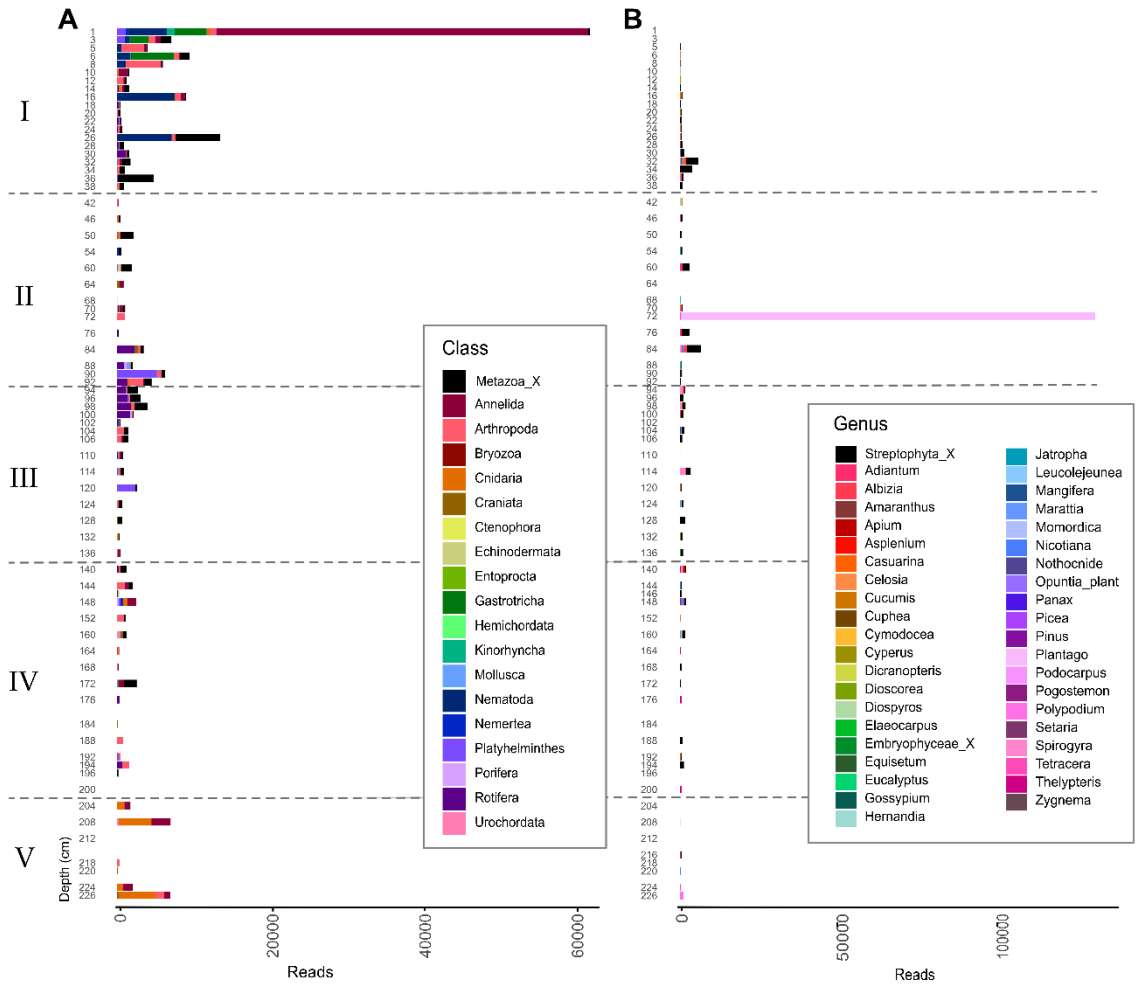
Supplementary Figure 4: Concentration of POC and TN along the core. Particulate Organic Carbon (POC) and Total Nitrogen (TN) were measured at high frequency along the core depth, exhibiting temporal variability. Sedimentary units are represented with the grey dotted line. A gap in the profile around 140 cm reflects missing samples due to technical issues.



Supplementary Figure 5: Mg/Ca ratios in Ammonia sp. and reconstructed sea surface temperature (SST) along the sediment core. Mg/Ca ratios were measured on ***Ammonia*** sp. tests (average of 5 specimens per sample) to track past sea surface temperatures. Ratios reflect tropical conditions (Mg/Ca ~70 mmol mol^{-1}) and were converted to SST using the calibration for ***Baculogypsina sphaerulata***. Values indicate tropical conditions, with reconstructed temperatures ranging from 16 to 27°C (mean = 22°C). A linear regression has been added to each profile to highlight temporal trends.

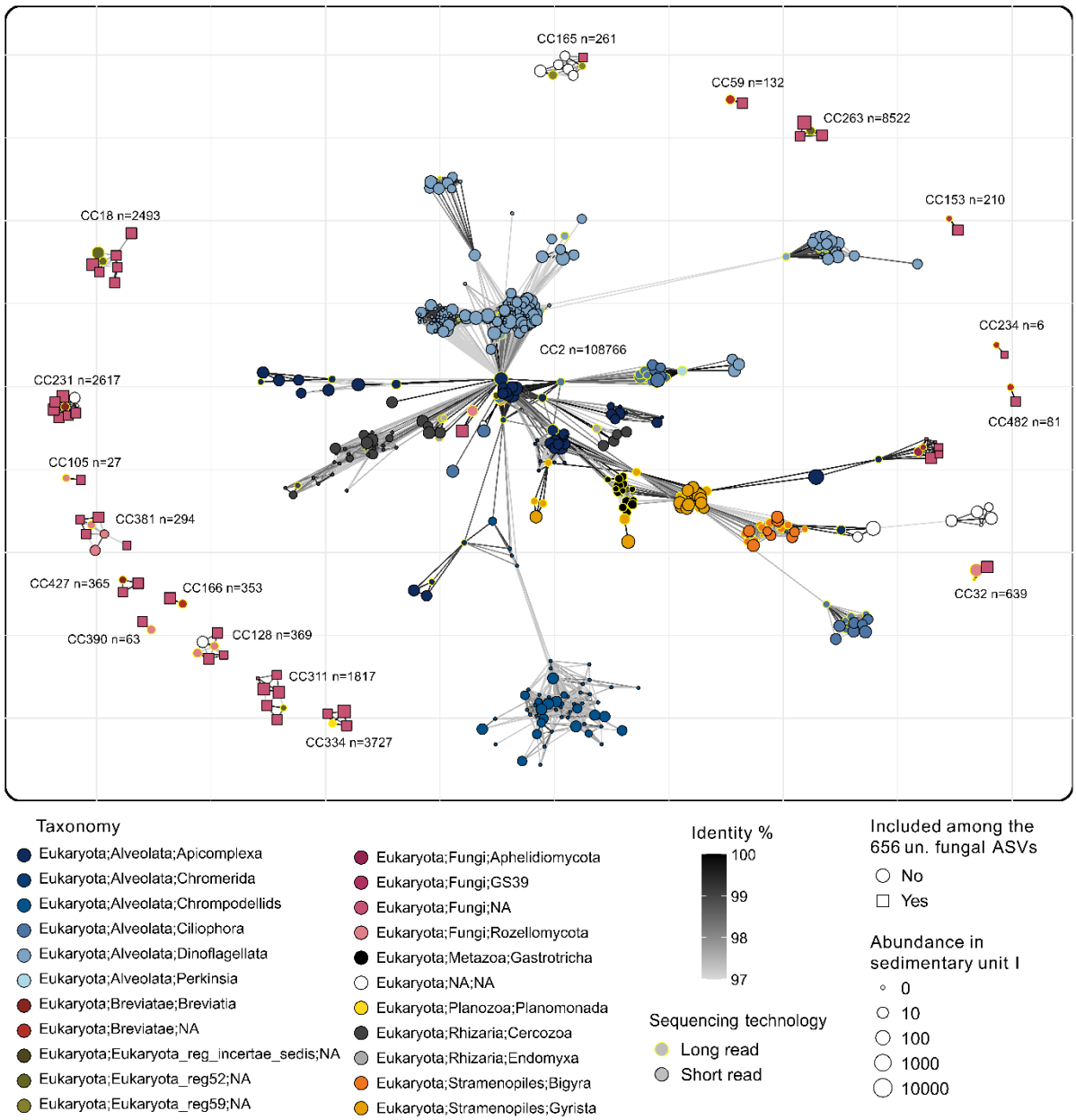


Supplementary Figure 6: Quantity and taxonomic composition of reads of Metazoa (at class level) and Streptophyta (at Genus level) removed from the raw dataset along the core. Black color represent ASV that could not be assigned at high precision, being metazoans or plants.

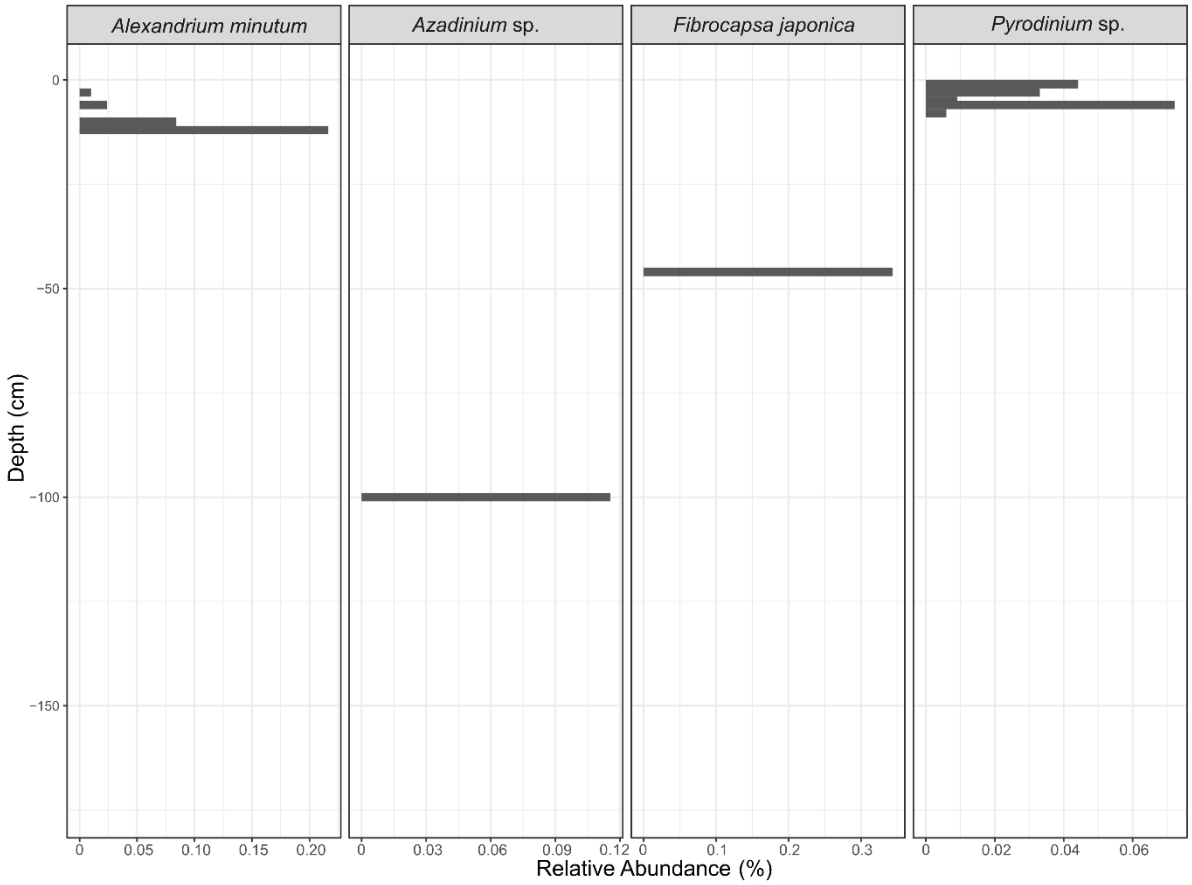


Supplementary Figure 7: Occurrence of foraminiferal genera along the core. Presence (dark box) and absence (light grey box) of the 25 foraminiferal genera identified along the sediment core (depth in cm). Genera are ordered based on the similarity of their distribution patterns along the sediment core, highlighting co-occurrence and turnover across sedimentary units. Bold genera are those showing non-random distribution presented in the main Fig. 5.

Supplementary Figure 8: Sequence similarity network. This network is generated from pairwise alignments of V4 short-read-derived ASVs (SR) initially affiliated with Unassigned Fungi (i.e., N=656) and V4 regions extracted from long-read sequences (LR). Each SR and LR are represented by a point and clustered based on their identity within Connected Components (CCs). Clustering enabled the refinement of taxonomic affiliation by leveraging sequence similarity thresholds. Each CC is identified by a number and characterised by the abundance in sedimentary unit I of its SR-derived ASVs. They are graphically represented with the "fr" layout, using ggraph (v2.2.1) and tydigraph (v1.3.1) R packages (Pedersen, 2022,2023, available at https://ggraph.data-maginist.com, https://github.com/thomasp85/ggraph, and https://tidygraph.data-imaginist.com, https://github.com/thomasp85/tidygraph respectively).

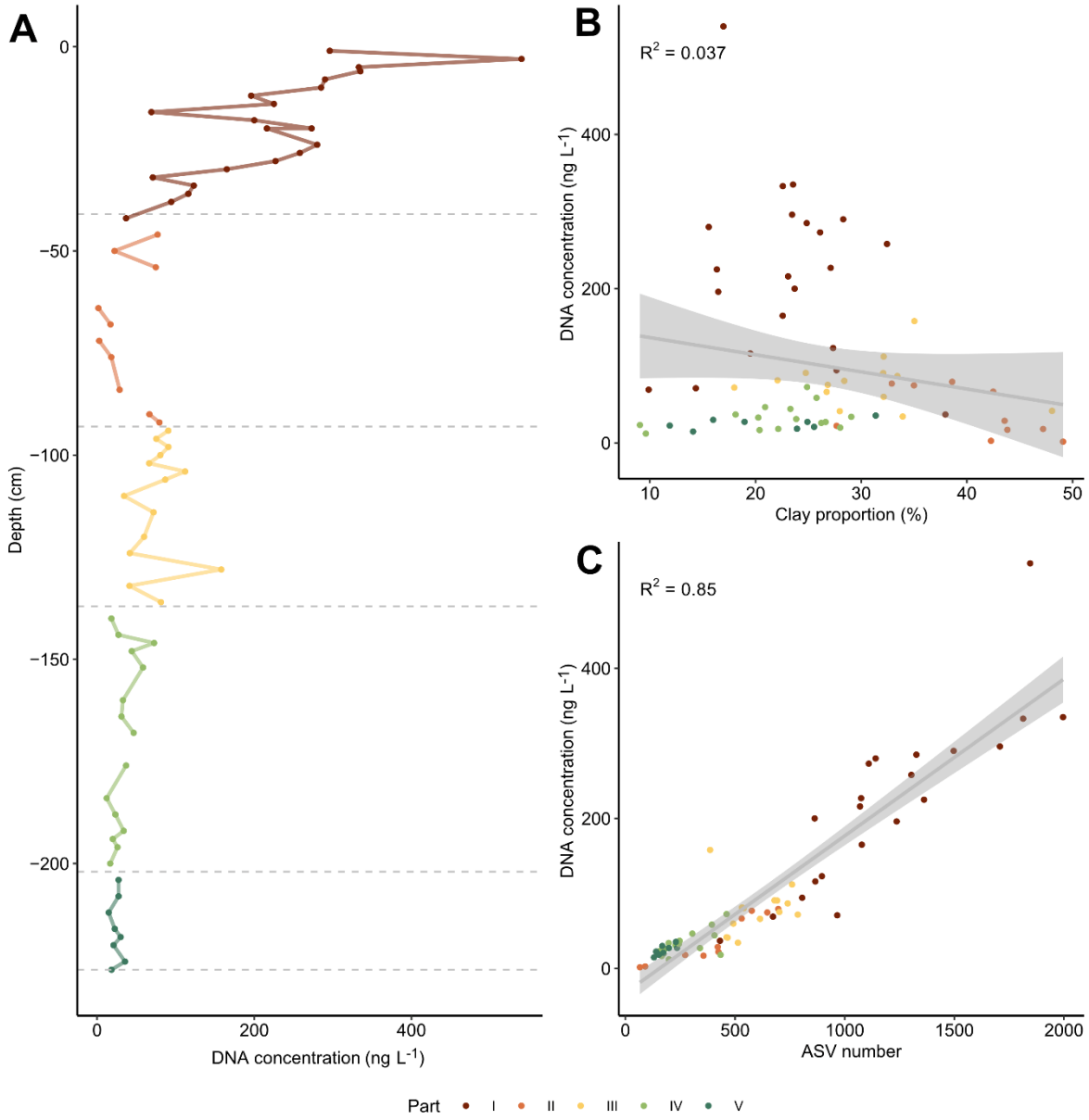


Supplementary Figure 9: Presence of four genera or species considered harmful. Their relative abundance of reads (among the all microeukaryote community) is represented along the core depth..

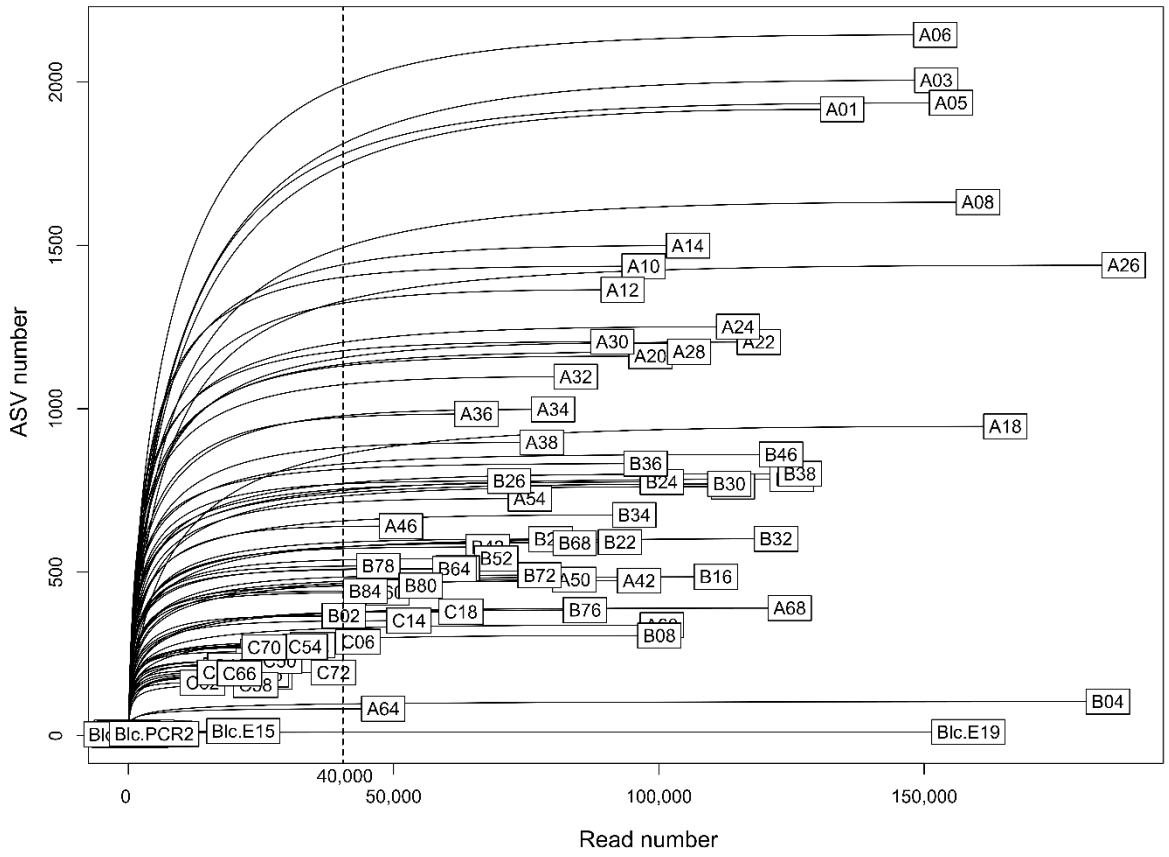


Supplementary Figure 10: Photography of the agent performing the DNA extraction in the controlled atmosphere laboratory (Credit: Hugues Lemonnier).

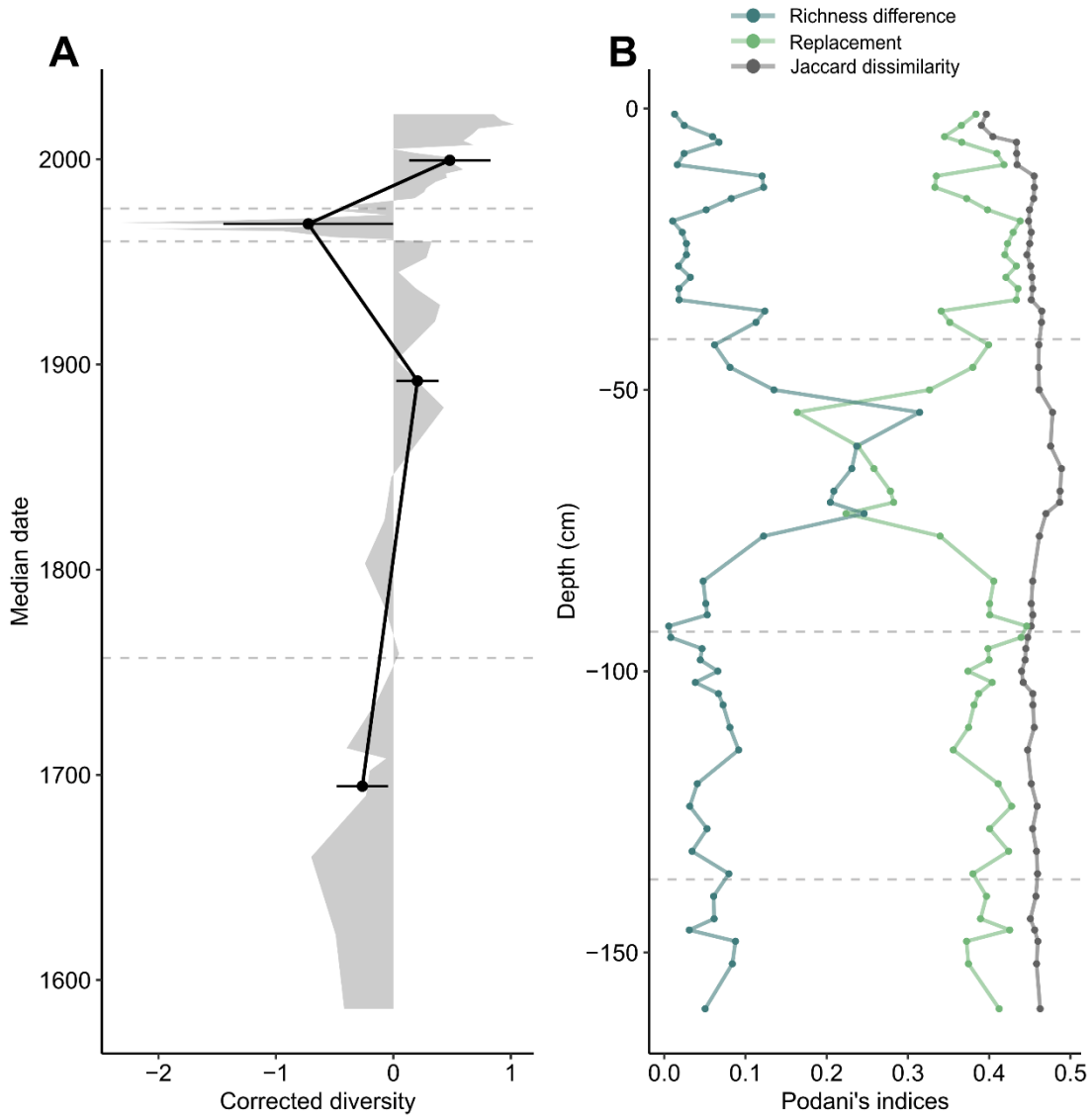
Supplementary Figure 11: Profile of DNA concentration along the sediment core. (A) DNA concentration as a function of sample depth along the core, and in relation to (B) clay proportion of the sediment and (C) ASV number of the raw dataset in the different sedimentary units.



Supplementary Figure 12: ASV richness as a function of sequencing depth for each sample after bioinformatics processing in the raw dataset. The dotted line indicates the rarefaction threshold of 40,000 reads per sample. Sample IDs beginning with "A," "B," or "C" correspond to sediment core samples (see Table S4 for sample depth and sedimentary unit), while IDs starting with "Blc." represent blank controls.



Supplementary Figure 13: Diversity analyses on the rarefied dataset, similar to Figure 2. (A) Microeukaryotic ASV richness (within-sample diversity). Grey shading shows observed richness for each sample, black dots indicate the mean richness per sedimentary unit, and error bars represent the standard deviation of the corrected mean richness. (B) Podani's indices (community structure changes). Three components are displayed: richness difference (blue; net change in ASV number), replacement (green; ASV turnover), and Jaccard dissimilarity (grey; overall compositional change). Each sample is compared with the two preceding samples to highlight short-term temporal shifts and periods of increased compositional variability. Grey dotted lines mark the boundaries of sedimentary units.



Supplementary Table 1: Table of age measurements. Age measures on different samples were used to constrain the model based on ^{14}C and ^{210}Pb methods, with the estimated SAR (sedimentation accumulation rate).

Sedimentary Unit	Depth (cm)	Age (yr Cal AD)				SAR (cm/yr)
		Type	min	max	mean	
I	2.5	^{210}Pb			2022.4	
I	9.5	^{210}Pb			2013.6	0.795
I	13.5	^{210}Pb			2005.0	0.465
I	30.5	^{210}Pb			1993.2	1.441
II	94	^{14}C	1957	1960	1959.0	1.857
III	100.5	^{210}Pb			1945.8	0.492
III	116.5	^{210}Pb			1930.0	1.013
IV	168	^{14}C	1474	1777	1614.0	0.163
V	218	^{14}C	1029	1292	1162.0	0.111

Supplementary Table 2: Genera showing relative abundance changes along the sediment core. The number of genera displaying changes in relative abundance between consecutive sedimentary units is shown for the 183 genera present in at least four samples. Genera with statistically significant changes (adjusted $p < 0.1$, Wilcoxon test) are indicated with an asterisk.

		IV vs III	***III vs II***	***II vs I***
Decline	Nb of genera	61 (0*)	71 (4*)	73 (2*)
	Percentage	33 (0*)	39 (2.2*)	40 (1*)
Increase	Nb genera	81 (1*)	78 (0*)	110 (2*)
	Percentage	44 (0.5*)	42 (0*)	60 (1.8*)

Supplementary Table 3: Contribution of each element (%) to the identified sediment sources. Two sources were distinguished (Ultramafic and Volcano-sedimentary), consistent with the known geology of the Thio watershed (see Fig. 6). Contributions were estimated using the PMF 5.0 model based on 20 measured metal concentrations and all grain size fractions.

Element	Ultramafic	Volcano-sedimentary
Li	7.6	92.4
Al	16.8	83.2
V	31.0	69.0
Cr	56.8	43.2
Mn	74.9	25.1
Fe	63.1	36.9
Co	75.9	24.1
Ni	81.2	18.8
Cu	29.4	70.6
Zn	56.1	43.9
As	16.9	83.1
Mo	64.6	35.4
Ag	13.7	86.3
Cd	66.1	33.9
Sn	14.7	85.3
Sb	16.1	83.9
Ba	11.9	88.1
Pb	16.4	83.6
U	14.3	85.7
Hg	33.1	66.9
Clay (0-3.9µm)	50.1	49.9
Very fine silt_(3.9-7.8µm)	42.0	58.0
Fine silt (7.8-15.6µm)	38.5	61.5
Medium silt (15.6-31.0µm)	28.6	71.4
Coarse silt (31-63µm)	16.5	83.5
Very fine sand (63-125µm)	6.1	93.9
Fine sand (125-250µm)	2.5	97.5

Medium sand (250-500µm)	3.6	96.4
Coarse sand (500-1000µm)	36.4	63.6
Very coarse sand (1000-2000µm)	91.5	8.5

Supplementary Table 4: Initial sample name associated with the sediment depth and the sedimentary unit. A, B, or C corresponds to the core section from after sampling, and the number corresponds to the sediment depth within this section.

Sample_ID	Depth (cm)	Sedimentary unit
A01	1	I
A03	3	I
A05	5	I
A06	6	I
A08	8	I
A10	10	I
A12	12	I
A14	14	I
A16	16	I
A18	18	I
A20	20	I
A22	22	I
A24	24	I
A26	26	I
A28	28	I
A30	30	I
A32	32	I
A34	34	I
A36	36	I
A38	38	I
A42	42	I
A46	46	II
A50	50	II
A54	54	II
A60	60	II
A64	64	II

A68	68	II
B02	70	II
B04	72	II
B08	76	II
B16	84	II
B20	88	II
B22	90	II
B24	92	II
B26	94	III
B28	96	III
B30	98	III
B32	100	III
B34	102	III
B36	104	III
B38	106	III
B42	110	III
B46	114	III
B52	120	III
B56	124	III
B60	128	III
B64	132	III
B68	136	III
B72	140	IV
B76	144	IV
B78	146	IV
B80	148	IV
B84	152	IV
C06	160	IV
C10	164	IV
C14	168	IV

C18	172	IV
C22	176	IV
C26	180	IV
C30	184	IV
C34	188	IV
C38	192	IV
C40	194	IV
C42	196	IV
C46	200	IV
C50	204	V
C54	208	V
C58	212	V
C62	216	V
C64	218	V
C66	220	V
C70	224	V
C72	226	V

Supplementary References

1. Aslani, F., Bahram, M., Geisen, S., Pent, M., Otsing, E., Tamm, H., Jones, A., Panagos, P., Köninger, J., Orgiazzi, A., et al. (2024). Land use intensification homogenizes soil protist communities and alters their diversity across Europe. Soil Biol. Biochem. *195*, 109459. https://doi.org/10.1016/j.soilbio.2024.109459.

2. Schwelm, A., Berney, C., Dixelius, C., Bass, D., and Neuhauser, S. (2016). The Large Subunit rDNA Sequence of *Plasmodiophora brassicae* Does not Contain Intra-species Polymorphism. Protist *167*, 544–554. https://doi.org/10.1016/j.protis.2016.08.008.

3. Mikryukov, V., Anslan, S., and Tedersoo, L. (2025). NextITS: a pipeline for metabarcoding eukaryotes with full-length ITS sequenced with PacBio. Version 1.0.0 (Zenodo). https://doi.org/10.5281/ZENODO.15074882 https://doi.org/10.5281/ZENODO.15074882.

4. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10. https://doi.org/10.14806/ej.17.1.200.

5. Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., et al. (2013). Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. Methods Ecol. Evol. *4*, 914–919. https://doi.org/10.1111/2041-210X.12073.

6. Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ *4*, e2584. https://doi.org/10.7717/peerj.2584.

7. Tedersoo, L., Hosseyni Moghaddam, M.S., Mikryukov, V., Hakimzadeh, A., Bahram, M., Nilsson, R.H., Yatsiuk, I., Geisen, S., Schwelm, A., Piwosz, K., et al. (2024). EUKARYOME: the rRNA gene reference database for identification of all eukaryotes. Database *2024*, baae043. https://doi.org/10.1093/database/baae043.

8. Csardi, G., and Nepusz, T. (2006). The Igraph Software Package for Complex Network Research. InterJournal *1695*, 1–9.

9. Dell'Anno, A., and Danovaro, R. (2005). Extracellular DNA Plays a Key Role in Deep-Sea Ecosystem Functioning. Science *309*, 2179–2179. https://doi.org/10.1126/science.1117475.

10. Torti, A., Jørgensen, B.B., and Lever, M.A. (2018). Preservation of microbial DNA in marine sediments: insights from extracellular DNA pools. Environ. Microbiol. *20*, 4526–4542. https://doi.org/10.1111/1462-2920.14401.

11. Torti, A., Lever, M.A., and Jørgensen, B.B. (2015). Origin, dynamics, and implications of extracellular DNA pools in marine sediments. Mar. Genomics *24*, 185–196. https://doi.org/10.1016/j.margen.2015.08.007.

12. Armbrecht, L. (2020). The Potential of Sedimentary Ancient DNA to Reconstruct Past Ocean Ecosystems. Oceanography *33*. https://doi.org/10.5670/oceanog.2020.211.

13. Zimmermann, H.H., Harðardóttir, S., and Ribeiro, S. (2024). Assessing the performance of short 18S rDNA markers for environmental DNA metabarcoding of marine protists. Environ. DNA *6*, e580. https://doi.org/10.1002/edn3.580.

14. Ramond, P., Sourisseau, M., Simon, N., Romac, S., Schmitt, S., Rigaut-Jalabert, F., Henry, N., de Vargas, C., and Siano, R. (2019). Coupling between taxonomic and functional diversity in protistan coastal communities. Environ. Microbiol. *21*, 730–749. https://doi.org/10.1111/1462-2920.14537.

15. Deng, H., He, C., Worden, A.Z., and Gong, J. (2024). Employing a triple metabarcoding approach to differentiate active, dormant and dead microeukaryotes in sediments. Environ. Microbiol. *26*, e16615. https://doi.org/10.1111/1462-2920.16615.

16. Siano, R., Lassudrie, M., Cuzin, P., Briant, N., Loizeau, V., Schmidt, S., Ehrhold, A., Mertens, K.N., Lambert, C., Quintric, L., et al. (2021). Sediment archives reveal irreversible shifts in plankton communities after World War II and agricultural pollution. Curr. Biol. *31*, 2682-2689.e7. https://doi.org/10.1016/j.cub.2021.03.079.

17. Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., Singer, D., Mayor, J., Bunge, J., Sernaker, S., et al. (2017). Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nat. Ecol. Evol. *1*, 1–8. https://doi.org/10.1038/s41559-017-0091.

18. de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. Science *348*, 1261605. https://doi.org/10.1126/science.1261605.

19. Leander, B.S. (2008). Marine gregarines: evolutionary prelude to the apicomplexan radiation? Trends Parasitol. *24*, 60–67. https://doi.org/10.1016/j.pt.2007.11.005.

20. Yahaya, Z.S., Izzaudin, N.A.I., and Razak, A.F.A. (2017). Parasitic Gregarine blattarum Found Infecting American Cockroaches, Periplaneta americana, in a Population in Pulau Pinang, Malaysia. Trop. Life Sci. Res. *28*, 145. https://doi.org/10.21315/tlsr2017.28.1.10.

21. Watson, M.E. (1916). Observations on Polycystid Gregarines from Arthropoda. J. Parasitol. *3*, 65–75. https://doi.org/10.2307/3271003.

22. del Campo, J., Heger, T.J., Rodríguez-Martínez, R., Worden, A.Z., Richards, T.A., Massana, R., and Keeling, P.J. (2019). Assessing the Diversity and Distribution of Apicomplexans in Host and Free-Living Environments Using High-Throughput Amplicon Data and a Phylogenetically Informed Reference Framework. Front. Microbiol. *10*. https://doi.org/10.3389/fmicb.2019.02373.