**Clitellate worms (Annelida) in late-glacial and Holocene sedimentary DNA records from the Polar Urals and northern Norway.**

Youri Lammers, Charlotte L. Clarke, Christer Erséus, Antony G. Brown, Mary E. Edwards, Ludovic Gielly, Haflidi Haflidason, Jan Mangerud, Emilia Rota, John Inge Svendsen and Inger Greve Alsos.

While there are extensive macro- and microfossil records of a range of plants and animals from Quaternary records, earthworms and their close relatives among annelids are not preserved as fossils, and therefore we have limited knowledge of their Quaternary distributions. 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 24,000-1,300 cal. years BP, and NE Norway (10,700-3,300 cal. years 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. years BP, and four times in the Norwegian core at 11,000 cal. years BP and between 3,600-3,300 cal. years 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.

**Keywords:** ancient DNA, Arctic, environmental DNA, metabarcoding, clitellates

*Youri Lammers (youri.lammers@uit.no), Tromsø Museum, University of Tromsø – The Arctic University of Norway, NO-9037 Tromsø, Norway*

*Charlotte L. Clarke (C.Clarke@soton.ac.uk), Geography and Environment, University of Southampton, Highfield, Southampton, SO17 1BJ, UK*

*Christer Erséus (christer.erseus@bioenv.gu.se), Department of Biological and Environmental Sciences, University of Gothenburg, P.O. Box 463, SE-405 30 Göteborg, Sweden*

*Antony G. Brown (Tony.Brown@soton.ac.uk), Geography and Environment, University of Southampton, Highfield, Southampton, SO17 1BJ, UK*

*Mary E. Edwards (M.E.Edwards@soton.ac.uk), 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 (ludovic.gielly@univ-grenoble-alpes.fr), Laboratoire d’Ecologie Alpine (LECA), Université Grenoble Alpes, C2 40700 38058, Grenoble Cedex 9, France*

*Haflidi Haflidason (Haflidi.Haflidason@uib.no), Department of Earth Science, University of Bergen, PO Box 7803, NO-5020 Bergen, Norway and Bjerknes Centre for Climate Research, Bergen, Norway*

*Jan Mangerud (Jan.Mangerud@uib.no), Department of Earth Science, University of Bergen, PO Box 7803, NO-5020 Bergen, Norway and Bjerknes Centre for Climate Research, Bergen, Norway*

*Emilia Rota (rota@unisi.it), Department of Physics, Earth and Environmental Sciences, University of Siena, Via P.A. Mattioli 4, IT-53100 Siena, Italy*

*John Inge Svendsen (John.Svendsen@uib.no), Department of Earth Science, University of Bergen, PO Box 7803, NO-5020 Bergen, Norway and Bjerknes Centre for Climate Research, Bergen, Norway*

*Inger Greve Alsos (inger.g.alsos@uit.no), Tromsø Museum, University of Tromsø – The Arctic University of Norway, NO-9037 Tromsø, Norway*

**Introduction**

The fact that earthworms (Clitellata: Megadrili) have an important function in cycling nutrients and structuring soils was famously recognised by Darwin (Darwin, 1881). Both earthworms as well as potworms (Clitellata: Enchytraeidae) are used as indicator species for various environmental issues in modern soils and aquatic systems as some are very tolerant to 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 particular pH (Edwards and Lofty, 1977; Beylich and 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), 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 and Willerslev, 2015; Domaizon et al., 2017).

After being released in 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 and 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 explored before, unlike modern soils yielded no results (Epp et al., 2012), suggesting that detecting 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 late-glacial 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 Schuchye is located in the northernmost Polar Ural Mountains of Arctic Russia (67°53’23.78”N, 66°18’52.59”E; 221 m a.s.l.; Figure 1). Bolshoye Schuchye is an elongated lake (12 km long, 1 km wide) located in a NW-SE oriented valley with a maximum water depth of 136 m in its central part (Svendsen et al., 2018) and up to a 160 m of lacustrine sediments in the central and northern parts (Haflidason et al., 2018). 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 km2 (Svendsen et al., 2018). The bedrock consists of Proterozoic-Cambrian basaltic and andesitic rocks, and Ordovician quartzite and phyllitic rocks in the E-NW and SW regions, respectively (Dushin et al., 2009). The Polar Urals remained mostly ice-free during the Late 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 oC (Solomina et al, 2010).

*Varanger Peninsula*

The lake (70o19’06.85”N, 30o01’43.84”E; Figure 1) caries the local Sami name ‘Uhca Rohči’, while it is unnamed on the 1:50,000 Norwegian Topographic Map (Norgeskart; <https://www.norgeskart.no>). 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, north-east 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. 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 and 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 and the peninsula was certainly free of glacial ice by 13,000-12,000 cal. years 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 oC (Norwegian Meteorological Institute; www.met.no).

**Methods**

*Polar Urals lake sediment*

Lake Bolshoye Schuchye was cored during several expeditions between 2007 and 2009. The 24 m long core (number 506-48) that was sampled for metabarcoding was obtained in July 2009 from the southern part of the lake (67°51’22.20”N, 66°21’30.07”E). The lake was cored 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 steel tubes. The full core was obtained by taking new segments from the same hole. All sections were stored and transported at above 0oC 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 lab contamination. Due to deformation near the top of each core segment, the samples form a non-continuous record, a second core was taken parallel to the first core at a 35cm offset to account for the deformations but was not sampled for this study. Dating was based on 26 AMS radiocarbon dates from plant macrofossils provided by the Poznan 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. (2018).

*Varanger Peninsula lake sediment*

The 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 to 1 m sections which were sealed to reduce the risk of contaminating the sediments. The core sections were kept at above 0 oC conditions in the field and during transport to avoid freezing of the sediments and were stored in a 4oC cold room at the Tromsø University Museum (TMU). Sampling of the core took place in a dedicated ancient DNA laboratory. The core was radiocarbon dated based on seven AMS radiocarbon dates on terrestrial plant macrofossils provided by the Poznan Radiocarbon Laboratory. Dates were calibrated using the terrestrial IntCal13 curve (Reimer et al., 2013), and the age model constructed using the Bayesian framework calibration software ‘Bacon’ (v2.2) (Blaauw and Christen, 2011), which was implemented in R (v3.2.4) (R Core Team, 2017). A full sedimentology and chronology is described by Clarke et al. (2018).

*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 the 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 ca. 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 samples 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 supress 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 Ural 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 2x125bp 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, 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*, 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 less 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 which 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 dataset. Common lab 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 were 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. years BP, with a total of 27,133 reads) and eight clitellate taxa including: two Enchytraeidae (*Enchytraeus norvegicus*, *Henlea perpusilla*), one Glossoscolecidae (*Pontoscolex corethrurus*) and six Lumbricidae (*Aporrectodea rosea*, *Dendrobaena octaedra*, *Bimastos norvegicus*, *Octolasion cyaneum* and *Octolasion tyrtaeum*) (Figure 2). The results also included seven 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, belonging to: *Rangifer tarandus* (four occurrences at 10,800 cal. years BP and three between 3,300 and 3,600 cal. years BP, with 44,979 reads), the spiny water flea (zooplanktonic cladoceran) *Bythotrephes longimanus* (six occurrences at 4,900; 5,600, 5,700; 6,300; 6,500 and 9,100 cal. years BP, sum 38,085 reads) and *Lumbriculus variegatus* (one occurrence at 10,800 cal. years BP with 227 reads) (Figure 3). 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 1,756 species (mean 175 sequences per family, SD=317.7) out of the 28 families listed in the NCBI taxonomy database. The weighted average number of mismatches in the forward and reverse primer was 2.07 (SD=0.05) and 2.04 (SD=0.24) respectively, with an average estimated amplicon length of 35.7bp (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 table for all clitellate families and species is provided in supplementary Table 1. 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, Figure 2. and three Varanger samples with two, three and four repeats, Figure 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 since *R. tarandus* has a circumpolar Eurasian distribution. It is known from western Norway at 13,500 cal. years BP from Blomvåg (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 years BP) as based on genetic data this area was probably its main glacial refugium (Flagstad and Roed, 2003; Yannic et al., 2014; Kvie et al., 2016).

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 loosing 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 dataset 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 would be included, the limited occurrences in the records (only a single sample) means that they are not useful for palaeoecological reconstructions as it requires reliably detection in the records.

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, age of the sediments or the size of the target amplicon. The amount of DNA deposited in the lake might be limited by the accessibility for for mammals, such as the steep slopes surrounding lake Bolshoye Schuchye. Alternatively, the plentiful water sources in the Komagdalen valley 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 worked with lake sediments from locations that had high mammalian concentrations, either due to migration routes (Pedersen et al., 2016), 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., 2018) and Polar Urals sites (Clarke et al., in prep.). Though the success for plants could be explained by the obvious higher biomass and thus DNA contribution to the sediments and a potential lower average fragment length, for example the plant data from the Varanger site had an average length of 44.3 bp (±15.6) (Clarke et al., 2018) 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 lab 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 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 Ural and Varanger core, 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 cites 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 lab contaminants by targeting a narrower taxonomic group. Other alternatives are to bypass the usage of primers altogether by either shotgun sequencing sediments extracts (Pedersen et al., 2016; Seersholm et al., 2016) or using DNA target capture (Slon et al., 2017).

*Presence of worms*

Off-target amplification of earthworms and other clitellates was observed in both the Polar Urals and 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 less 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 bp difference to account for the mismatches between the mammalian and earthworm primer binding site, 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 1,160 kg/km2 (Byzova et al., 1995) and Lumbricidae biomass in the northern Ural mountains is calculated to be 24,000 kg/km2 (Ermakov and Golovanova, 2010). Thus, the clitellates numbers are far higher than common herbivorous mammals such as the North American brown lemming (*Lemmus trimucronatus*) at 30 kg/km2 in the Canadian Arctic (Fauteux et al., 2015) or *R. tarandus* in central Norway at 165 kg/km2 (Finstad and 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 to be 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 core is 74 bp on average (Table 1) and the average amplicon length for all clitellate families is 35 bp (Supplementary Table 1). 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 2 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 (>1400m) elevations in southern Norway (Erséus et al., 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 and 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, current records extend up to southern Finland and northern Sweden (Terhivuo and Saura, 2006). In Norway it can be found to elevations of around 1,000 m in the south, and in lowland localities north of the Arctic Circle (Erséus et al., unpublished), but it is most often associated with human habitats. *Octolasion tyrtaeum* (also referred to as *Octolasion lacteum* (Shekhovtsov, Golovanova and Peltek, 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 and 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 and 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 (Erséus et al., unpublished data). Although none of these lumbricids are recorded in the Polar Urals today, it is not unlikely that they were there during the Holocene Hypsithermal or other warmer periods.

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 on 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 our results 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 lab, possibly due to the reagents used.

The expected clitellate diversity in the Polar Urals is high based on previous in lake diversity assessments. (Baturina et al., 2014), could record 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 species found on Varanger is an unidentified, probably undescribed, species. Elsewhere, it has been recorded from Greenland, high elevation sites (1,000-1,400 m) on the Scandinavian peninsula (Erséus et al., unpublished data) and California (Gustafsson et al., 2009). This suggests the species is at least partially cold-adapted and could occur in northern Norway. In addition to the *Lumbriculus* species, the zooplanktonic 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 3,304 – 10,759 cal yrs BP sediments) and 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 (Erséus et al., unpublished), indicates a high clitellate diversity (20-30 species). Both the previous metabarcoding study and in lake sampling indicates that the results obtained here are an underestimation of the true diversity.

Although not reported, reanalysis of the data presented by (Giguet-Covex et al., 2014) indicates that clitellate sequences were also recovered. However, the 8 species that could be identified (*Aporrectodea caligninosa*, *Chamaedrilus sphagnetorum*, *C. glandulosus*, *Dendrodrilus rubidus*, *Eiseniella tetraedra*, *Henlea perpusilla*, *Lumbricus meliboeus*, and *Tubifex tubifex*), did neither survive 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). The annelid results are likely worse than the results presented in this study due to the overall higher quality and success rate for mammalian DNA but confirms that the annelid bycatch in this study is not a fluke.

Overall scattered detection 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 lab 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 Schuchye (Polar Urals) are low in organic matter (1-5% LOI, see figure 4; Svendsen et al., 2018) and are essentially silt and clay, as clay is thought to attract and adsorb DNA in soils and lake sediments (Cai et al., 2006; Yankson and Steck, 2009). Also, given the thermal sensitivity of worms and the long record at this site (0-24,000 cal. years 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 (22,000-24,000 cal. years BP). This is also true for the Lumbricidae (*Aporrectodea rosea*, *Bimastos norvegicus*, *Dendrobaena octaedra*, *Octolasion cyaneum* and *O.* *tyrtaeum*), which occur in the Holocene and the late glacial. When aggregated, the DNA shows distinctly greater and more continuous values for the Lumbricidae in the Holocene but no trend in the Enchytraeidae (Figure 4). These records suggests 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 late-glacial 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 clitellates diversity and detection reliability. Once detection methods have been optimized, tracking clitellate communities through time in ancient sediments can yield valuable information and proxies for various environmental conditions, such as, temperature, soil moisture and acidity (Edwards and Lofty, 1977; Beylich and Graefe, 2009).

**Conclusion**

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. Clitellate worms, on the other hand, 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.

**Acknowledgements**

We thank Francisco Javier Ancin-Murguzur andJohn-Arne Kristiansen for field assistance in Varanger, Dmitry Nazarov and Aleksei Matioushkov for field assistance in the Polar Urals, Marie Kristine Føreid Merkel for lab assistance, Frédéric Boyer for raw sequence data handling, Dorothee Ehrich and Peter Heintzman for fruitful discussions.

**Funding**

This work was supported by the Norwegian Research Council (grant nos. 213692/F20 and 250963/F20 to Alsos and 255415 to Svendsen). Lammers was financed by an internal PhD position at the Tromsø University Museum. Erséus is supported by the Norwegian and Swedish Taxonomy Initiatives (Artsdatabanken and Artdatabanken) and the Adlerbert Foundation.

**Author contribution**

Lammers analysed the data and wrote the first version of the manuscript; Clarke, Brown and Alsos carried out fieldwork on Varanger and Svendsen, Haflidason and Mangerud in the Polar Urals; Clarke extracted the DNA; Gielly amplified the DNA and ran OBITools; Erséus and Rota provided taxonomic and ecological data; the two projects were devised by Alsos/Edwards and Svendsen/Mangerud/Haflidason. All authors contributed in various ways to the final version of the manuscript.

**References**

Alsos, I.G., Lammers, Y., Yoccoz, N.G., Jørgensen, T., Sjögren, P., Gielly, L., et al. 2018: Metabarcoding lake sediments: taphonomy and representation of contemporary vegetation in environmental DNA (eDNA) records. *PLOS ONE*, 264903.

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

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

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

Berman, D.I., Meshcheryakova, E.N. and 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. and 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., et al. 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., et al. 2007: The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS ONE* **2**, 1–9.

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

Blaauw, M. and 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., et al. 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. and Coissac, E. 2016: obitools: a unix-inspired software package for DNA metabarcoding. *Molecular ecology resources* **16**, 176–82.

Brown, S.P., Veach, A.M., Rigdon-Huss, A.R., Grond, K., Lickteig, S.K., Lothamer, K., et al. 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., et al. 2009: 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist* **184**, 449–456.

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

Cai, P., Huang, Q.Y. and Zhang, X.W. 2006: Interactions of DNA with clay minerals and soil colloidal particles and protection against degradation by DNase. *Environmental Science and Technology* **40**, 2971–2976.

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., et al. 2018: Holocene floristic diversity and richness in northeast Norway revealed by sedimentary ancient DNA (sedaDNA) and pollen. *Boreas*.

Darwin, C.R. 1881: *The Formation of Vegetable Mould, through the Action of Worms, with Observations on Their Habits*. John Murray, London.

Domaizon, I., Winegardner, A., Capo, E., Gauthier, J. and 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., et al. 2009: *State Geological Map of the Russian Federation 1:200000.*, 2nd ed. VSEGEI, St. Petersburg.

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

Epp, L.S., Boessenkool, S., Bellemain, E.P., Haile, J., Esposito, A., Riaz, T., et al. 2012: New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology* **21**, 1821–1833.

Ermakov, A.I. and 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. and 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., et al. 2010: An in silico approach for the evaluation of DNA barcodes. *BMC genomics* **11**, 434.

Ficetola, G.F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., et al. 2015: Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular ecology resources* **15**, 543–56.

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

Fjellanger, J. and 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**.

Flagstad, O. and 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., et al. 2014: Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nature communications* **5**, 3211.

Graham, R.W., Belmecheri, S., Choy, K., Culleton, B.J., Davies, L.J., Froese, D., et al. 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**, 9310–4.

Gustafsson, D.R., Price, D.A. and 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. and Elizaveta, L. 2018: The Last Glacial and Holocene Seismostratigraphy and sediment distribution of the Lake Bolshoye Shchuchye, Polar Ural, Arctic Russia. *Boreas*.

Haile, J., Froese, D.G., MacPhee, R.D.E., Roberts, R.G., Arnold, L.J., Reyes, A. V., et al. 2009: Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proceedings of the National Academy of Sciences* **106**, 22352–22357.

Hebert, P.D.N., Cywinska, A., Ball, S.L. and deWaard, J.R. 2003: Biological identifications through DNA barcodes. *Proceedings. Biological sciences / The Royal Society* **270**, 313–21.

Heethoff, M., Etzold, K. and 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. and 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. and 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. and Cetin, S.C. 2010: Effects of Earthworms on the Availability and Removal of Heavy Metals in Soil. pp. 369–388. Springer, Berlin, Heidelberg.

Kvie, K.S., Heggenes, J., Anderson, D.G., Kholodova, M. V., Sipko, T., Mizin, I., et al. 2016: Colonizing the High Arctic: Mitochondrial DNA Reveals Common Origin of Eurasian Archipelagic Reindeer (Rangifer tarandus). *PLOS ONE* **11**, e0165237.

Kwok, S., Kellogg, D., McKinney, N. and 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., et al. 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 Norwegia, Serie A* **7**, 41–46.

Mangerud, J., Briner, J.P., Goslar, T. and 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.

Pääbo, S., Poinar, H., Serre, D., Svante, P., Jaenicke-despr, V., Hebler, J., et al. 2004: Genetic Analyses from Ancient DNA. *Annu. Rev. Genet.* **38**, 645–679.

Pansu, J., De Danieli, S., Puissant, J., Gonzalez, J.-M., Gielly, L., Cordonnier, T., et al. 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., et al. 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., et al. 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., et al. 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.

Pedersen, M.W., Ginolhac, A., Orlando, L., Olsen, J., Andersen, K., Holm, J., et al. 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., Ruter, A., Schweger, C., Friebe, H., Staff, R.A., Kjeldsen, K.K., et al. 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. *Range and regularities in the distribution of earthworms of the USSR fauna.*

Porazinska, D.L., Giblin-Davis, R.M., Faller, L., Farmerie, W., Kanzaki, N., Morris, K., et al. 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.

Rasmussen, S.O., Bigler, M., Blockley, S.P., Blunier, T., Buchardt, S.L., Clausen, H.B., et al. 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., et al. 2009: Response to Comment by Goldberg et al. on “DNA from Pre-Clovis Human Coprolites in Oregon, North America.” *Science* **325**, 148 LP-148.

Rasmussen, L.M. and 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.

Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Ramsey, C.B., et al. 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. and 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. and Westcott, S.L. 2011: Reducing the effects of PCR amplification and sequencing Artifacts on 16s rRNA-based studies. *PLoS ONE* **6**.

Schnell, I.B., Bohmann, K. and 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., et al. 2016: DNA evidence of bowhead whale exploitation by Greenlandic Paleo-Inuit 4,000 years ago. *Nature Communications* **7**, 1–9.

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

Sipos, R., Székely, A.J., Palatinszky, M., Révész, S., Márialigeti, K. and 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., et al. 2017: Neandertal and Denisovan DNA from Pleistocene sediments. *Science* **356**, 605–608.

Solomina, O., Ivanov, M. and 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. and 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., et al. 2016: Deglaciation of Fennoscandia. *Quaternary Science Reviews* **147**, 91–121.

Stuiver, M., Reimer, P.J. and Reimer, R.W. 2018: CALIB 7.1 (http://calib.org/).

Sun, J., James, S.W., Jiang, J., Yao, B., Zhang, L., Liu, M., et al. 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., et al. 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., et al. 2018: Glacial and environmental changes the last 60,000 years in the Polar Ural Mountains, Arctic Russia, inferred from a high resolution lake record and observations from adjacent areas. *Boreas*.

Svendsen, J.I., Krüger, L.C., Mangerud, J., Astakhov, V.I., Paus, A., Nazarov, D., et al. 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. and Willerslev, E. 2012: Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular ecology* **21**, 2045–50.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., et al. 2007: Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic acids research* **35**, e14.

Taheri, S., Pelosi, C. and 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. and Saura, A. 2006: Dispersal and clonal diversity of North-European parthenogenetic earthworms. *Biological Invasions Belowground: Earthworms as Invasive Species* pp. 5–18. Springer.

Thomsen, P.F. and 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. and 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., et al. 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., et al. 2015: Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular ecology* **25**, 929–42.

Vistnes, I., Nellemann, C., Jordhøy, P. and 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., et al. 2003: Diverse plant and animal genetic records from holocene and pleistocene sediments. *Science* **300**, 791–795.

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

Wu, J.H., Hong, P.Y. and 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.

Yankson, K.K. and Steck, T.R. 2009: Strategy for extracting DNA from clay soil and detecting a specific target sequence via selective enrichment and real-time (quantitative) PCR amplification. *Applied and environmental microbiology* **75**, 6017–21.

Yannic, G., Pellissier, L., Ortego, J., Lecomte, N., Couturier, S., Cuyler, C., et al. 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., et al. 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.

**Figure legend**

Figure 1: The location of Lake Bolshoye Schuchye in the Polar Urals of Arctic Russia and Lake Uhca Rohči on the Varanger Peninsula, northeast Finnmark, Norway. The outer line represents the extend of the Eurasian ice sheet during the Marine Isotope Stage 2 (20,000-15,000 years BP). The inner white shaded area represents the ice sheet during the Younger Dryas (12,000 years BP).

Figure 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 lab 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 results for *Homo sapiens* is a combination of 7 different *H. sapiens* sequences for which the maximum number of repeats is plotted.

Figure 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 lab contaminants and were manually removed from the results. The results for *Homo sapiens* is a combination of 12 different *H. sapiens* sequences for which the maximum number of repeats is plotted.

Figure 4: Aggregated worm data from Lake Bolshoye Schuchye (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 Late Glacial Interstadial. HS and GS are respectively the Heinrich Stadial events and GS is the Greenland Stadials from Rasmussen et al. (2014).

**Table legend**

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.

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

**Supporting information**

Supplementary Table 1: Average mismatches between all clitellate families, species and the MamP007F – MamP007R mammal primers.

Supplementary Table 2: Metabarcoding results for the Polar Urals core, including the repeats, read abundances and sequence information.

Supplementary Table 3: Metabarcoding results for the Varanger core, including the repeats, read abundances and sequence information.