Clitellate worms (Annelida) in lateglacial and Holocene sedimentary DNA records from the Polar Urals and northern Norway

YOURI LAMMERS (iD), CHARLOTTE L. CLARKE (iD), CHRISTER ERSÉUS (iD), ANTONY G. BROWN, MARY E. EDWARDS, LUDOVIC GIELLY, HAFLIDI HAFLIDASON (iD), JAN MANGERUD, EMILIA ROTA (iD), JOHN INGE SVENDSEN (iD) AND INGER GREVE ALSOS (iD)

While there are extensive macro- and microfossil records of a range of plants and animals from the Quaternary, earthworms and their close relatives amongst annelids are not preserved as fossils and therefore the knowledge of their past distributions is limited. This lack of fossils means that clitellate worms (Annelida) are currently underused in palaeoecological research, even though they can provide valuable information about terrestrial and aquatic environmental conditions. Their DNA might be preserved in sediments, which offers an alternative method for detection. Here we analyse lacustrine sediments from lakes in the Polar Urals, Arctic Russia, covering the period 24 000–1300 cal. a BP, and NE Norway, covering 10 700–3300 cal. a BP, using a universal mammal 16S rDNA marker. While mammals were recorded using the marker (reindeer was detected twice in the Polar Urals core at 23 000 and 14 000 cal. a BP, and four times in the Norwegian core at 11 000 cal. a BP and between 3600–3300 cal. a BP), worm extracellular DNA 'bycatch' was rather high. In this paper we present the first reported worm detection from ancient DNA. Our results demonstrate that both aquatic and terrestrial clitellates can be identified in late-Quaternary lacustrine sediments, and the ecological information retrievable from this group warrants further research with a more targeted approach.

Youri Lammers (youri.lammers@uit.no) and Inger Greve Alsos, Tromsø Museum, University of Tromsø – The Arctic University of Norway, NO-9037 Tromsø, Norway; Charlotte L. Clarke and Antony G. Brown, Geography and Environment, University of Southampton, Highfield Southampton SO17 1BJ, UK; Christer Erséus, Department of Biological and Environmental Sciences, University of Gothenburg, P.O. Box 463, SE-405 30 Göteborg, Sweden; Antony G. Brown and Mary E. Edwards, Tromsø Museum, University of Tromsø – The Arctic University of Norway, NO-9037 Tromsø, Norway and Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK; Ludovic Gielly, Laboratoire d'Ecologie Alpine (LECA), Université Grenoble Alpes, C2 40700 38058, Grenoble Cedex 9, France; Haflidi Haflidason, Jan Mangerud and John I. Svendsen, Department of Earth Science, University of Bergen, PO Box 7803, NO-5020 Bergen, Norway and Bjerknes Centre for Climate Research, Bergen, Norway; Emilia Rota, Department of Physics, Earth and Environmental Sciences, University of Siena, Via P.A. Mattioli 4, IT-53100 Siena, Italy; received 30th March 2018, accepted 12th October 2018.

The fact that earthworms (Clitellata: Megadrili) have an important function in cycling nutrients and structuring soils was famously recognized by Darwin (Darwin 1881). Both earthworms as well as potworms (Clitellata: Enchytraeidae) are used as indicator species in various environmental studies of modern soils and aquatic systems as some are very tolerant of pollution while others are very sensitive (Karaca *et al.* 2010). In theory, they have high potential as indicators due to their known sensitivity to soil conditions including temperature, moisture status, soil texture and pH range (Edwards & Lofty 1977; Beylich & Graefe 2009). However, as soft-bodied organisms, worms rarely get preserved in sediments except as trace fossils and earthworm calcite granules (which can be radiocarbon dated; Canti 2003). Their limited preservation means that worms are currently underused in palaeoecology, even though they can provide valuable ecological information.

DNA barcoding has proven to be an important tool for the identification of species through the amplification and sequencing of small, yet informative, parts of the genome (Hebert *et al.* 2003). The barcoding process was revolutionized with the advent of next-generation sequencing, allowing complex samples such as environmental DNA to be barcoded (metabarcoding; Taberlet *et al.* 2012). Since then metabarcoding has been applied to a wide range of organisms, such as nematodes (Porazinska *et al.* 2009), plants (Taberlet *et al.* 2007; Parducci *et al.* 2017; Zimmermann *et al.* 2017), clitellate worms (Bienert *et al.* 2012; Epp *et al.* 2012; Pansu *et al.* 2015), amphibians and bony fishes (Valentini *et al.* 2015), fungi (Buée *et al.* 2009; Epp *et al.* 2012) and a range of other organisms (Thomsen & Willerslev 2015; Domaizon *et al.* 2017).

After being released into the environment by organisms, extracellular DNA degrades over time, but stabilized smaller fragments can persist over longer periods bound to fine-grained sediment particles or due to low temperatures (Pääbo *et al.* 2004; Willerslev *et al.* 2004; Barnes & Turner 2016). Thus, lake sediments in arctic or mountainous regions are prime locations for the recovery of ancient DNA (Parducci *et al.* 2012; Giguet-Covex *et al.* 2014; Pedersen *et al.* 2016). Metabarcoding of sedimentary ancient DNA (*seda*DNA; Haile *et al.* 2009) can provide valuable information about past environments and augment traditional methods such as pollen or macrofossils (Pedersen *et al.* 2013; Parducci *et al.* 2015;

Zimmermann *et al.* 2017) and is of particular interest for taxa that leave limited traces in the fossil record such as worms (Domaizon *et al.* 2017). Metabarcoding efforts targeting enchytraeid worms in ancient permafrost have been attempted before, but unlike those in modern soils, they yielded no results (Epp *et al.* 2012), suggesting that detection of worms in *seda*DNA is not as straightforward as for other taxa that have been explored.

In this study, we set out to analyse mammalian DNA from lateglacial and Holocene lake sediments for faunal reconstruction, but because of the low retrieval of mammalian DNA and the unexpected clitellate DNA barcoding 'bycatch', we explore the potential for DNA-based clitellate palaeoecology.

Study sites

Polar Urals

Lake Bolshoye Shchuchye is located in the northernmost Polar Ural Mountains of Arctic Russia (latitude 67°53′24″N, longitude 66°18′53″E, altitude 221 m a.s.l.; Fig. 1). Bolshoye Shchuchye is an elongated lake (12 km long, 1 km wide) located in a NW–SE orientated valley with a maximum water depth of 136 m in its central part (Svendsen *et al.* 2019) and up to 160 m of lacustrine sediments in the central and northern parts (Haflidason *et al.* 2019). The lake is flanked by steep rock faces but the terrain is more open towards its north side, resulting in a total catchment area of 215 km^2 (Svendsen *et al.* 2019). The bedrock consists of Proterozoic-Cambrian basaltic and andesitic rocks in the eastern and northwestern parts of the catchment and Ordovician quartzite and phyllitic rocks in the southwestern catchment (Dushin *et al.* 2009). The Polar Urals remained mostly ice-free during the Last Glacial Maximum (LGM), except for cirque glaciers or minor valley glaciers (Svendsen *et al.* 2004). Current climate conditions are cold and continental with mean summer temperatures of 7 °C (Solomina *et al.* 2010).

Varanger Peninsula

The Uhca Rohči lake (70°19′07″N, 30°01′44″E; Fig. 1) is unnamed on the 1:50 000 Norwegian Topographic Map (Norgeskart; https://www.norgeskart.no), but we use this name as the lake is located close to a river site with the local Sami name 'Uhca Rohči'. Uhca Rohči is a small lake (<1 ha) in a depression situated at 138 m a.s.l. within the river valley of Komagdalen on the Varanger Peninsula, northeast Finnmark, Norway. The Varanger Peninsula is a low-relief plateau (200–600 m a.s.l.) moulded from the Proterozoic paleic surface (pre-Quaternary erosion surface) by marine and glacial processes (Siedlecka & Roberts 1992). Relief is strongly controlled by rock type and structure, with the ridges being formed of Cambrian quartzites and sandstones, whilst valleys are eroded into

shales and mudstones. There is evidence that, having undergone uplift during the Pliocene, the area was subject to processes of erosion related to former sea levels, and glacial erosion (Fjellanger & Sørbel 2007). The lake's bedrock is composed of sandstone and mudstone (The Geological survey of Norway; www.ngu.no) and the Komagdalen valley was probably deglaciated by 15.4–14.2 ka BP; the peninsula was certainly free of glacial ice by 13 000–12 000 cal. a BP (Stokes *et al.* 2014; Hughes *et al.* 2016; Stroeven *et al.* 2016). The current climate is low Arctic with a mean summer temperature of 8.7 °C (Norwegian Meteorological Institute; www.met.no).

Material and methods

Polar Urals lake sediment

Lake Bolshoye Shchuchye was cored during several expeditions between 2007 and 2009. The 24-m-long core 506-48 that was sampled for metabarcoding was obtained in July 2009 from the southern part of the lake (67°51′22.2″N, 66°21′30.1″E). Coring was conducted with a UWITEC Piston Corer (http://www.uwitec.at) using 2-m-long by 10-cm-diameter PVC or 2-m-long by 9-cm-diameter steel tubes. The full core was obtained by taking consecutive segments from the same hole. All sections were stored and transported at above 0 °C to avoid freezing of the material. The core was subsampled within the Centre for Geobiology and Microbiology (University of Bergen) in a laminar flow cabinet and in the presence of subsampling controls (open water samples) in order to detect laboratory contamination. Due to deformation near the top of each core segment, the samples form a non-continuous record and a second core was taken parallel to the first core at a 35 cm offset to account for the deformations (Svendsen *et al.* 2019) but was not sampled for this study. Age determination was based on 26 AMS radiocarbon dates from plant macrofossils provided by the Poznań Radiocarbon Laboratory. Dates were calibrated using INTCAL13 (Reimer *et al.* 2013) and the online Calib program (Stuiver *et al.* 2018). A full chronology and sedimentology of this core is described by Svendsen *et al.* (2019).

Varanger Peninsula lake sediment

The Uhca Rohči lake was cored in February 2016 with a modified Nesje piston-corer (Nesje 1992), using a 4-m-long and 10-cm-diameter ABS polymer pipe. A 2.5-m-core was retrieved and cut in the field into 1-m sections, which were sealed to reduce the risk of contaminating the sediments. The core sections were kept at above 0 °C conditions in the field and during transport to avoid freezing of the sediments and were stored in a 4 °C cold room at the Tromsø University Museum (TMU). Sampling of the core took place in a dedicated ancient DNA laboratory. The age of the core was determined based on seven AMS radiocarbon
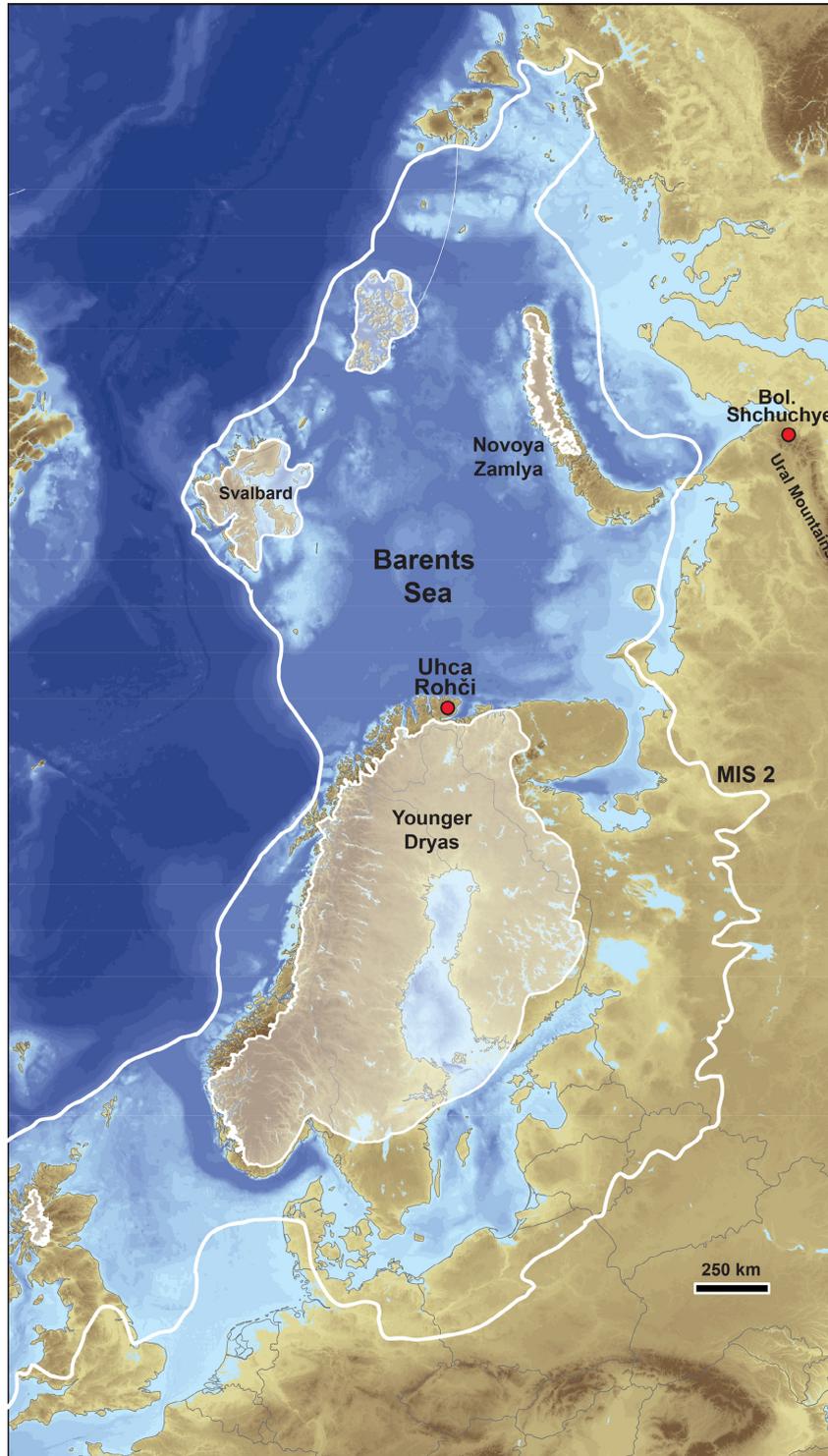
Fig. 1. The location of Lake Bolshoye Shchuchye in the Polar Urals of Arctic Russia and Lake Uhca Rohči on the Varanger Peninsula, northeast Finnmark, Norway. The outer white line represents the extension of the Eurasian ice sheet during Marine Isotope Stage 2 (20 000–15 000 a BP). The inner white shaded area represents the ice sheet during the Younger Dryas (12 800–11 400 a BP). [Colour figure can be viewed at www.boreas.dk]

dates on terrestrial plant macrofossils provided by the Poznań Radiocarbon Laboratory. Dates were calibrated using the terrestrial INTCAL13 curve (Reimer *et al.* 2013), and the age model was constructed using the Bayesian framework calibration software 'Bacon' (v2.2; Blaauw & Christen 2011), which was implemented in R (v3.2.4; R Core Team 2017). A full sedimentology and chronology of this core is described by Clarke *et al.* (2019).

DNA extraction

For the Polar Urals site, 153 lake sediment samples, 17 extraction controls and three subsampling controls underwent DNA extraction. DNA from the Varanger site was extracted from 77 sediment samples and nine extraction controls. All extractions were done at the Tromsø University Museum ancient DNA lab, using PowerMax soil DNA isolation kit (MOBIO Laboratories, Carlsbad, CA, USA), following the manufacturer's protocol with minor modifications by Alsos *et al.* (2016).

PCR amplification and sequencing

PCR reactions were carried out in a dedicated PCR room for ancient DNA at the Laboratoire d'Ecologie Alpine (Université Grenoble Alpes, France), using the MamP007F and MamP007R primers that target a ~70-bp-long part of the mammalian mitochondrial 16S rDNA (Giguet-Covex *et al.* 2014). Both forward and reverse primers had the same unique 8-bp tag on the 5′ end to allow sample multiplexing (Binladen *et al.* 2007; Valentini *et al.* 2009). In addition to the forward and reverse primers, the human blocking primer MamP007_B_Hum1 was added to suppress the amplification of human material (Giguet-Covex *et al.* 2014). The PCR reactions for each lake were carried out at different times to avoid cross-contamination of material. The Polar Urals samples included additional nine PCR negatives (excluding template DNA) and four PCR positives (including the marsupial *Didelphis marsupialis*, not found in Europe). The Varanger samples included six PCR negatives. For each sample, eight PCR repeats were carried out following a previously described PCR protocol (Giguet-Covex *et al.* 2014). PCR products were cleaned and pooled following the methods described by Alsos *et al.* (2016). Libraries (four for the Polar Urals core and two for the Varanger core) were prepared using the PCR free 'MetaFAST' library preparation protocol at Fasteris SA, Switzerland, and sequenced on an Illumina HiSeq 2500 at 2 × 125 bp paired-end sequencing.

DNA sequence analysis

The sequence data were analysed with the OBITOOLS software package (Boyer *et al.* 2016), using default settings unless otherwise specified. Paired-end data were merged with the *illuminapairedend* function and alignments with a score lower than 40 were removed. Data were demultiplexed with *ngsfilter* based on the known PCR tags. Identical sequences were merged with *obiuniq* and singleton sequences and those shorter than 10 bp were removed. Sequences were corrected for PCR and sequencing errors with *obiclean* with a 'head' to 'internal' ratio of 0.05. The remaining sequences were identified by comparing them to the EMBL nucleotide database (r133) with *ecotag*.

The identified sequences were further filtered in R (v3.4.2; R Core Team 2017) with a custom R script.

Sequence occurrences that had fewer than 10 reads for a repeat were removed, to account for low-level sequence errors that survived the *obiclean* step and tag switching (Schnell *et al.* 2015). Only sequences that had a 100% match to reference data were kept. Furthermore, sequences had to be present in at least one sediment sample with two or more repeats. If that condition was met, single occurrences for other sediment samples were kept in. Finally, a sequence could only be present in the control samples with at most one repeat; if a sequence was found in a control sample with two or more repeats it was removed from the total data set. Common laboratory contaminants, such as human, *Homo sapiens*, pig, *Sus scrofa*, and chicken, *Gallus gallus* (Leonard *et al.* 2007), were manually removed from the list of sequences that survived filtering.

In silico primer analysis

The *ecoPCR* program (Ficetola *et al.* 2010) was used to calculate the mismatches between clitellate (=oligochaete) worms and the MamP007F – MamP007R primers. The target taxonomic group was set to NCBI TAXID 6381 (referred to as subclass Oligochaeta), the maximum number of mismatches in the primer to five, the amplicon size range to 10–100 bp and the EMBL r133 nucleotide release as database. For each clitellate family and species with available data in the EMBL release, the following were calculated: mean length of the amplicon, mean number of mismatches in each primer and the presence of mismatches in the last three bases of the primer 3′ end, which can hinder amplification (Kwok *et al.* 1990; Wu *et al.* 2009).

The same procedure was repeated for the following families that could be observed in the metabarcoding results: Cervidae (TAXID 9850), Hominidae (TAXID 9604), Phasianidae (TAXID 9005), Suidae (TAXID 9821) and Cercopagididae (TAXID 77756), with the exception that an amplicon size range of 25–150 bp was used to account for the longer expected fragment length.

Results

Polar Urals samples

A total of 80 983 160 raw reads was obtained for the four Polar Urals sequence libraries, which could be assigned to 68 521 unique sequences. Post-identification filtering reduced the number of sequences to 17, representing 1 123 241 reads. The sequences belonged to reindeer (*Rangifer tarandus* – two occurrences in the core at 23 000 and 14 000 cal. a BP, with a total of 27 133 reads) and eight clitellate taxa: two Enchytraeidae (*Enchytraeus norvegicus*, *Henlea perpusilla*), one Glossoscolecidae (*Pontoscolex corethrurus*) and five Lumbricidae (*Aporrectodea rosea*, *Dendrobaena octaedra*, *Bimastos norvegicus*, *Octolasion cyaneum* and *Octolasion tyrtaeum*) (Fig. 2, Table S2). The results also included seven
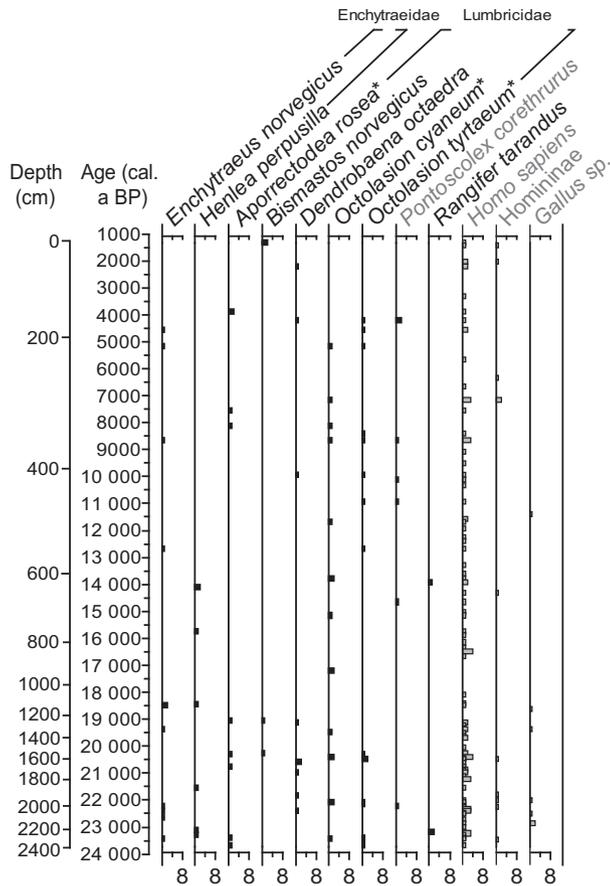
Fig. 2. Metabarcoding results for the Polar Urals core. The width of the bars indicates the number of PCR repeats. The grey taxa were assumed to be laboratory contaminants and were manually removed from the results. Non-cold tolerant taxa or taxa not found in the Polar Urals are indicated with an asterisk. The result for *Homo sapiens* is a combination of seven different *H. sapiens* sequences for which the maximum number of repeats is plotted.

Hominidae sequences (six assigned to *Homo sapiens* and one to Hominidae) and one *Gallus* sequence that survived the filtering criteria and were manually removed.

Species that did not survive filtering include steppe bison (*Bison priscus*, 100% match), Arctic lemming (*Dicrostonyx torquatus,* 98% match), rock ptarmigan (*Lagopus muta,* 99% match) and mountain hare (*Lepus timidus,* 98% match); these species are expected in the region, but none of them occurred in more than one sample and one repeat and thus did not survive our filtering criteria.

Varanger Peninsula samples

We obtained 52 562 858 raw reads for the two Varanger Peninsula (Uhca Rohči) libraries that represented 22 461 unique sequences. After R filtering, 18 sequences remained, representing 877 555 reads, which belong to: *Rangifer tarandus* (four occurrences at 10 800 cal. a BP and three between 3300 and 3600 cal. a BP, with 44 979 reads), the spiny water flea (cercopagidid cladoceran) *Bythotrephes longimanus* (six occurrences at 4900, 5600, 5700, 6300,

6500 and 9100 cal. a BP, sum 38 085 reads) and the oligochaete *Lumbriculus variegatus* (one occurrence at 10 800 cal. a BP with 227 reads) (Fig. 3, Table S3). A total of 12 *Homo sapiens*, one *Sus* and one *Gallus* sequences survived filtering and were manually removed.

Several worm taxa did not survive filtering, including *Dendrobaena octaedra*, *Tubifex tubifex* and a *Limnodrilus* sequence that could not be identified to species level. None of these taxa were detected in multiple repeats for a sample, but they are taxa that can be expected to occur in the Varanger area today.

In silico primer analysis

Primer matches between the mammal primer and annelid sequences could be calculated for 22 clitellate families and 1756 species (mean 175 sequences per family, SD = 317.7) out of the 28 families listed in the NCBI taxonomy database. The weighted average numbers of mismatches in the forward and reverse primer were 2.07 (SD = 0.05) and 2.04 (SD = 0.24), respectively, with an average estimated amplicon length of 35.7 bp (SD = 0.65; excluding primers).

The results for the clitellate families and the species that were detected in the metabarcoding results are displayed in Table 1, along with the mammalian and avian results. A full account of all clitellate families and species is provided in Table S1. The mismatch overview here is limited by the available clitellate data on EMBL, and some mismatch numbers might be over- or underestimated for some families depending on sampling and sequencing biases or depth.

Discussion

Mammal records

Rangifer tarandus was the only mammal in the Polar Urals and Varanger lake sediments that was detected in several PCR replicates (one Polar Urals sample with two repeats, Fig. 2, and three Varanger samples with two, three and four repeats, Fig. 3). *R. tarandus* was detected in a limited number of samples, furthermore, replicability was poor, with at most four out of eight PCR repeats. The limited presence is surprising, as *R. tarandus* has a circumpolar Eurasian distribution. It is known from western Norway at 13 500 cal. a BP from the village Blomvåg 30 km northwest of Bergen (Lie 1986; Mangerud *et al.* 2017) and it would be expected that *R. tarandus* was one of the first species immigrating north and west into Varanger after the ice receded after the LGM. Likewise, it is not surprising that *R. tarandus* was present in the Urals to the northeast of the Eurasian–Fennoscandian ice sheet during the Late Weichselian (24 000–15 000 a BP) as this area was probably its main glacial refugium based on genetic data (Flagstad & Roed 2003; Yannic *et al.* 2014; Kvie *et al.* 2016).
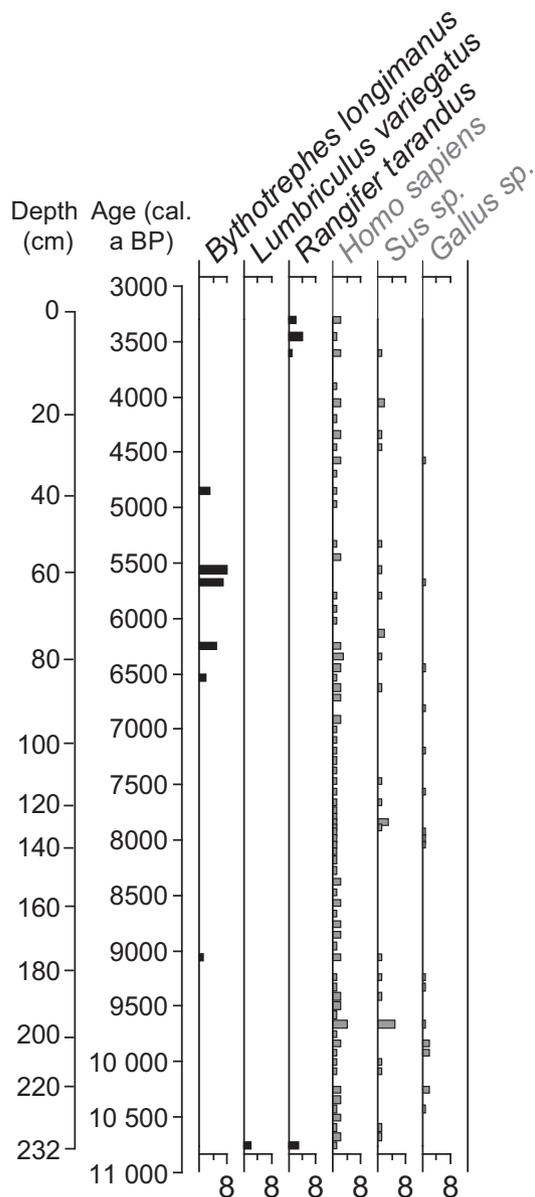
Fig. 3. Metabarcoding results for the Varanger core. The width of the bars indicates the number of PCR repeats. The grey taxa were assumed to be laboratory contaminants and were manually removed from the results. The result for *Homo sapiens* is a combination of 12 different *H. sapiens* sequences for which the maximum number of repeats is plotted.

The other mammals detected, *Bison priscus*, *Dicrostonyx torquatus* and *Lepus timidus*, are all likely for the sites in the period studied but were filtered out because they could only be observed in one sample and with one PCR repeat out of eight. There is always a trade-off between losing assumed true positives and keeping false negatives when setting a cut-off level for filtering (Ficetola *et al.* 2015). Lowering the cut-off level to include these taxa would increase our data set with many records that we suspected to be false positives. While probability statistics may be used to inform the likelihood of a record to represent a true positive, they require an independent

record for calibration (Alsos *et al.* 2018). Thus, without records of bones, detection when there are low read numbers and few PCR repeats should be interpreted with caution. Furthermore, even if the filtered taxa were included, the limited occurrences in the records (only a single sample) suggest that the approach used here lacks the capability to reliably detect taxon presence, and thus is not appropriate for palaeoecological reconstructions.

The poor detection of mammals may either be explained by low DNA concentrations in extracts due to lack of template material, potentially caused by the low amounts of mammalian DNA deposited in the lakes, the age of the sediments or the size of the target amplicon. The amount of DNA deposited in the lake might be limited by accessibility for mammals, such as the steep slopes surrounding Lake Bolshoye Shchuchye. Alternatively, the plentiful water sources in the Komagdalen valley on Varanger Peninsula could have resulted in deposition of mammalian DNA over a large region, effectively diluting it in the process. Ancient DNA fragments found in lake sediments are of a relatively small length (Pedersen *et al.* 2015) and it is possible that the longer fragment required for the amplification of mammal material (*R. tarandus* requires a fragment of 111 bp, including primers) is too restricted in older sediments, especially considering the low biomass of mammals compared to other groups such as plants or invertebrates. Metabarcoding studies that successfully targeted ancient mammal DNA either worked with frozen material from localities affected by permafrost (Willerslev *et al.* 2003; Haile *et al.* 2009; Boessenkool *et al.* 2012) where conditions possibly preserved longer fragments (Pääbo *et al.* 2004; Willerslev *et al.* 2004), or with lake sediments from locations that had high mammalian concentrations, either due to migration routes (Pedersen *et al.* 2016), due to a waterhole (Graham *et al.* 2016) or due to human influence (Giguet-Covex *et al.* 2014). Thus, a combination of low mammal DNA concentration and long target fragment length may have caused the poor detection of mammals.

It is unlikely that failed DNA extractions are responsible for the poor mammal results, as the same DNA extracts were used for the metabarcoding of plants with the *g-h* universal plant primers (Taberlet *et al.* 2007) and produced successful results for both the Varanger (Clarke *et al.* 2019) and Polar Urals sites (C. L. Clarke, pers. comm. 2018). The success for plants could be explained by the obvious higher biomass and thus DNA contribution to the sediments and a potentially lower average fragment length required for amplification; for example the plant data from the Varanger site had an average length of 44.3 bp (±15.6) (Clarke *et al.* 2019) compared to the 73 bp of *Rangifer tarandus*.

The limited amount of mammal template material in the sediment extracts may have led to the amplification of laboratory contaminants and off-target species. *Homo sapiens* was by far the most dominant species in the filtered results for both the Polar Urals and Varanger

Table 1. The amplicon lengths and mismatches between the taxa and families that were detected in the metabarcoding results and the MamP007F–MamP007R mammal primers. * = this taxon is in reality a species complex.

Family	Species	Number of sequences	Average amplicon length (bp)	Forward primer		Reverse primer	
				Average mismatches	% 3′-end mismatches	Average mismatches	% 3′-end mismatches
Clitellate							
Acanthodrilidae		71	35.52 (±1.06)	2.06 (±0.29)	0	1.89 (±0.36)	0
Almidae		36	36.94 (±0.74)	2	0	2	0
Enchytraeidae		239	33.61 (±0.97)	2.05 (±0.25)	1.67	1.82 (±0.96)	1.67
	Enchytraeus norvegicus	1	34	2	0	1	0
	Henlea perpusilla	1	34	2	0	2	0
Eudrilidae		6	34.67 (±1.7)	2	0	2.67 (±0.94)	33.33
Glossoscolecidae		16	34.00 (±1.8)	2.38 (±0.48)	0	2.75 (±1.03)	12.5
	*Pontoscolex corethrurus**	7	33.86 (±1.64)	2	0	3.29 (±0.7)	14.29
Hormogastridae		585	35.38 (±1.34)	2.03 (±0.2)	0	1.98 (±0.19)	0.17
Lumbricidae		1037	35.74 (±1.22)	2.07 (±0.26)	2.41	1.97 (±0.2)	0.19
	*Aporrectodea rosea**	143	35.76 (±0.56)	2.23 (±0.42)	0.7	2.01 (±0.08)	0
	*Bimastos norvegicus**	9	37	2.22 (±0.63)	0	2	0
	*Dendrobaena octaedra**	15	35.73 (±0.77)	2	0	2	0
	Octolasion cyaneum	1	36	2	0	2	0
	*Octolasion tyrtaeum**	3	36.66 (±0.94)	2	0	2	0
Lumbriculidae		73	36.55 (±1.15)	2.03 (±0.16)	0	1.67 (±0.52)	0
	*Lumbriculus variegatus**	30	36.87 (±0.34)	2	0	2	0
Megascolecidae		763	35.86 (±1.21)	2.14 (±0.59)	0.66	1.86 (±0.4)	0.26
Moniligastridae		77	35.78 (±1.3)	2.13 (±0.41)	0	1.84 (±0.58)	2.6
Sparganophilidae		21	34.86 (±0.35)	2	0	2	0
Tubificidae		907	36.21 (±1.98)	2.06 (±0.29)	1.87	2.42 (±0.96)	20.84
Mammalia							
Cervidae		302	73.07 (±0.51)	0.03 (±0.27)	0.7	0.03 (±0.21)	0.7
	Rangifer tarandus	9	73	0	0	0	0
Hominidae		39 080	72.13 (±4.61)	0.07 (±0.58)	0.8	0.07 (±0.57)	0.9
	Homo sapiens	38 492	72.1 (±4.26)	0.06 (±0.53)	0.6	0.6 (±0.53)	0.8
Suidae		428	76.93 (±12.95)	0.66 (±1.65)	0.72	0.65 (±1.64)	10
	Sus sp.	400	76.99 (±13.4)	0.7 (±1.7)	0.78	0.69 (±1.68)	10.8
Aves							
Phasianidae		236	76.09 (±10.08)	2.23 (±0.81)	9.75	1.34 (±1.06)	5.5
	Gallus gallus	126	77.59 (±11.23)	2.33 (±0.98)	9.52	1.44 (±1.24)	8.7

samples before human DNA sequences were manually removed (767 186 out of 1 123 241 reads and 706 027 out of 877 555 reads for the Polar Urals and Varanger cores, respectively). Chicken, *Gallus gallus* (both sites), and pig, *Sus scrofa* (Varanger only), made up the remaining contaminants. Chicken could be amplified in both samples due to the limited differences between the mammalian primers used and the binding sites for chicken (Table 1). The amplification of *H. sapiens* was possible even in the presence of a human blocking primer, which is further indication that there was a limited amount of non-human template material available in the DNA extracts (Boessenkool *et al.* 2012).

The problems with the mammal primer presented here support the case for the exploration of alternative primers or methods for the detection of mammals in ancient sediments, especially where template material is probably low. Several metabarcoding primer sets have been suggested for mammals, with the shortest sets amplifying a mitochondrial 16S fragment of 68–71 bp (Rasmussen *et al.* 2009) or 60–84 bp (Giguet-Covex *et al.* 2014), both

of which might be too long for reliable amplification of low concentration mammal material in ancient lake sediments. Alternative primer sets might yield better results if they target a shorter fragment or do not amplify common laboratory contaminants by targeting a narrower taxonomic group. Other alternatives are to bypass the usage of primers altogether by either shotgun sequencing sediment extracts (Pedersen *et al.* 2016; Seersholm *et al.* 2016) or by using DNA target capture methods (Slon *et al.* 2017).

Presence of worms

Off-target amplification of earthworms and other clitellates was observed in both the Polar Urals and the Varanger samples. Such amplification can be expected when there is limited target template available in the DNA extracts (Sipos *et al.* 2007; Schloss *et al.* 2011; Brown *et al.* 2015). The *in silico* amplification of clitellates with MamP007F and MamP007R primers revealed that 17 families and 849 species have a low number of mismatches (two or fewer outside the primer 3′ end) and that these

could potentially be amplified if there is limited competing mammal template available.

Metabarcoding potential of the mitochondrial 16S region targeted in this study has previously been demonstrated for earthworms with specific primers (Bienert *et al.* 2012). A comparison between the mammalian primers used in this study and the earthworm primers developed by Bienert *et al.* (2012) is given in Table 2. The forward primers are highly similar, with only a two base pair difference to account for the mismatches between the mammalian and earthworm primer binding sites; the reverse primer is shifted by four bases, but is otherwise comparable, once more indicating that the used mammal primers can amplify worms.

Another factor is the potential amount of DNA present in the sediment for various groups of organisms. Enchytraeidae biomass in Svalbard is estimated to be 1160 kg km^{-2} (Byzova *et al.* 1995) and Lumbricidae biomass in the northern Ural mountains is calculated to be 24 000 kg km^{-2} (Ermakov & Golovanova 2010). Thus, the clitellate numbers are far higher than common herbivorous mammals such as the North American brown lemming (*Lemmus trimucronatus*) at 30 kg km^{-2} in the Canadian Arctic (Fauteux *et al.* 2015) or *R. tarandus* in central Norway at 165 kg km^{-2} (Finstad & Prichard 2000; Vistnes *et al.* 2001). These rough biomass numbers give an indication that worms can produce vastly more DNA than the relatively sparse mammals, meaning that the clitellate DNA has a higher chance of being captured in the sediments. The difference in DNA production and contribution to the sediments, along with the additional problems of mammalian DNA described above, make worms more likely to be detected via metabarcoding.

Additionally, the clitellate amplicon length is considerably shorter than that of the mammalian taxa. The amplicon (excluding primer binding sites) for the mammals detected in the Polar Urals and Varanger cores is 74 bp on average (Table 1) and the average amplicon length for all clitellate families is 35 bp (Table S1). The shorter clitellate amplicon length increases the potential amount of template material in highly fragmented *seda*DNA compared to the longer, and thus rarer mammalian target material. The downside of a shorter amplicon is the potential loss of taxonomic resolution, a problem that is difficult to estimate given the limited reference material available for clitellates.

Four worm species that are reported to be cold tolerant were recorded in the Polar Urals samples, and these could be expected to survive in the region. The enchytraeid *Henlea perpusilla* (six samples, one sample with two

repeats) is found throughout Europe and is capable of surviving in the Arctic (Birkemoe *et al.* 2000). *Enchytraeus norvegicus* (10 samples, one sample with two repeats) is also known to have a broad range, extending from sea level in the Mediterranean (Rota *et al.* 2014) to colder temperate zones (Rota 1995) and at high (>1400 m a.s.l.) elevations in southern Norway (C. Erséus, unpublished data). The cosmopolitan lumbricid *Dendrobaena octaedra* (one sample with two repeats) has frost-tolerant populations in Finland, Greenland and Magadan Oblast, eastern Russia (Rasmussen & Holmstrup 2002). *Bimastos norvegicus* (three samples, one sample with two repeats) is part of the taxonomically difficult *Bimastos rubidus* (syn. *Dendrodrilus rubidus*) species complex, which is abundant in Scandinavia and European Russia, and is reported as freeze resistant. However, the known distribution today does not extend to the Ural region (Berman *et al.* 2010).

The remaining three lumbricid earthworms are less likely to be present in the northern Polar Urals, although they all show wide altitudinal ranges at lower latitudes. *Octolasion cyaneum* (12 samples, one sample with two repeats) is native to central and western Europe, and current records extend up to southern Finland and northern Sweden (Terhivuo & Saura 2006). In Norway it can be found to elevations of around 1000 m a.s.l. in the south, and in lowland localities north of the Arctic Circle (C. Erséus, unpublished data), but it is most often associated with human habitats. *Octolasion tyrtaeum* (also referred to as *Octolasion lacteum* (Shekhovtsov *et al.* 2014); 16 samples, one sample with two repeats) is a species complex with two cryptic lineages (Heethoff *et al.* 2004); it occurs in Europe, with populations extending to central Finland (Terhivuo & Saura 2006) and the taiga forests of European Russia (Perel 1979). *Aporrectodea rosea* (eight samples, one sample with two repeats) is also a species complex with a range that extends northwards from the Mediterranean towards central Finland (Terhivuo & Saura 2006) and the Middle Urals (Perel 1979). Tiunov *et al.* (2006) associate its occurrences in the northern part of the European Russian plain with cultivated soil (e.g. vegetable gardens), secondary deciduous forests and river valleys. The species found in the Urals is the one referred to as *A. rosea* L1 in the BOLD database, and this also occurs north of the Arctic Circle in Norway (C. Erséus, unpublished data). Although none of these lumbricids is recorded in the Polar Urals today, it is not unlikely that they were there during the Holocene Hypsithermal or other warmer periods.

Table 2. Comparison of the mammalian MamP007F – MamP007R primers with the ewB – ewC earthworm primers developed by Bienert *et al.* (2012).

	Forward	Reverse
Mammalian	5′–CGAGAAGACCCTATGGAGCT–3′	5′–CCGAGGTCRCCCCAACC–3′
Earthworm	5′–CAAGAAGACCCTATAGAGCTT–3′	5′–GGTCGCCCCAACCGAAT–3′

The glossoscolecid earthworm *Pontoscolex corethrurus* (six samples, one with two repeats) is a species complex with a circum-tropical distribution, native to South and Central America, but introduced in tropical and subtropical regions worldwide (Taheri *et al.* 2018). The family has no relatives in temperate environments and the genetic distance to other clitellate families rules out misidentification due to amplification or sequencing errors. The closest annelid sequence in GenBank belongs to the Asian *Amynthas glabrus* (Megascolecidae) recorded in China (Sun *et al.* 2017) and Japan (Blakemore 2003), at 77% sequence identity and with an edit distance of 8. The closest species in the results presented here is *Octolasion tyrtaeum* (Lumbricidae) at 63% sequence identity and an edit distance of 15. The most likely explanation for the detection of *Pontoscolex corethrurus* is contamination in the laboratory, possibly due to the reagents used.

The expected clitellate diversity in the Polar Urals is high based on previous lake diversity assessments. Baturina *et al.* (2014) recorded 30 aquatic species in the region. Unfortunately, little is known about the terrestrial clitellate diversity in the Polar Urals, making it difficult to assess how much of the diversity is captured in this study and what potential improvements can be made.

Only one annelid sequence was recorded at the Varanger site, representing a species in the *Lumbriculus variegatus* (one sample with two repeats) species complex. This complex has a current cosmopolitan distribution, but the particular lineage found on Varanger is an unidentified, probably undescribed, species; in its short (36 bp) 16S barcode it is 100% identical to the form of *L. variegatus* referred to as clade III by Gustafsson *et al.* (2009). Elsewhere, it has been recorded from Greenland, high-elevation sites (1000–1400 m a.s.l.) on the Scandinavian Peninsula (S. Martinsson & C. Erséus, unpublished data) and a lake at >3000 m a.s.l. in California (Gustafsson *et al.* 2009). This suggests that the complex is at least partially cold-adapted and could occur in northern Norway. In addition to the *Lumbriculus* species, the cercopagidid cladoceran *Bythotrephes longimanus* (six samples, five with two or more repeats) was detected, a species that is native to northern Europe and previously recorded on the Varanger Peninsula (Hessen *et al.* 2011).

Previous metabarcoding efforts of modern sediments on the Varanger Peninsula with enchytraeid specific primers targeting the mitochondrial 12S region resulted in identifications of *Cognettia sphagnetorum* and *Mesenchytraeus armatus* (Epp *et al.* 2012). Neither of these species could be detected in the results presented here. The discrepancy can be explained by the different primers used (the enchytraeid-specific primers can be expected to perform better than mammal primers used in this study), the age of the sediments (modern sediments compared to 3304–10 759 cal. a BP sediments), the type of sediment and how it retains DNA (heath and meadow plots compared to lake sediments) and local variation in clitellate diversity.

In-lake sampling of northern Norwegian lakes (C. Erséus & M. Klinth, unpublished data) indicates a high clitellate diversity (20–30 species). Both the previous metabarcoding study and in-lake sampling indicate that the results obtained here are an underestimation of the true diversity.

Although not reported before, after re-analysing the data presented by Giguet-Covex *et al.* (2014) we noted that clitellate sequences were also recovered. However, the eight species that could be identified (*Aporrectodea caliginosa*, *Chamaedrilus sphagnetorum*, *C. glandulosus*, *Dendrodrilus rubidus*, *Eiseniella tetraedra*, *Henlea perpusilla*, *Lumbricus meliboeus* and *Tubifex tubifex*) neither survived the filtering criteria applied by the authors (amplicon length shorter than 50 bp or identified as non-mammalian) nor the criteria used in this study (each taxon was only detected in a single repeat). Their annelid results are probably worse than the results presented in this study, due to the overall higher quality and success rate for mammalian DNA, but confirm that the annelid bycatch in this study is not a fluke.

The overall scattered detections of worm sequences in the Polar Urals and Varanger samples are most likely due to the non 'worm-specific' primers used, hindering, but not completely preventing, the amplification of the material. Furthermore, the detection of the four unexpected earthworm species warrants an explanation. These species might represent true positives, which have not been recorded in the region and represent past distributions during warmer periods. Alternatively, they could be artefacts of limited DNA reference material and might be misidentified to the wrong species or a consequence of amplification or sequencing errors. Finally, the observed worm sequences could be the results of contamination, either in the field or during sampling, extracting and amplification of the DNA. The laboratory standards used along with the negative controls give some confidence that these results are true detections, but contamination cannot be fully ruled out and is a likely explanation for the tropical *Pontoscolex corethrurus*.

Palaeoenvironmental implications of the worm detections

The sediments of Bolshoye Shchuchye (Polar Urals) are low in organic matter (1–5% LOI, see Fig. 4, based on Svendsen *et al.* 2019) and are essentially silt and clay. Given the thermal sensitivity of worms and the long record at this site (0–24 000 cal. a BP), we might expect a temporal pattern in the worm occurrence. At the species level this is not the case with the two Enchytraeidae (*Enchytraeus norvegicus* and *Henlea perpusilla*) occurring in both warm periods, such as the Holocene, and cold periods, including Heinrich Stadial 2 (24 000–22 000 cal. a BP). This is also true for the lumbricid earthworms (*Aporrectodea rosea*, *Bimastos norvegicus*, *Dendrobaena octaedra*, *Octolasion cyaneum* and *O. tyrtaeum*), which occur in the Holocene and the Lateglacial. When aggregated, the DNA shows distinctly greater and more continuous values for the Lumbricidae in the Holocene but no trend in the Enchy-
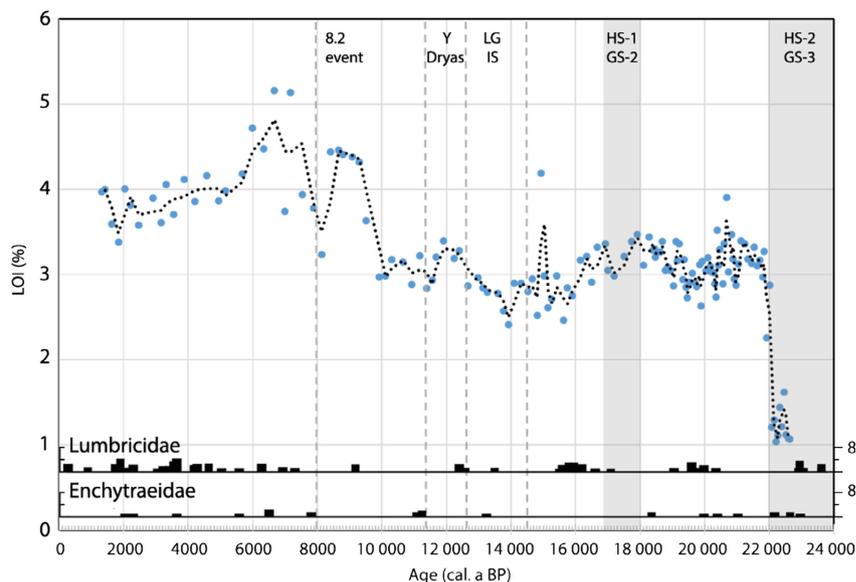
Fig. 4. Aggregated worm data from Lake Bolshoye Shchuchye (Polar Urals) with the core LOI and a two-period moving average (dotted line). Y Dryas is the Younger Dryas and LG IS is the Lateglacial interstadial. HS and GS are the Heinrich Stadial events and the Greenland Stadials from Rasmussen *et al.* (2014), respectively. [Colour figure can be viewed at www.boreas.dk]

traeidae (Fig. 4). These records suggest that both the soils and the lake sediments remained biologically active over the last 24 000 years, and that soils were almost certainly not set to zero biologically during the LGM or the Lateglacial stadials, when cold, dry conditions prevailed and the vegetation was predominantly tundra-steppe (Svendsen *et al.* 2014). However, the results do suggest higher rates of worm activity, and thus more soil formation, during the Holocene than during the Late Weichselian.

Potential for annelids in ancient DNA

Although the results for the worms detected in this study are not optimal, due to mismatched primers, overall poor metabarcoding results, low number of 'bycatch' taxa and perhaps some contamination of the samples, they indicate that earthworms and other clitellates can be identified in ancient sediments up to 24 000 years old. The use of a more optimized primer targeting short barcode regions in annelids, as has been done for the mitochondrial 16S and 12S regions (Bienert *et al.* 2012; Epp *et al.* 2012; Pansu *et al.* 2015), should increase clitellate diversity and detection reliability. Once detection methods have been optimized, tracking clitellate communities through time in ancient sediments may yield valuable information and proxies for various environmental conditions, such as temperature, soil moisture and acidity (Edwards & Lofty 1977; Beylich & Graefe 2009).

Conclusions

The results presented in this study show that the detection of mammalian material in ancient lake sediments in the Sub-Arctic via 16S metabarcoding is possible, but not without problems. The highly fragmented nature of *seda*DNA means that amplification of long fragments of low biomass taxa is problematic and might benefit from alternative identification methods. By contrast, clitellate worms look like a more promising group for metabarcoding in older late-Quaternary sediments. Although a previous attempt to retrieve enchytraeid material from permafrost sediments failed (Epp *et al.* 2012), the combination of suitable primers for targeting short fragments, high biomass (for earthworms in particular) and DNA contribution to the sediments warrants further investigation in the group and the possible effects of age and sediment types on metabarcoding success.

Data availability

The merged forward and reverse reads for both the Varanger and Polar Urals core, the used primer and tag sequences per sample, R script for the OBITOOLS filtering and the filtered OBITOOLS output are available on Dryad: https://doi.org/10.5061/dryad.g0f4hv0.

References

Alsos, I. G., Lammers, Y., Yoccoz, N. G., Jørgensen, T., Sjögren, P., Gielly, L. & Edwards, M. E. 2018: Metabarcoding lake sediments: taphonomy and representation of contemporary vegetation in environmental DNA (eDNA) records. *PLoS ONE 13*, e0195403, https://doi.org/10.1371/journal.pone.0195403.

Alsos, I. G., Sjögren, P., Edwards, M. E., Landvik, J. Y., Gielly, L., Forwick, M., Coissac, E., Brown, A. G., Jakobsen, L. V., Foreid, M. K. & Pedersen, M. W. 2016: Sedimentary ancient DNA from Lake Skartjorna, Svalbard: assessing the resilience of arctic flora to Holocene climate change. *Holocene 26*, 627–642.

Barnes, M. A. & Turner, C. R. 2016: The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics 17*, 1–17.

Baturina, M., Timm, T. & Loskutova, O. 2014: Oligochaete (Annelida, Clitellata) communities in lakes of the Ural Mountains (Russia). *Zoosymposia 9*, 77–94.

Berman, D. I., Meshcheryakova, E. N. & Leirikh, A. N. 2010: Egg cocoons of the earthworm *Dendrodrilus rubidus tenuis* (Lumbricidae, Oligochaeta) withstand the temperature of liquid nitrogen. *Doklady Biological Sciences 434*, 347–350.

Beylich, A. & Graefe, U. 2009: Investigations of annelids at soil monitoring sites in Northern Germany: reference ranges and time-series data. *Soil Organisms 81*, 175–196.

Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J. J. & Taberlet, P. 2012: Tracking earthworm communities from soil DNA. *Molecular Ecology 21*, 2017–2030.

Binladen, J., Gilbert, M. T. P., Bollback, J. P., Panitz, F., Bendixen, C., Nielsen, R. & Willerslev, E. 2007: The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS ONE 2*, e197, https://doi.org/10.1371/journal.pone.0000197.

Birkemoe, T., Coulson, S. J. & Sømme, L. 2000: Life cycles and population dynamics of enchytraeids (Oligochaeta) from the High Arctic. *Canadian Journal of Zoology 78*, 2079–2086.

Blaauw, M. & Christen, J. A. 2011: Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Analysis 6*, 457–474.

Blakemore, R. J. 2003: Japanese earthworms (Annelida: Oligochaeta): a review and checklist of species. *Organisms Diversity & Evolution 3*, 241–244.

Boessenkool, S., Epp, L. S., Haile, J., Bellemain, E., Edwards, M., Coissac, E., Willerslev, E. & Brochmann, C. 2012: Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. *Molecular Ecology 21*, 1806–1815.

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.

Brown, S. P., Veach, A. M., Rigdon-Huss, A. R., Grond, K., Lickteig, S. K., Lothamer, K., Oliver, A. K. & Jumpponen, A. 2015: Scraping the bottom of the barrel: are rare high throughput sequences artifacts? *Fungal Ecology 13*, 221–225.

Buée, M., Reich, M., Murat, C., Morin, E., Nilsson, R. H., Uroz, S. & Martin, F. 2009: 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist 184*, 449–456.

Byzova, J. B., Uvarov, A. V. & Petrova, A. D. 1995: Seasonal changes in communities of soil invertebrates in tundra ecosystems of Hornsund, Spitsbergen. *Polish Polar Research 16*, 245–266.

Canti, M. G. 2003: Earthworm activity and archaeological stratigraphy: a review of products and processes. *Journal of Archaeological Science 30*, 135–148.

Clarke, C., Edwards, M. E., Brown, A. G., Gielly, L., Lammers, Y., Heintzman, P. D., Ancin-Murguzur, F. J., Bråthen, K. A., Goslar, T. & Alsos, I. G. 2019: Holocene floristic diversity and richness in northeast Norway revealed by sedimentary ancient DNA (sedaDNA) and pollen. *Boreas 48*, 299–316.

Darwin, C. R. 1881: *The Formation of Vegetable Mould, Through the Action of Worms, with Observations on their Habits.* 326 pp. John Murray, London.

Domaizon, I., Winegardner, A., Capo, E., Gauthier, J. & Gregory-Eaves, I. 2017: DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity. *Journal of Paleolimnology 58*, 1–21.

Dushin, V. A., Serdyukova, O. P., Malyugin, A. A., Nikulina, I. A., Kozmin, V. S., Burmako, P. L., Abaturova, I. V. & Kozmina, L. I. 2009: *State Geological Map of the Russian Federation 1:200000. Sheet Q-42-I.* VSEGEI, St. Petersburg.

Edwards, C. A. & Lofty, J. R. 1977: *Biology of Earthworms.* 333 pp. Chapman and Hall, London.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E. & Brochmann, C. 2012: New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology 21*, 1821–1833.

Ermakov, A. I. & Golovanova, E. V. 2010: Species composition and abundance of earthworms in the tundra biocenoses of Denezhkin Kamen' Mountain (Northern Urals). *Contemporary Problems of Ecology 3*, 10–14.

Fauteux, D., Gauthier, G. & Berteaux, D. 2015: Seasonal demography of a cyclic lemming population in the Canadian Arctic. *Journal of Animal Ecology 84*, 1412–1422.

Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P. & Pompanon, F. 2010: An in silico approach for the evaluation of DNA barcodes. *BMC Genomics 11*, 434, https://doi.org/10.1186/1471-2164-11-434.

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

Finstad, G. L. & Prichard, A. K. 2000: Growth and body weight of free-range reindeer in western Alaska. *Rangifer 20*, 221–227.

Fjellanger, J. & Sørbel, L. 2007: Origin of the palaeic landforms and glacial impact on the Varanger Peninsula, northern Norway. *Norwegian Journal of Geology/Norsk Geologisk Forening 87*, 223–238.

Flagstad, O. & Roed, K. H. 2003: Refugial origins of reindeer (*Rangifer tarandus* L.) inferred from mitochondrial DNA sequences. *Evolution 57*, 658–670.

Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P.-J., Griggo, C., Gielly, L., Domaizon, I., Coissac, E., David, F., Choler, P., Poulenard, J. & Taberlet, P. 2014: Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature Communications 5*, 3211, https://doi.org/10.1038/ncomms4211.

Graham, R. W., Belmecheri, S., Choy, K., Culleton, B. J., Davies, L. J., Froese, D., Heintzman, P. D., Hritz, C., Kapp, J. D., Newsom, L. A., Rawcliffe, R., Saulnier-Talbot, É., Shapiro, B., Wang, Y., Williams, J. W. & Wooller, M. J. 2016: Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska. *Proceedings of the National Academy of Sciences of the United States of America 113*, 9310–9314.

Gustafsson, D. R., Price, D. A. & Erséus, C. 2009: Genetic variation in the popular lab worm *Lumbriculus variegatus* (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. *Molecular Phylogenetics and Evolution 51*, 182–189.

Haflidason, H., Lundekvam, J., Gyllencreutz, R., Svendsen, J. I., Gladysh, S. & Elizaveta, L. 2019: The Last Glacial and Holocene

seismostratigraphy and sediment distribution of the Lake Bolshoye Shchuchye, Polar Ural, Arctic Russia. *Boreas* 48, 452–469.

Haile, J., Froese, D. G., MacPhee, R. D. E., Roberts, R. G., Arnold, L. J., Reyes, A. V., Rasmussen, M., Nielsen, R., Brook, B. W., Robinson, S., Demuro, M., Gilbert, M. T. P., Munch, K., Austin, J. J., Cooper, A., Barnes, I., Moller, P. & Willerslev, E. 2009: Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proceedings of the National Academy of Sciences of the United States of America* 106, 22352–22357.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & DeWaard, J. R. 2003: Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* 270, 313–321.

Heethoff, M., Etzold, K. & Scheu, S. 2004: Mitochondrial COII sequences indicate that the parthenogenetic earthworm *Octolasion tyrtaeum* (Savigny 1826) constitutes of two lineages differing in body size and genotype. *Pedobiologia* 48, 9–13.

Hessen, D. O., Bakkestuen, V. & Walseng, B. 2011: The ecological niches of *Bythotrephes* and *Leptodora*: lessons for predicting long-term effects of invasion. *Biological Invasions* 13, 2561–2572.

Hughes, A. L. C., Gyllencreutz, R., Lohne, Ø. S., Mangerud, J. & Svendsen, J. I. 2016: The last Eurasian ice sheets – a chronological database and time-slice reconstruction, DATED-1. *Boreas* 45, 1–45.

Karaca, A., Kizilkaya, R., Turgay, O. C. & Cetin, S. C. 2010: Effects of earthworms on the availability and removal of heavy metals in soil. *In* Sherameti, I. & Varma, A. (eds.): *Soil Heavy Metals*, 369–388. Springer, Berlin.

Kvie, K. S., Heggenes, J., Anderson, D. G., Kholodova, M. V., Sipko, T., Mizin, I. & Røed, K. H. 2016: Colonizing the High Arctic: mitochondrial DNA reveals common origin of Eurasian archipelagic reindeer (*Rangifer tarandus*). *PLoS ONE* 11, e0165237, https://doi.org/10.1371/journal.pone.0165237.

Kwok, S., Kellogg, D., McKinney, N. & Spasic, D. 1990: Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. *Nucleic Acids Research* 18, 999–1005.

Leonard, J. A., Shanks, O., Hofreiter, M., Kreuz, E., Hodges, L., Ream, W., Wayne, R. K. & Fleischer, R. C. 2007: Animal DNA in PCR reagents plagues ancient DNA research. *Journal of Archaeological Science* 34, 1361–1366.

Lie, R. 1986: Animal bones from the Late Weichselian in Norway. *Fauna Norwegica, Serie A 7*, 41–46.

Mangerud, J., Briner, J. P., Goslar, T. & Svendsen, J. I. 2017: The Bølling-age Blomvåg Beds, western Norway: implications for the Older Dryas glacial re-advance and the age of the deglaciation. *Boreas* 46, 162–184.

Nesje, A. 1992: A piston corer for lacustrine and marine sediments. *Arctic and Alpine Research* 24, 257–259.

Pääbo, S., Poinar, H., Serre, D., Svante, P., Jaenicke-despr, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. & Hofreiter, M. 2004: Genetic analyses from ancient DNA. *Annual Review of Genetics* 38, 645–679.

Pansu, J., De Danieli, S., Puissant, J., Gonzalez, J.-M., Gielly, L., Cordonnier, T., Zinger, L., Brun, J.-J., Choler, P., Taberlet, P. & Cécillon, L. 2015: Landscape-scale distribution patterns of earthworms inferred from soil DNA. *Soil Biology and Biochemistry* 83, 100–105.

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

Parducci, L., Jørgensen, T., Tollefsrud, M. M., Elverland, E., Alm, T., Fontana, S. L., Bennett, K. D., Haile, J., Matetovici, I., Suyama, Y., Edwards, M. E., Andersen, K., Rasmussen, M., Boessenkool, S., Coissac, E., Brochmann, C., Taberlet, P., Houmark-Nielsen, M., Larsen, N. K., Orlando, L., Gilbert, M. T. P., Kjær, K. H., Alsos, I. G. & Willerslev, E. 2012: Glacial survival of boreal trees in northern Scandinavia. *Science* 335, 1083–1086.

Parducci, L., Väliranta, M., Salonen, J. S., Ronkainen, T., Matetovici, I., Fontana, S. L., Eskola, T., Sarala, P. & Suyama, Y. 2015: Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences* 370, 20130382, https://doi.org/10.1098/rstb.2013.0382.

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

Pedersen, M. W., Overballe-Petersen, S., Ermini, L., Der Sarkissian, C., Haile, J., Hellstrom, M., Spens, J., Thomsen, P. F., Bohmann, K., Cappellini, E., Schnell, I. B., Wales, N. A., Carøe, C., Campos, F., Schmidt, A. M. Z., Gilbert, M. T. P., Hansen, A. J., Orlando, L. & Willerslev, E. 2015: Ancient and modern environmental DNA. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 20130383, https://doi.org/10.1098/rstb.2013.0383.

Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., Mendoza, M. L. Z., Beaudoin, A. B., Zutter, C., Larsen, N. K., Potter, B. A., Nielsen, R., Rainville, R. A., Orlando, L., Meltzer, D. J., Kjær, K. H. & Willerslev, E. 2016: Postglacial viability and colonization in North America's ice-free corridor. *Nature* 537, 45–49.

Perel, T. S. 1979: *Range and Regularities in the Distribution of Earthworms of the USSR Fauna*. 268 pp. Nauka, Moscow.

Porazinska, D. L., Giblin-Davis, R. M., Faller, L., Farmerie, W., Kanzaki, N., Morris, K., Powers, T. O., Tucker, A. E., Sung, W. & Thomas, W. K. 2009: Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Molecular Ecology Resources* 9, 1439–1450.

R Core Team. 2017: *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.r-project.org/.

Rasmussen, L. M. & Holmstrup, M. 2002: Geographic variation of freeze-tolerance in the earthworm *Dendrobaena octaedra*. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology* 172, 691–698.

Rasmussen, S. O., Bigler, M., Blockley, S. P., Blunier, T., Buchardt, S. L., Clausen, H. B., Cvijanovic, I., Dahl-Jensen, D., Johnsen, S. J., Fischer, H., Gkinis, V., Guillevic, M., Hoek, W. Z., Lowe, J. J., Pedro, J. B., Popp, T., Seierstad, I. K., Steffensen, J. P., Svensson, A. M., Vallelonga, P., Vinther, B. M., Walker, M. J. C., Wheatley, J. J. & Winstrup, M. 2014: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy. *Quaternary Science Reviews* 106, 14–28.

Rasmussen, M., Cummings, L. S., Gilbert, M. T. P., Bryant, V., Smith, C., Jenkins, D. L. & Willerslev, E. 2009: Response to comment by Goldberg et al. on "DNA from Pre-Clovis Human Coprolites in Oregon, North America". *Science* 325, 148 LP-148, https://doi.org/10.1126/science.1167672.

Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Ramsey, C. B., Buck, C. E., Cheng, H., Edwards, R. L., Friedrich, M., Grootes, P. M., Guilderson, T. P., Haflidason, H., Hajdas, I., Hatté, C., Heaton, T. J., Hoffmann, D. L., Hogg, A. G., Hughen, K. A., Kaiser, K. F., Kromer, B., Manning, S. W., Niu, M., Reimer, R. W., Richards, D. A., Scott, E. M., Southon, J. R., Staff, R. A., Turney, C. S. M. & van der Plicht, J. 2013: IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* 55, 1869–1887.

Rota, E. 1995: Italian Enchytraeidae (Oligochaeta). I. *Italian Journal of Zoology* 62, 183–231.

Rota, E., Caruso, T. & Bargagli, R. 2014: Community structure, diversity and spatial organization of enchytraeids in Mediterranean urban holm oak stands. *European Journal of Soil Biology* 62, 83–91.

Schloss, P. D., Gevers, D. & Westcott, S. L. 2011: Reducing the effects of PCR amplification and sequencing artifacts on 16s rRNA-based studies. *PLoS ONE* 6, e27310, https://doi.org/10.1371/journal.pone.0027310.

Schnell, I. B., Bohmann, K. & Gilbert, M. T. P. 2015: Tag jumps illuminated – reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources* 15, 1289–1303.

Seersholm, F. V., Pedersen, M. W., Søe, M. J., Shokry, H., Mak, S. S. T., Ruter, A., Raghavan, M., Fitzhugh, W., Kjær, K. H., Willerslev, E., Meldgaard, M., Kapel, C. M. O. & Hansen, A. J. 2016: DNA evidence of bowhead whale exploitation by Greenlandic Paleo-Inuit 4,000 years ago. *Nature Communications* 7, https://doi.org/10.1038/ncomms13389.

Shekhovtsov, S. V., Golovanova, E. V. & Peltek, S. E. 2014: Genetic diversity of the earthworm *Octolasion tyrtaeum* (Lumbricidae, Annelida). *Pedobiologia 57*, 245–250.

Siedlecka, A. & Roberts, D. 1992: *The Bedrock Geology of Varanger Peninsula, Finnmark, North Norway: An Excursion Guide.* 45 pp. Norges Geologiske Undersøkelse, Trondheim.

Sipos, R., Székely, A. J., Palatinszky, M., Révész, S., Márialigeti, K. & Nikolausz, M. 2007: Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiology Ecology 60*, 341–350.

Slon, V., Hopfe, C., Weiß, C. L., Mafessoni, F., De La Rasilla, M., Lalueza-Fox, C., Rosas, A., Soressi, M., Knul, M. V., Miller, R., Stewart, J. R., Derevianko, A. P., Jacobs, Z., Li, B., Roberts, R. G., Shunkov, M. V., De Lumley, H., Perrenoud, C., Gušić, I., Kućan, Ž., Rudan, P., Aximu-Petri, A., Essel, E., Nagel, S., Nickel, B., Schmidt, A., Prüfer, K., Kelso, J., Burbano, H. A., Pääbo, S. & Meyer, M. 2017: Neandertal and Denisovan DNA from Pleistocene sediments. *Science 356*, 605–608.

Solomina, O., Ivanov, M. & Bradwell, T. 2010: Lichenometric studies on moraines in the Polar Urals. *Geografiska Annaler: Series A, Physical Geography 92*, 81–99.

Stokes, C. R., Corner, G. D., Winsborrow, M. C. M., Husum, K. & Andreassen, K. 2014: Asynchronous response of marine-terminating outlet glaciers during deglaciation of the Fennoscandian Ice Sheet. *Geology 42*, 455–458.

Stroeven, A. P., Hättestrand, C., Kleman, J., Heyman, J., Fabel, D., Fredin, O., Goodfellow, B. W., Harbor, J. M., Jansen, J. D., Olsen, L., Caffee, M. W., Fink, D., Lundqvist, J., Rosqvist, G. C., Strömberg, B. & Jansson, K. N. 2016: Deglaciation of Fennoscandia. *Quaternary Science Reviews 147*, 91–121.

Stuiver, M., Reimer, P. J. & Reimer, R. W. 2018: *CALIB 7.1.* Available at: http://calib.org/.

Sun, J., James, S. W., Jiang, J., Yao, B., Zhang, L., Liu, M., Qiu, J. & Hu, F. 2017: Phylogenetic evaluation of *Amynthas* earthworms from South China reveals the initial ancestral state of spermathecae. *Molecular Phylogenetics and Evolution 115*, 106–114.

Svendsen, J. I., Alexanderson, H., Astakhov, V. I., Demidov, I., Dowdeswell, J. A., Funder, S., Gataullin, V., Henriksen, M., Hjort, C., Houmark-Nielsen, M., Hubberten, H. W., Ingólfsson, Ó., Jakobsson, M., Kjær, K. H., Larsen, E., Lokrantz, H., Lunkka, J. P., Lyså, A., Mangerud, J., Matiouchkov, A., Murray, A., Möller, P., Niessen, F., Nikolskaya, O., Polyak, L., Saarnisto, M., Siegert, C., Siegert, M. J., Spielhagen, R. F. & Stein, R. 2004: Late Quaternary ice sheet history of northern Eurasia. *Quaternary Science Reviews 23*, 1229–1271.

Svendsen, J. I., Færseth, L. M. B., Gyllencreutz, R., Haflidason, H., Henriksen, M., Hovland, M. N., Lohne, Ø. S., Mangerud, J., Nazarov, D., Regnéll, C. & Schaefer, J. M. 2019: Glacial and environmental changes over the last 60,000 years in the Polar Ural Mountains, Arctic Russia. *Boreas 48*, 407–431.

Svendsen, J. I., Krüger, L. C., Mangerud, J., Astakhov, V. I., Paus, A., Nazarov, D. & Murray, A. 2014: Glacial and vegetation history of the Polar Ural Mountains in northern Russia during the Last Ice Age, Marine Isotope Stages 5–2. *Quaternary Science Reviews 92*, 409–428.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. 2012: Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology 21*, 2045–2050.

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

Taheri, S., Pelosi, C. & Dupont, L. 2018: Harmful or useful? A case study of the exotic peregrine earthworm morphospecies *Pontoscolex corethrurus. Soil Biology and Biochemistry 116*, 277–289.

Terhivuo, J. & Saura, A. 2006: Dispersal and clonal diversity of North-European parthenogenetic earthworms. *In* Hendrit, P. F. (ed.): *Biological Invasions Belowground: Earthworms as Invasive Species*, 5–18. Springer, Dordrecht.

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

Tiunov, A. V., Hale, C. M., Holdsworth, A. R. & Vsevolodova-Perel, T. S. 2006: Invasion patterns of Lumbricidae into the previously earthworm-free areas of northeastern Europe and the western Great Lakes region of North America. *Biological Invasions 8*, 1223–1234.

Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E., Coissac, E., Pompanon, F., Gielly, L., Cruaud, C., Nascetti, G., Wincker, P., Swenson, J. E. & Taberlet, P. 2009: New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Molecular Ecology Resources 9*, 51–60.

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.-M., Peroux, T., Crivelli, A. J., Olivier, A., Acqueberge, M., Le Brun, M., Møller, P. R., Willerslev, E. & Dejean, T. 2015: Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology 25*, 929–942.

Vistnes, I., Nellemann, C., Jordhøy, P. & Strand, O. 2001: Wild reindeer: impacts of progressive infrastructure development on distribution and range use. *Polar Biology 24*, 531–537.

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D. A. & Cooper, A. 2003: Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science 300*, 791–795.

Willerslev, E., Hansen, A. J. & Poinar, H. N. 2004: Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends in Ecology and Evolution 19*, 141–147.

Wu, J. H., Hong, P. Y. & Liu, W. T. 2009: Quantitative effects of position and type of single mismatch on single base primer extension. *Journal of Microbiological Methods 77*, 267–275.

Yannic, G., Pellissier, L., Ortego, J., Lecomte, N., Couturier, S., Cuyler, C., Dussault, C., Hundertmark, K. J., Irvine, R. J., Jenkins, D. A., Kolpashikov, L., Mager, K., Musiani, M., Parker, K. L., Røed, K. H., Sipko, T., Þórisson, S. G., Weckworth, B. V., Guisan, A., Bernatchez, L. & Côté, S. D. 2014: Genetic diversity in caribou linked to past and future climate change. *Nature Climate Change 4*, 132–137.

Zimmermann, H., Raschke, E., Epp, L., Stoof-Leichsenring, K., Schirrmeister, L., Schwamborn, G. & Herzschuh, U. 2017: The history of tree and shrub taxa on Bol'shoy Lyakhovsky Island (New Siberian Archipelago) since the last interglacial uncovered by sedimentary ancient DNA and pollen data. *Genes 8*, 273, https://doi.org/10.3390/genes8100273.

Supporting Information

Additional Supporting Information may be found in the online version of this article at http://www.boreas.dk.

Table S1. Average mismatches between all clitellate families, species and the MamP007F – MamP007R mammal primers.

Table S2. Metabarcoding results for the Polar Urals core, including the repeats, read abundances and sequence information.

Table S3. Metabarcoding results for the Varanger core, including the repeats, read abundances and sequence information.