Supplementary Material

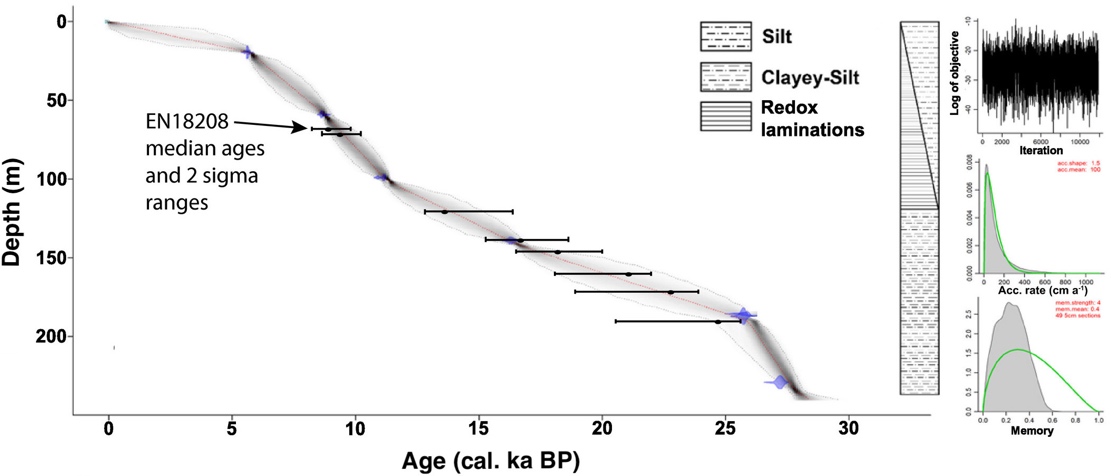
# Supplementary Data

Supplementary Data 1. The final dataset, with DNA sequences, relative abundances and taxonomic assignments (Extended Excel file).

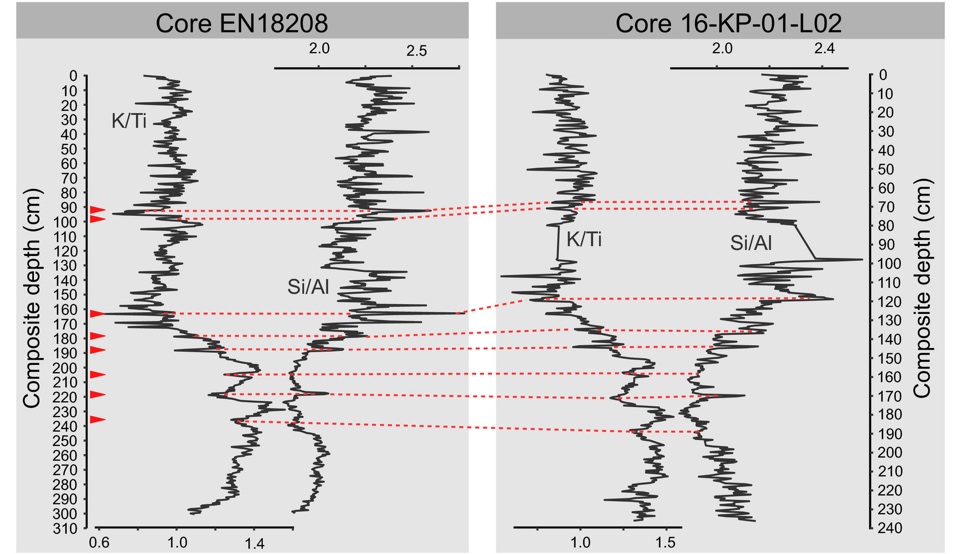
# Supplementary Figures and Tables

## Supplementary Figures

**A**



**B**

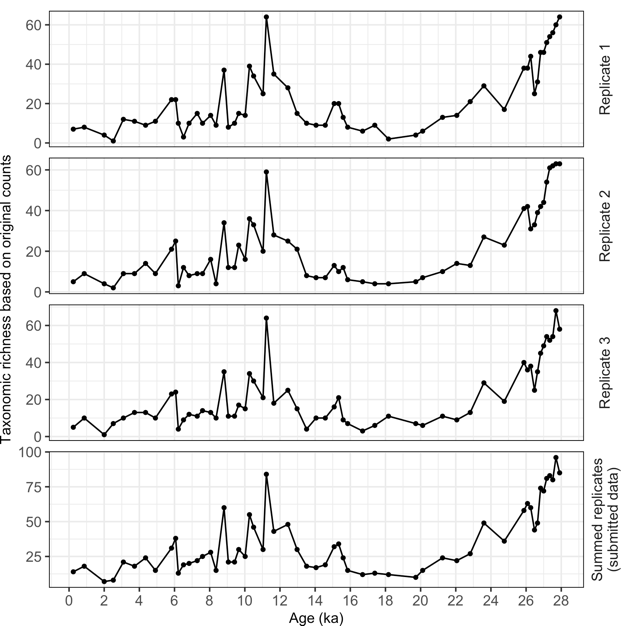
****

**Supplementary Figure S1** Age-depth model for sediment core 16-KP-01-L02 (A). Correlation between core EN18208 and 16-KP-01-L02 based on X-ray ﬂuorescence analysis data (Vyse et al., 2020) to compare median ages and sigma ranges (B). The figure has been published in Andreev et al., (2021).

A close up of a map

Description automatically generated

**Supplementary Figure S2** Taxonomic and phylogenetic diversity in the plant assemblages from sedimentary ancient DNA. The measurements from left to right are abundance-weighted net relatedness index (NRI\_abun), occurrence-based net relatedness index (NRI\_occ), occurrence-based nearest taxon index (NTI), occurrence-based local contributions to beta diversity (LCBD), and phylogenetic occurrence-based local contributions to beta diversity (pLCBD). NRI, NTI, and LCBD values are coloured in dark blue when they are statistically significant (*p* < 0.05) and light blue when they are not significant (*p* > 0.05).



**Supplementary Figure S3** Taxonomic richness using data from individual PCR replicate batches (replicate 1, 2, and 3; top three panels) and summed counts of the three replicates (submitted data; fourth panel).

**A**

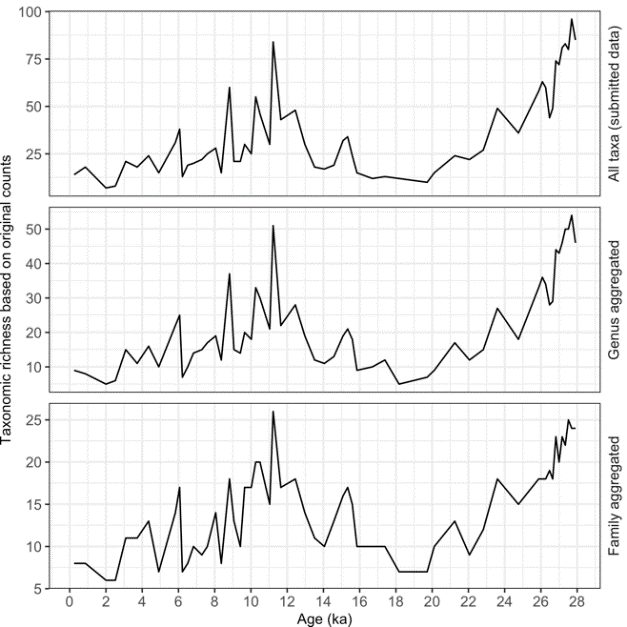
A picture containing object, fence

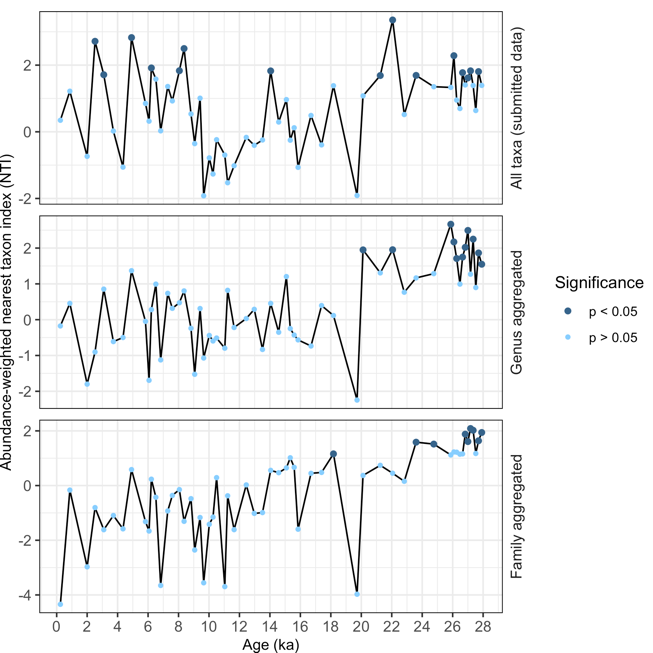
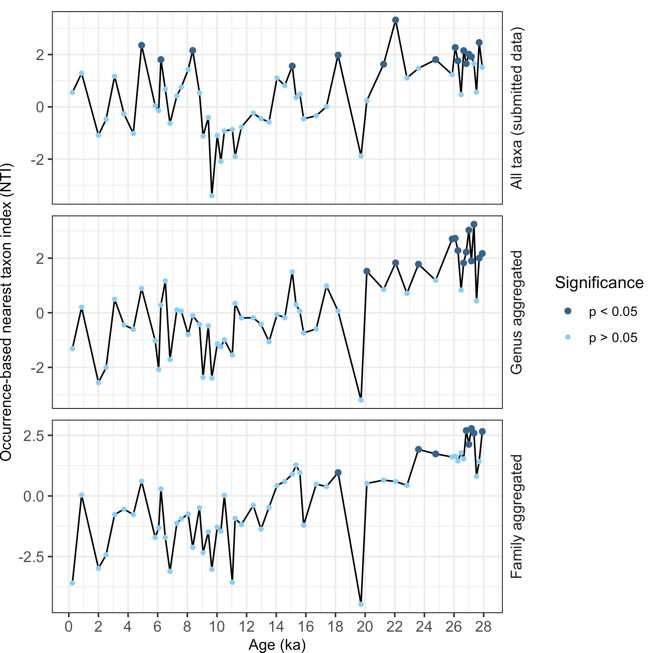
Description automatically generated

**B**A picture containing comb

Description automatically generated

**Supplementary Figure S4** Abundance-weighted **(A)** and occurrence-based **(B)** species contributions to beta diversity (SCBD) in the plant assemblages of sedimentary ancient DNA from subarctic Far East Russia.

**A**

**B** **C**

**Supplementary Figure S5** Taxonomic **(A)** and phylogenetic diversity **(B)** and **(C)** by taxonomic rank. The abundance- **(B)** and occurrence- based **(C)** nearest taxon index (NTI) were used to indicate the phylogenetic alpha diversity. The sequences were assigned by ArctBorBryo with 100%. All taxa (top panels) – richness of all 158 taxa with taxonomic levels at family and lower levels. Genus aggregated (middle panels) – aggregated richness of taxa within genus and lower levels. Family aggregated (bottom panels) – aggregated richness of taxa within family and lower levels.

**A**Graphical user interface

Description automatically generated **B**Graphical user interface, chart

Description automatically generated

**Supplementary Figure S6** Abundance-weighted **(A)** and occurrence-based **(B)** nearest taxon index (NTI) of all plants (158 taxa, submitted data; top panels), angiosperms (136 taxa; middle panels), and non-angiosperms (22 taxa; bottom panels) in the dataset. Non-angiosperms occurred only sporadically in the core.

## Supplementary Tables

**Supplementary Table S1**Radiocarbon and Intcal13 calibrated ages derived from bulk TOC samples from Lake Ilirney. The ages were used for age-depth modeling. The table is adapted from Andreev et al., (2021).

|  |  |  |  |
| --- | --- | --- | --- |
| **Sample ID** | **Depth (cm)** | **14C age and error (a BP)** | **Reservoir corrected, calibrated age (cal. a BP) with minimum and maximum age range** |
| EN18214 surface | 0-0.5 | 1721±28 | - |
| Core 16-KP-01-L02 Long 3 | 19–20 | 6605±29 | 5630 (5365–5796) |
| Core 16-KP-01-L02 Long 3 | 58–60 | 9562±34 | 8733 (8519–9045) |
| Core 16-KP-01-L02 Long 3 | 98–100 | 11414±37 | 11188 (10817–11399) |
| Core 16-KP-01-L02 Long 3 | 138–140 | 15259±47 | 16268 (15550–16764) |
| Core 16-KP-01-L02 Long 3 | 186–188 | 23102±72 | 25453 (23940–26125) |
| Core 16-KP-01-L02 Long 3 | 228–230 | 24546±62 | 27619 (27119–27998) |

**Supplementary Table S2** Further data filtering of the sequences. The raw data represents the output of the OBITools pipeline. Each unique DNA sequence refers to an amplicon sequence variant (ASV). The output of the last filtering step is used as the final dataset for statistical analyses.

|  |  |  |  |
| --- | --- | --- | --- |
| **Filtering step** | **Number of ASVs** | **Number of reads** | **% of raw data** |
| Raw data produced by OBITools | 12816 | 45,103,727 | 100 |
| Keep the 100% identity assigned by either ArctBorBryo or EMBL | 371 | 32,631,335 | 72.35 |
| Discard sequences that occurred less than three times in all PCR batches including replicates | 229 | 32,518,879 | 72.10 |
| Keep the identity assigned by ArctBorBryo as 100% | 186 | 32,235,387 | 71.47 |
| Keep only the sequences that are assigned to family and below family levels | 181 | 32,226,526 | 71.45 |
| Exclude sequences of contaminants (*Convallaria majalis*, *Pinus*) | 179 | 23,044,793 | 51.10 |
| Remove DNA extraction blanks and PCR NTCs | 179 | 21,929,271 | 48.62 |
| Exclude sequences of aquatic submerges (*Stuckenia*, *Potamogeton*, *Myriophyllum*) | 174 | 21,758,767 | 48.24 |
| Exclude sequences of moss | 158 | 21,697,725 | 48.11 |

**Supplementary Table S3** PCR was performed on three replicates and in total 21 batches. Five to nine samples were included in one PCR batch. Each PCR batch contained one DNA extraction blank (blank) and one PCR negative control (NTC). In total 1,115,522 reads were removed due to blanks and NTCs, which accounted for 2.5% of the raw data. More than 80% of the reads in blanks and NTCs were assigned to *Dryas* and Saliceae. They both occurred in the extraction blanks of three PCR batches (PCR batch 1, 2 and 21; marked in red).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Samples in the DNA extraction batch**  **(age in cal yr BP)** | **PCR batch number** | **Replicate number** | **Reads in Blank or NTC** | **% of *Dryas* reads** | **% of Saliceae reads** |
| 26477, 26650, 26825, 27000, 27172, 27344, 27517, 27698, 27907 | PCR\_batch\_1 | Blank\_Replicate\_1 | 331722 | 83.23 | 0.56 |
|  | NTC\_Replicate\_1 | 0 | 0 | 0 |
| PCR\_batch\_2 | Blank\_Replicate\_2 | 429226 | 84.74 | 5.69 |
|  | NTC\_Replicate\_2 | 0 | 0 | 0 |
| PCR\_batch\_3 | Blank\_Replicate\_3 | 67 | 0 | 100 |
|  | NTC\_Replicate\_3 | 0 | 0 | 0 |
| 243, 3092, 3723, 4354, 4914, 5830, 864, 2001, 2518 | PCR\_batch\_4 | Blank\_Replicate\_1 | 31 | 32.26 | 67.74 |
|  | NTC\_Replicate\_1 | 0 | 0 | 0 |
| PCR\_batch\_5 | Blank\_Replicate\_2 | 39 | 0 | 100 |
|  | NTC\_Replicate\_2 | 148 | 0 | 0 |
| PCR\_batch\_6 | Blank\_Replicate\_3 | 32 | 0 | 100 |
|  | NTC\_Replicate\_3 | 0 | 0 | 0 |
| 6065, 6217, 6512, 6828, 7291, 7594, 8056, 8364, 8818 | PCR\_batch\_7 | Blank\_Replicate\_1 | 0 | 0 | 0 |
|  | NTC\_Replicate\_1 | 16 | 0 | 100 |
| PCR\_batch\_8 | Blank\_Replicate\_2 | 0 | 0 | 0 |
|  | NTC\_Replicate\_2 | 0 | 0 | 0 |
| PCR\_batch\_9 | Blank\_Replicate\_3 | 0 | 0 | 0 |
|  | NTC\_Replicate\_3 | 49 | 0 | 100 |
| 11225, 11649, 9062, 9414, 9652, 10020, 10264, 10499, 11037 | PCR\_batch\_10 | Blank\_Replicate\_1 | 0 | 0 | 0 |
|  | NTC\_Replicate\_1 | 0 | 0 | 0 |
| PCR\_batch\_11 | Blank\_Replicate\_2 | 0 | 0 | 0 |
|  | NTC\_Replicate\_2 | 0 | 0 | 0 |
| PCR\_batch\_12 | Blank\_Replicate\_3 | 0 | 0 | 0 |
|  | NTC\_Replicate\_3 | 0 | 0 | 0 |
| 12448, 12973, 13513, 14051, 14577, 15082, 15337, 15593, 15850 | PCR\_batch\_13 | Blank\_Replicate\_1 | 0 | 0 | 0 |
|  | NTC\_Replicate\_1 | 0 | 0 | 0 |
| PCR\_batch\_14 | Blank\_Replicate\_2 | 0 | 0 | 0 |
|  | NTC\_Replicate\_2 | 0 | 0 | 0 |
| PCR\_batch\_15 | Blank\_Replicate\_3 | 0 | 0 | 0 |
|  | NTC\_Replicate\_3 | 0 | 0 | 0 |
| 16699, 17392, 18173, 19723, 20116, 21247, 22057, 22821 | PCR\_batch\_16 | Blank\_Replicate\_1 | 98 | 0 | 100 |
|  | NTC\_Replicate\_1 | 0 | 0 | 0 |
| PCR\_batch\_17 | Blank\_Replicate\_2 | 101 | 0 | 100 |
|  | NTC\_Replicate\_2 | 0 | 0 | 0 |
| PCR\_batch\_18 | Blank\_Replicate\_3 | 50 | 0 | 100 |
|  | NTC\_Replicate\_3 | 0 | 0 | 0 |
| 23599, 24759, 25876, 26080, 26262 | PCR\_batch\_19 | Blank\_Replicate\_1 | 39 | 0 | 100 |
|  | NTC\_Replicate\_1 | 0 | 0 | 0 |
| PCR\_batch\_20 | Blank\_Replicate\_2 | 20 | 0 | 100 |
|  | NTC\_Replicate\_2 | 0 | 0 | 0 |
| PCR\_batch\_21 | Blank\_Replicate\_3 | 353884 | 0 | 100 |
|  | NTC\_Replicate\_3 | 0 | 0 | 0 |

**Supplementary Table S4**Unique plant taxa detected in each zone. Zones are evaluated from on a depth-constrained CONISS cluster analysis using the plant assemblage composition data.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Zone 1**  **(28–19 ka)** | **Zone 2**  **(18–14 ka)** | **Zone 3**  **(14–0 ka)** |
|  | **Forbs:**  *Monolepis asiatica*  *Erysimum* sp.  *Transberingia bursifolia*  *Eremogone capillaris*  *Minuartia* sp.  *Silene linnaeana*  *Silene* sp.  *Claytonia eschscholtzii*  *Claytonia* sp.  *Pedicularis oederi*  *Armeria scabra*  *Delphinium* sp.  *Ranunculus glacialis*  *Ranunculus sulphureu*  *Micranthes calycina*  *Saxifraga* sp.  **Herbs:**  *Phalaris* *arundinacea*  *Poinae*  *Puccinellia* sp. | (No unique taxa) | **Trees and erect shrubs:**  *Ulmus* *glabra*  *Sambucus* sp.  *Ribes* sp.  **Dwarf shrubs:**  *Andromeda* *polifolia*  *Chamaedaphne* *calyculata*  *Vaccinium* *ovalifolium*  *Vaccinium* sp.  *Comarum* *palustre*  Fragariinae  *Spiraea* sp.  **Forbs:**  Gnaphalieae  *Pedicularis* *lapponica*  *Hippuris* sp.  *Aconogonon* *tripterocarpum*  *Ranunculus* *reptans*  *Ranunculus* sp. 1  *Galium* *boreale*  **Herbs:**  *Carex* *vaginata*  *Dryopteris* *fragrans*  *Equisetum* sp.  Lycopodiaceae |
| Total number of unique taxa | 19 | 0 | 21 |

**Supplementary Table S5** Pearson correlation of taxonomic richness using data from individual PCR replicate batches (replicate 1, 2, and 3), summed counts of three replicates (submitted data).

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlation** | **Richness of replicate 1** | **Richness of replicate 2** | **Richness of replicate 3** |
| Richness using summed replicates  (submitted data) | r = 0.98  (*p* < 0.001) | r = 0.98  (*p* < 0.001) | r = 0.98  (*p* < 0.001) |

**Supplementary Table S6** Pearson correlation of taxonomic richness and phylogenetic diversity by taxonomic rank. The abundance- and occurrence- based nearest taxon index (NTI) were used to indicate the phylogenetic alpha diversity. The sequences were assigned by ArctBorBryo with 100%. All taxa - richness of all 158 taxa with taxonomic levels at family and lower levels. Genus aggregated - aggregated richness of taxa within genus and lower levels. Family aggregated - aggregated richness of taxa within family and lower levels.

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlation** | **28–0 ka** | **28–14 ka** | **14–0 ka** |
| **Taxonomic richness** |  |  |  |
| All taxa ~ Genus aggregated | r = 0.99 (*p* < 0.001) | r = 0.99 (*p* < 0.001) | r = 0.99 (*p* < 0.001) |
| All taxa ~ Family aggregated | r = 0.93 (*p* < 0.001) | r = 0.95 (*p* < 0.001) | r = 0.92 (*p* < 0.001) |
| **Abundance-weighted NTI** | | | |
| All taxa ~ Genus aggregated | r = 0.66 (*p* < 0.001) | r = 0.81 (*p* < 0.001) | r = 0.55 (*p* < 0.005) |
| All taxa ~ Family aggregated | r = 0.57 (*p* < 0.001) | r = 0.72 (*p* < 0.001) | r = 0.45 (*p* < 0.05) |
| **Occurrence-based NTI** | | | |
| All taxa ~ Genus aggregated | r = 0.76 (*p* < 0.001) | r = 0.77 (*p* < 0.001) | r = 0.62 (*p* < 0.001) |
| All taxa ~ Family aggregated | r = 0.68 (*p* < 0.001) | r = 0.69 (*p* < 0.001) | r = 0.45 (*p* < 0.05) |

**Supplementary Table S7** Pearson correlation between nearest taxon index (NTI) of all plants (158 taxa, submitted data) and angiosperms (136 taxa) in the dataset. Both abundance- and occurrence- based data were analysed.

|  |  |  |  |
| --- | --- | --- | --- |
| **Correlation** | **28-0 ka** | **28-14 ka** | **14-0 ka** |
| Abundance-weighted data | | | |
| NTI of all taxa ~ NTI of angiosperms | r = 0.72 (*p* < 0.001) | r = 0.83 (*p* < 0.001) | r = 0.79 (*p* < 0.001) |
| Occurrence-based data | | | |
| NTI of all taxa ~ NTI of angiosperms | r = 0.64 (*p* < 0.001) | r = 0.81 (*p* < 0.001) | r = 0.66 (*p* < 0.001) |