Received: 8 June 2016 | Revised: 13 April 2017 | Accepted: 25 April 2017

DOI: 10.1111/1755-0998.12689

RESOURCE ARTICLE

WILEY MOLECULAR ECOLOGY RESOURCES

A comparison of sedimentary DNA and pollen from lake sediments in recording vegetation composition at the Siberian treeline

Bastian Niemeyer[1,2] | Laura S. Epp[1] | Kathleen R. Stoof-Leichsenring[1] | Luidmila A. Pestryakova[3] | Ulrike Herzschuh[1,2,4]

[1]Periglacial Research Section, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany

[2]Institute of Earth and Environmental Science, University of Potsdam, Potsdam-Golm, Germany

[3]Department for Geography and Biology, North-Eastern Federal University of Yakutsk, Yakutsk, Russia

[4]Institute of Biochemistry and Biology, University of Potsdam, Potsdam-Golm, Germany

Correspondence
Bastian Niemeyer and Ulrike Herzschuh, Periglacial Research Section, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany.
Emails: Bastian.Niemeyer@awi.de and Ulrike.Herzschuh@awi.de

Funding information
Bundesministerium für Bildung und Forschung, Grant/Award Number: 01DJ14003; Deutsche Forschungsgemeinschaft, Grant/Award Number: EP98/2-1 to L.S.E.; Russian Foundation for Basic Research, Grant/Award Number: 13-05-00327

Abstract

Reliable information on past and present vegetation is important to project future changes, especially for rapidly transitioning areas such as the boreal treeline. To study past vegetation, pollen analysis is common, while current vegetation is usually assessed by field surveys. Application of detailed sedimentary DNA (sedDNA) records has the potential to enhance our understanding of vegetation changes, but studies systematically investigating the power of this proxy are rare to date. This study compares sedDNA metabarcoding and pollen records from surface sediments of 31 lakes along a north–south gradient of increasing forest cover in northern Siberia (Taymyr peninsula) with data from field surveys in the surroundings of the lakes. sedDNA metabarcoding recorded 114 plant taxa, about half of them to species level, while pollen analyses identified 43 taxa, both exceeding the 31 taxa found by vegetation field surveys. Increasing *Larix* percentages from north to south were consistently recorded by all three methods and principal component analyses based on percentage data of vegetation surveys and DNA sequences separated tundra from forested sites. Comparisons of the ordinations using PROCRUSTES and PROTEST analyses show a significant fit among all compared pairs of records. Despite similarities of sedDNA and pollen records, certain idiosyncrasies, such as high percentages of *Alnus* and *Betula* in all pollen and high percentages of *Salix* in all sedDNA spectra, are observable. Our results from the tundra to single-tree tundra transition zone show that sedDNA analyses perform better than pollen in recording site-specific richness (i.e., presence/absence of taxa in the vicinity of the lake) and perform as well as pollen in tracing vegetation composition.

KEYWORDS
environmental DNA, metabarcoding, pollen, Siberia, *trn*L marker, vegetation

1 | INTRODUCTION

The reliable assessment of vegetation changes on small spatial scales and through time is a pressing research target in the light of current global ecosystem perturbations. Environmental DNA (eDNA), such as DNA archived in lake sediments (sedimentary DNA or sedDNA) is increasingly used to reconstruct past vegetation (Alsos et al., 2016; Anderson-Carpenter et al., 2011; Giguet-Covex et al., 2014; Jørgensen et al., 2012; Parducci et al., 2013), but the degree of reliability of this tool has not been fully assessed (Parducci et al., 2014). Studies comparing the taxonomic overlap of pollen and DNA in lake sediment core records suggest that the DNA stems from highly local

sources (Boessenkool et al., 2014; Jørgensen et al., 2012; Pedersen et al., 2014), and a comparison of lake sediment core DNA and historical planting records has ascertained that the chronology of the sedDNA is reliable and not confounded by leaching through the sediment (Sjögren et al., 2016). For eDNA extracted from terrestrial soils, it has been confirmed that it mirrors the vegetation growing on the soils (Yoccoz et al., 2012), but a similar analysis of DNA from lake sediments across different vegetation types is lacking to date. Such analyses, however, are indispensable to prove that it is not lake-to-lake differences in the taphonomy but the site-to-site difference in the vegetation that dominates the sedDNA signal. Such proof represents a prerequisite to apply sedDNA to lake sediment records with the purpose of millennial-scale vegetation reconstruction as lakes typically undergo long-term changes in their hydrologic and sedimentation regimes.

The northern Siberian boreal treeline ecotone, covering an area of northern taiga to southern tundra, is one of the most prominent ecosystem boundaries and vegetation gradients on earth, even visible from space. It is known for its particular sensitivity to temperature change (Lloyd, Bunn, & Berner, 2011; MacDonald, Kremenetski, & Beilman, 2008), and a northward shift is expected in response to global warming (ACIA 2004). This change will potentially have significant effects not only on northern ecosystems and biodiversity, but also on feedback to the climate system due to albedo changes (Bonan, 2008). The reconstruction of past positions of the treeline is therefore a pressing ecological research target and has been addressed by numerous palaeoecological studies (e.g., Anderson, Lozhkin, & Brubaker, 2002; Andreev, Schirrmeister, et al., 2002; Andreev, Siegert, et al., 2002; Binney, Gething, Nield, Sugita, & Edwards, 2011; Blyakharchuk & Sulerzhitsky, 1999; Frost & Epstein, 2014; Hahne & Melles, 1997; Niemeyer, Herzschuh, & Pestryakova, 2015; Pellatt, Smith, Mathewes, & Walker, 1998). However, a reliable reconstruction of past treeline changes still presents a methodological challenge. Even current vegetation assessments based on vegetation field surveys are challenging in such remote, high latitude areas, which can typically be investigated only for a limited amount of time during the very short growing season.

Traditionally, treeline transitions are studied using pollen analyses of lake sediments (Gervais, MacDonald, Snyder, & Kremenetski, 2002; Moser & MacDonald, 1990; Pellatt et al., 1998; Seppä, Nyman, Korhola, & Weckström, 2002), but unfortunately, the relationship between sedimentary pollen and surrounding vegetation is not straightforward. Siberia encompasses Earth's largest boreal treeline area, which is formed solely by larch (*Larix* sp.). This poses a particular problem to pollen-based vegetation reconstructions, because, compared to other woody taxa, larch pollen is only produced in low amounts (Li et al., 2015; Sjögren, van der Knaap, Huusko, & van Leeuwen, 2008). It is also poorly dispersed and even more poorly preserved in sediments and thus strongly underrepresented in pollen records (Binney et al., 2011; Niemeyer, Klemm, Pestryakova, & Herzschuh, 2015; Niemeyer, Herzschuh, et al., 2015). To guide the inference of treeline transitions in fossil pollen records, modern pollen signals from lakes located along forest tundra transects have been investigated (Kay, 1979; Moser & MacDonald, 1990; Pelankova et al., 2008; Seppä et al., 2002; Spear, 1993; Velichko, Andreev, & Klimanov, 1997). An adequately reliable record of small-scale vegetation changes within the forest tundra ecotone was only possible when plant macrofossils could be used as an additional proxy (Birks & Birks, 2000), but these are sparsely recorded in lake sediments (Birks, 2001).

Employing DNA metabarcoding (Taberlet et al., 2012), sedDNA can be used to identify plants to a high taxonomic resolution (Sønstebø et al., 2010; Taberlet et al., 2007), especially when specific sequence reference databases can be employed for sequence identification, as is possible for arctic and boreal plants (e.g., Soininen et al., 2015; Sønstebø et al., 2010; Willerslev et al., 2014). Although earlier studies report a lower diversity retrieved by sedDNA (Boessenkool et al., 2014; Parducci et al., 2014; Paus et al., 2015; compiled by Birks & Birks, 2016), the putatively local source area of sedDNA (Boessenkool et al., 2014; Jørgensen et al., 2012; Parducci et al., 2014) suggests that sedDNA analysis could be a particularly appropriate tool to assess small-scale vegetation changes, as in the forest tundra ecotone.

Here we investigate and compare the potential of sedDNA and physical pollen signals to record the present position of the treeline in Siberia, in order to assess their comparative power to track historical changes. In contrast to earlier studies, which have compared pollen and sedDNA merely by assessing the overlap in taxa retrieved (Boessenkool et al., 2014; Jørgensen et al., 2012; Pedersen et al., 2014), we here conduct an explicit statistical comparison of ordinations of the multivariate data sets. We investigated records from modern lake sediments from the treeline area in northern Siberia (Russia) and aimed (i) to investigate whether the results of three different vegetation assessment techniques (vegetation field surveys, sedDNA metabarcoding, pollen analyses) are comparable in tracing vegetation richness and (ii) to assess whether sedDNA analysis is capable of providing additional information to pollen analysis in future palaeovegetation studies.

2 | MATERIALS AND METHODS

2.1 | Study area

The study area, located on the southern Taymyr peninsula (Krasnoyarsk Krai, Russia; Figure 1), extends along a north–south transect (71.1–72.7°N and 105.9–100.8°E) spanning approximately 250 km between the most distant lakes. The sites ranged from tundra areas mainly characterized by herbaceous taxa, via single-tree tundra with open vegetation and a few larch trees, to forest tundra areas characterized by higher abundances of tree and shrub taxa. This lowland area was not glaciated during the Last Glacial (Bigelow et al., 2003; Hahne & Melles, 1997; Möller, Bolshiyanov, & Bergsten, 1999).

The landscape, which is underlain by continuous permafrost, is flat and characterized by a large number of thermokarst lakes. Active layer depths up to 80 cm were measured in the field. Mean July and January temperatures are 14.0°C and 31.1°C, respectively, and the

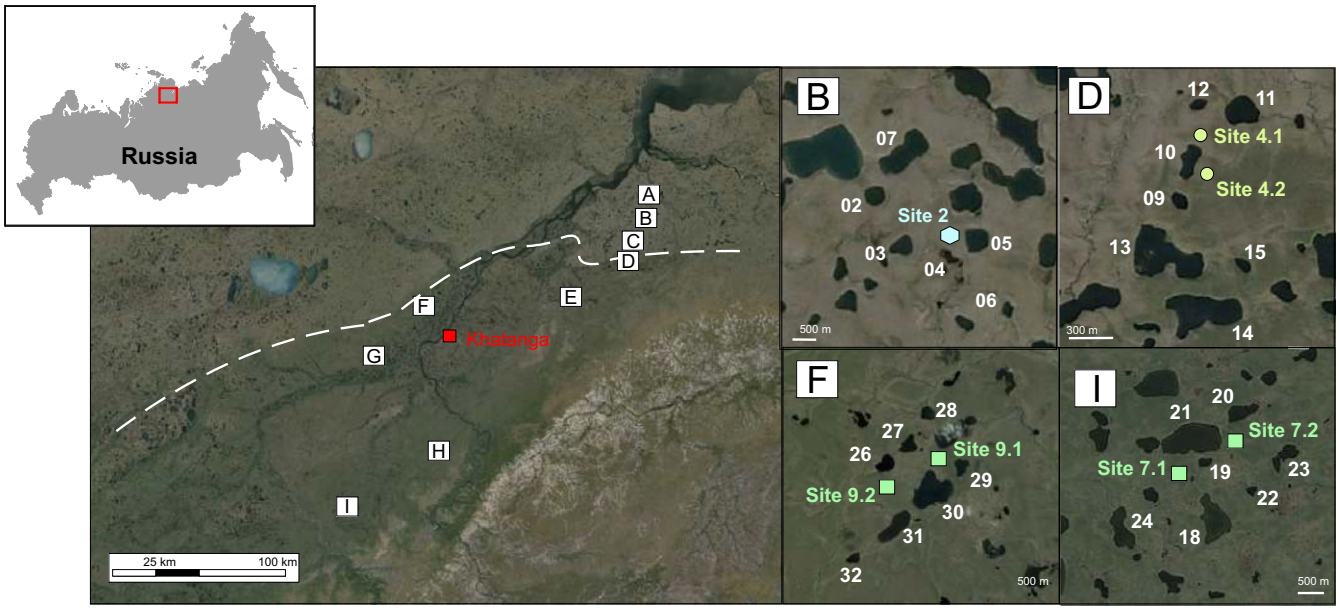

FIGURE 1 Location of the vegetation field survey sites and of the sampled lakes across a north–south transect across the treeline ecotone in the Khatanga region (main image) and its location in Russia (inset). The border between tundra and forest tundra is indicated by the dashed line. The locations of lakes sampled as single lakes (A, C, E, G, H) or as part of a lake group (B, D, F, I) are indicated. Images adapted from Google Earth

mean annual precipitation is 252 mm (Khatanga weather station 71.969°N and 102.467°E, WMO-ID: 20891; 1906–2014; Global Historical Climatology Network, Menne, Durre, Vose, Gleason, & Houston, 2012). The area is covered by a thin layer of snow in winter. The southernmost study sites are situated within *Larix* forest tundra (mainly composed of *L. gmelinii* (Rupr.) Rupr.) (Abaimov, 2010), while northernmost sites are located in the tundra where some larch krummholz patches are present but shrubs and herbs dominate the vegetation.

2.2 | Vegetation field surveys and lake sediment sampling

Fieldwork was conducted in July and August 2013. In total, 32 medium-sized lakes (Table 1; radii range: 50–150 m; median: 80 m) surrounded by regionally characteristic and homogenous vegetation were selected for sampling with the help of satellite images and after visual inspection from a helicopter. The vegetation surveys focused on seven sites, each associated with a specific cluster of lakes (Table 1). The vegetation survey sites represent the three major vegetation types present in our study area: tundra, single-tree tundra and forest tundra. For most of the survey sites, plots of 20 × 20 m (400 m^2) with 1 × 1 m subplots comprising the typical vegetation were assessed by estimating individual taxon cover (%) and are located between 100 and 2,000 m from the lake shores. To ensure that spatial variability within a 2 km radius has only a small influence on the vegetation composition and richness, especially for the patchy vegetation areas of the tundra to single-tree tundra border, we carried out two comparative vegetation field surveys at three of the study sites (4, 7 and 9; Table 1, Figure 1). Because of greater heterogeneity, the northernmost vegetation survey was conducted in 47 1 × 1 m plots distributed within a 0.25-km^2 area. Vascular plant cover was averaged across each survey site and the percentage of each taxon relative to the total vascular plant sum was calculated. Our tight fieldwork schedule in this remote area precluded an extensive vegetation investigation. Nevertheless, we conducted the surveys in a way that the time spent on them was approximately estimated to be equivalent to the time spent on collecting and later processing the sediment samples in the laboratory.

Lake sediments were sampled from a rubber boat using an Ekman-Birge bottom sampler that collects approximately the uppermost 10 cm of undisturbed surface sediment. Samples for DNA and pollen analyses were retrieved from the uppermost 1 cm of surface sediment, which is considered to reflect the last 5–10 years as inferred by Niemeyer, Herzschuh, et al. (2015), and transferred by sterile spatula into plastic bottles (Nalgene). These bottles had been previously cleaned in a dedicated ancient DNA laboratory at the Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI) in Potsdam, Germany, by consecutively using sodium hypochlorite, ethanol and DNA-free water, and subjecting to UV-light exposure at 10 cm distance for 15 min. All samples were stored cool and dark during transportation and were finally stored at 7°C at AWI Potsdam until further analysis.

2.3 | Pollen analysis

Pollen preparation followed the standard procedure described in Faegri and Iversen (1989). From each subsample, 1 ml of sediment was taken using a syringe. Before treatment, a *Lycopodium* spore tablet (Batch nr. 1031; $n = 20{,}848$) was added to each of the

TABLE 1 Name, location and information on surrounding vegetation of the sampled lakes

Lake name	Latitude (°N)	Longitude (°E)	Vegetation type	Lake cluster	Nearest vegetation survey site (site number)
13-TY-01	72.6678	105.8807	Tundra	A	–
13-TY-02	72.5531	105.7175	Tundra	B	2
13-TY-05	72.5481	105.7505	Tundra	B	2
13-TY-03	72.5479	105.7258	Tundra	B	2
13-TY-04	72.5473	105.7421	Tundra	B	2
13-TY-06	72.5416	105.7635	Tundra	B	2
13-TY-08	72.4883	105.6483	Tundra	C	–
13-TY-12	72.4129	105.4477	Single-tree tundra	D	4.1
13-TY-11	72.4118	105.4636	Single-tree tundra	D	4.1
13-TY-10	72.4062	105.4425	Single-tree tundra	D	4.2
13-TY-09	72.4008	105.4399	Single-tree tundra	D	4.2
13-TY-15	72.3924	105.4658	Single-tree tundra	D	4.2
13-TY-13	72.3916	105.4435	Single-tree tundra	D	4.2
13-TY-14	72.3866	105.4566	Single-tree tundra	D	4.2
13-TY-16	72.1810	104.4879	Single-tree tundra	E	–
13-TY-28	72.1549	102.1013	Forest tundra	F	9.1
13-TY-27	72.1533	102.0754	Forest tundra	F	9.1
13-TY-26	72.1493	102.0560	Forest tundra	F	9.2
13-TY-29	72.1484	102.1160	Forest tundra	F	9.1
13-TY-30	72.1420	102.0905	Forest tundra	F	9.1
13-TY-31	72.1360	102.0680	Forest tundra	F	9.2
13-TY-32	72.1278	102.0344	Forest tundra	F	9.2
13-TY-25	71.8865	101.2170	Forest tundra	G	–
13-TY-17	71.4031	102.2826	Forest tundra	H	–
13-TY-20	71.1122	100.8529	Forest tundra	I	7.2
13-TY-21	71.1067	100.8229	Forest tundra	I	7.2
13-TY-23	71.1038	100.8760	Forest tundra	I	7.2
13-TY-19	71.1020	100.8272	Forest tundra	I	7.1
13-TY-22	71.0971	100.8539	Forest tundra	I	7.2
13-TY-24	71.0959	100.7975	Forest tundra	I	7.1
13-TY-18	71.0927	100.8310	Forest tundra	I	7.1

The related lake cluster and nearest field vegetation survey site are also indicated. The dashes (–) indicate lakes which do not have a vegetation field survey site in their vicinity.

subsamples and diluted by hydrochloric acid (HCl; 10%). The subsequent procedure included treatment by a potassium hydroxide (KOH; 10%) water bath to remove organic material, followed by sieving of macrofossils in the samples (200 μm mesh size), hot hydrofluoric acid (42%) to dilute the minerogenic content and a two minute acetolysis (consisting of acetic anhydrite and sulphuric acid 9:1) treatment to colour the pollen grains for microscopic identification. To remove small pollutants from the samples, we used an ultrasonic bath (sieve size: 7 μm) for a maximum of 30-s, after having tested that this ultrasound treatment did not damage large pollen grains, such as *Larix*. Pollen samples were analysed using a Zeiss Axioskop 40 light microscope at 400× magnification. Identification was based on relevant literature (Beug, 2004; Moore, Webb, & Collinson, 1991; Savelieva, Raschke, & Titova, 2013) and the pollen reference collection of the AWI in Potsdam. At least 498 terrestrial pollen (arboreal and nonarboreal) grains were counted in each sample and included in the pollen sum. Taxa which were present in only one sample and had a maximum of two pollen grains per sample were excluded from the data set. Three randomly selected samples were analysed twice, and comparison of respective spectra yielded negligible differences.

2.4 | Sedimentary DNA analysis

2.4.1 | Molecular genetic laboratory work

Extractions were performed in a dedicated DNA isolation and pre-PCR laboratory at the Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research in Potsdam, Germany, working

under a UV cabinet (AirClean600 PCR Working station; Star Lab), which is only used for DNA extractions of environmental samples. The DNA isolation laboratory is physically separated from the post-PCR area to prevent contamination of DNA samples with PCR products. Approximately 8 g of sediment was taken from the samples and processed using the PowerMax® Soil DNA Isolation Kit (MoBio Laboratories, Inc., USA). For lysis, the samples were added to the Powerbead solution with 0.8 mg proteinase K and 1.2 ml of C1 buffer, vortexed thoroughly and incubated overnight in a shaker at 56°C. The successive steps of the extraction protocol were carried out according to the manufacturer's instructions, using 1.6 ml elution buffer for the final elution step. In total, we extracted 32 samples in four batches of seven or nine samples with one extraction blank per batch. Extraction blanks were processed in the same way as the samples. PCRs were set up for each of the batches separately, each including a negative control. PCR set-up was performed under a dedicated UV cabinet in the pre-PCR laboratory. PCRs were set up in 25 μl volumes containing 1.25 U Platinum® Taq High Fidelity DNA Polymerase (Invitrogen), 1× PCR buffer, 2 mM $MgSO_4$, 1 mM dNTPs, 0.2 mM of each primer, 0.8 mg/ml bovine serum albumin and 3 μl DNA extract and carried out in a Biometra Professional thermocycler. PCR conditions were 94°C for five minutes, followed by 50 cycles of 94°C for 30 s, 50°C for 30 s for annealing, 68°C for 30 s and a final elongation at 72°C for 10 min. Because we aimed to assess the power of DNA metabarcoding to track historical changes in vegetation, we chose the short marker amplified using the primers *trn*L *g* and *trn*L *h* (Taberlet et al., 2007).We also performed the PCRs with a high number of cycles, as it is customary in aDNA studies. Both primers carried unique 8-bp tags on the 5′ end that varied from each other in at least five base pairs and were elongated by NNN tagging (De Barba et al., 2014) to improve generating clusters on the Illumina sequencer. All 32 extracted samples were processed in four PCR runs, including the appropriate extraction blank and a negative control per PCR. One sample did not yield any PCR product and was therefore excluded from further analysis. In total, our set-up included four PCR negative controls and four extraction blanks, and these were negative in all PCR runs that were analysed further. Each PCR was repeated and 15 μl of the two PCR products from the same sample were merged. This pooling was performed to increase the richness and the probability of tracing rare sequences. However, a recent study has shown that the main vegetation components can already be retrieved from a single PCR (Alsos et al., 2016). Pooled samples were purified using the MinElute PCR Purification Kit (Qiagen). The DNA concentration was quantified with the Qubit® 2.0 fluorometer (Invitrogen), using the broad range DNA assay, and the purified PCR products were pooled equimolarly. Library preparation and sequencing (2 × 100 bp, paired-end reads) was performed by an external sequencing service (Fasteris SA, Switzerland), using 10% of an Illumina HiSeq lane. Libraries were prepared according to the MetaFast protocol, which omits a PCR step in the library build. Samples were sequenced on two individual runs, and the data were later combined into a single data set.

2.4.2 | Analysis of sequence data

To filter, sort and assign taxonomic information to the resulting DNA sequences, we used the OBITOOLS package (http://metabarcoding.org/obitools/doc/welcome.html; Boyer et al., 2016). We used *illuminapairedend* to align forward and reverse reads, *obigrep* to remove sequences shorter than 10 bp in length, *obiuniq* to merge identical reads (coverage of 100% required) and *obiclean* to remove probable PCR errors. Taxonomic assignments were performed using the program ECOTAG and were based on a sequence reference database for Arctic and Boreal vascular plants and bryophytes (containing 1,664 vascular plants and 486 bryophyte species, published by Soininen et al. (2015), Sønstebø et al. (2010) and Willerslev et al. (2014)). Additionally, we used a sequence reference database constructed from the embl standard sequences release 117 (Kanz et al., 2005, http://www.ebi.ac.uk/ena) as described in Epp et al. (2015).

We based our further analyses on the sequences that showed a 100% match (best identity value 1.0 in the ECOTAG analysis) to sequences of plants occurring in Arctic regions and deposited in one of the sequence reference databases mentioned above. This is a conservative approach, but after applying this threshold, the record still provided 219 unique DNA sequences, of which 164 appeared in three or more samples, represented by several hundreds to thousands of reads per sequence. Some of them belonged to the same taxon, and reads of sequences that were identified to the same taxonomic unit were combined. Finally, 114 taxa were included in further data handling (Appendix S1). Sequences with full identity to sequences deposited in embl, but belonging to taxa that are considered highly unlikely to be authentic (e.g., because they stem from food plants or occur only in a very different geographical area; e.g., Malae and Triticeae; 106 and 66 sequence reads, respectively) and sequences of taxa that appeared in the blanks (*Festuca* and Pooideae; 14,804 and 12 sequence reads, respectively) were discarded from all samples. We also checked how many sequences and corresponding taxa were retrieved from three or more independent samples, and we consider these sequences to be more reliably present in the data set than sequences not retrieved independently. Before comparison to pollen and vegetation field data, all sequences belonging to aquatic flora (except for Cyperaceae, as they are also common elements of the arctic terrestrial flora) and bryophytes were extracted from the record, keeping only terrestrial plants for comparison.

2.5 | Statistical analyses

The multivariate data sets of seven vegetation survey sites and 31 lakes (for pollen and sedDNA data) were used for statistical analyses. To allow visual comparison of the different records and changes of abundance, percentage bar plots of vegetation, pollen counts, and sedDNA composition were produced using the TILIA software (Grimm, 1992). To realistically estimate the taxonomic richness in pollen and sedDNA and to test whether the pollen and sequence counts were sufficient to capture the diversity in each sample, the raw count data

were subjected to rarefaction analyses (Birks & Line, 1992). Only those taxa that occurred with a minimum of 0.5% in at least five samples were kept for the numerical analyses of the pollen and sedDNA data, as this restricts the analyses to the taxa that were most reliably retrieved from the samples and which are considered part of the dominant vegetation. For the vegetation field surveys, the same threshold was applied, but an additional taxon, which mainly occurred at the northernmost site, but there reached a cover of over 20%, was included too. All recorded DNA and pollen percentages, which met our criteria for inclusion in the statistical analyses, were square-root-transformed prior to computational analyses to limit the influence of outliers.

Principal component analysis (PCA; Bennett, 1996; Legendre, Oksanen, & ter Braak, 2011; ter Braak, 1983, 1994) was performed to investigate the major structure in the individual data sets. The transformation of the vegetation field data to pollen types in the PCA caused problems in the differentiation of the explanatory variables, because with the transformed data, the number of taxa (11) only slightly exceeds the number of sites (7). To circumvent this problem, here we did not use the transformed taxa, and we only explored the first two axes of the PCA as they reflect nearly all variables (Figure 2). To allow direct comparison of the records, vegetation and sedDNA data were transformed into the same taxonomic units as the pollen taxa data (e.g., *Pedicularis capitata* was transformed to *Pedicularis*) and a PCA was rerun for each record.

To compare the fit between site and species scores of the first four PCA axes for pollen and sedDNA, as well as vegetation against pollen and DNA records, PROCRUSTES rotation analyses was performed (Peres-Neto & Jackson, 2001) and the significance of similarity between the tested data sets was assessed by PROTEST (Jackson, 1995). All analyses were carried out using the R software (version L3.0.3; R Development Core Team, 2014) using packages "VEGAN" (Oksanen et al., 2011) and "RIOJA" (Juggins, 2009). All data sets are available at http://pangaea.de (please see Data availability section for details). The implementation of all three vegetation assessment approaches (sampling and laboratory work) required a comparable amount of time, approximately estimated at 8 hrs per sample or site.

3 | RESULTS

3.1 | Vegetation field survey

The vegetation field surveys noted 31 different vascular plant taxa, reflecting three different vegetation types (Figure 2). The tundra site is particularly characterized by a high *Dryas* (*D. octopetala*) cover of about 25% and a lack of alder (*Alnus alnobetula*), with only very few krummholz individuals of larch (*L. gmelinii*). Single-tree tundra is mainly characterized by up to 15% of *Cassiope tetragona*, Cyperaceae, *Betula nana* and *Eriophorum* sp. The forest tundra sites in contrast have a *Larix* cover of between 14% and 20% and are also characterized by *Alnus*, even though this occurs with varying abundance. In contrast to the tundra site, where *Salix glauca* obtains a cover of about 4%, *Salix* spp. only has a small cover at the other sites (median 0.2%). *Vaccinium* (*V. uliginosum* and *V. vitis-idaea*) occurs steadily at all sites with a cover of between 6% and 11%.

The first two axes of the PCA explain 83.6% of the variance. The PCA biplot of the vegetation survey data reveals that tundra vegetation is characterized by a distinct vegetation composition, in particular by high *Dryas* and *Betula* values (Figure 3). The plant taxa obtained by the vegetation survey could be transformed into 23 terrestrial pollen taxa, whereof 11 taxa surpassed the threshold to be included in further statistical analyses. They are presented in the reduced bar plot (Appendix S2) and a reduced PCA biplot (Appendix S3).

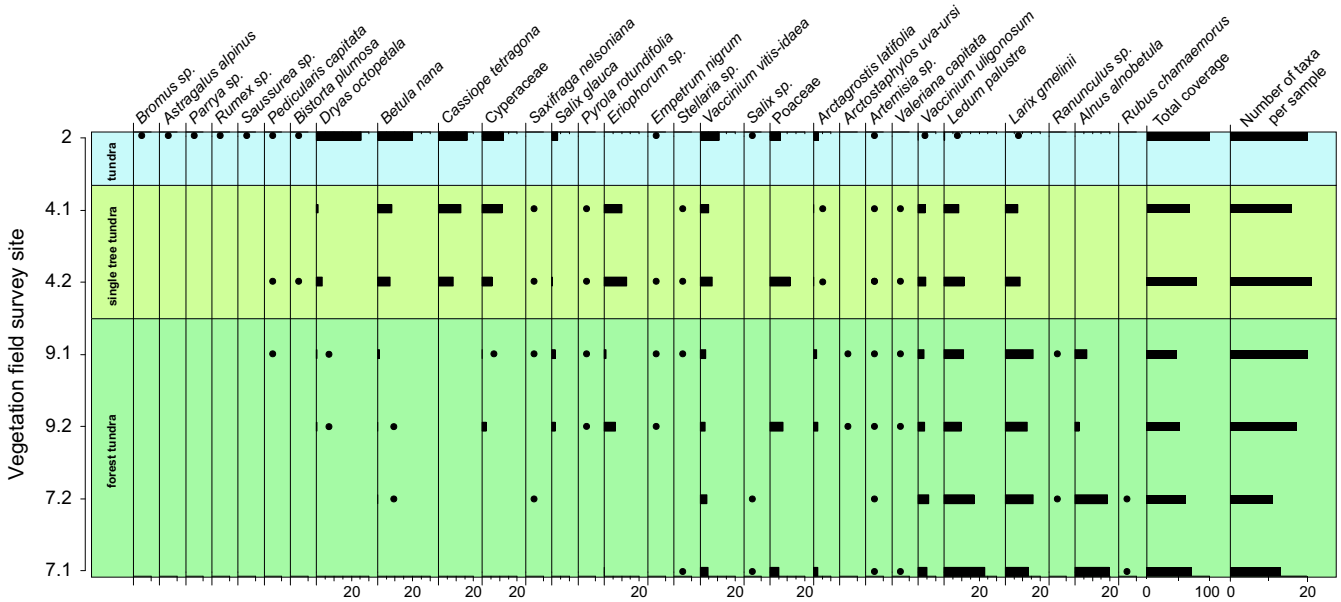

FIGURE 2 Vegetation cover as revealed by field surveys of 400-m^2 plots from tundra (top) to forest tundra (bottom). Black dots indicate percentages below 0.5%. Taxa are sorted according to their PCA taxa scores

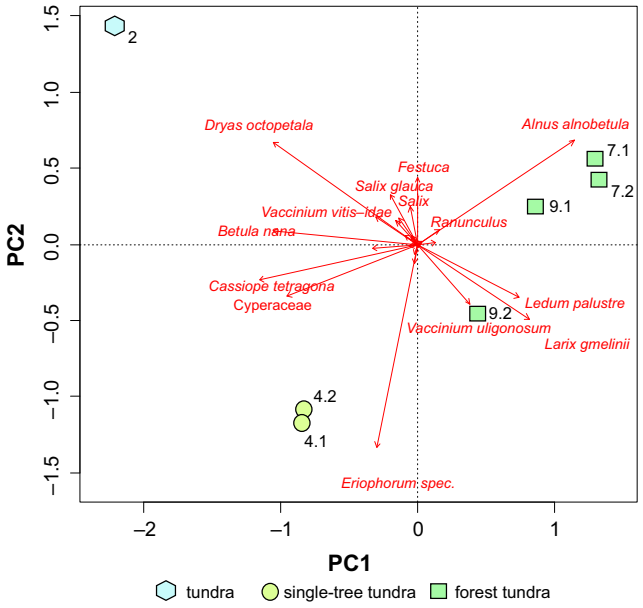

FIGURE 3 The first two axes of a principal component analysis of the vegetation field survey data, which cumulatively explain 90% of the variance. The plot indicates that the vegetation composition of the tundra site is distinct from those of the single-tree tundra and forest tundra sites. Taxa close to the centre of the plot are not labelled to enhance the readability of the plot

3.2 | Pollen data

In total, 43 different terrestrial vascular plant taxa were identified by pollen (Appendix S4), the majority of them to genus level (33), some to family level (8) and only two to species level. Of these, nineteen pollen taxa met the criteria to be included in the statistical analyses (Figure 4). Together, these taxa represent more than 95% of the investigated pollen grains in each sample. Rarefied pollen richness is about 25 taxa (relative to a pollen sum of 498 grains) in each sample, without any obvious trend across the transect. Rarefaction curves of all samples do not show an asymptotic response within a count sum of between 498 (indicated by the solid line) and 830 (Appendix S5).

Overall, *Betula*, *Alnus* and Cyperaceae are the dominant pollen taxa. *Betula* and *Alnus* have an increasing trend from north to south, while Cyperaceae decreases. Furthermore, Poaceae and *Vaccinium* occur in all samples with at least 15% and 4%, respectively. *Larix* pollen percentages are always below 5% and are very low at most tundra and single-tree tundra sites. Samples with a particularly high Pinaceae pollen content are found at tundra and single-tree tundra sites. These overall trends become more obvious in the biplot of the first two PCA axes (Figure 5), which together explain 52% of the variance in the data set. The pollen composition of tundra lake sediments is distinct from the pollen composition of lake sediments originating from single-tree tundra and forest tundra with respect to their higher percentages of Cyperaceae, *Artemisia* and Rosaceae. All spectra from tundra and most single-tree tundra lakes are located in the left part of the biplot, while most spectra from forest tundra lakes appear in the right part of the plot.

3.3 | sedDNA data

The sum of obtained terrestrial seed plant sequence reads in the single samples ranges between 3,722 and 257,881 after filtering. In total, 114 terrestrial seed plant taxa were identified by the sedDNA analyses, of which about 45% could be identified to species level. Of

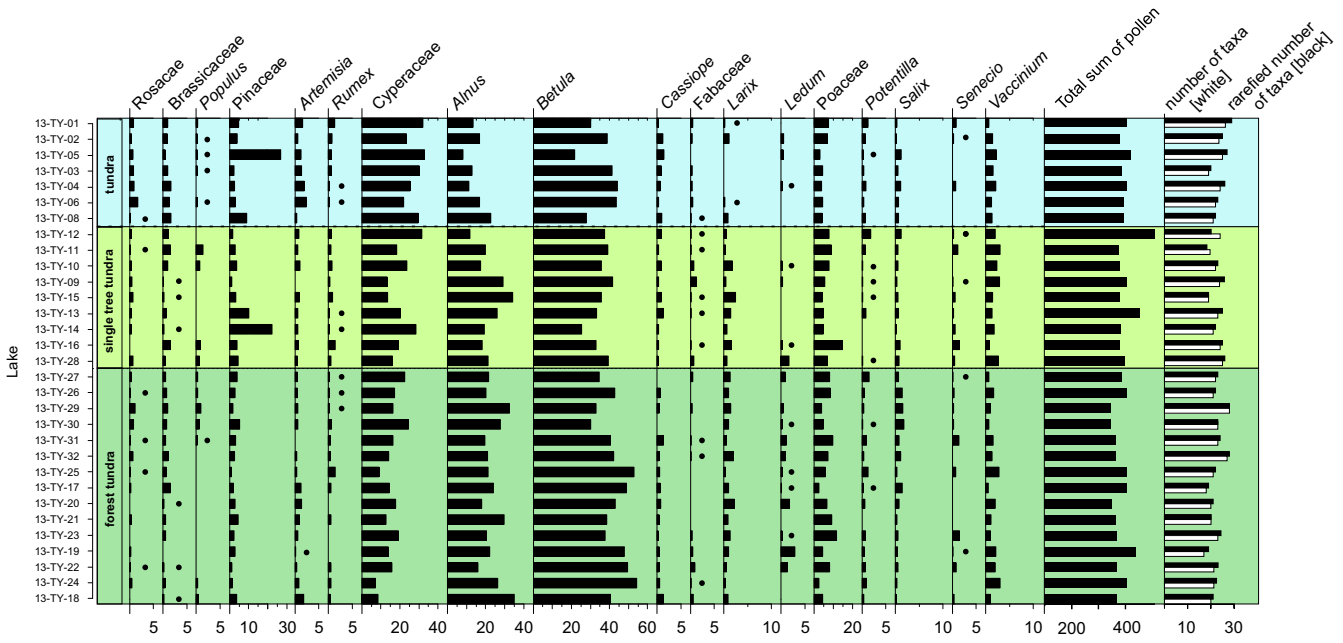

FIGURE 4 Bar plot of pollen spectra from 31 lakes located along a north–south transect in the Siberian treeline area. Pollen percentages of those taxa included in the statistical analyses are shown and are arranged according to their species scores. Single black dots identify occurrences of <0.5% of the total sum of pollen

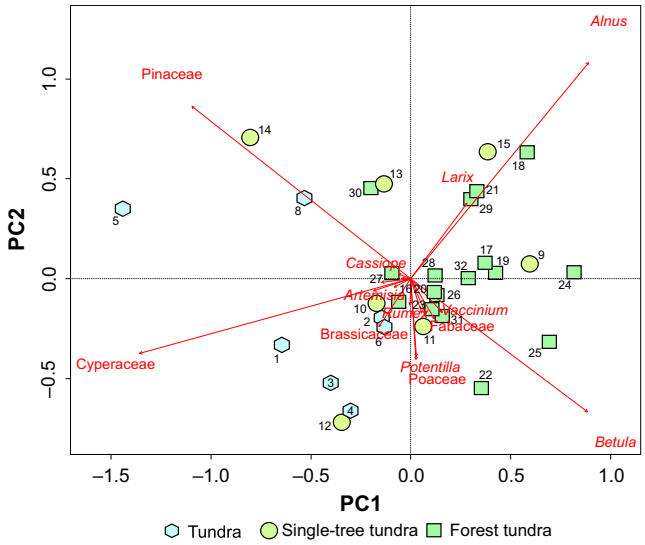

FIGURE 5 The first two axes of a principal component analysis of the pollen percentages that together explain 52% of the total variance. Samples from the tundra sites are mostly located separately from single-tree tundra and forest tundra sites. Taxa located close to the centre of the plot are not displayed

these, 95 were retrieved from three or more independent samples. Rarefied richness per sample varied between 54 and 14 taxa relative to a sequence number of 3,722, with no obvious trend across the transect. Approximately one-third of the lake samples do not show saturation in the rarefaction curves (Appendix S6).

Overall, the relative composition of taxa is rather variable even between nearby lakes (Figure 6). *Salix* is the dominant taxon in most spectra, but it shows no clear trend from forest tundra to tundra sites. *Betula* has high percentages, particularly in some tundra lakes. The percentages for *Carex aquatilis*, *Carex* spp. and especially *Dryas* are, on average, higher in the tundra samples than in samples from more southerly sites. *Larix* is below 1% in the samples from tundra lakes. *Larix* and Rosoideae percentages are lower in the single-tree tundra than in the forest tundra. *Alnus* percentages are below 10% (except for lake 13-TY-27), but are, on average, highest in the single-tree tundra lakes. Percentages of *Ledum* increase from single-tree to forest tundra.

When transformed to pollen taxonomic levels, 47 taxa were identified by the sedDNA approach (Appendix S7), of which 14 taxa surpass the threshold to be included in the statistical analyses (see bar plot in Appendix S8 and biplot in Appendix S9). The ordination biplot of the first two axes of the sedDNA data (Figure 7), explaining 46% of the data variance, distinguishes between samples from tundra lakes in the left lower corner and single-tree and forest tundra lakes in the upper right and upper left. Spectra from single-tree and forest tundra form a joint cluster in the biplot.

3.4 | Comparison of the ordinations by PROCRUSTES rotation and PROTEST

Results of the PROCRUSTES rotation analyses and associated PROTEST (Table 2) indicate significant accordance between the score of the sites for pollen and sedDNA ($m^{12} = 0.5949$, $p = .001$). Site-specific residuals show no trend along the transect. The pairwise comparison of the scores of the taxa among all three records likewise yields significant fits. The best fit, as indicated by a low *p*-value, is found between vegetation and pollen species scores. Taxa residuals of *Salix*, Rosaceae, *Ranunculus*, *Betula* and *Larix* are always above average (Figures 8 and 9), showing weak similarity for these taxa in all tested data set combinations.

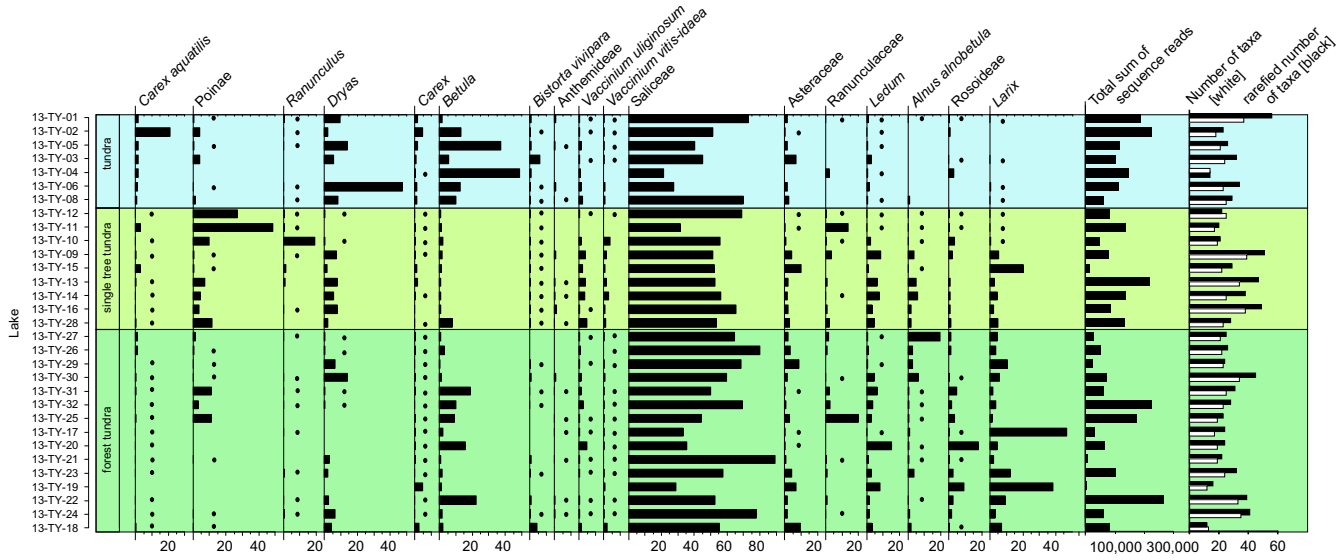

FIGURE 6 Bar plot of DNA spectra from 31 lakes located along a north–south transect in the Siberian treeline area. sedDNA sequence percentages matching the threshold applied for statistical analyses are shown and arranged according their species scores in ordination. Black dots indicate sequence percentages <0.5% of the total sum of DNA

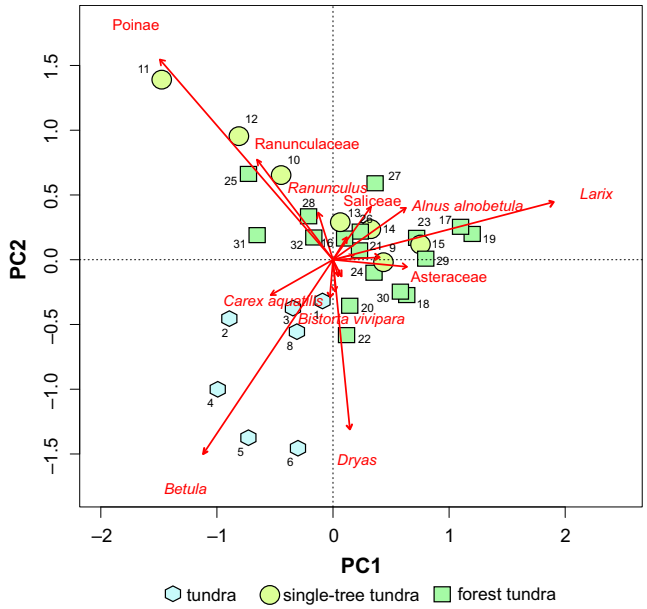

FIGURE 7 The first two axes of a principal component analysis of the DNA sequence percentages that together explain 46% of the total variance. Taxa located close to the centre of the plot are not displayed. Samples from the tundra sites are clearly separated from single-tree tundra and forest tundra sites

TABLE 2 Results of PROCRUSTES and PROTEST analyses indicate significant fit among all pairwise comparisons of data sets obtained by vegetation field survey, pollen and sedDNA analyses

	p-value	*r*	m^{12}	rmse
Pollen and sedDNA: sites	.001	.5923	0.6491	0.3405
Pollen and sedDNA: taxa	.005	.5088	0.7411	0.2604
Vegetation and sedDNA: taxa	.004	.5502	0.6973	0.2991
Vegetation and pollen: taxa	.001	.6848	0.5311	0.3602

p-value, likelihood of the relationship by chance; *r*, correlation between the two ordination results; m^{12}, PROCRUSTES rotation sum of squares; rmse, root mean square error.

4 | DISCUSSION

4.1 | Comparison of the vegetation field survey, pollen and sedDNA approaches in recording vegetation richness and composition

Overall we found a good fit between the results obtained by the three different methods—vegetation surveys, pollen and sedDNA analyses—that were used to record vegetation composition in the vicinity of lakes located across a forest tundra transect in north-central Siberia. All methods recorded a rather similar ensemble of the 10–15 most common taxa. Furthermore, all methods were able to differentiate typical tundra from forest tundra sites. However, a detailed comparison yielded certain peculiarities of each approach.

The results of the vegetation field surveys reliably reflect the general vegetation composition pattern of denser vegetation in the forest tundra compared with tundra, which is characterized by high *Ledum* cover (Zyryanova, Abaimov, Daimaru, & Matsuura, 2010), *Dryas* as a major element in various tundra plant community associations (Matveyeva, 1994) and increasing *Larix* cover from north to south. In general, the vegetation composition recorded by the field surveys clearly reflects the three vegetation types in the study area as also observed from satellite images and from helicopter.

The 31 vascular seed plant taxa in our *vegetation field surveys*, represent only about one-sixth of the taxa recorded by extensive (100 km^2) vegetation surveys conducted on the Taymyr peninsula (Matveyeva, 1994; Zyryanova et al., 2010) and in northern Yakutia (Kuznetsova et al., 2010), although the diversity is not fully comparable as we did not include bryophytes, which show a high diversity in north Siberia (Zibulski, Herzschuh, & Pestryakova, 2016). The bias mainly originates from the limited area of our surveys (missing the spatial heterogeneity) and from the once-only sampling strategy (missing the seasonal and interannual heterogeneity). The latter issue meant that some taxa could not be identified to species level because they were not, for example, collected in a flowering state. Thus, we assume that richness is strongly underestimated by our vegetation field surveys, although the temporal effort of the field surveys is comparable with that taken for the pollen and sedDNA approaches.

Pollen analyses of lake sediments yielded 43 taxa, which is slightly higher than other modern (Pisaric, MacDonald, Cwynar, & Velichko, 2001: 20 taxa; Klemm et al., 2013: 32 taxa) or fossil studies (Hahne & Melles, 1997: 22 taxa; Andreev, Siegert, et al., 2002: 38 taxa; Niemeyer, Herzschuh, et al., 2015: 36 taxa) performed in the area. Our pollen approach yielded approximately twice as many taxa as recorded by the vegetation field survey even though pollen data are mostly limited to the taxonomic level of genera, which should cause an underestimation of species richness by pollen analyses. Moreover, the rarefaction curves indicate that our pollen counts were not sufficient to capture the entire richness, as none of the rarefaction curves reach saturation.

Pollen spectra across all vegetation types are dominated by shrub taxa, in particular by *Alnus* and *Betula*. This is in accordance with previous studies by Klemm et al. (2013) and Pisaric et al. (2001), who investigated forest tundra transects located further east. The high percentages of herbs (Cyperaceae, Brassicaceae, *Artemisia*, Rosaceae) that we found are also in agreement with results from these previous studies. Other studies corroborate our finding that high Pinaceae percentages characterize tundra sites although the nearest pine stands are located several hundreds of kilometres to the south (Klemm et al., 2013; Tchebakova, Rehfeld, & Parfenova, 2005). These results are reasonable, as *Pinus* is a common element of the far-distance wind-transported component of pollen spectra in nonforested areas where pollen productivity of local and regional vegetation is low (Campbell, McDonald, Flannigan, & Kringayark, 1999; Hjelmroos, 1991).

In agreement with the vegetation surveys, percentages of *Alnus* and *Larix* are higher in the forest tundra lakes than in the tundra lakes. However, with respect to their relative abundance in the

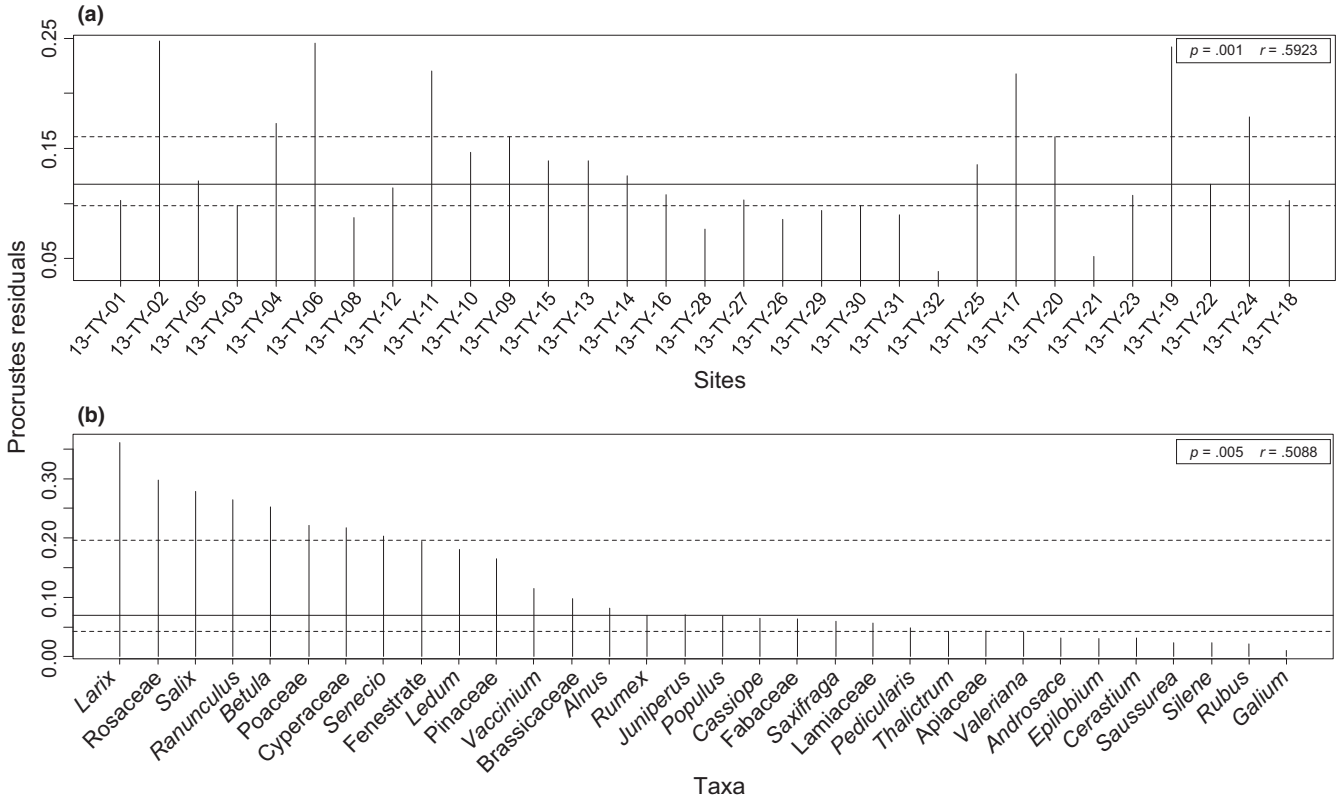

FIGURE 8 PROCRUSTES residuals plot of comparison (a) between PCA site scores of pollen and sedDNA (residuals were ordered according to their position in the transect) and (b) between respective PCA taxa scores (ordered according to the residual scores). Dashed and solid lines are the first, second and third quartiles. The *p*-value indicates the likelihood of the relationship occurring by chance and the *r*-value reflects the correlation between the two ordination results by superimposition

record distinct biases exist, which probably originate mostly from differences in the pollen productivity of the individual taxa (Davis, 1963; Sugita, 1994). Niemeyer, Klemm, et al. (2015) found for this region that, relative to Poaceae as a reference, *Alnus* is almost three times overrepresented while *Larix* is underrepresented by a factor of 0.16 in lake sediments compared to the surrounding vegetation. This is in agreement with results of the PROCRUSTES comparison of species scores, which yield above-average residuals for these two taxa. Other taxa, such as *Ledum*, Poaceae and *Vaccinium* also show strong differences in their representation within pollen spectra compared to vegetation, indicating only medium fit of the ordination results.

This study is the first, to our knowledge, to compare sedDNA data from modern lake sediments along a transect of changing vegetation to data retrieved from vegetation field surveys and pollen analyses. The sedDNA approach yielded more than three times the number of taxa found by the vegetation field survey, even when considering only taxa that were retrieved from three independent samples. This holds true despite both approaches requiring a comparable amount of time. sedDNA analyses also retrieved more taxa than pollen analyses. Even after conversion of sedDNA taxa to the identification level reached by pollen, sedDNA richness is higher than pollen richness. In this respect our results contradict previous studies which reveal a lower richness in the (ancient) sedDNA records compared with pollen or macrofossil records from the same samples (Boessenkool et al., 2014; Parducci et al., 2014). We note that these studies were conducted on ancient material, in which DNA preservation might have been lower, and that they were conducted using different sequencing technologies (Roche 454 platform or Sanger sequencing of clones, while we used Illumina sequencing). Additionally, some of the earlier studies used long fusion primers in the PCR set-up, which included the adapters used on the sequencing platform, and which might diminish the efficiency of the PCR. In terms of richness, our data are, however, comparable to those reported by Willerslev et al. (2014).

The significant fit between the PCA species scores of sedDNA and vegetation field survey data, as well as PCA scores of sedDNA and pollen data, indicates that a similar composition of major taxa was recorded by each of the three approaches. Interestingly, all methods are able to identify tundra composition although the ensembles of taxa were slightly different in the respective analyses: while the vegetation field survey identified *D. octopetala*, *B. nana* and *C. tetragona* as most characteristic taxa of tundra records, the sedDNA identified *Dryas*, *Betula* and Cyperaceae (such as *Carex vaginata* or *Eriophorum*) and the pollen record identified Cyperaceae and Pinaceae as the main taxa, instead. Furthermore, all three approaches fail to distinguish between records from single-tree and forest tundra.

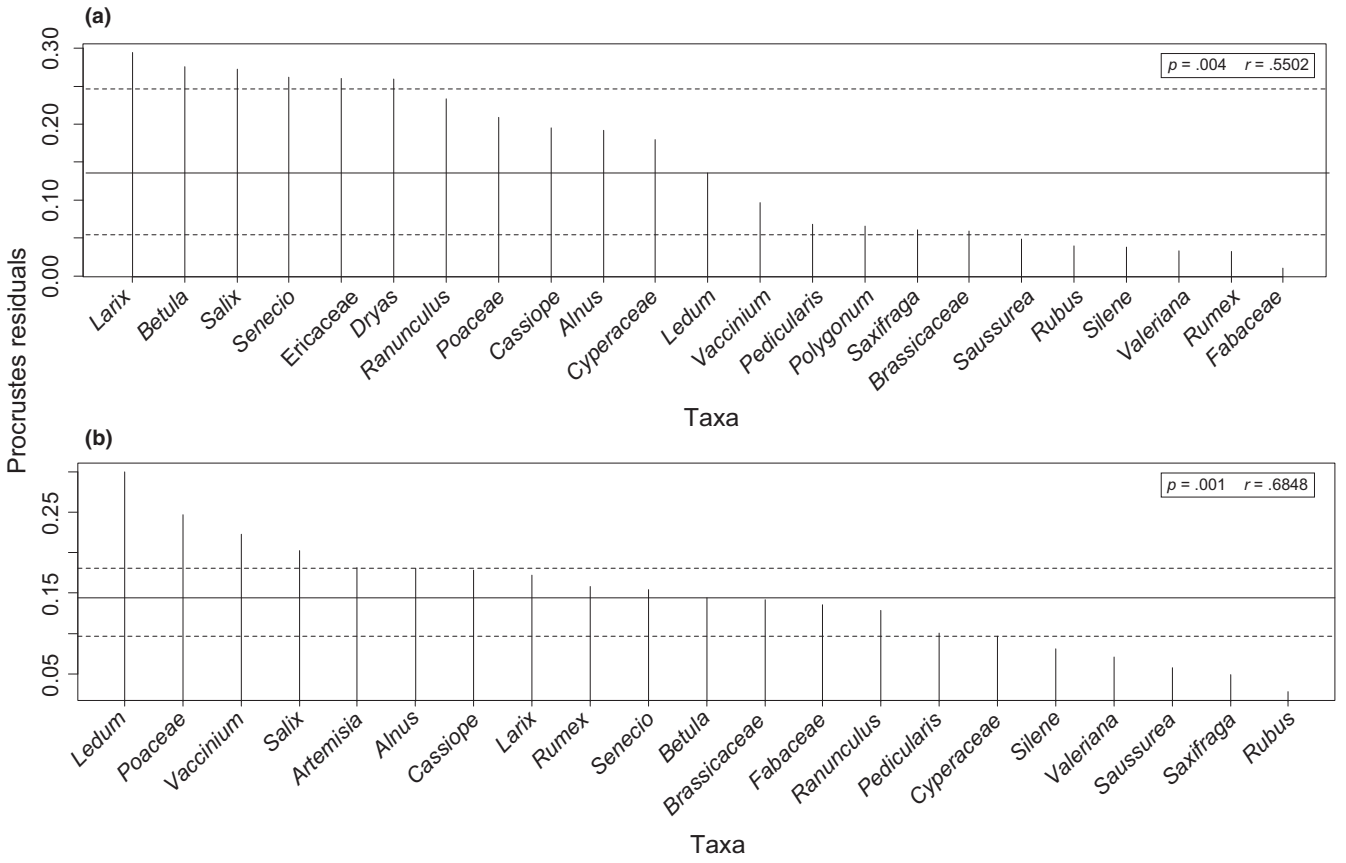

FIGURE 9 PROCRUSTES residuals plot of the comparison of PCA taxa scores of (a) sedDNA and vegetation field data and (b) of pollen and vegetation field data. Both are presented as residuals of taxa scores reflecting dissimilarity between the tested data sets. Dashed lines are the first and third quartiles, while the solid line is the second quartile

There are particular differences in the composition of the three records. For example, pollen data are unique in having high percentages of extra-regional plant taxa and steadily high percentages of *Betula*, *Alnus*, Cyperaceae or Poaceae, whereas sedDNA spectra are unique in recording high *Salix* percentages. Pollen-based *Larix* percentages are lower than those recorded by the sedDNA approach and the field vegetation surveys. Furthermore, *Larix* and some other recorded taxa show high residuals in the PROCRUSTES and PROTEST analyses, which indicate weak agreement between the pollen and the sedDNA records for these taxa (Figure 8). However, the residuals are related neither to location along the transect nor to lake size.

4.2 | Comparative assessment of the value of sedDNA and pollen for inferring vegetation transitions in the treeline area

Tracing the invasion of trees into tundra areas is currently a pressing research target in palaeoecology, not only for Siberia (e.g., Devi et al., 2008; Lloyd et al., 2011; MacDonald et al., 2008) but also for northern America (Lloyd, 2005; Moore, 2004) and Europe (Eronen, Lindholm, & Saastamoinen, 1999; Seppä et al., 2002). In our data, both the pollen and the sedDNA approaches allow a separation between samples from largely treeless vegetation and vegetation holding trees based on *Larix* percentages. However, both proxy methods only provide a semiquantitative estimate of *Larix* in the vegetation. For example, *Larix* occurred with a median of 1% in the pollen record from forest tundra sites and it obtained a relative abundance of 4.7% in the sedDNA spectra, while it has a relative *Larix* abundance of 23.9% in the vegetation. Furthermore, neither *trn*L-based metabarcoding of DNA nor pollen analyses can identify *Larix* to species level, and accordingly, the migration history of the different *Larix* species, subspecies and hybrids (Abaimov, 2010) cannot be reconstructed with these approaches.

Based on our results we assume that differences between sedDNA and pollen analyses in the qualitative record (i.e., taxa richness) are mainly related to methodological differences with respect to diversity retrieval from the sedimentary plant remains record, while taphonomic differences among the two methods are probably subordinate. First, retrieved species richness depends to a certain extent on the taxonomic resolution of the marker and, second, upon the comprehensiveness of the databases. The former means, that even though the P6 loop of the *trn*L intron is known to identify genera and species (i.e., 77% and 33%, respectively, for arctic plants; Sønstebø et al., 2010), it does not, like some markers on noncoding chloroplast DNA regions, resolve closely related species or intraspecific variation (Shaw et al., 2005). For example, the marker is

not able to resolve the two closely related *Larix* species, that is, *L. gmelinii* and *L. sibirica* described for the eastern and western Taymyr region. A better resolution could be achieved using more variable markers, which will then be less universal than the *trn*L fragment. Some metabarcoding studies have used a suite of additional markers targeting smaller taxonomic groups in addition to the *trn*L fragment (Baamrane et al., 2012; Willerslev et al., 2014), which could be explored in future sedDNA studies in treeline areas. In contrast, a comparable increase in the taxonomic resolution of pollen is not expected in the future.

Richness recorded by metabarcoding approaches is, to some extent, dependent on the comprehensiveness of taxa recorded within reference databases for DNA-based identification (Jørgensen et al., 2012; Parducci et al., 2014; Pedersen et al., 2014; Zhang et al., 2012). For example, employing the reference database for arctic and boreal plants, which includes many northern plant taxa (Soininen et al., 2015; Sønstebø et al., 2010; Willerslev et al., 2014), allows a higher number of taxa to be assigned compared to that of the EMBL database. The restriction of using only taxa which are 100% identical to taxa in the reference database might additionally lower the recorded richness.

Other errors can occur during PCR amplification and sequencing, such as punctual false implemented bases or chimeric sequences (Pääbo et al., 2004), but these can be processed by bioinformatic tools such as the OBITOOLS program (Boyer et al., 2016), which we used during sedDNA data processing.

Irrespective of the method applied, our results do not suggest richness changes across the Siberian treeline vegetation types, but this is probably only for methodological reasons. As both marker specificity and DNA reference databases will increase in the future, the added value of performing sedDNA in palaeovegetation studies will also increase, particularly when richness reconstruction is targeted.

Based on our results we assume that differences in the quantitative composition between pollen and sedDNA spectra originate both from different taphonomies and from methodological differences. With respect to taphonomy, differences in productivity, source and preservation are crucial. First, one original pollen grain represents an analytical unit in pollen analyses, although the percentage of a single taxon in the pollen spectra does not directly indicate its cover in the vegetation. Nevertheless, pollen productivity estimates (PPE) have been developed that can be used to correct for these biases (e.g., Hellman, Gaillard, Broström, & Sugita, 2008; Soepboer, Sugita, & Lotter, 2010; Twiddle, Jones, Caseldine, & Sugita, 2012), but many of them have large errors, particularly for *Larix* (Niemeyer, Klemm, et al., 2015). Productivity effects on the relationship between sedDNA and vegetation are currently largely unknown as equivalent taphonomic studies have not been carried out, and it is unclear if comparable analogue corrections can be established in the future. For example, an analytical sedDNA unit represents an artificial sequence read, which was, in our case, produced by amplification of a single template molecule from a mixed-template sample in 50 PCR cycles. The amplification of the different components of the mixed-template sample is not equally efficient, and this bias generally complicates—or possibly renders impossible—estimating a correction of the original DNA amounts in the sediment. Consequently, currently we cannot estimate sedDNA productivity similar to PPE.

Second, recording treeline transitions by DNA stored in the sediment requires an understanding of the original source location of the plant remains providing the DNA. For example, tree density in circum-arctic treeline areas is highly heterogeneous on small spatial scales, due to its patchiness. Hence, methods recording the vegetation from a small source area should show a higher variability among samples from nearby lakes than those representing larger source areas. Furthermore, it is assumed that the pollen load in intermediate-sized lakes (i.e., radius 50–150 m) is dominated by the extra-local or regional pollen component, while the area in the direct vicinity only contributes up to 20% of the total pollen load of the study area (Jacobson & Bradshaw, 1981). This, in our case, leads to rather homogeneous pollen records for single-tree and forest tundra, as all sampling was performed on intermediate-sized lakes, capturing pollen load from a regional origin, rather than a local signal. This effect is to some extent dependent on the transportation abilities of the different taxa. For example, comparatively light grains of *Alnus* (fall speed: 0.0021 m/s; Eisenhut, 1961) are probably transported from regional sources into tundra lakes while the comparatively large and heavy *Larix* grains (0.126 m/s; Eisenhut, 1961) are preferentially contributed from more nearby sources. Accordingly, a *Larix* migration into tundra regions would be traced by pollen only if the tree density around the lakes is already rather high, while an approaching pine treeline, for example, could be detected even if it is still tens of kilometres away from the sampled lake, due to its better transportability of pollen (fall speed approximately 0.031 m/s; Eisenhut, 1961).

In contrast, leaves, twigs and seeds are assumed to represent the major sources of sedDNA (Willerslev, 2003) rather than pollen (Boessenkool et al., 2014; Haile et al., 2007; Parducci et al., 2014; Yoccoz et al., 2012). As inferred by modern analogue studies, these macrofossil remains originate from within an area of only a few tens of metres around the lake (Greatrex, 1983; Pisaric, 2002; Zhao, Sayer, Birks, Hughes, & Peglar, 2006). Our record shows a more local source of sedDNA compared with the pollen load, too. This is obvious from the below-threshold *Pinus* occurrence in the sedDNA record despite constant presence of Pinaceae within the pollen record. This implies that sedDNA analyses should be better than pollen analyses for reconstructing the vegetation in the direct vicinity of the lake. Nonetheless, in our data both the sedDNA and pollen approaches reflected the vegetation signal similarly well.

Third, pre- and postdepositional preservation of pollen and DNA may affect the compositional signals. For example, pollen taxa differ in their vulnerability for distortion (Hall, 1981) due to mechanical forces in sediments or during sample treatment (Havinga, 1967). In particular, the large and thin *Larix* pollen grains are probably less robust than, for example, those of *Betula* and *Alnus* (Cushing, 1967; Havinga, 1967), which may be a cause of the underrepresentation of *Larix* in the pollen record. Information about species-specific DNA preservation is so far entirely lacking. Accordingly, it is impossible to

conclude whether the overrepresentation of *Salix*, for example, in the sedDNA spectra is a source signal (i.e., because *Salix* grows close to the lake rims but is only abundant in small amounts in the vegetation survey sites; Figure 2), is due to PCR bias (i.e., exponentially amplifying already high numbers of *Salix* sequences), or is a preservation signal (i.e., willow leaves may originally be particularly rich in DNA, which may be more resistant to decomposition). Due to this currently limited proxy understanding, the advantages of using DNA analyses in palaeoecological research cannot be fully appreciated, yet (Birks & Birks, 2016; Parducci et al., 2013; Thomsen & Willerslev, 2015).

In addition, the sedDNA results are biased by experimental procedures, as each step (DNA extraction, PCR, sequencing) acts as a filter (Jahn, Zetzsche, Reinhardt, & Gemeinholzer, 2007). Bias can be particularly strong during PCR, due to the imperfect match of the primers (Bellemain et al., 2010) and due to preferential initial amplification of the more common templates. Also, errors occurring during PCR amplification and sequencing, such as false incorporation of nucleotides or the formation of chimeric sequences (Pääbo et al., 2004), can lead to a false high diversity, potentially introducing false positives (Deakin et al., 2014; Ficetola, Taberlet, & Coissac, 2016) and causing the numbers of nonerroneous sequences to become skewed (Thomsen & Willerslev, 2015). Studies focusing on the presence or absence of certain indicator species therefore require good planning and replication of results (Ficetola et al., 2015, 2016). For vegetation reconstruction, a specific test of the replication level needed to reach reliable conclusions is still lacking, but both our comparison of established techniques to DNA metabarcoding, as well as another recent study (Alsos et al., 2016), indicates that the main components of the vegetation are retrieved reliably without replication.

Nonetheless, with respect to tracing treeline vegetation in space our study indicates that sedDNA is more powerful than pollen in recording taxa richness and performs comparably well in recording genus-level vegetation composition. Both our own study and other diversity assessment studies using modern environmental DNA (e.g., Yoccoz et al., 2012) as well as ancient DNA studies (Alsos et al., 2016; Jørgensen et al., 2012; Parducci et al., 2013; Sønstebø et al., 2010; Willerslev et al., 2007) highlight the advantage of high-throughput DNA sequencing analyses to infer vegetation diversity. Finally, our study supports the view that the metabarcoding approach provides information that enhances palaeoecological understanding, especially for the differentiation of tundra and nontundra areas on a spatial scale, because it is often complementary to insights gained by traditional approaches (Alsos et al., 2016; Jørgensen et al., 2012; Parducci et al., 2013). Advantages of the sedDNA approach that will become even stronger in the future are a higher taxonomic resolution (Table 3), its applicability to a variety of archive types such as permafrost, terrestrial, marine or lacustrine sediments (see overview in Thomsen & Willerslev, 2015), a better potential of methodological standardization among different laboratories, better efficiency in sample throughput (Sønstebø et al., 2010) and improvements to reduce PCR bias, such as DNA capture approaches (Perry, Marioni, Melsted, & Gilad, 2010) and metagenomic shotgun analyses (Jones et al., 2015; Pedersen et al., 2016). Future optimisations aside, our study of lake sedDNA from the treeline area in Siberia indicates that sedDNA has already a high potential as a circum-arctic palaeoecological proxy for small-scale spatial investigations of vegetation type differentiation.

TABLE 3 Overview of taxa recorded using the three different approaches

Approach	Total number of taxa	Family level	Genus level	Species level
Vegetation	31	0	12	19
Pollen	42	10	30	2
DNA	114	12	50	52

Differences in richness appear strongest in total number of taxa and the taxonomic level of species. The data reflect the results retrieved by vegetation field surveys and lacustrine surface samples. In total, vegetation data are based on seven vegetation field surveys, and pollen and sedDNA data were retrieved from 31 different lakes.

5 | CONCLUSIONS

The objectives of this study were to compare sedDNA of lake sediments to the established methods of vegetation field surveys and pollen analyses in the Siberian treeline area, which represents a distinct spatial vegetation gradient. We in particular compared the methods with respect to their capability of recording vegetation richness and composition using a comparable amount of time for processing each of the three different assessment methodologies, and to assess the proxy value of sedDNA analyses in palaeoecological treeline studies.

The main result of our study is that sedDNA analyses capture a higher richness and at a better taxonomic resolution than vegetation field surveys or pollen analyses. However, none of the three approaches yields richness trends along the transect. With the development and application of more specific markers (also for bryophytes) along with the construction of respective DNA reference databases, an even higher richness will probably be inferred from such samples in the future, which are likely to track even minor richness trends. We see a reflection of the treeline, and we conclude that sedDNA is highly useful to analyse vegetation change along vegetation gradients.

With respect to the quantitative and qualitative composition, all approaches are capable of differentiating spectra of tundra sites from those originating from single-tree/forest tundra sites. At the same time, each method shows certain idiosyncrasies, such as recording a far-distance transported component by the pollen approach and an overrepresentation of *Salix* in sedDNA. For pollen spectra, taxon-specific productivity, source or preservation signals are largely understood, but comparable knowledge is currently lacking for sedDNA.

We conclude that sedDNA can be used as a standalone proxy to investigate presence/absence of taxa in treeline areas and to clearly differentiate between tundra and nontundra areas. Hence, sedDNA can well be applied as a supplementary approach to traditional pollen-based investigations when the macrofossil record is scarce. However, before using sedDNA for quantitative reconstructions of vegetation composition, a more comprehensive understanding of sedDNA needs to be developed.

ACKNOWLEDGEMENTS

We thank Ruslan Gorodnichev and Alexey Pestryakov for assisting us with lake sampling; Xenia Schreiber, Alexey Kolmogorov and Stefan Kruse for assisting us with the vegetation field surveys; and Jan Lütje and Claudia Havel for supporting us during sample extraction and preparation for genetic analysis. Furthermore, we acknowledge the Alfred Wegener Institute Helmholtz Centre of Polar and Marine Research for providing the facilities to conduct DNA and pollen analysis. This research was performed within the joint Russian–German *Arctic Ecological Network* (*Arc-EcoNet*) supported by the Federal Ministry of Education and Research (BMBF grant no. 01DJ14003) and the Russian Foundation for Basic Research (RFBR grant no. 13-05-00327). Laura S. Epp is supported by the German Research Council (grant EP98/2-1 to L.S.E.).

AUTHOR CONTRIBUTION

All authors developed the research strategy. B.N. performed the analyses of pollen (guided by U.H.) and sedDNA (guided by L.E. and K.S.-L.) and applied the statistics (guided by U.H.). B.N. and U.H. wrote a first version of the manuscript that all other authors commented on.

DATA ACCESSIBILITY

The vegetation survey data (https://doi.org/10.1594/pangaea.860657), the pollen records (https://doi.org/10.1594/pangaea.860629) and the sedDNA sequence data (https://doi.org/10.1594/pangaea.860628) can be retrieved via http://pangaea.de.

REFERENCES

Abaimov, A. P. (2010). Geographical distribution and genetics of *Siberian larch* species. In A. Osawa, O. A. Zyryanova, Y. Matsuura, T. Kajimoto, & R. W. Wein (Eds.), *Permafrost ecosystems* (pp. 41–58). Dordrecht, the Netherlands: Springer Netherlands.

ACIA (2004). *Impacts of a warming arctic: Arctic climate impact assessment*. Cambridge: Cambridge University Press.

Alsos, I. G., Sjogren, P., Edwards, M. E., Landvik, J. Y., Gielly, L., Forwick, M., … Pedersen, M. W. (2016). Sedimentary ancient DNA from Lake Skartjorna, Svalbard: Assessing the resilience of arctic flora to Holocene climate change. *The Holocene*, *26*, 627–642.

Anderson, P. M., Lozhkin, A. V., & Brubaker, L. B. (2002). Implications of a 24,000-yr palynological record for a Younger Dryas cooling and for boreal forest development in northeastern Siberia. *Quaternary Research*, *57*, 325–333.

Anderson-Carpenter, L. L., McLachlan, J. S., Jackson, S. T., Kuch, M., Lumibao, C. Y., & Poinar, H. N. (2011). Ancient DNA from lake sediments: Bridging the gap between paleoecology and genetics. *BMC Evolutionary Biology*, *11*, 30.

Andreev, A. A., Schirrmeister, L., Siegelt, C., Bobrov, A. A., Demske, D., Seiffert, M., & Hubberten, H.-W. (2002). Paleoenvironmental changes in northeastern Siberia during the Late Quaternary – Evidence from pollen records of the Bykovsky Peninsula. *Polarforschung*, *70*, 13–25.

Andreev, A. A., Siegert, C., Klimanov, V. A., Derevyagin, A. Y., Shilova, G. N., & Melles, M. (2002). Late Pleistocene and Holocene vegetation and climate on the Taymyr Lowland, northern Siberia. *Quaternary Research*, *57*, 138–150.

Baamrane, M. A. A., Shehzad, W., Ouhammou, A., Abbad, A., Naimi, M., Coissac, E., … Znari, M. (2012). Assessment of the food habits of the Moroccan Dorcas Gazelle in M'Sabih Talaa, West Central Morocco, using the trnL approach. *PLoS One*, *7*, e35643.

Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., & Kauserud, H. (2010). ITS as an environmental DNA barcode for fungi: An in silico approach reveals potential PCR biases. *BMC Microbiology*, *10*, 189.

Bennett, K. D. (1996). Determination of the number of zones in a biostratigraphical sequence. *New Phytologist*, *132*, 155–170.

Beug, H.-J. (2004). *Leitfaden der Pollenbestimmung*. München, Germany: Verlag Dr. Friedrich Pfeil.

Bigelow, N. H., Brubaker, L. B., Edwards, M. E., Harrison, S. P., Prentice, I. C., Anderson, P. M., … Volkova, V. S. (2003). Climate change and Arctic ecosystems: 1. Vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present. *Journal of Geophysical Research*, *108*, 8170.

Binney, H. A., Gething, P. W., Nield, J. M., Sugita, S., & Edwards, M. E. (2011). Tree line identification from pollen data: Beyond the limit? *Journal of Biogeography*, *38*, 1792–1806.

Birks, H. H. (2001). Plant macrofossils. In J. P. Smol, H. J. B. Birks, & W. M. Last (Eds.), *Tracking environmental change using lake sediments*, *vol. 3: Terrestrial, algal and siliceous indicators* (pp. 49–74). Dordrecht, the Netherlands: Kluwer Academic Press.

Birks, H. H., & Birks, H. J. B. (2000). Future uses of pollen analysis must include plant macrofossils. *Journal of Biogeography*, *27*, 31–35.

Birks, H. J. B., & Birks, H. H. (2016). How have studies of ancient DNA from sediments contributed to the reconstruction of Quaternary floras? *New Phytologist*, *209*, 499–506.

Birks, H. J. B., & Line, J. M. (1992). The use of rarefaction analysis for estimating palynological richness from Quaternary pollen-analytical data. *The Holocene*, *2*, 1–10.

Blyakharchuk, T. A., & Sulerzhitsky, L. D. (1999). Holocene vegetational and climatic changes in the forest zone of Western Siberia according to pollen records from the extrazonal palsa bog Bugristoye. *The Holocene*, *9*, 621–628.

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

Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, *320*, 1444–1449.

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*, 176–182.

Campbell, I. D., McDonald, K., Flannigan, M. D., & Kringayark, J. (1999). Long-distance transport of pollen into the Arctic cause and effect in evolution. *Nature*, *399*, 29–30.

Cushing, E. J. (1967). Evidence for differential pollen preservation in late Quaternary sediments in Minnesota. *Review of Palaeobotany and Palynology*, *4*, 87–101.

Davis, M. B. (1963). On the theory of pollen analysis. *American Journal of Science*, *261*, 897–912.

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, *14*, 306–323.

Deakin, C. T., Deakin, J. J., Ginn, S. L., Young, P., Humphreys, D., Suter, C. M., … Hallwirth, C. V. (2014). Impact of next-generation sequencing error on analysis of barcoded plasmid libraries of known complexity and sequence. *Nucleic Acids Research*, *42*, e129.

Devi, N., Hagedorn, F., Moiseev, P., Bugmann, H., Shiyatov, S., Mazepa, V., & Rigling, A. (2008). Expanding forests and changing growth forms of Siberian larch at the Polar Urals treeline during the 20th century. *Global Change Biology*, *14*, 1581–1591.

Eisenhut, G. (1961). *Untersuchung über die Morphologie und Ökologie der Pollenkörner heimischer und fremdländischer Waldbäume*. Hamburg, Germany: Parey.

Epp, L. S., Gussarova, G., Boessenkool, S., Olsen, J., Haile, J., Schröder-Nielsen, A., … 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.

Eronen, M., Lindholm, M., & Saastamoinen, S. (1999). Variable Holocene climate, treeline dynamics and changes in natural environments in northern Finnish Lapland. *Global Change Science*, *1*, 377–387.

Faegri, K., & Iversen, J. (1989). *Textbook of pollen analysis*. Chichester, UK: John Wiley & Sons.

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., … Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*, 543–556.

Ficetola, G. F., Taberlet, P., & Coissac, E. (2016). How to limit false positives in environmental DNA and metabarcoding? *Molecular Ecology Resources*, *16*, 604–607.

Frost, G. V., & Epstein, H. E. (2014). Tall shrub and tree expansion in Siberian tundra ecotones since the 1960s. *Global Change Biology*, *20*, 1264–1277.

Gervais, B. R., MacDonald, G. M., Snyder, J. A., & Kremenetski, C. V. (2002). Pinus sylvestris treeline development and movement on the Kola Peninsula of Russia: Pollen and stomate evidence. *Journal of Ecology*, *90*, 627–638.

Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P.-J., Griggo, C., Gielly, L., … Taberlet, P. (2014). Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature Communications*, *5*, 3211.

Greatrex, A. P. A. (1983). Interpretation of macrofossil assemblages from surface sampling of macroscopic plant remains in mire communities. *Journal of Ecology*, *71*, 773–791.

Grimm, E. C. (1992). Tilia and Tilia-graph: Pollen spreadsheet and graphics programs. In: *Program and abstracts* (pp. 56). 8th International Palynological Congress, Aix-en-Provence, France.

Hahne, J., & Melles, M. (1997). Late- and post-glacial vegetation and climate history of the south-western Taymyr Peninsula, central Siberia, as revealed by pollen analysis of a core from Lake Lama. *Vegetation History and Archaeobotany*, *6*, 1–8.

Haile, J., Holdaway, R., Oliver, K., Bunce, M., Gilbert, M. T. P., Nielsen, R., … Willerslev, E. (2007). Ancient DNA chronology within sediment deposits: Are paleobiological reconstructions possible and is DNA leaching a factor? *Molecular Biology and Evolution*, *24*, 982–989.

Hall, S. A. (1981). Deteriorated pollen grains and the interpretation of Quaternary pollen diagrams. *Review of Palaeobotany and Palynology*, *32*, 193–206.

Havinga, A. J. (1967). Palynology and pollen preservation. *Review of Palaeobotany and Palynology*, *2*, 81–98.

Hellman, S. E. V., Gaillard, M., Broström, A., & Sugita, S. (2008). Effects of the sampling design and selection of parameter values on pollen-based quantitative reconstructions of regional vegetation: A case study in southern Sweden using the REVEALS model. *Vegetation History and Archaeobotany*, *17*, 445–459.

Hjelmroos, M. (1991). Evidence of long-distance transport of Betula pollen. *Grana*, *30*, 215–228.

Jackson, D. A. (1995). PROTEST: A PROcrustean Randomization TEST of community environment concordance. *Ecoscience*, *2*, 297–303.

Jacobson, G. L., & Bradshaw, R. H. W. (1981). The selection of sites for paleovegetational studies. *Quaternary Research*, *16*, 80–96.

Jahn, R., Zetzsche, H., Reinhardt, R., & Gemeinholzer, B. (2007). Diatoms and DNA barcoding: A pilot study on an environmental sample. In R. Jahn, & W.-H. Kusber (Eds.), *Proceedings of the 1st Central European Diatom Meeting 2007* (pp. 63–68). Berlin, Germany: Botanic Garden and Botanical Museum Berlin-Dahlem (FU-Berlin).

Jones, M. B., Highlander, S. K., Anderson, E. L., Li, W., Dayrit, M., Klitgord, N., … Venter, J. C. (2015). Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proceedings of the National Academy of Sciences USA*, *112*, 14024–14029.

Jørgensen, T., Haile, J., Möller, P., Andreev, A., Boessenkool, S., Rasmussen, M., … Willerslev, E. (2012). A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Molecular Ecology*, *21*, 1989–2003.

Juggins, S. (2009). *rioja: Analysis of quaternary science data*. *R package version 0.5-6*. Retrieved from http://cran. r-project. org/package= rioja.

Kanz, C., Aldebert, P., Althorpe, N., Baker, W., Baldwin, A., Bates, K., … Apweiler, R. (2005). The EMBL nucleotide sequence database. *Nucleic Acids Research*, *33*, 29–33.

Kay, P. A. (1979). Multivariate statistical estimates of Holocene vegetation and climate change, forest-tundra transition zone, NWT, Canada. *Quaternary Research*, *11*, 125–140.

Klemm, J., Herzschuh, U., Pisaric, M. F. J., Telford, R. J., Heim, B., & Pestryakova, L. A. (2013). A pollen-climate transfer function from the tundra and taiga vegetation in Arctic Siberia and its applicability to a Holocene record. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *386*, 702–713.

Kuznetsova, L. V., Zakharova, V. I., Sosina, N. K., Nikolin, E. G., Ivanova, E. I., Sofronova, E. V., & Kopyrina, L. I. (2010). Flora of Yakutia: Composition and ecological structure. In E. I. Troeva, M. M. Cherosov, A. P. Isaev & N. S. Karpov (Eds.), *The far north: Plant biodiversity and ecology of yakutia* (pp. 385). Dordrecht, the Netherlands: Springer Netherlands.

Legendre, P., Oksanen, J., & ter Braak, C. J. F. (2011). Testing the significance of canonical axes in redundancy analysis. *Methods in Ecology and Evolution*, *2*, 269–277.

Li, Y., Nielsen, A. B., Zhao, X., Shan, L., Wang, S., Wu, J., & Zhou, L. (2015). Pollen production estimates (PPEs) and fall speeds for major tree taxa and relevant source areas of pollen (RSAP) in Changbai Mountain, northeastern China. *Review of Palaeobotany and Palynology*, *216*, 92–100.

Lloyd, A. H. (2005). Ecological histories from Alaskan tree lines provide insight into future change. *Ecology*, *86*, 1687–1695.

Lloyd, A. H., Bunn, A. G., & Berner, L. (2011). A latitudinal gradient in tree growth response to climate warming in the Siberian taiga. *Global Change Biology*, *17*, 1935–1945.

MacDonald, G. M., Kremenetski, K. V., & Beilman, D. W. (2008). Climate change and the northern Russian treeline zone. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*, 2285–2299.

Matveyeva, N. V. (1994). Floristic classification and ecology of tundra vegetation of the Taymyr Peninsula, northern Siberia. *Journal of Vegetation Science*, *5*, 813–828.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., & Houston, T. G. (2012). An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*, *29*, 897–910.

Möller, P., Bolshiyanov, D. Y., & Bergsten, H. (1999). Weichselian geology and palaeoenvironmental history of the central Taymyr Peninsula, Siberia, indicating no glaciation during the last global glacial maximum. *Boreas*, *28*, 92–114.

Moore, P. D. (2004). Ecology: Hope in the hills for tundra? *Nature*, *432*, 159–160.

Moore, P. D., Webb, J. A., & Collinson, M. E. (1991). *Pollen analysis*. Oxford, UK: Blackwell Scientific Publications.

Moser, K. A., & MacDonald, G. M. (1990). Holocene vegetation change at treeline north of Yellowknife, northwest Territories, Canada. *Quaternary Research*, *34*, 227–239.

Niemeyer, B., Herzschuh, U., & Pestryakova, L. A. (2015). Vegetation and lake changes on the southern Taymyr peninsula, northern Siberia, during the last 300 years inferred from pollen and Pediastrum green algae records. *The Holocene*, *25*, 596–606.

Niemeyer, B., Klemm, J., Pestryakova, L. A., & Herzschuh, U. (2015). Relative pollen productivity estimates for common taxa of the northern Siberian Arctic. *Review of Palaeobotany and Palynology*, *221*, 71–82.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., … Wagner, H. (2011). *vegan: Community ecology package. R package version 2.0-2*. Retrieved from https://CRAN.R-project.org/package=vegan

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., … Hofreiter, M. (2004). Genetic analyses from ancient DNA. *Annual Review of Genetics*, *38*, 645–679.

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

Parducci, L., Väliranta, M., Salonen, J. S., Ronkainen, T., Matetovici, I., Fontana, S. L., … Suyama, Y. (2014). Proxy comparison in ancient peat sediments: Pollen, macrofossil and plant DNA. *Philosophical Transactions of the Royal Society B*, *370*, 20130382.

Paus, A., Boessenkool, S., Brochmann, C., Epp, L. S., Fabel, D., Haflidason, H., & Linge, H. (2015). Lake Store Finnsjøen – 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., Overballe-Petersen, S., Ermini, L., Sarkissian, C. D., Haile, J., Hellstrom, M., … Willerslev, E. (2014). Ancient and modern environmental DNA. *Philosophical Transactions of the Royal Society B*, *370*, 20130383.

Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., … Willerslev, E. (2016). Postglacial viability and colonization in North America's ice-free corridor. *Nature*, *537*, 1–15.

Pelankova, B., Kuneš, P., Chytrý, M., Jankovska, V., Ermakov, N., & Svobodova-Svitavska, H. (2008). The relationships of modern pollen spectra to vegetation and climate along a steppe-forest-tundra transition in southern Siberia, explored by decision trees. *The Holocene*, *18*, 1259–1271.

Pellatt, M. G., Smith, M. J., Mathewes, R. W., & Walker, I. R. (1998). Palaeoecology of postglacial treeline shifts in the northern Cascade Mountains, Canada. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *141*, 123–138.

Peres-Neto, P. R., & Jackson, D. A. (2001). How well do multivariate data sets match? The advantages of a procrustean superimposition approach over the Mantel test. *Oecologia*, *129*, 169–178.

Perry, G. H., Marioni, J. C., Melsted, P., & Gilad, Y. (2010). Genomic-scale capture and sequencing of endogenous DNA from feces. *Molecular Ecology*, *19*, 5332–5344.

Pisaric, M. F. J. (2002). Long-distance transport of terrestrial plant material by convection resulting from forest fires. *Journal of Paleolimnology*, *28*, 349–354.

Pisaric, M. F. J., MacDonald, G. M., Cwynar, L. C., & Velichko, A. A. (2001). Modern pollen and conifer stomates from north-central Siberian lake sediments: Their use in interpreting late Quaternary fossil pollen assemblages. *Arctic, Antarctic, and Alpine Research*, *33*, 19–27.

R Development Core Team (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Savelieva, L. A., Raschke, E. A., & Titova, D. V. (2013). *Photographic Atlas of plants and pollen of the Lena River Delta*. Saint-Petersburg, Russia: St. Petersburg State University.

Seppä, H., Nyman, M., Korhola, A., & Weckström, J. (2002). Changes of treelines and alpine vegetation in relation to post-glacial climate dynamics in northern Fennoscandia based on pollen and chironomid records. *Journal of Quaternary Science*, *17*, 287–301.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B, Liu, W., Miller, J., Siripun, K. C., … Small, R. L. (2005). The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, *92*, 142–166.

Sjögren, P., van der Knaap, W. O., Huusko, A., & van Leeuwen, J. F. N. (2008). Pollen productivity, dispersal, and correction factors for major tree taxa in the Swiss Alps based on pollen-trap results. *Review of Palaeobotany and Palynology*, *152*, 200–210.

Sjögren, P., Edwards, M. E., Gielly, L., Langdon, C. T., Croudace, I. W., Merkel, M. K. F., … Alsos, I. G. (2016). Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland. *New Phytologist*, *213*(2), 929–941.

Soepboer, W., Sugita, S., & Lotter, A. F. (2010). Regional vegetation-cover changes on the Swiss Plateau during the past two millennia: A pollen-based reconstruction using the REVEALS model. *Quaternary Science Reviews*, *29*, 472–483.

Soininen, E. M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., … Yoccoz, N. G. (2015). Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLoS One*, *10*, e0115335.

Sønstebø, J. H., Gielly, L., Brysting, A. K., Elven, R., Edwards, M., Haile, J., … Brochmann, C. (2010). Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Molecular Ecology Resources*, *10*, 1009–1018.

Spear, R. W. (1993). The palynological record of Late-Quaternary arctic tree-line in northwest Canada. *Review of Palaeobotany and Palynology*, *79*, 99–111.

Sugita, S. (1994). Pollen representation in Quaternary sediments: Theory and method in patchy vegetation. *Journal of Ecology*, *82*, 881–897.

Taberlet, P., Coissac, E., Gielly, L., Miquel, C., Brochmann, C., Valentini, A., … Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, *35*, 1–8.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793.

Tchebakova, N. M., Rehfeld, G. E., & Parfenova, E. I. (2005). Impacts of climate change on the distribution of Larix spp. and Pinus sylvestris and their climatypes in Siberia. *Mitigation and Adaptation Strategies for Global Change*, *11*, 861–882.

ter Braak, C. J. F. (1983). Principal components biplots and alpha and beta diversity. *Ecology*, *64*, 454–462.

ter Braak, C. J. F. (1994). Canonical community ordination. Part I: Basic theory and linear methods. *Ecoscience*, *1*, 127–140.

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18.

Twiddle, C. L., Jones, R. T., Caseldine, C. J., & Sugita, S. (2012). Pollen productivity estimates for a pine woodland in eastern Scotland: The influence of sampling design and vegetation patterning. *Review of Palaeobotany and Palynology*, *174*, 67–78.

Velichko, A. A., Andreev, A. A., & Klimanov, V. A. (1997). Climate and vegetation dynamics in the Tundra and forest zone during the Late Glacial and Holocene. *Quaternary International*, *41/42*, 71–96.

Willerslev, E. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*, *300*, 791–795.

Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M. B., Brand, T. B., … Collins, M. J. (2007). Ancient biomolecules from deep ice cores reveal a forested southern Greenland. *Science*, *317*, 111–114.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., … Taberlet, P. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*, *506*, 47–51.

Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., … Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, *21*, 3647–3655.

Zhang, A. B., Muster, C., Liang, H. B., Zhu, C. D., Crozier, R., Wan, P., … Ward, R. D. (2012). A fuzzy-set-theory-based approach to analyse species membership in DNA barcoding. *Molecular Ecology*, *21*, 1848–1863.

Zhao, Y., Sayer, C. D., Birks, H. H., Hughes, M., & Peglar, S. M. (2006). Spatial representation of aquatic vegetation by macrofossils and pollen in a small and shallow lake. *Journal of Paleolimnology*, *35*, 335–350.

Zibulski, R., Herzschuh, U., & Pestryakova, L. A. (2016). Vegetation patterns along micro-relief and vegetation type transects in polygonal landscapes of the Siberian Arctic. *Journal of Vegetation Science*, *27*, 377–386.

Zyryanova, O. A., Abaimov, A. P., Daimaru, H., & Matsuura, Y. (2010). Floristic diversity and its geographical background in Central Siberia. In: A. Osawa, O. A. Zyryanova, Y. Matsuura, T. Kajimoto & W. W. Ross (Eds.), *Permafrost ecosystems: Siberian larch forests* (pp. 502). Dordrecht, the Netherlands: Springer Netherlands.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Niemeyer B, Epp LS, Stoof-Leichsenring KR, Pestryakova LA, Herzschuh U. A comparison of sedimentary DNA and pollen from lake sediments in recording vegetation composition at the Siberian treeline. *Mol Ecol Resour*. 2017;17:e46–e62. https://doi.org/10.1111/1755-0998.12689