Title

PalaeoChip Arctic1.0: An optimised eDNA targeted enrichment approach to reconstructing past environments

Summary

Ancient environmental DNA has been established as a viable biomolecular proxy for tracking taxonomic presence through time in a local environment, even in the total absence of primary tissues. It is thought that sedimentary ancient DNA (sedaDNA) survives through mineral binding. And while these organo-mineral complexes likely facilitate long-term preservation, they also challenge our ability to release and isolate target molecules. Two limitations in sedaDNA extraction impede many palaeoenvironmental reconstructions: the post-extraction carryover of enzymatic inhibitors, and sedaDNA loss when attempting to reduce inhibitor co-elution. Here, we present an optimised eDNA targeted enrichment approach for reconstructing past environments. Our new extraction protocol with targeted enrichment averages a 14.6-fold increase in on-target plant and animal DNA compared to a commercial soil extraction kit, and a 22.6-fold increase compared to a PCR metabarcoding approach. To illustrate the effectiveness of the PalaeoChip Arctic1.0 protocol, we present results of plant and animal presence from permafrost samples and discuss new potential evidence for the late survival (ca. 9685 BP) of mammoth (*Mammuthus sp.*) and horse (*Equus sp.*) in the Klondike Region of Yukon, Canada. This approach translates to a more diverse and sensitive dataset with increased sequencing efficiency of ecologically informative sedaDNA.

Authors and Affiliations

Tyler J. Murchie[1,2*], Melanie Kuch[1,2], Ana Duggan[1,2], Marissa L. Ledger[3], Kévin Roche[4,5], Jennifer Klunk[1,6], Emil Karpinski[1,6], Dirk Hackenberger[1,7], Tara Sadoway[1,8], Ross MacPhee[9], Duane Froese[10], Hendrik Poinar[1,2,7*]

[1]McMaster Ancient DNA Centre, McMaster University, Canada.
[2]Department of Anthropology, McMaster University, Canada.
[3]Department of Archaeology, University of Cambridge, United Kingdom.
[4]University of Bourgogne Franche-Comté, France.
[5]University of Bordeaux Montaigne, France.
[6]Department of Biology, McMaster University, Canada.
[7]Department of Biochemistry, McMaster University, Canada.
[8]University Health Network, Toronto.
[9]Division of Vertebrate Zoology/Mammalogy, American Museum of Natural History.
[10]Earth and Atmospheric Sciences, University of Alberta.
[*]corresponding authors

Author contributions:

TJM: Research design, wet-lab processing, analysis, figure design, writing
MK: Research design, wet-lab processing, manuscript editing
AD: Bioinformatic assistance, analysis, manuscript editing
MLL: Laboratory assistance, writing, manuscript editing
KR: Laboratory assistance
JK: Laboratory and research design assistance
EK: Research design assistance, initial experiments with 4°C spin, manuscript editing
DH: Initial experiments with 4°C spin
TS: Metabarcoding and preliminary analyses of cores, core collection, manuscript editing
RM: Research design, manuscript editing, collaborative principal investigator
DF: Research design, analysis, manuscript editing, core collection, collaborative principal investigator
HP: Research design, analysis, manuscript editing, McMaster Ancient DNA Centre principal investigator

Introduction

Means of recovering and analyzing ecologically informative sedimentary ancient DNA (sedaDNA) have improved substantially thanks to ongoing developments in high-throughput sequencing (HTS) technologies (Taberlet et al., 2018). SedaDNA molecules have been successfully recovered to evaluate the 'local' (Parducci et al., 2017, p. 930; Rawlence et al., 2014, p. 616) diachronic presence of animals (Giguet-Covex et al., 2014; Graham et al., 2016; Haile et al., 2009; Pedersen et al., 2016; Slon et al., 2017), plants (Alsos et al., 2015; Anderson-Carpenter et al., 2011; Epp et al., 2015; Niemeyer et al., 2017; Willerslev et al., 2014), fungi (Bellemain et al., 2013), microbiota (Ahmed et al., 2018; D'Costa et al., 2011), and eukaryotic parasites (Søe et al., 2018) from a diverse range of depositional settings. It is thought that much of sedaDNA survives in the absence of primary tissues through the formation of organo-mineral complexes (Arnold et al., 2011; Blum et al., 1997; Gardner and Gunsch, 2017; Greaves and Wilson, 1970; Lorenz and Wackernagel, 1987a, 1987b; Morrissey et al., 2015; Ogram et al., 1988) as extracellular genetic material binds to common constituents of sediments such as humics (Crecchio and Stotzky, 1998), calcite (Cleaves et al., 2011), clays (Cai et al., 2006; Goring and Bartholomew, 1952; Greaves and Wilson, 1969), and silica (Bezanilla et al., 1995; Lorenz and Wackernagel, 1987a). Soil minerals have been found to stabilize a fraction of environmental DNA, allowing those molecules to resist decomposition (Morrissey et al., 2015), but strong mineral binding can also result in marginal sedaDNA release (Alvarez et al., 1998; Saeki et al., 2010). Extracellular mineral-bound sedaDNA is recovered in bulk in the form of disseminated biomolecules from a diverse range of organisms. This fact typically prohibits genomic reconstructions of single individuals, but it can allow for identifying the presence (and to a lesser extent, absence and potentially even relative abundance) of taxa at ecologically informative taxonomic ranks. The method shows the most promise in reconstructing palaeoflora (Anderson-Carpenter et al., 2011; Niemeyer et al., 2017; Pedersen et al., 2016; Sjögren et al., 2016; Willerslev et al., 2014), or the differential presence of a particular taxon through time (Graham et al., 2016). A single library can be used to identify sedaDNA from multiple domains simultaneously with shotgun sequencing or can be targeted to amplify or enrich for specific taxa of interest.

Despite rapid advances in ancient DNA (aDNA) techniques, two extraction related challenges persist that can limit the ability to fully utilize sedimentary genetic archives: 1) the

carryover of enzymatic inhibitors with techniques designed to maximize the recovery of aDNA characteristic molecules, and 2) the loss, due to overly vigorous inhibitor removal techniques, of ecologically informative sedaDNA that otherwise might be amenable to adapter ligation or amplification. To some degree PCR metabarcoding can mitigate inhibition through dilutions or additional purifications (McKee et al., 2015), with the addition of reagents such as bovine serum albumin (BSA) (Garland et al., 2010; Kreader, 1996), or with very high polymerase concentrations (Alsos et al., 2015). However, metabarcoding can be vulnerable to differential amplification rates due to variable molecular abundance per taxa, unequal damage, variability in metabarcode amplification efficiency, and PCR conditions (Bellemain et al., 2010; Kanagawa, 2003; Krehenwinkel et al., 2018; Nichols et al., 2018; Sze and Schloss, 2019). These factors compound downstream biases in taxonomic determinations, especially if there was substantial loss of low-abundance molecules during inhibitor removal from taxa with comparatively low biomass turnover (Yoccoz et al., 2012, p. 3651).

This study evaluates various inhibitor removal treatments for their ability to reduce the carryover of enzymatic inhibitors in sedaDNA extracts while maximally retaining endogenous palaeoenvironmental DNA that can successfully undergo library adapter ligation. Our aim is to minimize the need for excessive PCR amplification on purified eluates to mitigate the propagation of stochastic biases. Four previously studied (D'Costa et al., 2011; Mahony, 2015; Sadoway, 2014) open-air Yukon permafrost core exposures (Table 1, Figure 1) were chosen to experimentally optimize sedaDNA extraction (detailed in the supplementary materials). We based these modifications on our in-house lysis solution and high-volume binding buffer silica-spin column extraction method as per Dabney et al. (2013). Thereafter, an optimized protocol was selected to evaluate taxonomic assignments between shotgun sequenced and target enriched datasets with this extraction method, as compared with shotgun and enriched libraries extracted using the DNeasy PowerSoil DNA extraction kit (QIAGEN) following manufacturer specifications. For each of the four core sections, sediments were subsampled and homogenized, then split into three 250 mg replicates for both extraction methods. We also compared this sequence data with previously sequenced PCR metabarcoding data on the same core sections. Our sedaDNA modified Dabney et al. (2013) extraction protocol is described here (Figure 2, SET-E). Experimentation with various inhibition removal techniques are detailed in the supplementary materials as SET-A through SET-D.

Results

Quantitative PCRs (qPCR) on the adapted libraries show an up to 7.0-fold increase in total adapted DNA among the four core samples (average 3.6-fold increase) with our sedaDNA modified Dabney et al. (2013) extraction protocol, and an up to 5.6-fold increase in 'endogenous' *trnL* library adapted chloroplast DNA (average 3.0-fold increase) (see Figure 3). Inhibition indices for our sedaDNA modified Dabney extractions were lower than PowerSoil (average = 0.75 versus 0.95 for PowerSoil, see methods section 4 and Figure E14 for a description of the 'inhibition index'), but this low-level constituent of latent polymerase inhibitors did not impede enzymes for adapter ligation as these samples quantify much higher than PowerSoil extracts post-library prep. Our ongoing experiments with a diverse range of other sediments suggest that extracts with inhibition indices over ~0.3 are still amenable to library preparation, although potentially with reduced adapter ligation efficiency (see section SET-D in the supplementary online materials for a discussion of extract qPCR inhibition).

Mapped, *BLASTn* aligned, and LCA-assigned reads extracted with our modified Dabney protocol show an average 14.6-fold increase in bait on-target, map filtered reads over PowerSoil extractions, and a 22.6-fold increase in map-filtered reads over samples targeted with PCR metabarcoding (Table 1). This large fold increase in on-target molecules of total reads translates to a broader range of taxa identified, and a higher proportion of ecologically informative reads sequenced overall (Figures 4–7). Taxa with sufficiently high LCA-assigned read counts also show characteristic aDNA deamination patterns and fragment length distributions with *mapDamage* (Table E2; Figures E8–E12).

Discussion

Our novel 4°C centrifuge inhibitor removal procedure paired with Dabney et al. (2013) aDNA purifications and targeted enrichment consistently outperformed a sedaDNA commercial extraction kit across all extraction replicates, as well as outperforming a PCR metabarcoding approach. These results demonstrate the viability of targeted enrichment for taxonomically diverse environmental samples from open-air sites without the necessity of PCR metabarcoding and the associated compounding biases therein. These data also demonstrate the significantly improved efficiency of ecologically informative sequencing with RNA capture enrichment compared with a shotgun approach. Deep shotgun sequencing to library exhaustion would be ideal as it is the least taxonomically biasing approach. However, until data storage,

computational power, database completeness, and sequencing costs are improved, deep sequencing strategies are often unachievable for most users except for those with immense computational and sequencing resources.

Overcoming enzymatic inhibitors

Of interest for further research is the interaction between SDS and the 4°C spin for inhibitor precipitation (see section "SET-D. SDS and sarkosyl" in Appendix A). We suspect that the efficiency of inhibitor precipitation with this method could be further optimized as experiments suggest that the presence of SDS in the lysis buffer, which we hypothesize leads to the formation of micelles through constant agitation during the spin, significantly contributes to the precipitation of humics and other inhibitors at low temperatures. This technique is unlikely to be optimal for all forms of sedaDNA inhibition however, as it has been observed that identifying the specific inhibitory substances involved is critical to mitigating the compound specific mechanisms that affect enzymatic reactions (Opel et al., 2010). Further research, potentially with mass spectrometry, is needed to identify the inhibitor constituents of sedaDNA target samples in order to improve the inhibitor precipitation we observed while maximizing sedaDNA retention.

SedaDNA authenticity

Damage profiles for taxa with sufficiently high read counts ($\gtrsim$200 reads at minimum map quality 30) consistently show characteristic aDNA deamination patterns and short fragment length distributions. We observed when mapping to the mitogenome that taxa with $\lesssim$200 reads typically have insufficient mapping coverage to confidently identify damage patterns, making it difficult to authenticate rare taxa with low read counts in this dataset. The inflated number of reads mapping to specific taxa compared with the read counts that were *bwa* mapped to our curated baits, *BLASTn* aligned, and *MEGAN* LCA-assigned are suggestive that our quality filtering steps are sufficiently conservative to dramatically reduce the noise characteristic of metagenomic datasets (Eisenhofer et al., 2019; Lu and Salzberg, 2018), but may also strip out some potentially informative (but less confidently assigned) reads. Our pre-*BLASTn* map-filtering approach allows for a much more streamlined analysis with confidently LCA-assigned taxa in the mapped dataset and less confident LCA-assigned taxa in the unmapped reads.

Extraction and library preparation blanks do not contain map-filtered reads (Table 1). The unmapped LCA-assigned reads for these blanks are predominantly adapter contaminated

sequences (Figure E13). None of the ecologically informative taxa identified in the metagenomic comparisons appear in the blanks, suggesting patterns observed in our sediment samples are authentic and not the result of contamination.

Palaeoecology

This study is intended as a proof of concept to demonstrate the viability of targeted enrichment for complex environmental datasets. Additional ongoing research is intended to utilize these methods and complementary palaeoecological techniques on Yukon lake sediment and permafrost cores from the Klondike Plateau (in addition to the metagenomic data from this paper) to temporally track ecological change during the Pleistocene/Holocene transition in Eastern Beringia. However, it is worth briefly contextualizing these broad taxonomic trends here for authenticity purposes. The Bear Creek (BC 4-2B, 30,000 cal-BP [D'Costa et al., 2011]) and older Lucky Lady II core sections (LLII 12-217-8, 15,865 cal-BP [Mahony, 2015]) both date to a period in which Eastern Beringia is thought to have been largely a herb tundra biome, dominated by exposed mineral surfaces, prostrate willows, grasses, forbs (nongraminoid herbs), and occupied by diverse and abundant megafauna (Dyke, 2005; Mahony, 2015). Our data reflects this environmental setting, particularly in the case of the Bear Creek core which has been demonstrated previously to exhibit remarkable preservation (D'Costa et al., 2011). We identified a similar series of mammal species compared to previous work on the same core by D'Costa et al. (2011), but with additional taxa (e.g. caribou, *Rangifer tarandus*) and more specific taxonomic assignments (e.g. potentially yak, *Bos mutus*/*grunniens*). Results from the younger Lucky Lady II core section (LLII 12-84-3, 13,205 cal-BP [Sadoway, 2014]) indicate an expansion of Western Beringian birch shrub tundra (Dyke, 2005), reflected by a decrease in grasses and a proportional increase in *Betula* and *Salix*. Analysis of the most recent core sample (MM12-118b, 9,685 cal-BP [Mahony, 2015]), suggests that there had been a shift in the forest/tundra ecotone in the Yukon before the development of the boreal (taiga) forest, which first established in southern Yukon by ~9,000 cal-BP (Dyke, 2005). Our data shows a proportional increase in conifers, particularly spruce (*Picea*).

The mammalian constituents also display a marked change, dwindling moving forward in time into the Holocene (Figures 4–6), but perhaps less sharply than commonly thought. For example, we recovered genetic evidence of both woolly mammoth (*Mammuthus primigenius*) and horse (*Equus sp.*) in the Upper Goldbottom core dated to ~9,685 cal-BP (Mahony, 2015).

Previous radiocarbon analyses appeared to indicate that horses disappeared from high-latitude northwestern North America relatively early, ca. 12,500 ^{14}C BP ("last appearance date" 13,125 cal-BP, based on AMNH 134BX36 from Upper Cleary Creek [Guthrie, 2003]). This ~3,500 year difference implies the existence of a substantial *ghost range* (Haile et al., 2009) (i.e., spatiotemporal range extending beyond its last appearance age as indicated by directly dated materials) that cannot (yet) be corroborated by the macrofossil record for *Equus*, but consistent with previous ancient eDNA results from central Alaska (Haile et al., 2009). However, in the absence of additional information, it is difficult to assess whether this signal may be considered chronostratigraphically reliable, or may have been affected by factors such as leaching, cryo- or bioturbation, or reworking (redeposition) (Arnold et al., 2011), altering the relative positions of sequenced sedaDNA organomineral complexes. In the case of the mammoth reads, after merging the sequenced data from the three Upper Goldbottom core (MM12-118b) extraction replicates, coverage was insufficient (low read count) to reliably assess characteristic aDNA damage patterns (Figure E12), although there is arguably some indication of damage with the merged *mapDamage* and FLD profile. We hope that with subsequent sedaDNA analyses from this region (with geographic targets selected to minimize the probability of containing stratigraphically allochthonous sedaDNA) we will be more able to closely evaluate the temporal authenticity of the megafaunal *ghost ranges* hinted at here with this dataset.

Limitations of comparison

There are several caveats to keep in mind when assessing our comparison of protocols and the application PalaeoChip Arctic1.0. First, the lysing stage of our PowerSoil and Modified Dabney protocols were not equivalent in duration or reagents. We followed manufacturer specifications for PowerSoil, but the lysis stage of extraction with equivalent kits can be increased in duration and augmented with additional reagents to theoretically increase DNA yield (Niemeyer et al., 2017). Further, a recently released update to the PowerSoil kit, the DNeasy PowerSoil Pro, claims to have an 8-fold increase in DNA yield compared with comparative commercial kits (it is unclear what the fold increase over standard PowerSoil is with this updated kit). Our experiments with the PowerSoil inhibitor removal solution C3 found consistently low DNA retention compared with our longer duration 4°C spin as an inhibitor removal technique (see section SET-B in Appendix A). The PowerSoil inhibitor removal solution is effective at rapidly precipitating enzymatic inhibitors, but this study suggests that it is

overly aggressive and consistently precipitated viable sedaDNA in the process (see Figure S7 in Appendix A). We suspect that a longer lysis stage with PowerSoil would increase overall yields, but would not mitigate the substantial losses associated with overly aggressive humic precipitation when utilizing solution C3. We found that the 4°C spin is sufficiently effective at removing enzymatic inhibition to allow for successful adapter ligation, even if the extracts were not as inhibitor free as PowerSoil (Figure 3).

Second, metabarcoding is not directly equivalent to enrichment when comparing taxonomic coverage and LCA-assigned read counts. Metabarcoding is not limited by pre-synthesized baits, but rather by primer design (binding site conservation, database completeness, etc.). Mapping our data back to the baits does strip out potentially taxonomically informative hits *a priori*. To mitigate this, we have included a non-mapped comparative variant in Appendix A (see section SET-E). We observe that mapping to the curated baits (which have low complexity and non-diagnostic regions masked or removed) substantially reduces the number of low confidence (potential false-positive) spurious hits. Although, it is still worth looking through non-mapped data for positive, bait off-target assignments. This component highlights a limitation of enrichment overall, that one's taxonomic recovery is limited by the quality of the relevant bait-set, whereas PCR metabarcoded data is limited by biased PCR amplifications (low abundance taxa being swamped out by the over-amplification of high abundance or comparatively undamaged molecules) as well as primer sequence conservation (or lack thereof), and database completeness for design and alignment.

However, taxonomic alignments using amplicon sequence data may improve in the future as reference databases are updated. For example, *Sus scrofa* is not included in our targeted baits as members of this family (which includes peccaries or New World pigs) are not considered to have been present in the subarctic during the Pleistocene/Holocene transition (see Appendix B, note that flat-headed and collared peccary [*Pecari tajacu* and *Platygonus peccary* respectively] are included in the bait-set). However, Sadoway's (2014) metabarcoding approach was able to amplify suid reads in core BC 4-2B (Figure 7), which pass mapped-LCA filtering. Although some species of peccaries were able to live close to the ice margin in mid-continental North America during the late Pleistocene (Kurten and Anderson, 1980), there is no palaeontological evidence for their presence at that time in the Yukon (Stuart, 2015). In this case, one suspects that those metabarcoding reads are amplified contaminants. Regardless of specific instances like

this, it is reasonable to employ *a priori* limitations on a capture-based approach for filtering out taxa (either intentionally or unintentionally) prior to sequencing. The false-negative error potential of enrichment with sedaDNA can be mitigated with improved iterations of the bait-set (e.g. PalaeoChip Arctic2.0, Mediterranean1.5) as genetic reference databases are updated.

It should be noted that the PCR metabarcoding data utilized in this analysis was not extracted with our sedaDNA optimized strategy. It is likely that a metabarcoding approach may also be further improved by utilizing a 4°C inhibitor removal procedure to maximize sedaDNA retention. We acknowledge that this is an unfair limitation of our comparison, as the extraction methodology between our metabarcoding and enriched data are not equivalent. This topic is worthy of further research but lies outside the scope of our analysis here, which is intended to establish the viability of enrichment for sedaDNA contexts, and to report on a new inhibitor removal technique that may be further optimized. We intend to expand on PalaeoChip Arctic1.0 with further target sequences for regionally specific vegetation, mammals, insects, fungi, and microbiota. PalaeoChip is also intended to be optimized for other non-arctic regions.

Conclusion

We believe the experiments outlined in this report through the development of a Holarctic PalaeoChip, in this case Arctic1.0, clearly demonstrate the utility of our novel inhibitor removal technique paired with high volume Dabney et al. (2013) purifications for overcoming enzymatic inhibitors, which likewise enables the viability of targeted enrichment for sedaDNA analyses. This approach can simultaneously capture complex environmental DNA without excessive PCR amplifications (and the associated costs of using high concentrations of polymerase to overcome co-eluted enzymatic inhibitors). By increasingly iterating on the taxonomic breadth of our complex environmental baits, and by further optimizing enrichment and sedaDNA extraction conditions, this technique can continue to improve the sequenced fraction of on-target molecules without deep shotgun sequencing, or potentially biased PCR amplifications. We believe our extraction method and enrichment strategy are of relevance not only for ancient DNA research, but also potentially for modern environmental DNA monitoring applications where the DNA copy number is high and targeted capture is of even more relevance for improving the analytic and cost efficiency of high-throughput metagenomic sequencing.

Methods

All supplementary procedures, tables, and figures reported here with an 'S' prefix (e.g. Table S1, Figure S1) are included in the supplementary online materials in Appendix A.

Lab setting

Laboratory work was conducted in clean rooms at the McMaster Ancient DNA Centre, which are subdivided into dedicated facilities for sample preparation, stock solution setup, and DNA extraction through library preparation. Post-indexing and enrichment clean rooms are in a physically isolated facility, while high-copy PCR workspaces are in separate building with a one-way workflow progressing from low-copy to high-copy facilities. Each dedicated workspace is physically separated with air pressure gradients between rooms to reduce exogenous airborne contamination. Prior to all phases of laboratory work, dead air hoods and workspaces were cleaned using a 6% solution of sodium hypochlorite (commercial bleach) followed by a wash with Nanopure purified water (Barnstead) and 30 minutes of UV irradiation at >100 mJ/cm^2.

1. *Subsampling*

Metal sampling tools were cleaned with commercial bleach, rinsed with Nanopure water immediately thereafter (to reduce rusting), and heated overnight in an oven at ~130°C. Once the tools had cooled, work surfaces were cleaned with bleach and Nanopure water and covered with sterile lab-grade tin foil. Sediment cores previously split into disks (D'Costa et al., 2011, p. SI. 4–5; Sadoway, 2014, chap. 1) and stored at −20°C had the upper ~1 mm of external sediment chiselled off to create a fresh sampling area free of exogenous contaminants for a hollow cylindrical drill bit. The drill bit (diameter 0.5 cm) was immersed in liquid nitrogen prior to sampling and a drill press was used to repeatedly subsample the disk sections(D'Costa et al., 2011, fig. S3). Sediment was pushed out of the drill bit using a sterile nail and the bottom 1–2 mm of sediment from the bit was removed before dislodging the remaining sample. This exterior core portion was carefully removed as it has a higher chance of containing sedaDNA from other stratigraphic contexts due to coring and core splitting. A bulk set of subsampled sediment from the same core disk was homogenized by stirring in a 50 mL falcon tube and stored at −20°C for subsequent extractions. This process was repeated individually for each core.

2. Physical disruption, chemical lysis, and extraction

DNeasy PowerSoil DNA Extraction kit samples were extracted following manufacturer specifications; purified DNA was eluted twice with 25 µL EBT buffer (10 mM Tris-Cl, 0.05% Tween-20). Samples extracted with the sedaDNA modified Dabney et al. (2013) procedure were processed as follows (see Figure 2):

Lysis

1) 500 µL of a digestion solution (see Table S1) initially without proteinase K was added to PowerBead tubes (already containing garnet beads and 750 µl 181 mM NaPO4 and 121 mM guanidinium isothiocyanate).
2) 250 mg of homogenized sediment was added to each PowerBead tube.
3) PowerBead tubes were vortexed at high speed for 15 minutes, then centrifuged briefly to remove liquid from the lids.
4) 15.63 µL of proteinase K (stock 20 mg/mL) was added to each tube to reach a proteinase K concentration of 0.25 mg/mL in the digestion and PowerBead solution (1.25 mL).
5) Tubes were finger vortexed to disrupt sediment and beads that had pelleted in step 3.
6) PowerBead tubes were securely fixed in a hybridization oven set to 35°C and rotated overnight for ~19 hours, ensuring that the digestion solution, sediment, and PowerBeads were moving with each oscillation.
7) PowerBead tubes were removed from the oven and centrifuged at 10,000 x g for 5 minutes (the maximum speed recommended for PowerBead tubes). Supernatant was pipetted into a MAXYMum Recovery 2 mL tube and stored at −20°C.

Purification

8) Digestion supernatant was thawed, briefly centrifuged, pipetted into 16.25 mL (13 volumes) of high-volume Dabney binding buffer (see Table S2) in a 50 mL falcon tube and mixed.
9) Falcon tubes were spun at 4500 rpm in a refrigerated centrifuge set to 4°C for 20 hours overnight.
10) After centrifugation, falcon tubes were carefully removed and the supernatant was decanted, taking care to not disturb the darkly coloured pellet that had formed during the cold spin at the base of the tube.
11) The binding buffer was passed through a high-volume silica-column (High Pure Extender Assembly, Roche Diagnostics) and extraction proceeded as per Dabney et al. (2013).
12) Purified DNA was eluted off the silica columns twice with 25 µL EBT.

Prior to all subsequent experiments, both the modified Dabney and PowerSoil extracts were centrifuged at 16,000 x g for 5 minutes to pellet remaining co-eluted inhibitors.

3. Library preparation

Doubled stranded libraries were prepared as described in Meyer and Kircher (2010) with modifications from Kircher et al. (2012) and a modified end-repair reaction to account for the

lack of uracil excision (Table S3). Samples were purified after blunt-end repair with a QIAquick PCR Purification Kit (QIAGEN) (to maximally retain small fragments) and after adapter ligation (Table S4) with a MinElute PCR Purification Kit (QIAGEN), both using manufacturers protocols. Library preparation master mix concentrations can be found in Tables S3-S6.

4. QPCR: Inhibition spike tests, total quantification, and indexing

A positive control spike qPCR assay (Enk et al., 2016; King et al., 2009) was used to assess the relative impact of DNA independent inhibitors (co-eluted substances such as humics that inhibit enzyme function) on the enzymatic amplification efficiency of a spiked amplicon in the presence of template sedaDNA derived from variable lysing and extraction methods (Table S7). We suspected that enzymes in library preparation would be inhibited similarly to AmpliTaq Gold polymerase in qPCR. Shifts in the qPCR amplification slope of our spiked oligo with AmpliTaq Gold (due to co-eluted inhibitors in sedaDNA extracts) could then be quantified and used to infer the likelihood of failed adapter ligation due to enzymatic inhibitors (rather than a lack of sedaDNA). Admittedly, AmpliTaq Gold is not a 1:1 stand-in for inhibition sensitivity during blunt-end repair and adapter ligation, as AmpliTaq is among the most sensitive polymerases to inhibition induced reductions in amplification efficiency (Al-Soud and Radstrom, 1998), and due to qPCR specific inhibition such as the reduction in florescence despite successful amplification (Sidstedt et al., 2015). Our experiments do suggest that these enzymes have a very roughly commensurate inhibition sensitivity, insofar as eluates completely inhibited during this spike test are unlikely to successfully undergo library adapter ligation.

To quantify the co-eluted inhibition affecting each spiked amplification, we compared the qPCR slope of an oligo-spiked sedaDNA extract (1 µL of sample eluate spiked with 1 µL of a 49-bp oligo [1000 copies {E^3}], see Table S7) with the qPCR slope of 1 µL E^3 oligo standard in 1 µL of EBT. Average C_q and max relative fluorescence units (RFU) for each PCR replicate were calculated, as was the hill slope of the amplification curve by fitting a variable-slope sigmoidal dose-response curve to the raw fluorescence data using GraphPad Prism v. 7.04 (based on King et al. [2009]). The E^3 oligo-spiked averages (C_q, RFU, and sigmoidal hillslope) were divided by the corresponding E^3 oligo standard amplification value, then averaged together to generate an 'inhibition index' per PCR replicate, which were averaged again across PCR replicates to determine an extract's inhibition index, ranging from 0–1. In this case, 0 indicates a completely inhibited reaction (no measurable increase in RFU), and 1 indicates a completely

uninhibited reaction relative to the spiked E^3 oligo-standard (see Figure E14). Anything above 0.9 (the bottom range for blanks and standards of differing starting quantities) is considered essentially uninhibited insofar as *Taq* polymerase inhibition is concerned.

Most total DNA quantifications in SET-A to SET-D_2 (detailed in the supplementary materials) used the short amplification primer sites on the library adapters and were compared against the same library prepared 49-bp oligo standard used in the spike tests (see Table S8). This assay was also modified in some instances to quantify the 'endogenous' chloroplast constituent of adapted molecules by pairing the *trn*L P6-loop forward primer-g (Taberlet et al., 2007) with the reverse P7R library adapter primer (IS8, see Table S9). Enk et al. (2013) demonstrated that a single-locus qPCR assay can be used to predict on-target ancient DNA high-throughput sequencing read counts. Previous analyses (D'Costa et al., 2011; Sadoway, 2014) indicated that ancient vegetation was the most consistently abundant fraction of the biomolecules in these cores, and as such could serve as a rough proxy for assessing aDNA retention for successfully library adapted molecules between various inhibitor removal strategies. For all qPCR results reported here, standard curve metrics are included in the associated captions. Ideal standard curve values are: $R^2 = 1$, slope = -3.3 (or between -3.1 and -3.5), efficiency = 90–105%.

The extraction triplicate with the highest DNA concentrations for each of the four cores and two extraction methods (as based on the short amplification qPCR) was indexed for shotgun sequencing (8 samples + 3 blanks). All extraction triplicates (24 samples + 3 blanks) were indexed separately thereafter for targeted enrichment.

5. Enrichment: PalaeoChip, Arctic1.0

The PalaeoChip, Arctic1.0 RNA hybridization bait-sets were designed to target whole mtDNA of extinct and extant Quaternary animals (focused primarily on megafauna; number of taxa ≈ 180), and high latitude plant cpDNA based on curated reference databases developed by Sønstebø et al. (2010), Soininen et al. (2015), and Willerslev et al. (2014), initially targeting *trn*L (n ≈ 1650 taxa) (see Appendix B for taxonomic list). This list was queried with the *NCBI Mass Sequence Downloader* software (Pina-Martins and Paulo, 2015) to recover additional nucleotide data from GenBank (Benson et al., 2018) for *trn*L, as well as adding targets for *mat*K and *rbc*L. These three regions were selected as they are among the most sequenced and taxonomically informative portions of the chloroplast genome (Hollingsworth et al., 2011). Baits were designed

in collaboration with Arbor Biosciences to 80 bp with ~3x flexible tiling density, clustered with >96% identity and >83% overlap, and baits were removed with >25% soft-masking (to reduce low complexity baits with a high chance of being off-target in complex environmental samples). Bait sequences were queried with *BLASTn* against the NCBI database on a local computer cluster using a July 2018 database, then inspected in *MEGAN* (Huson et al., 2016, 2007). Baits with a mismatched taxonomic target and *BLASTn* alignment were queried again using a web-blast script (Camacho et al., 2009; NCBI Resource Coordinators, 2018) to determine if these mismatches were due to local database incongruities. Mismatches were again extracted with *MEGAN*, individually inspected, then removed from the bait-set if determined to be insufficiently specific.

Hybridization and bait mixes were prepared to the concentrations in Table S11. For each library, 7 µL of template was combined with 2.95µL of Bloligos (blocking oligios which prevent the hybridization between library adapter sequences). The hybridization and bait mixes were combined and pre-warmed to 60°C, before being combined with the library-Bloligo mixture. The final reaction was incubated for 24 hours at 55°C for bait-library hybridization.

The next day, beads were dispensed (540 µL total between two tubes), washed three times with 200 µL of binding buffer for each tube, then suspension in 270 µL of binding buffer per tube and aliquoted into PCR strips. Baits were captured using 20 µL of the bead suspension per library, incubated at 55°C for 2.5 minutes, finger vortexed and spun down, and incubated for another 2.5 minutes. Beads were pelleted and the supernatant (the non-captured library fraction) was removed and stored at −20°C. The beads were resuspended in 180 µL of 55°C Wash Buffer X and washed four times following the MYbaits V4 protocol. Beads were eluted in 15 µL EBT, PCR reamplified (Table S6), then purified over MinElute following manufacturer's protocols in 15 µL EBT.

6. Post-indexing total quantification, pooling, size-selection, and sequencing

Enriched and shotgun samples were quantified using the long-amplification total library qPCR assay (Table S10). Enriched and shotgun libraries were equal-molar pooled separately. The two pools were size-selected with gel excision following electrophoresis for molecules ranging from 150 bp to 600 bp. Gel plugs were purified using the QIAquick Gel Extraction Kit (QIAGEN), according to manufacturer's protocol, then sequenced on an Illumina HiSeq 1500

with a 2 x 90 bp paired-end protocol at the Farncombe Metagenomics Facility (McMaster University, ON).

7. PCR Metabarcoding

Sadoway (2014) previously worked with these and many other open-air Yukon permafrost core samples using a metabarcoding approach. These libraries had been extracted in duplicate with guanidinium protocols (Boom et al., 1990; D'Costa et al., 2011) from 250 mg of the same core sections, purified with silica (Höss and Pääbo, 1993), and eluted twice (Handt et al., 1996). Extensive inhibition at the time, as detected using similar qPCR spike tests developed by King et al. (2009), necessitated a tenfold extract dilution, which were then amplified in duplicate for each primer set, targeting: *rbc*L (CBOL Plant Working Group, 2009; Hollingsworth, 2011; Willerslev et al., 2003), *trnL* (Taberlet et al., 2007), *16S rRNA* (Höss et al., 1996), *and 12S rRNA* (Kuch et al., 2002), each following cited PCR conditions. The locus *cytochrome b* (*cyt-b*) was also targeted using a set of degenerate primers designed with FastPCR (Kalendar et al., 2011; Sadoway, 2014). *Cyt-b* amplifications were found to be most efficient in 20 µL reactions using AmpliTaq Gold (0.05U/µL), 1X PCR Buffer II, 2.5 mM MgCl2, 0.25 mM dNTPs, 0.5X Evagreen, 250 nM (forward/reverse primers) when cycled with a 3 minute denaturation at 95°C, 45 cycles of 95°C for 30 seconds, and 60°C for 30 seconds (Sadoway, 2014). QPCR products were purified with 10K AcroPrep Pall plates (Pall Canada Direct Ltd., Mississauga, ON, Canada) using a vacuum manifold. QPCR assays were used to pool each amplicon set in equimolar concentrations, which were library prepared and dual-index following the same Illumina protocols as described above (Kircher et al., 2012; Meyer and Kircher, 2010). Each sample extract had its own unique combination of forward and reverse indices, which were sequenced on a HiSeq 1500 Rapid Run (2 x 100bp, Illumina Cambridge Ltd, Essex, UK) at the Farncombe Metagenomics Facility (McMaster University, ON) to approximately 100,000 reads each. These PCR metabarcoded libraries were processed for this paper with an identical bioinformatic pipeline as described below.

8. Bioinformatics

Reads from all library sets (enriched, shotgunned, and PCR metabarcoded) were demultiplexed with *bcl2fastq* (v 1.8.4), converted to bam files with *fastq2bam* (https://github.com/grenaud/BCL2BAM2FASTQ), then trimmed and merged with using *leeHom* (Renaud et al., 2014) using ancient DNA specific parameters (--ancientdna). Reads were then

aligned to a concatenated reference of the animal and plant bait-set with *network-aware-BWA* (Li and Durbin, 2009) (https://github.com/mpieva/network-aware-bwa) with a maximum edit distance of 0.01 (-n 0.01), allowing for a maximum two gap openings (-o 2), and with seeding effectively disabled (-l 16500). Mapped reads that were either merged or unmerged but properly paired were extracted with *libbam* (https://github.com/grenaud/libbam), collapsed based on unique 5' and 3' positions with *biohazard* (https://bitbucket.org/ustenzel/biohazard), and restricted to a minimum length of 24bp. Both mapped and non-map filtered reads were string deduplicated using the *NGSXRemoveDuplicates* module of *NGSeXplore* (https://github.com/ktmeaton/NGSeXplore), then queried with *BLASTn* to return the top 100 alignments (-num_alignments 100 -max_hsps 1) against a July 2018 version of the NCBI Nucleotide database on a local computer cluster. Non-map filtered libraries were treated identically, although only returned the top 10 alignments to mitigate unwieldy (>10 gb) file sizes. Sequencing summary counts are in Table 1.

Blast and fasta files for each sample (unmapped and mapped variants) were passed to *MEGAN* (Huson et al., 2016, 2007) using the following LCA parameters: min-score = 50 (default), max expected (e-value) = 1.0E-5, minimum percent identity = 95% (allows 1 base mismatch at 24 bp, 2 at 50 bp, and 3 at 60 bp to account for cytosine deamination and other aDNA characteristic damage or sequencing errors), top percent consideration of hits based on bit-score = 15% (allows for slightly more conservative taxonomic assignments than the 10% default based on trial and error), minimum read support = 3 (number of unique reads aligning to an NCBI sequence for that taxon to be considered for LCA), minimum complexity = 0.3 (default minimum complexity filter), and utilizing the LCA weighted algorithm at 80% (two rounds of analysis that purportedly increases specificity but doubles run time over the native algorithm). Metagenomic profiles were compared in *MEGAN* using absolute and normalized read counts (normalized in *MEGAN* to the lowest read count, keeping at least 1 read per taxon). Libraries were not subsampled to an equal depth prior to processing; McMurdie and Holmes (2014) have demonstrated that this rarefying approach is the most ineffective means of accounting for unequally sequenced metagenomic data. However, in an attempt to fairly account for biases inherent in comparing capture enriched and amplicon data of unequal sequencing depths, we have included four versions of the bubble chart taxonomic comparisons per core: 1) absolute counts mapped filtered to our baits (in the main text) with total reads and proportional LCA-

assigned totals at the bottom of each column, 2) normalized counts mapped to baits, 3) absolute counts for all reads (not map filtered), and 4) normalized counts for all reads (not map filtered). The latter three comparisons are included in Appendix A.

To visualize the taxonomic variability between these replicates, comparative trees in *MEGAN* were uncollapsed to the 'order' rank (meaning that all lower taxonomic assignments within that 'order' are summed); animalia was then fully uncollapsed (as the read counts were more manageable compared with plant assignments). Viridiplantae clades were collapsed to higher ranks (higher than 'order') in some cases for summarized visualizations (otherwise there were too many leaves to display at once in a single figure, even if when only showing summaries by 'order'). Thereafter, all leaves were selected and visualized with logarithmically scaled bubble charts; additional higher LCA-assigned animalia ranks were also selected where taxonomically informative (for example, reads that could only be conservatively LCA-assigned to Elephantidae or *Mammuthus sp.*, but which in this context likely represent hits to *Mammuthus primigenius* [woolly mammoth]). Low abundance (<3 reads), non-informative and non-target clades (e.g. bacteria, fungi, or LCA-assignments to high ranks) were excluded for visualization purposes. Reads within viridiplantae frequently hit low taxonomic ranks (family, genus, and occasionally species) but this resolution is not shown here to facilitate our goal of an inter-method comparison.

Taxa with high blast and LCA-assigned read-counts were also selected to evaluate damage patterns and fragment length distributions (FLD) (Table E2). Enriched libraries were mapped to reference genomes of either the LCA-assigned organism itself (e.g. *Mammuthus primigenius*) or a phylogenetically closely related organism (e.g. *Equus caballus*) if there was no species call or if a reliable reference genome does not yet exist. Mapping followed the aforementioned parameters and software, with an additional map-quality filter to ≥30 with *samtools* (https://github.com/samtools/samtools) and passed to *mapDamage* (Jónsson et al., 2013) (v 2.0.3, https://ginolhac.github.io/mapDamage/). Plant chloroplast DNA references were reduced to the target barcoding loci (*trn*L, *rbc*L, and *mat*K), each separated by 100 Ns. Mitochondrial reference genomes were used for animal taxa of interest.

552 **Acknowledgements**

553 We wish to thank the Arctic Institute of North America, the Canadian Institutes of Health
554 Research, the Garfield Weston Foundation, the Natural Sciences and Engineering Research
555 Council of Canada, McMaster University, Polar Knowledge Canada (POLAR), and the Social
556 Sciences and Humanities Research Council of Canada for each funding various components of
557 this research.

558

559

560

561

562 **References**

563 Ahmed, E., Parducci, L., Unneberg, P., Ågren, R., Schenk, F., Rattray, J.E., Han, L.,
564 Muschitiello, F., Pedersen, M.W., Smittenberg, R.H., Yamoah, K.A., Slotte, T., Wohlfarth,
565 B., 2018. Archaeal community changes in Lateglacial lake sediments: Evidence from
566 ancient DNA. Quat. Sci. Rev. 181, 19–29. doi:10.1016/j.quascirev.2017.11.037

567 Al-Soud, W.A., Radstrom, P., 1998. Capacity of nine thermostable DNA polymerases to mediate
568 DNA amplification in the presence of PCR-inhibiting samples. Appl. Environ. Microbiol.
569 64, 3748–3753.

570 Alsos, I.G., Sjo gren, P., Edwards, M.E., Landvik, J.Y., Gielly, L., Forwick, M., Coissac, E.,
571 Brown, A.G., Jakobsen, L. V., Foreid, M.K., Pedersen, M.W., 2015. Sedimentary ancient
572 DNA from Lake Skartjorna, Svalbard: Assessing the resilience of arctic flora to Holocene
573 climate change. The Holocene. doi:10.1177/0959683615612563

574 Alvarez, A.J., Khanna, M., Toranzos, G.A., Stotzky, G., 1998. Amplification of DNA bound on
575 clay minerals. Mol. Ecol. 7, 775–778.

576 Anderson-Carpenter, L.L., McLachlan, J.S., Jackson, S.T., Kuch, M., Lumibao, C.Y., Poinar,
577 H.N., 2011. Ancient DNA from lake sediments: bridging the gap between paleoecology and
578 genetics. BMC Evol. Biol. 11, 1–15. doi:10.1186/1471-2148-11-30

579 Arnold, L.J., Roberts, R.G., Macphee, R.D.E., Haile, J.S., Brock, F., M??ller, P., Froese, D.G.,
580 Tikhonov, A.N., Chivas, A.R., Gilbert, M.T.P., Willerslev, E., 2011. Paper II - Dirt, dates
581 and DNA: OSL and radiocarbon chronologies of perennially frozen sediments in Siberia,
582 and their implications for sedimentary ancient DNA studies. Boreas 40, 417–445.
583 doi:10.1111/j.1502-3885.2010.00181.x

584 Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., Kauserud, H., 2010. ITS as
585 an environmental DNA barcode for fungi: an in silico approach reveals potential PCR
586 biases. BMC Microbiol. 10, 189. doi:10.1186/1471-2180-10-189

587 Bellemain, E., Davey, M.L., Kauserud, H., Epp, L.S., Boessenkool, S., Coissac, E., Geml, J.,
588 Edwards, M., Willerslev, E., Gussarova, G., Taberlet, P., Haile, J., Brochmann, C., 2013.
589 Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from
590 arctic permafrost. Environ. Microbiol. 15, 1176–89. doi:10.1111/1462-2920.12020

591 Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Ostell, J., Pruitt, K.D., Sayers,
592 E.W., 2018. GenBank. Nucleic Acids Res. 46, D41–D47. doi:10.1093/nar/gkx1094

593 Bezanilla, M., Manne, S., Laney, D.E., Lyubchenko, Y.L., Hansma, H.G., 1995. Adsorption of
594 DNA to Mica, Silylated Mica, and Minerals: Characterization by Atomic Force
595 Microscopy. Langmuir 11, 655–659. doi:10.1021/la00002a050

596 Blum, S.A.E., Lorenz, M.G., Wackernagel, W., 1997. Mechanism of retarded DNA degradation
597 and prokaryotic origin of DNases in nonsterile soils. Syst. Appl. Microbiol. 20, 513–521.
598 doi:10.1016/S0723-2020(97)80021-5

599 Boom, R., Sol, C.J., Salimans, M.M., Jansen, C.L., Wertheim-Van Dillen, P.M., van der
600 Noordaa, J., 1990. Rapid and simple method for purification of nucleic acids.
601 J.Clin.Microbiol. 28, 495–503. doi:10.1556/AMicr.58.2011.1.7

602 Cai, P., Huang, Q., Zhang, X., Chen, H., 2006. Adsorption of DNA on clay minerals and various
603 colloidal particles from an Alfisol. Soil Biol. Biochem. 38, 471–476.
604 doi:10.1016/j.soilbio.2005.05.019

605 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L.,

606 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10, 1–9.
607 doi:10.1186/1471-2105-10-421
608 CBOL Plant Working Group, 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. U. S.
609 A. 106, 12794–7. doi:10.1073/pnas.0905845106
610 Clark, P.. U., Mix, A.. C., 2002. Ice sheets and sea level of the Last Glacial Maximum. Quat. Sci.
611 Rev. 21, 1–7.
612 Clark, P.U., 2009. The Last Glacial Maximum. Science (80-.). 325, 710–714.
613 doi:10.1126/science.1172873
614 Cleaves, H.J., Crapster-Pregont, E., Jonsson, C.M., Jonsson, C.L., Sverjensky, D.A., Hazen,
615 R.A., 2011. The adsorption of short single-stranded DNA oligomers to mineral surfaces.
616 Chemosphere 83, 1560–1567. doi:10.1016/j.chemosphere.2011.01.023
617 Crecchio, C., Stotzky, G., 1998. Binding of DNA on humic acids: Effect of transformation of
618 Bacillus subtilis and resistance to DNase. Soil Biol. Biochem. 30, 1061–1067.
619 D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W.L., Schwarz, C., Froese, D.,
620 Zazula, G., Calmels, F., Debruyne, R., Golding, G.B., Poinar, H.N., Wright, G.D., 2011.
621 Antibiotic resistance is ancient. Nature 477, 457–461. doi:10.1038/nature10388
622 Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C.,
623 Garcia, N., Paabo, S., Arsuaga, J.-L., Meyer, M., 2013. Complete mitochondrial genome
624 sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments.
625 Proc. Natl. Acad. Sci. 110, 15758–15763. doi:10.1073/pnas.1314445110
626 Dyke, A.S., 2005. Late Quaternary Vegetation History of Northern North America Based on
627 Pollen, Macrofossil, and Faunal Remains. Géographie Phys. Quat. 59, 211.
628 doi:10.7202/014755ar
629 Dyke, A.S., 2004. An outline of the deglaciation of North America with emphasis on central and
630 northern Canada. Dev. Quat. Sci. 2, 373–424.
631 Ehlers, J., Gibbard, P.L., Hughes, P.D. (Eds.), 2011. Quaternary Glaciations-Extent and
632 Chronology: 1st Edition. Elsevier.
633 Eisenhofer, R., Minich, J.J., Marotz, C., Cooper, A., Knight, R., Weyrich, L.S., 2019.
634 Contamination in Low Microbial Biomass Microbiome Studies: Issues and
635 Recommendations. Trends Microbiol. 27, 105–117. doi:10.1016/j.tim.2018.11.003
636 Enk, J., Devault, A., Widga, C., Saunders, J., Szpak, P., Southon, J., Rouillard, J.-M., Shapiro,
637 B., Golding, G.B., Zazula, G., Froese, D., Fisher, D.C., Macphee, R.D.E., Poinar, H., 2016.
638 Mammuthus population dynamics in Late Pleistocene North America: Divergence,
639 Phylogeogrpaphy and Introgression. Front. Ecol. Evol. 4, 1–13.
640 doi:10.3389/fevo.2016.00042
641 Enk, J., Rouillard, J.M., Poinar, H., 2013. Quantitative PCR as a predictor of aligned ancient
642 DNA read counts following targeted enrichment. Biotechniques 55, 300–309.
643 doi:10.2144/000114114
644 Epp, L.S., Gussarova, G., Boessenkool, S., Olsen, J., Haile, J., Schr??der-Nielsen, A., Ludikova,
645 A., Hassel, K., Sten??ien, H.K., Funder, S., Willerslev, E., Kj??r, K., Brochmann, C., 2015.
646 Lake sediment multi-taxon DNA from North Greenland records early post-glacial
647 appearance of vascular plants and accurately tracks environmental changes. Quat. Sci. Rev.
648 117, 152–163. doi:10.1016/j.quascirev.2015.03.027
649 Gardner, C.M., Gunsch, C.K., 2017. Adsorption capacity of multiple DNA sources to clay

650 minerals and environmental soil matrices less than previously estimated. Chemosphere 175,
651 45–51. doi:10.1016/j.chemosphere.2017.02.030
652 Garland, S., Baker, A., Phillott, A.D., Skerratt, L.F., 2010. BSA reduces inhibition in a
653 TaqMan® assay for the detection of Batrachochytrium dendrobatidis. Dis. Aquat. Organ.
654 92, 113–116. doi:10.3354/dao02053
655 Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P.-J., Griggo, C., Gielly, L., Domaizon, I.,
656 Coissac, E., David, F., Choler, P., Poulenard, J., Taberlet, P., 2014. Long livestock farming
657 history and human landscape shaping revealed by lake sediment DNA. Nat. Commun. 5,
658 3211. doi:10.1038/ncomms4211
659 Goring, C.A.I., Bartholomew, W.V., 1952. Adsorption of mononucleotides, nucleic acids, and
660 nucleoproteins by clays. Soil Sci. doi:10.1097/00010694-195208000-00005
661 Graham, R.W., Belmecheri, S., Choy, K., Culleton, B.J., Davies, L.J., Froese, D., Heintzman,
662 P.D., Hritz, C., Kapp, J.D., Newsom, L.A., Rawcliffe, R., Saulnier-Talbot, É., Shapiro, B.,
663 Wang, Y., Williams, J.W., Wooller, M.J., 2016. Timing and causes of mid-Holocene
664 mammoth extinction on St. Paul Island, Alaska. Proc. Natl. Acad. Sci. 113, 9310–9314.
665 doi:10.1073/pnas.1604903113
666 Greaves, M.P., Wilson, M.J., 1970. The degradation of nucleic acids and montmorillonite-
667 nucleic-acid complexes by soil microorganisms. Soil Biol. Biochem. 2, 257–268.
668 doi:10.1016/0038-0717(70)90032-5
669 Greaves, M.P., Wilson, M.J., 1969. The adsorption of nucleic acids by montmorillonite. Soil
670 Biol. Biochem. 1, 317–323. doi:10.1016/0038-0717(69)90014-5
671 Guthrie, R.D., 2003. Rapid body size decline in Alaskan Pleistocene horses before extinction.
672 Nature 426, 169–171.
673 Haile, J., Froese, D.G., Macphee, R.D.E., Roberts, R.G., Arnold, L.J., Reyes, A. V, Rasmussen,
674 M., Nielsen, R., Brook, B.W., Robinson, S., Demuro, M., Gilbert, M.T.P., Munch, K.,
675 Austin, J.J., Cooper, A., Barnes, I., Möller, P., Willerslev, E., 2009. Ancient DNA reveals
676 late survival of mammoth and horse in interior Alaska. Proc. Natl. Acad. Sci. U. S. A. 106,
677 22352–22357. doi:10.1073/pnas.0912510106
678 Handt, O., Krings, M., Ward, R.H., Pääbo, S., 1996. The retrieval of ancient human DNA
679 sequences. Am. J. Hum. Genet. 59, 368–76.
680 Hollingsworth, P.M., 2011. Refining the DNA barcode for land plants. Proc. Natl. Acad. Sci.
681 108, 19451–19452. doi:10.1073/pnas.1116812108
682 Hollingsworth, P.M., Graham, S.W., Little, D.P., 2011. Choosing and using a plant DNA
683 barcode. PLoS One 6. doi:10.1371/journal.pone.0019254
684 Höss, M., Dilling, A., Currant, A., Pääbo, S., 1996. Molecular phylogeny of the extinct ground
685 sloth Mylodon darwinii. Proc. Natl. Acad. Sci. 93, 181–185. doi:10.1073/pnas.93.1.181
686 Höss, M., Pääbo, S., 1993. DNA extraction from Pleistocene bones by purification method.
687 Nucleic Acids Res. 21, 3913–3914.
688 Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data.
689 Genome Res. 17, 377–386. doi:10.1101/gr.5969107
690 Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., Tappu,
691 R., 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-
692 Scale Microbiome Sequencing Data. PLoS Comput. Biol. 12.
693 doi:10.1371/journal.pcbi.1004957

694 Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. MapDamage2.0:
695 Fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics
696 29, 1682–1684. doi:10.1093/bioinformatics/btt193
697 Kalendar, R., Lee, D., Schulman, A.H., 2011. Java web tools for PCR, in silico PCR, and
698 oligonucleotide assembly and analysis. Genomics 98, 137–144.
699 doi:10.1016/j.ygeno.2011.04.009
700 Kanagawa, T., 2003. Bias and artifacts in multitemplate polymerase chain reactions (PCR). J.
701 Biosci. Bioeng. 96, 317–323. doi:http://dx.doi.org/10.1016/S1389-1723(03)90130-7
702 King, C.E., Debruyne, R., Kuch, M., Schwarz, C., Poinar, H.N., 2009. A quantitative approach to
703 detect and overcome PCR inhibition in ancient DNA extracts. Biotechniques 47, 941–949.
704 doi:10.2144/000113244
705 Kircher, M., Sawyer, S., Meyer, M., 2012. Double indexing overcomes inaccuracies in multiplex
706 sequencing on the Illumina platform. Nucleic Acids Res. 40, 1–8. doi:10.1093/nar/gkr771
707 Kreader, C.A., 1996. Relief of Amplification Inhibition in PCR with Bovine Serum Albumin or
708 T4 Gene 32 Protein. Appl. Environ. Microbiol. 62, 1102–1106.
709 Krehenwinkel, H., Fong, M., Kennedy, S., Huang, E.G., Noriyuki, S., Cayetano, L., Gillespie,
710 R., 2018. The effect of DNA degradation bias in passive sampling devices on
711 metabarcoding studies of arthropod communities and their associated microbiota. PLoS One
712 13, 1–14. doi:10.1371/journal.pone.0189188
713 Kuch, M., Rohland, N., Betancourt, J.L., Latorre, C., Steppan, S., Poinar, H.N., 2002. Molecular
714 analysis of a 11 700-year-old rodent midden from the Atacama Desert, Chile. Mol. Ecol. 11,
715 913–924. doi:10.1046/j.1365-294X.2002.01492.x
716 Kurten, B., Anderson, E., 1980. Pleistocene Mammals of North America. Columbia University
717 Press.
718 Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler
719 transform. Bioinformatics 25, 1754–60. doi:10.1093/bioinformatics/btp324
720 Lorenz, M.G., Wackernagel, W., 1987a. Adsorption of DNA to sand and variable degradation
721 rates of adsorbed DNA. Appl. Environ. Microbiol. 53, 2948–2952.
722 Lorenz, M.G., Wackernagel, W., 1987b. Adsorption of DNA to sand and variable degradation of
723 adsorbed DNA. Appl. Environ. Microbiol. 53, 2948–2952.
724 Lu, J., Salzberg, S.L., 2018. Removing Contaminants from Metagenomic Databases. PLoS
725 Comput. Biol. 261859. doi:10.1101/261859
726 Mahony, M.E., 2015. 50,000 years of paleoenvironmental change recorded in meteoric waters
727 and coeval paleoecological and cryostratigraphic indicators from the Klondike goldfields,
728 Yukon, Canada. University of Alberta. doi:10.1145/3132847.3132886
729 McKee, A.M., Spear, S.F., Pierson, T.W., 2015. The effect of dilution and the use of a post-
730 extraction nucleic acid purification column on the accuracy, precision, and inhibition of
731 environmental DNA samples. Biol. Conserv. 183, 70–76. doi:10.1016/j.biocon.2014.11.031
732 McMurdie, P.J., Holmes, S., 2014. Waste not, want not: why rarefying microbiome data is
733 inadmissible. PLoS Comput. Biol. 10, e1003531. doi:10.1371/journal.pcbi.1003531
734 Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed
735 target capture and sequencing. Cold Spring Harb. Protoc. 5. doi:10.1101/pdb.prot5448
736 Morrissey, E.M., McHugh, T.A., Preteska, L., Hayer, M., Dijkstra, P., Hungate, B.A., Schwartz,
737 E., 2015. Dynamics of extracellular DNA decomposition and bacterial community

composition in soil. Soil Biol. Biochem. 86, 42–49. doi:10.1016/j.soilbio.2015.03.020

NCBI Resource Coordinators, 2018. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 46, 8–13. doi:10.1093/nar/gkx985

Nichols, R. V., Vollmers, C., Newsom, L.A., Wang, Y., Heintzman, P.D., Leighton, M., Green, R.E., Shapiro, B., 2018. Minimizing polymerase biases in metabarcoding. Mol. Ecol. Resour. 18, 927–939. doi:10.1111/1755-0998.12895

Niemeyer, B., Epp, L.S., Stoof-Leichsenring, K.R., Pestryakova, L.A., Herzschuh, U., 2017. A comparison of sedimentary DNA and pollen from lake sediments in recording vegetation composition at the Siberian treeline. Mol. Ecol. Resour. 17, e46–e62. doi:10.1111/1755-0998.12689

Ogram, A., Sayler, G., Gustin, D., Lewis, R., 1988. DNA adsorption to soils and sediments. Environ. Sci. Technol. 22, 982–984.

Opel, K.L., Chung, D., McCord, B.R., 2010. A study of PCR inhibition mechanisms using real time PCR. J. Forensic Sci. 55, 25–33. doi:10.1111/j.1556-4029.2009.01245.x

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 Phytol. 214, 924–942. doi:10.1111/nph.14470

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. doi:10.1038/nature19085

Pina-Martins, F., Paulo, O.S., 2015. NCBI Mass Sequence Downloader–Large dataset downloading made easy. SoftwareX 5, 80–83. doi:10.1016/j.softx.2016.04.007

Rawlence, N.J., Lowe, D.J., Wood, J.R., Young, J.M., Churchman, G.J., Huang, Y.T., Cooper, A., 2014. Using palaeoenvironmental DNA to reconstruct past environments: Progress and prospects. J. Quat. Sci. 29, 610–626. doi:10.1002/jqs.2740

Renaud, G., Stenzel, U., Kelso, J., 2014. LeeHom: Adaptor trimming and merging for Illumina sequencing reads. Nucleic Acids Res. 42, e141. doi:10.1093/nar/gku699

Sadoway, T.R., 2014. A Metagenomic Analysis of Ancient Sedimentary DNA Across the Pleistocene-Holocene Transition. McMaster University.

Saeki, K., Sakai, M., Wada, S.I., 2010. DNA adsorption on synthetic and natural allophanes. Appl. Clay Sci. 50, 493–497. doi:10.1016/j.clay.2010.09.015

Sidstedt, M., Jansson, L., Nilsson, E., Noppa, L., Forsman, M., Rådström, P., Hedman, J., 2015. Humic substances cause fluorescence inhibition in real-time polymerase chain reaction. Anal. Biochem. 487, 30–37. doi:10.1016/j.ab.2015.07.002

Sjögren, P., Edwards, M.E., Gielly, L., Langdon, C.T., Croudace, I.W., Merkel, M.K.F., Fonville, T., Alsos, I.G., 2016. Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland. New Phytol. 1–13. doi:10.1111/nph.14199

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 (80-.). 356, 605–608. doi:10.1126/science.aam9695

Søe, M.J., Nejsum, P., Seersholm, F.V., Fredensborg, B.L., Habraken, R., Haase, K., Hald, M.M., Simonsen, R., Højlund, F., Blanke, L., Merkyte, I., Willerslev, E., Kapel, C.M.O., 2018. Ancient DNA from latrines in Northern Europe and the Middle East (500 BC–1700 AD) reveals past parasites and diet. PLoS One 13, e0195481. doi:10.1371/journal.pone.0195481

Soininen, E.M., Gauthier, G., Bilodeau, F., Berteaux, D., Gielly, L., Taberlet, P., Gussarova, G., Bellemain, E., Hassel, K., Stenoien, H.K., Epp, L., Schroder-Nielsen, A., Brochmann, C., Yoccoz, N.G., 2015. Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. PLoS One 10, 1–18. doi:10.1371/journal.pone.0115335

Sønstebø, J.H., Gielly, L., Brysting, A.K., Elven, R., Edwards, M., Haile, J., Willerslev, E., Coissac, E., Rioux, D., Sannier, J., Taberlet, P., Brochmann, C., Sonstebo, J.H., Gielly, L., Brysting, A.K., Elven, R., Edwards, M., Haile, J., Willerslev, E., Coissac, E., Rioux, D., Sannier, J., Taberlet, P., Brochmann, C., 2010. Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. Mol. Ecol. Resour. 10, 1009–1018. doi:10.1111/j.1755-0998.2010.02855.x

Stuart, A.J., 2015. Late Quaternary megafaunal extinctions on the continents: a short review. Geol. J. 50, 414–433. doi:10.1002/gj

Sze, M.A., Schloss, P.D., 2019. The impact of DNA polymerase and number of rounds of amplification in PCR on 16S rRNA gene sequence data. bioRxiv 565598. doi:10.1101/565598

Taberlet, P., Bonin, A., Zinger, L., Coissac, E., 2018. Environmental DNA: For Biodiversity Research and Monitoring. Oxford University Press, Oxford, U.K.

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 Res. 35, e14. doi:10.1093/nar/gkl938

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M.E., Lorenzen, E.D., Vestergard, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L.S., Pearman, P.B., Cheddadi, R., Murray, D., Brathen, K.A., Yoccoz, N., Binney, H., Cruaud, C., Wincker, P., Goslar, T., Alsos, I.G., Bellemain, E., Brysting, A.K., Elven, R., Sonstebo, J.H., Murton, J., Sher, A., Rasmussen, M., Ronn, R., Mourier, T., Cooper, A., Austin, J., Moller, P., Froese, D., Zazula, G., Pompanon, F., Rioux, D., Niderkorn, V., Tikhonov, A., Savvinov, G., Roberts, R.G., MacPhee, R.D.E., Gilbert, M.T.P., Kjaer, K.H., Orlando, L., Brochmann, C., Taberlet, P., 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. Nature 506, 47–51. doi:10.1038/nature12921

Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shaprio, B., Bunce, M., Wiuf, C., Gilichinsky, D.A., Cooper, A., 2003. Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments. Science (80-.). 300, 791–795. doi:10.1126/science.1084114

Yoccoz, N.G., Bråthen, K.A., Gielly, L., Haile, J., Edwards, M.E., Goslar, T., Von Stedingk, H., Brysting, A.K., COISSAC, E., Pompanon, F., Sønstebø, J.H., Miquel, C., Valentini, A., De Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., Rasmussen, M., Gilbert, M.T.P., Orlando, L., Brochmann, C., Willerslev, E., Taberlet, P., 2012. DNA from soil mirrors plant taxonomic and growth form diversity. Mol. Ecol. 21, 3647–3655.

827 doi:10.1111/j.1365-294X.2012.05545.x

828

829

830

831

832 **Main Text Tables**

833

Table 1 Sample IDs for experiment SET-E with read filtering summaries.

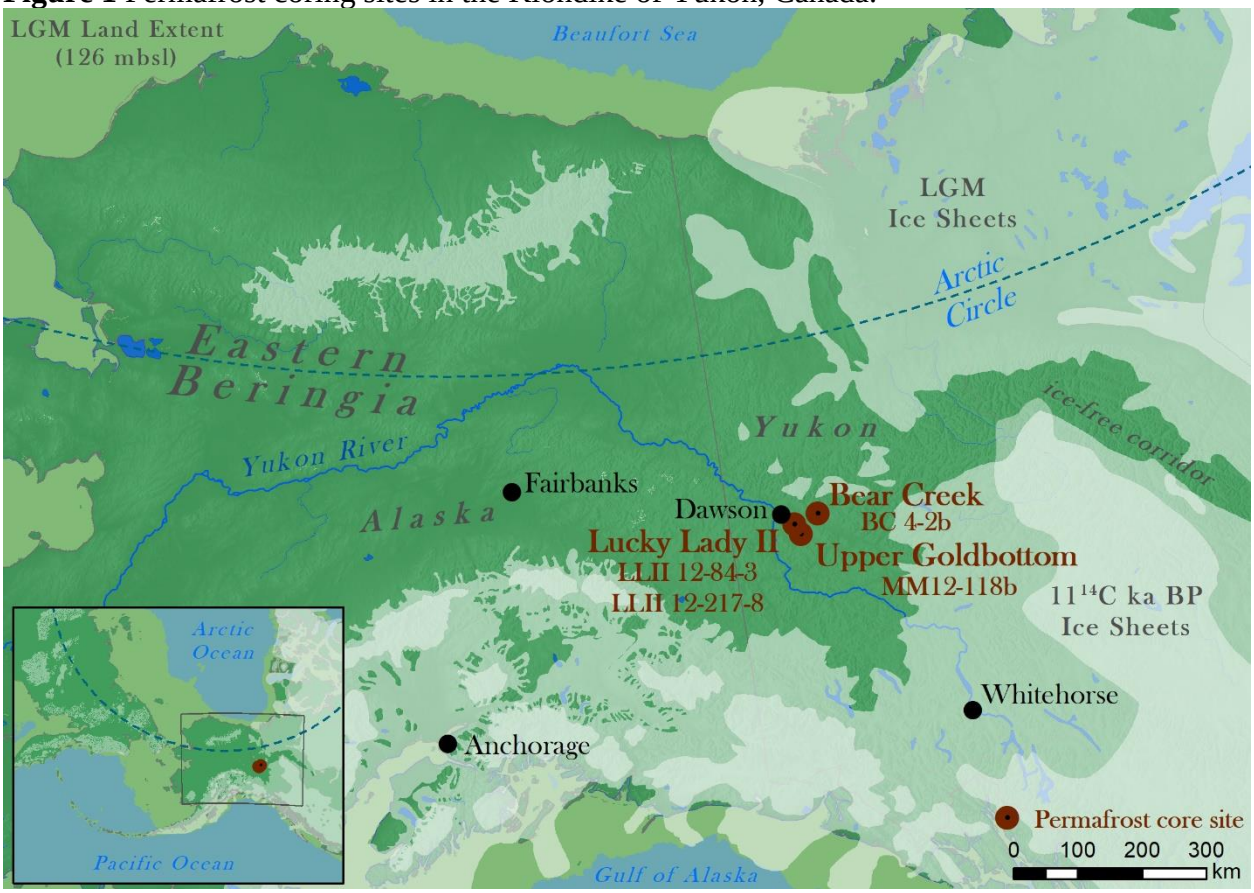
Library	Core Section	Section Age	Site	SedaDNA Extraction	Targeting Strategy	Total Reads	Mapped-to-Baits, Unique, & ≥ 24bp	BLASTn Hits LCA-Assigned	Assigned of Total	On-Target Fold Increase Over	
										PowerSoil	Metabarcoding
L-SET-256-En						6,292,874	14,408	12,812	0.20%		
L-SET-257-En	MM12-118b	9,685 cal-BP	Upper Goldbottom	PowerSoil	Enriched	675,550	4,615	4,105	0.61%		
L-SET-258-En						820,382	4,885	4,354	0.53%		
L-SET-259-En						515,685	2,916	2,563	0.50%		
L-SET-260-En	LLII 12-84-3	13,205 cal-BP	Lucky Lady II	PowerSoil	Enriched	692,434	4,028	3,470	0.50%		
L-SET-261-En						1,060,105	2,760	2,368	0.22%		
L-SET-262-En						534,677	500	452	0.08%		
L-SET-263-En	LLII 12-217-8	15,865 cal-BP	Lucky Lady II	PowerSoil	Enriched	873,741	348	312	0.04%		
L-SET-264-En						418,730	717	648	0.15%		
L-SET-265-En						682,544	1,788	1,590	0.23%		
L-SET-266-En	BC 4-2B	30,000 cal-BP	Bear Creek	PowerSoil	Enriched	868,112	2,087	1,770	0.20%		
L-SET-267-En						857,863	2,540	2,249	0.26%		
L-SET-BK22-En		Extraction Blank		PowerSoil	Enriched	102,752	0	0	0.00%		
L-SET-268-En						1,704,149	140,049	122,282	7.18%		
L-SET-269-En	MM12-118b	9,685 cal-BP	Upper Goldbottom	Modified Dabney	Enriched	1,782,291	122,815	104,760	5.88%	15.7	24.7
L-SET-270-En						1,269,901	117,267	102,542	8.07%		
L-SET-271-En						1,837,939	137,192	119,031	6.48%		
L-SET-272-En	LLII 12-84-3	13,205 cal-BP	Lucky Lady II	Modified Dabney	Enriched	1,343,928	126,649	110,609	8.23%	19.3	59.9
L-SET-273-En						1,267,881	128,565	112,855	8.90%		
L-SET-274-En						867,275	8,235	6,600	0.76%		
L-SET-275-En	LLII 12-217-8	15,865 cal-BP	Lucky Lady II	Modified Dabney	Enriched	996,802	7,025	5,614	0.56%	7.7	1.2
L-SET-276-En						352,658	3,560	2,802	0.79%		
L-SET-277-En						1,254,744	72,913	62,778	5.00%		
L-SET-278-En	BC 4-2B	30,000 cal-BP	Bear Creek	Modified Dabney	Enriched	1,294,040	43,451	37,878	2.93%	15.8	4.7
L-SET-279-En						1,544,949	54,013	47,613	3.08%		
L-SET-BK23-En		Extraction Blank		Modified Dabney	Enriched	1,186	2	0	0.00%	**Average**	**Average**
L-SET-BK24-En		Library Blank			Enriched	677	0	0	0.00%	**14.6**	**22.6**
L-SET-256-SG	MM12-118b	9,685 cal-BP	Upper Goldbottom	PowerSoil	Shotgun	1,717,174	55	22	0.00%		
L-SET-259-SG	LLII 12-84-3	13,205 cal-BP	Lucky Lady II	PowerSoil	Shotgun	1,517,583	53	17	0.00%		
L-SET-264-SG	LLII 12-217-8	15,865 cal-BP	Lucky Lady II	PowerSoil	Shotgun	1,497,940	19	6	0.00%		
L-SET-265-SG	BC 4-2B	30,000 cal-BP	Bear Creek	PowerSoil	Shotgun	1,338,467	34	5	0.00%		
L-SET-268-SG	MM12-118b	9,685 cal-BP	Upper Goldbottom	Modified Dabney	Shotgun	2,202,687	544	152	0.01%		
L-SET-272-SG	LLII 12-84-3	13,205 cal-BP	Lucky Lady II	Modified Dabney	Shotgun	2,122,805	573	145	0.01%		
L-SET-274-SG	LLII 12-217-8	15,865 cal-BP	Lucky Lady II	Modified Dabney	Shotgun	1,992,150	283	40	0.00%		
L-SET-279-SG	BC 4-2B	30,000 cal-BP	Bear Creek	Modified Dabney	Shotgun	8,593,408	816	133	0.00%		
L-SET-BK22-SG		Extraction Blank		PowerSoil	Shotgun	2,756,360	3	0	0.00%		
L-SET-BK23-SG		Extraction Blank		Modified Dabney	Shotgun	1,748,595	3	0	0.00%		
L-SET-BK24-SG		Library Blank			Shotgun	2,841,911	4	0	0.00%		
GB1	MM12-118b	9,685 cal-BP	Upper Goldbottom	D'Costa et al. 2011	Metabarcoding	109,233	320	311	0.28%		
LL3	LLII 12-84-3	13,205 cal-BP	Lucky Lady II	D'Costa et al. 2011	Metabarcoding	738,708	1047	971	0.13%		
LL1	LLII 12-217-8	15,865 cal-BP	Lucky Lady II	D'Costa et al. 2011	Metabarcoding	77,373	450	448	0.58%		
BC	BC 4-2B	30,000 cal-BP	Bear Creek	D'Costa et al. 2011	Metabarcoding	172,330	1359	1348	0.78%		

Core section ages as per D'Coasta et al. (2011), Mahony (2015), and Sadoway (2014). See Tables S17 and S18 in the SOM for a summary of major clade LCA-assignments.

837 **Main Text Figures**

838

839 **Figure 1** Permafrost coring sites in the Klondike of Yukon, Canada.

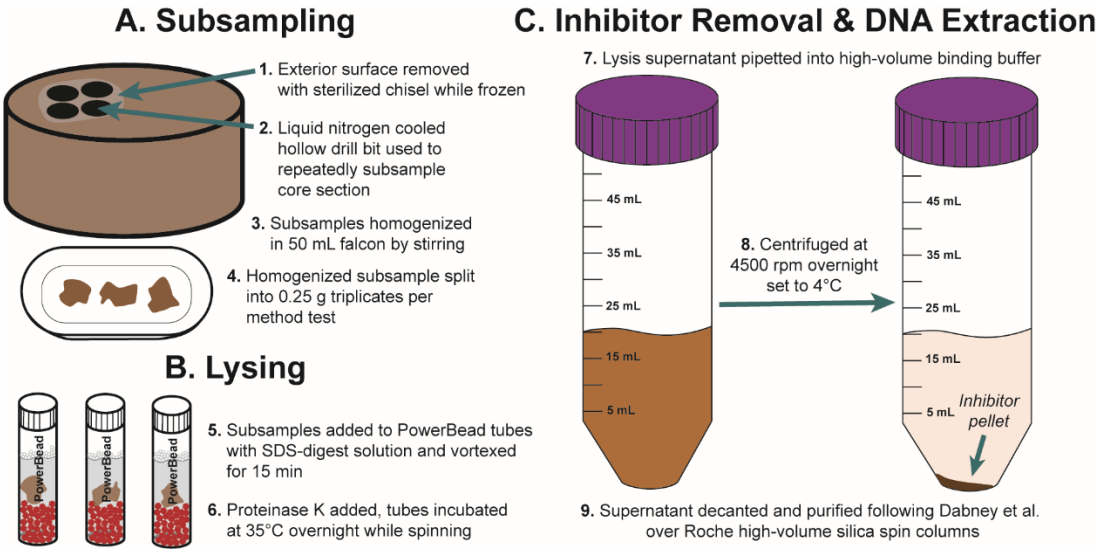


840

841 Ice sheet data from Dyke (2004) and Ehlers et al. (2011). Sea level at Last Glacial Maximum (LGM,
842 26.5–19 ky BP) (Clark, 2009) set to 126 meters below sea level (msbl) based on midpoint between
843 maximum and minimum eustatic sea level estimation models in Clark and Mix (2002).

844

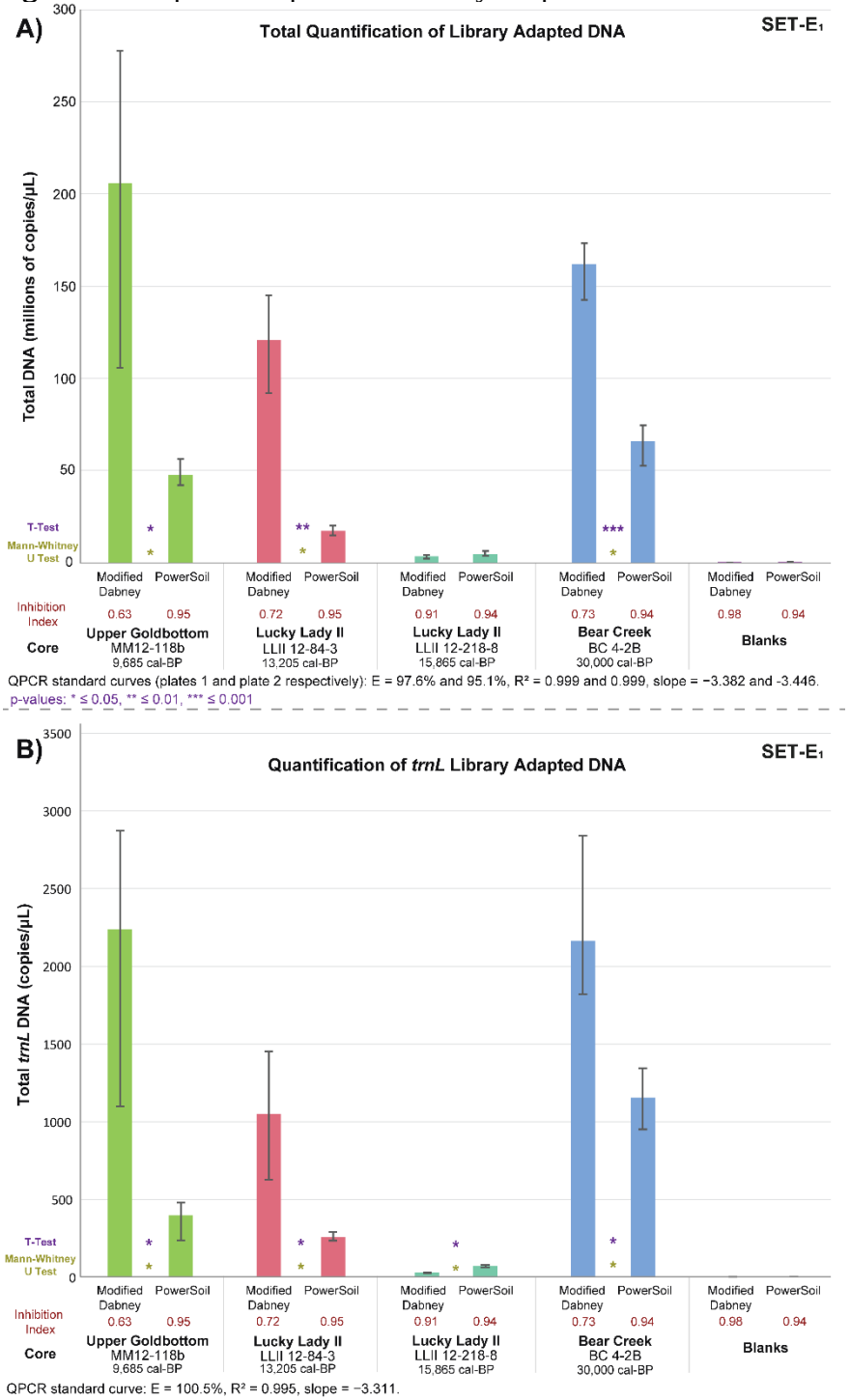
845 **Figure 2** SedaDNA modified Dabney et al. (2013) extraction workflow.



846

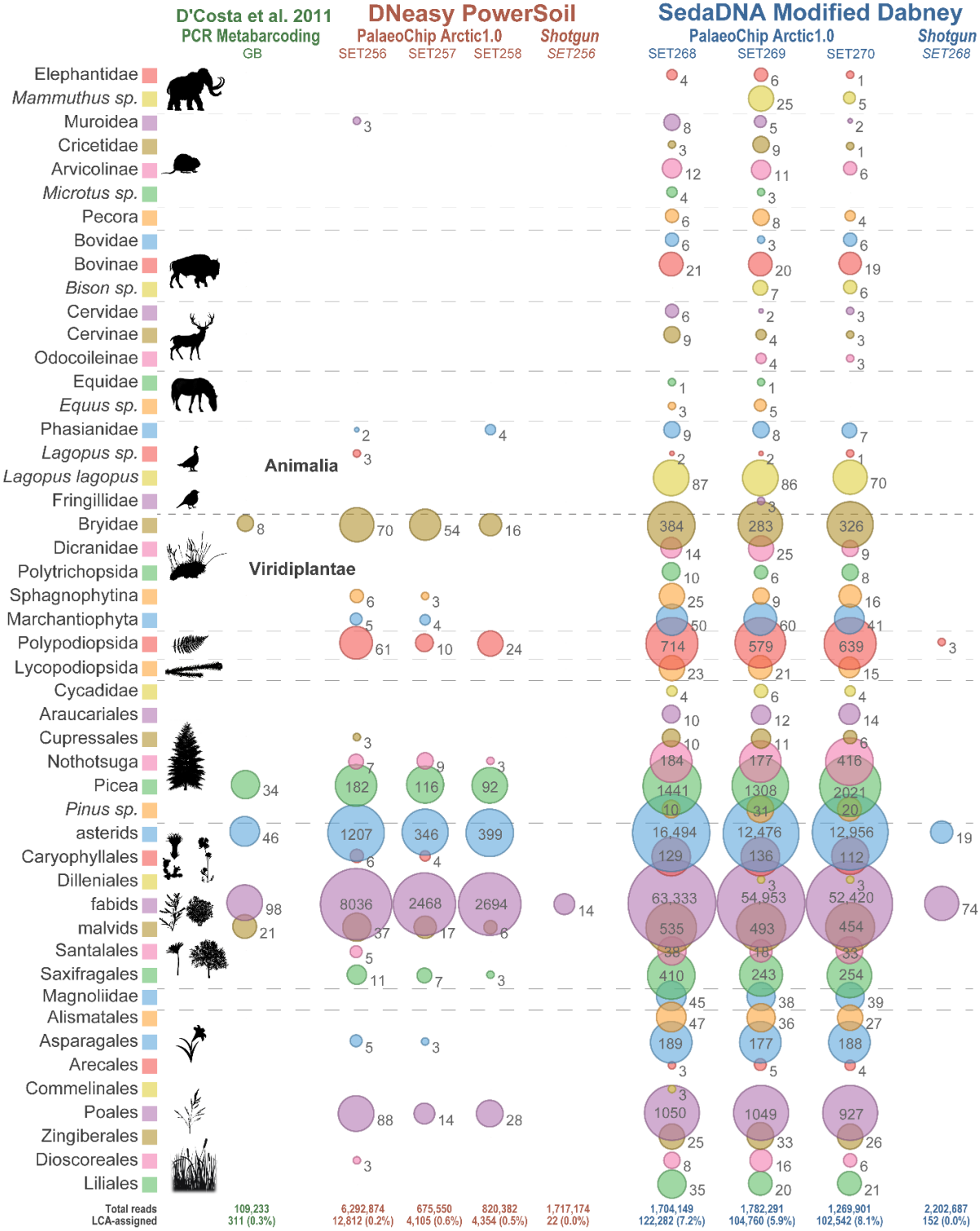
847 See the methods section for further details on extraction, double-stranded library preparation, capture
848 enrichment, qPCR assays, and the bioinformatic workflow.
849

Figure 3 Comparative qPCRs of library adapted DNA.



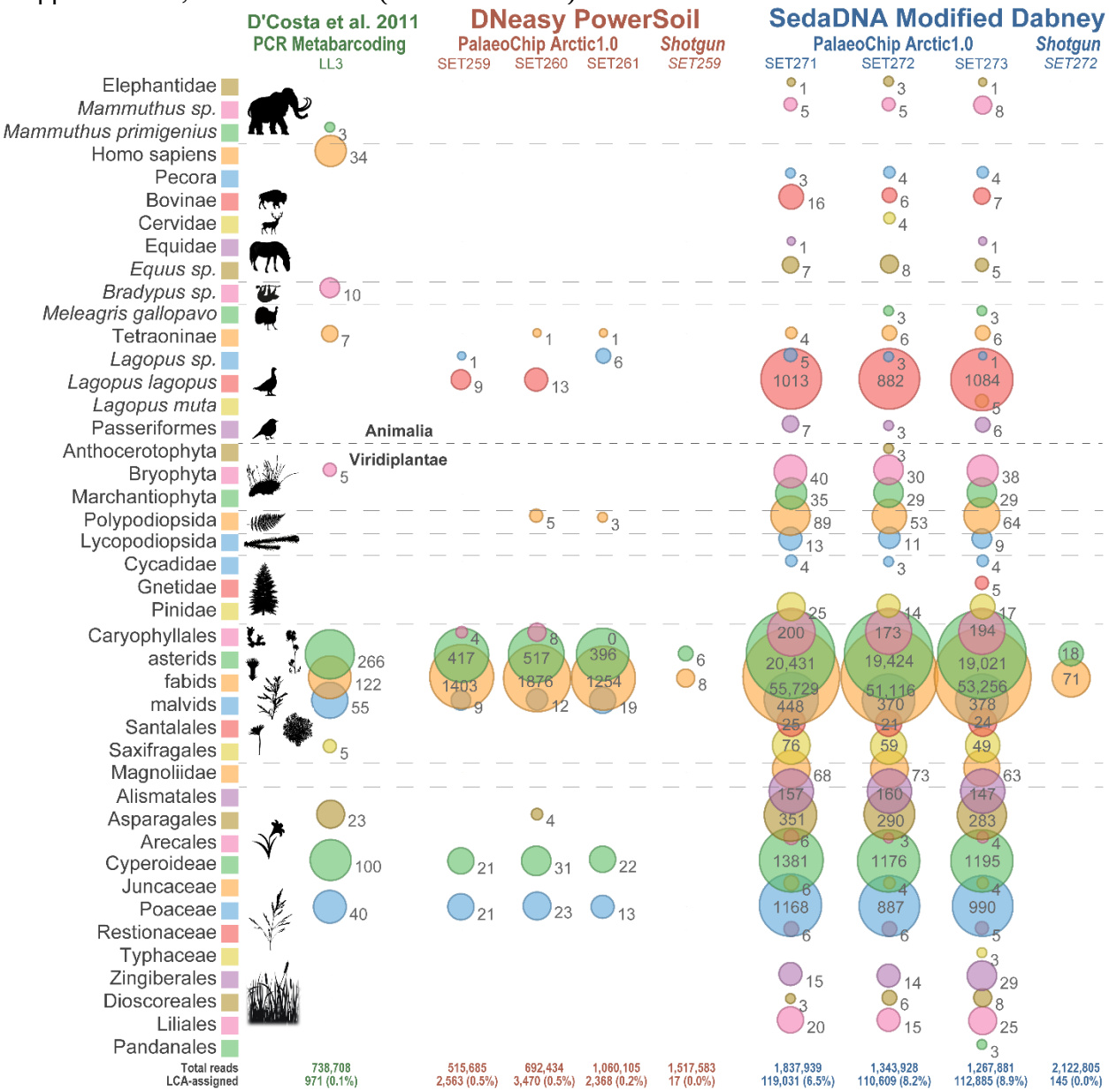
Each bar represents the average DNA copy number of the extraction-core triplicate, with the error bars indicating the maximum and minimum range. DNA concentration for each SET sample was averaged across PCR triplicates. Inhibition index refers to inhibition spike test (Table S7) on extracts prior to library prep. **A)** Total DNA quantification comparing both extraction methods by core; see Table S8 for PCR specifications. **B)** Total *trnL* adapted DNA; see Table S9 for PCR specifications. P-values calculated with both a two-sample t-test (parametric) and Mann-Whitney U test (non-parametric) (two-sided). The large range for modified Dabney extraction core MM12-118B is driven by a single lower copy number extraction replicate. Core LLII 12-217-8 consistently has low DNA recovery, but also a low co-elution of inhibition.

Figure 4 Metagenomic comparison of Upper Goldbottom permafrost core MM12-118b, reads mapped to baits, absolute counts (non-normalized).



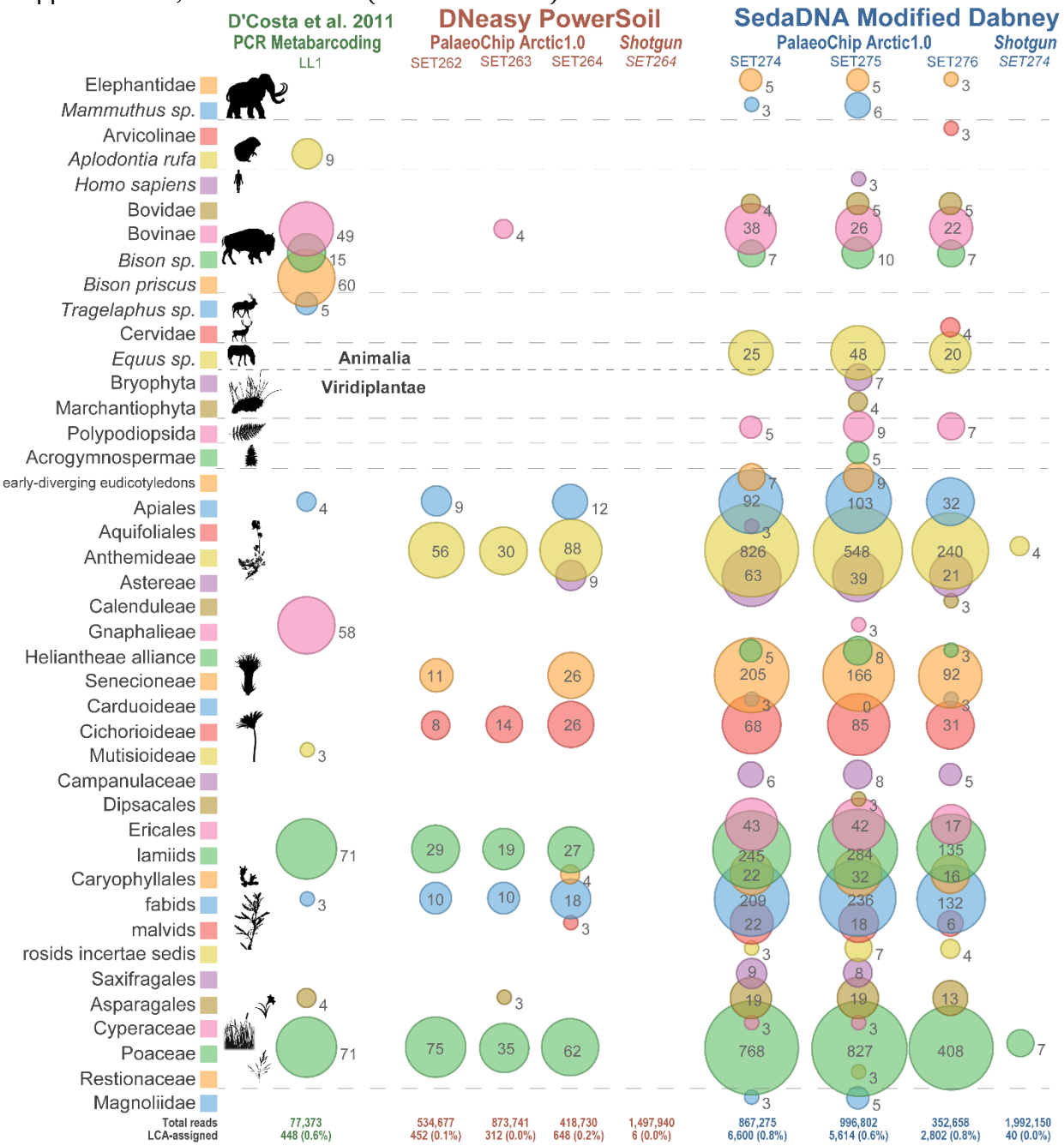
Core slice dated to 9,685 cal-BP (Mahony, 2015; Sadoway, 2014). Values indicate total reads assigned to that taxon node for Animalia, and a clade summation of reads for Viridiplantae. See Table 1 for read summaries.

Figure 5 Metagenomic comparison of Lucky Lady II permafrost core LLII-12-84-3, reads mapped to baits, absolute counts (non-normalized).



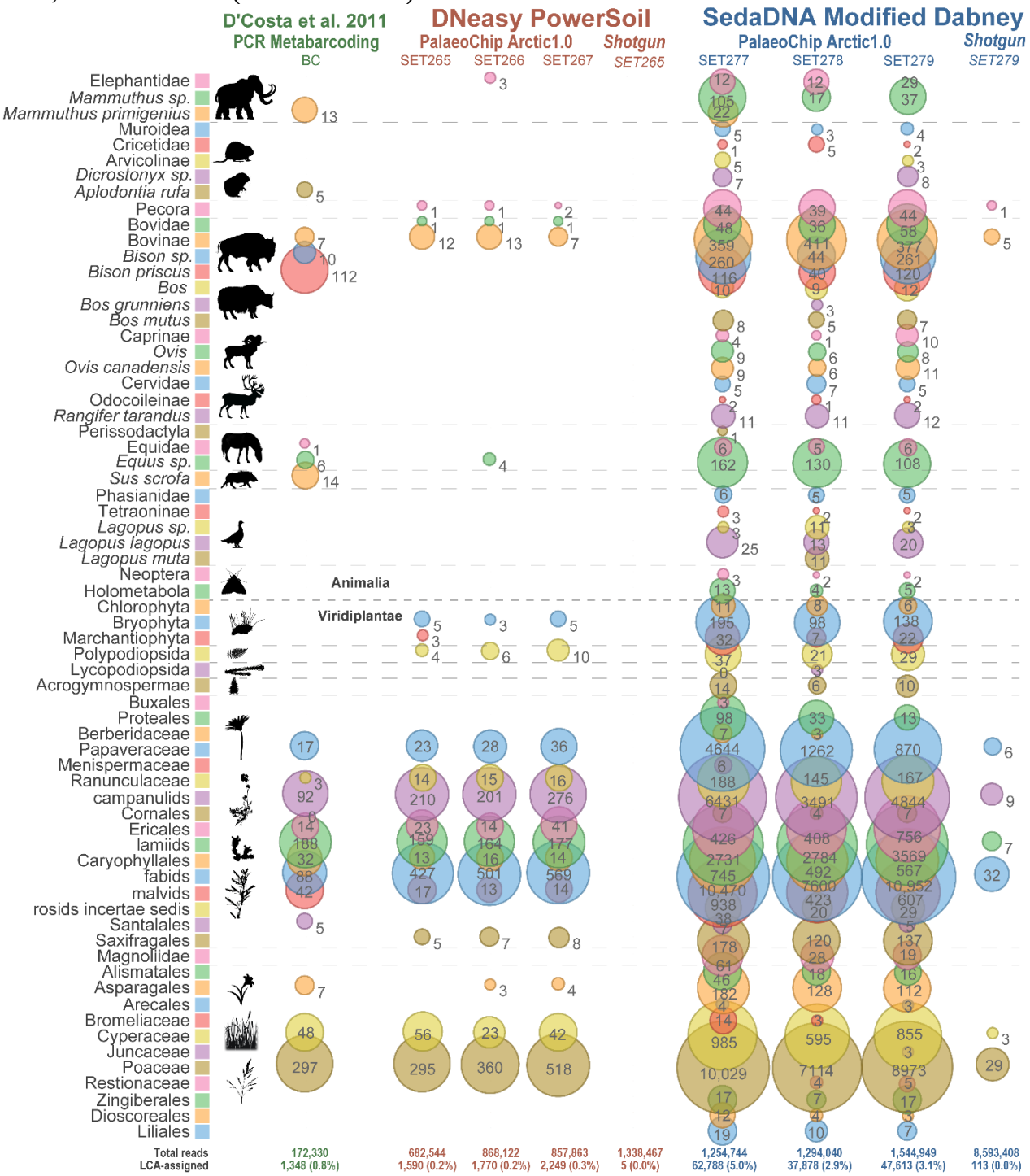
Core slice dated to 13,205 cal-BP (Sadoway, 2014). Values indicate total reads assigned to that taxon node for Animalia, and a clade summation of reads for Viridiplantae. See Table 1 for read summaries.

Figure 6 Metagenomic comparison of Lucky Lady II permafrost core LLII-12-217-8, reads mapped to baits, absolute counts (non-normalized).



Core slice dated to 15,865 cal-BP (Sadoway, 2014). Values indicate total reads assigned to that taxon node for Animalia, and a clade summation of reads for Viridiplantae. See Table 1 for read summaries.

878 **Figure 7** Metagenomic comparison of Bear Creek permafrost core BC 4-2B, reads mapped to
879 baits, absolute counts (non-normalized).



880
881 Core slice dated to ~30,000 cal-BP (D'Costa et al., 2011; Mahony, 2015; Sadoway, 2014).
882 Values indicate total reads assigned to that taxon node for Animalia, and a clade summation of
883 reads for Viridiplantae. See Table 1 for read summaries.
884

Main Text Extended Tables

Table E2 Taxon specific mapping summary at a minimum length of 24 bp and mapping quality of 30.

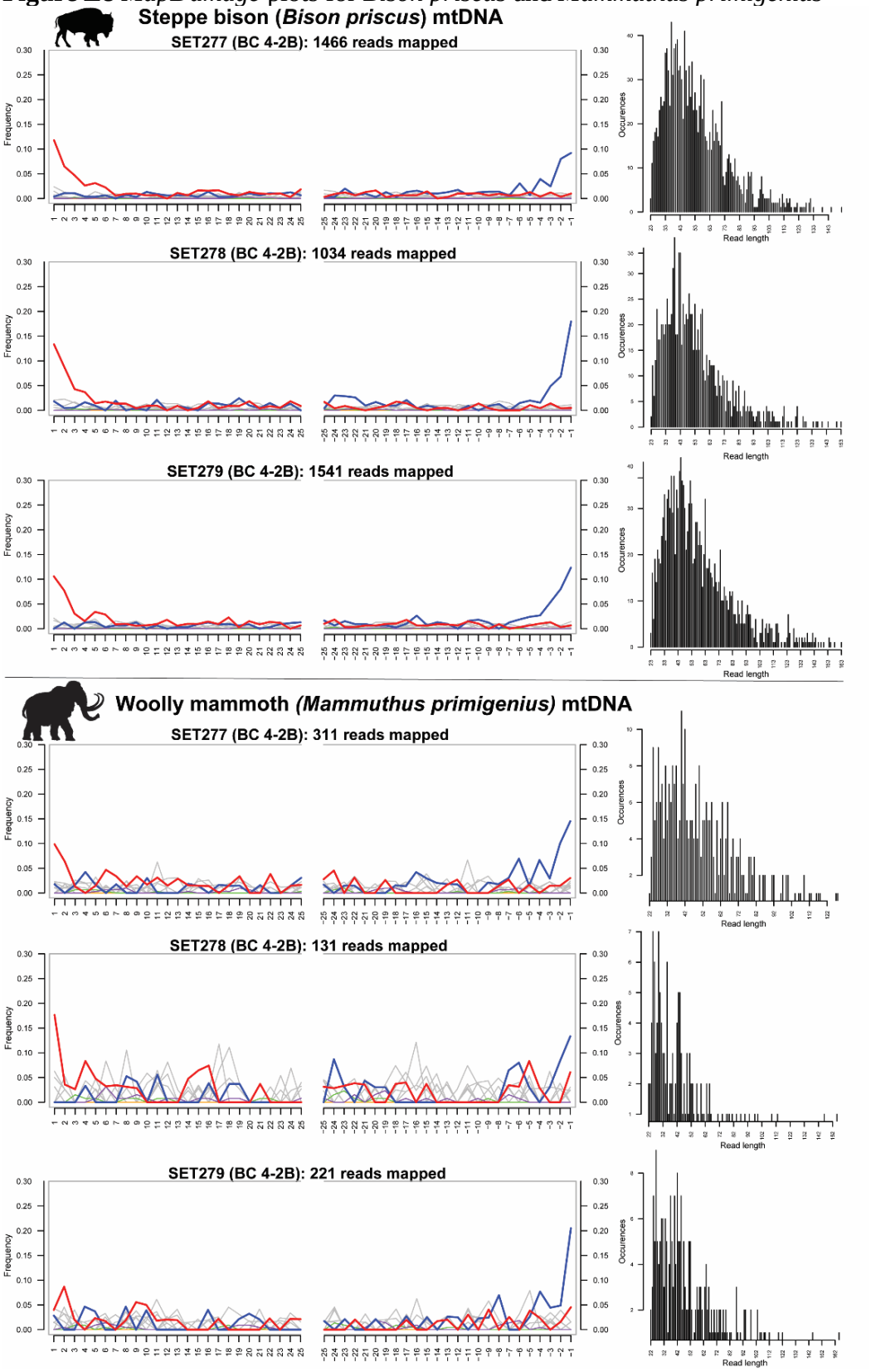
Library	Core	Extraction	*Bison priscus* NC_027233	*Equus caballus* NC_001640	*Mammuthus primigenius* NC_007596	*Lagopus lagopus* NC_035568	*Picea glauca* NC_028594	*Poa palustris* NC_027484	*Salix interior* NC_024681	*Artemisia frigida* NC_020607
L-SET-256-En			6	0	3	14	1,723	0	37,569	3,645
L-SET-257-En	MM12-118b	PowerSoil	1	0	0	7	1,279	298	15,097	1,371
L-SET-258-En			2	0	1	24	1,212	263	14,894	1,495
L-SET-259-En			2	4	2	110	130	218	4,236	1,023
L-SET-260-En	LLII 12-84-3	PowerSoil	5	1	5	74	196	294	5,214	1,209
L-SET-261-En			1	0	1	63	136	196	3,419	917
L-SET-262-En			5	4	3	3	17	227	113	1,080
L-SET-263-En	LLII 12-217-8	PowerSoil	6	2	0	1	20	106	93	850
L-SET-264-En			9	7	3	1	33	171	169	1,671
L-SET-265-En			37	4	2	7	52	920	1,518	939
L-SET-266-En	BC 4-2B	PowerSoil	35	8	8	10	95	953	1,453	977
L-SET-267-En			26	7	2	13	79	1,477	1,864	1,149
L-SET-268-En			103	47	44	**245**	13,524	11,502	**141,195**	34,802
L-SET-269-En	MM12-118b	Modified Dabney	106	45	83	201	12,396	**10,791**	113,197	29,313
L-SET-270-En			104	32	37	178	**14,575**	10,480	112,797	30,262
L-SET-271-En			74	49	59	1,798	**10,170**	13,523	92,828	**37,685**
L-SET-272-En	LLII 12-84-3	Modified Dabney	82	59	67	**1,611**	9,921	12,811	85,995	36,168
L-SET-273-En			78	51	61	**1,950**	9,718	12,694	**96,063**	36,377
L-SET-274-En			89	58	17	21	533	**1,551**	1,731	**6,462**
L-SET-275-En	LLII 12-217-8	Modified Dabney	80	81	31	21	444	1,484	1,426	4,455
L-SET-276-En			74	43	13	14	231	727	745	2,226
L-SET-277-En			**1,466**	**427**	**311**	127	**4,907**	**20,006**	**30,198**	**26,042**
L-SET-278-En	BC 4-2B	Modified Dabney	**1,034**	**338**	**131**	123	3,082	13,724	20,619	15,035
L-SET-279-En			**1,541**	**370**	**221**	113	3,770	16,781	26,821	18,734
L-SET-BK22-En	Ext. Blank	PowerSoil	0	0	0	0	0	0	0	0
L-SET-BK23-En	Ext. Blank	Modified Dabney	0	0	0	1	0	0	0	0
L-SET-Bk24-En	Library Blank		0	0	0	0	0	0	0	0

MapDamage profiles for highlighted cells (see Figures E8–E12).

889 **Main Text Extended Figures**

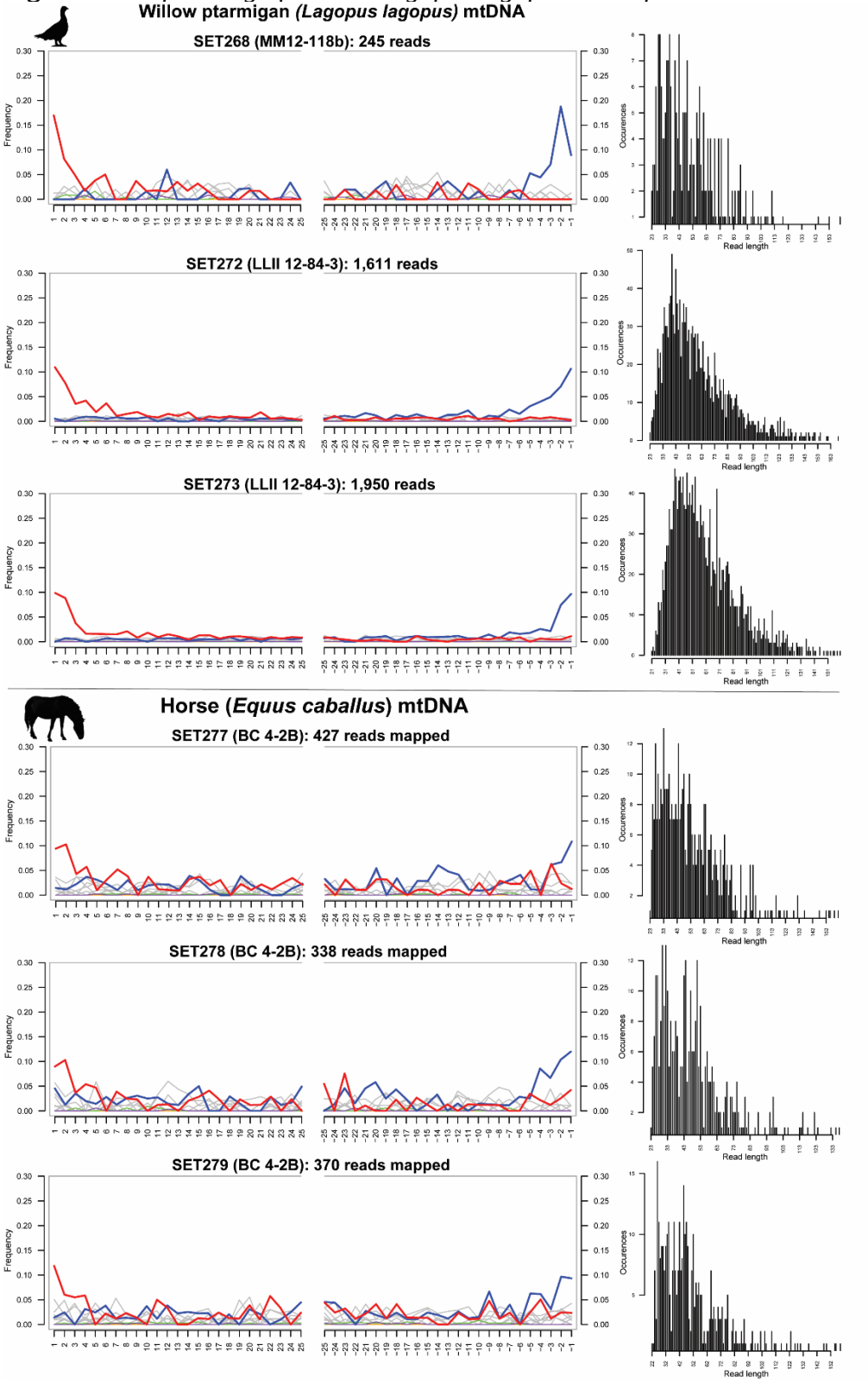
890

891 **Figure E8** *MapDamage* plots for *Bison priscus* and *Mammuthus primigenius*



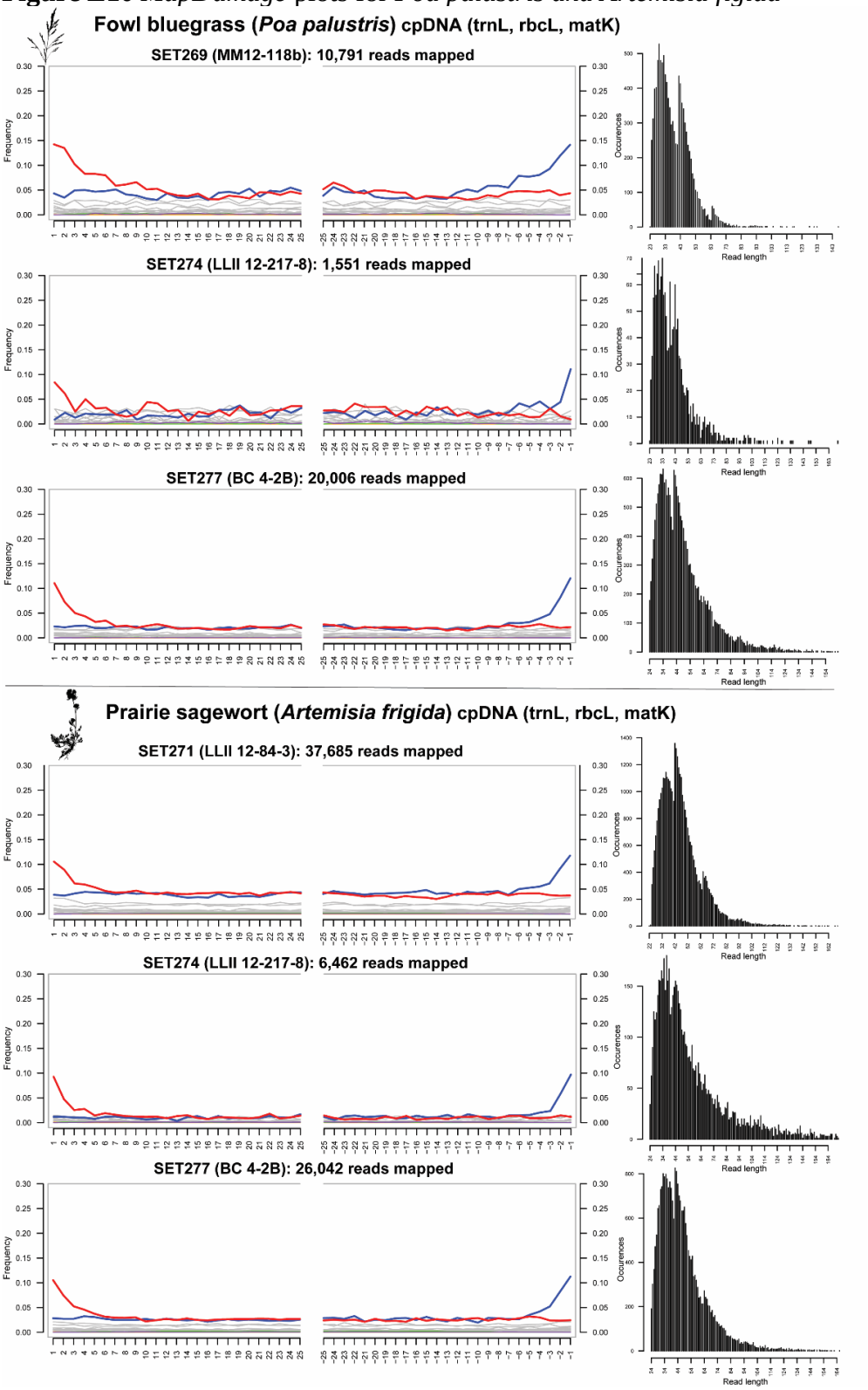
892
893 Minimum length = 24 bp, minimum mapping quality = 30.

894 **Figure E9** *MapDamage* plots for *Lagopus lagopus* and *Equus caballus*



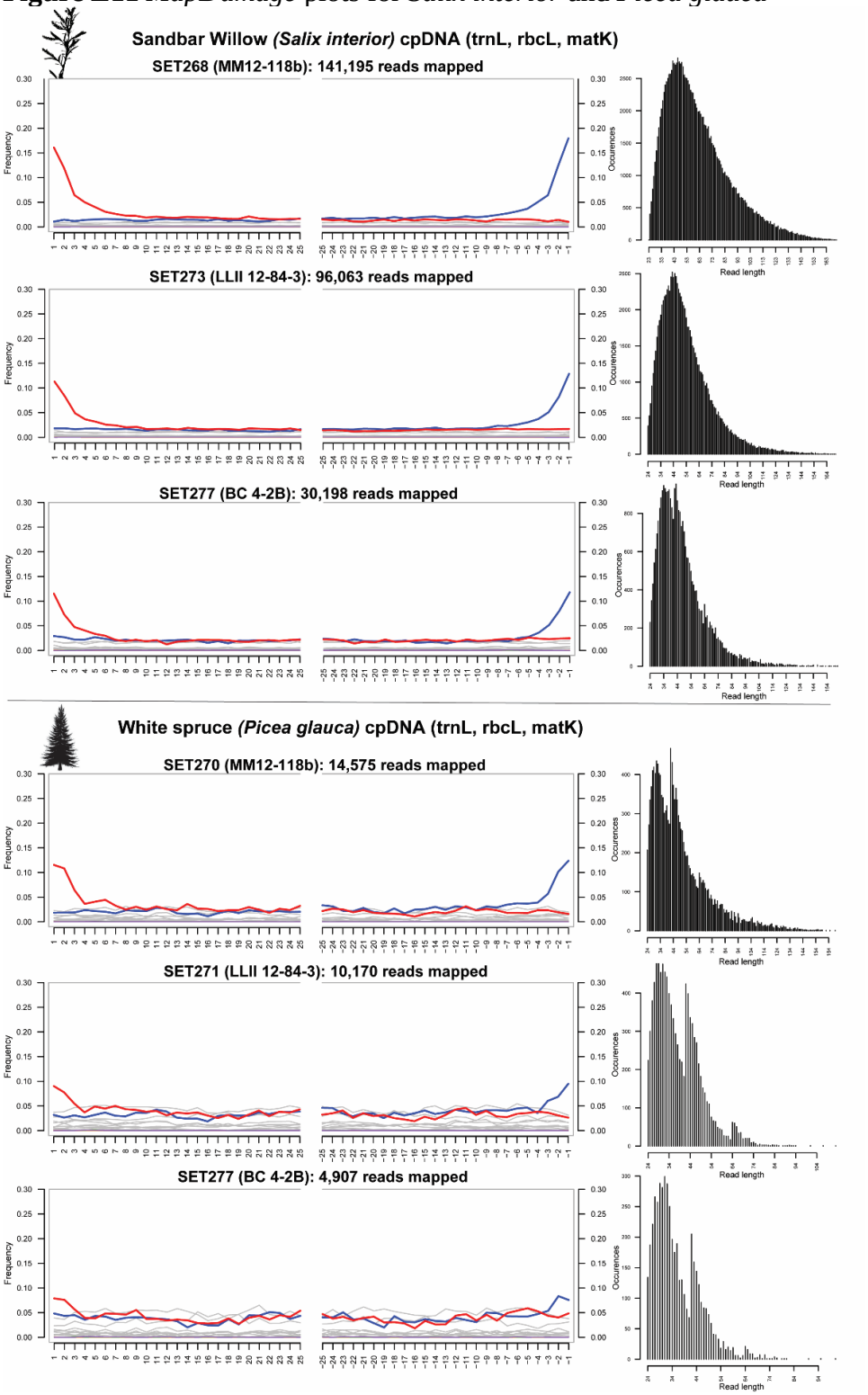
895
896 Minimum length = 24 bp, minimum mapping quality = 30.

Figure E10 *MapDamage* plots for *Poa palustris and Artemisia figida*



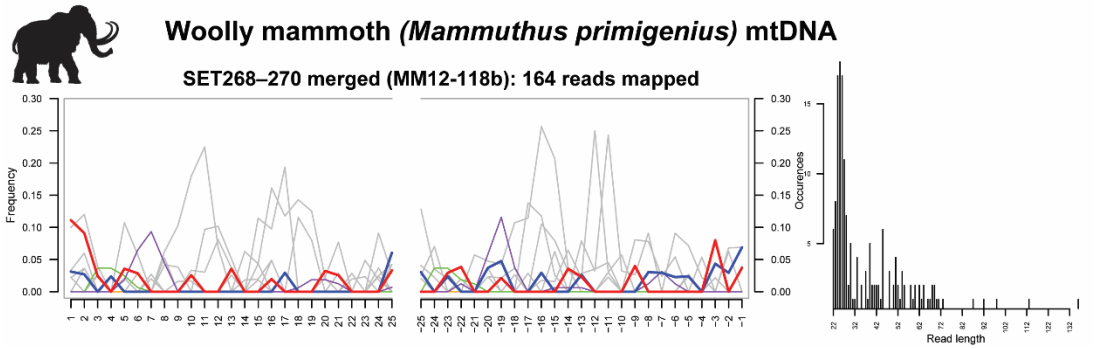
Minimum length = 24 bp, minimum mapping quality = 30. We suspect that the biomodial distribution of the fragment length distributions is due to non-specific mapping of closely related taxa in conserved regions of these cpDNA barcoding loci.

902 **Figure E11** *MapDamage* plots for *Salix interior* and *Picea glauca*



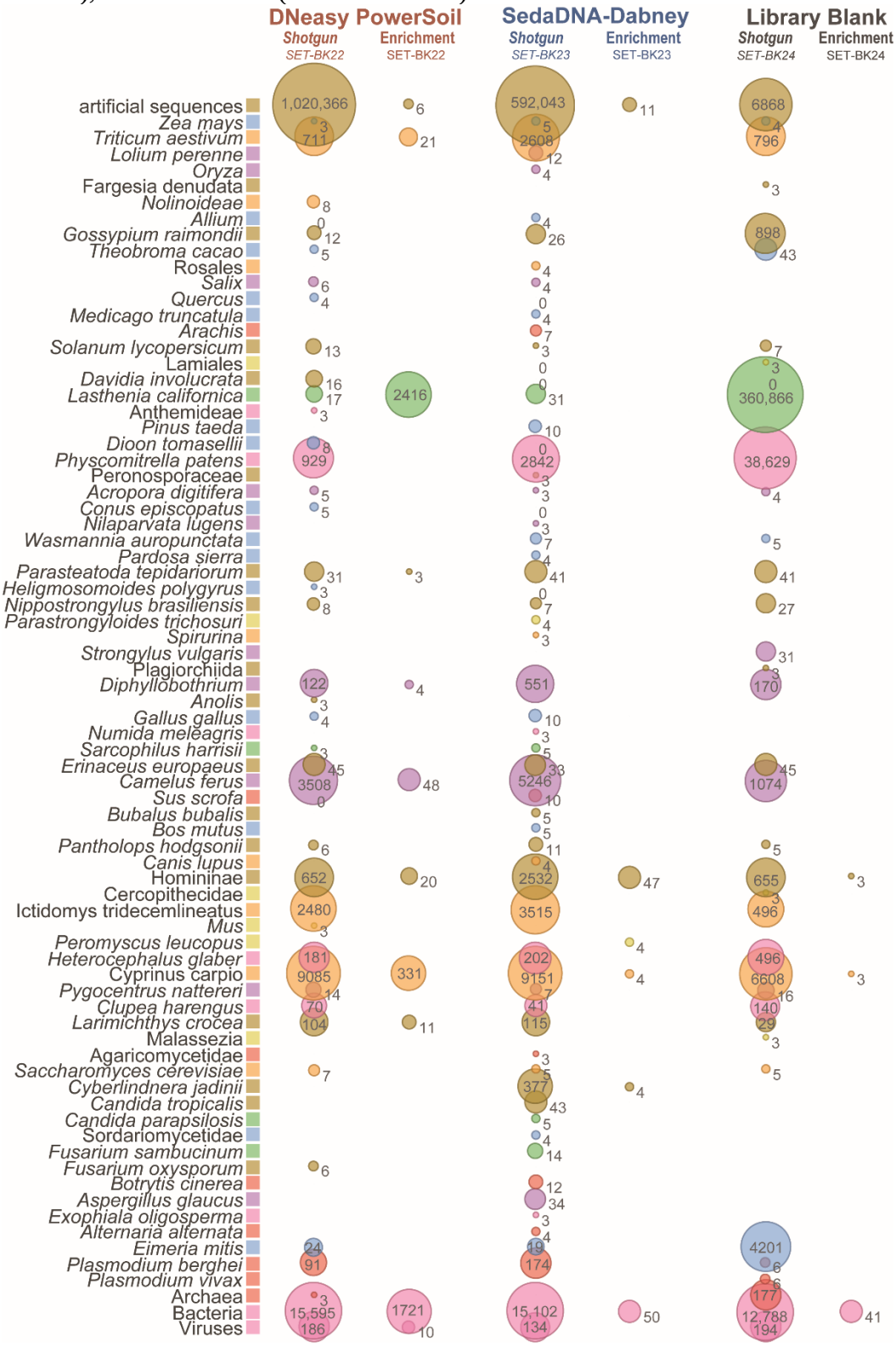
903
904 Minimum length = 24 bp, minimum mapping quality = 30. We suspect that the biomodial distribution of the
905 fragment length distributions is due to non-specific mapping of closely related taxa in conserved regions of these
906 cpDNA barcoding loci.

907 **Figure E12** *MapDamage* MM12-118b merged replicates plot for *Mammuthus primigenius*.



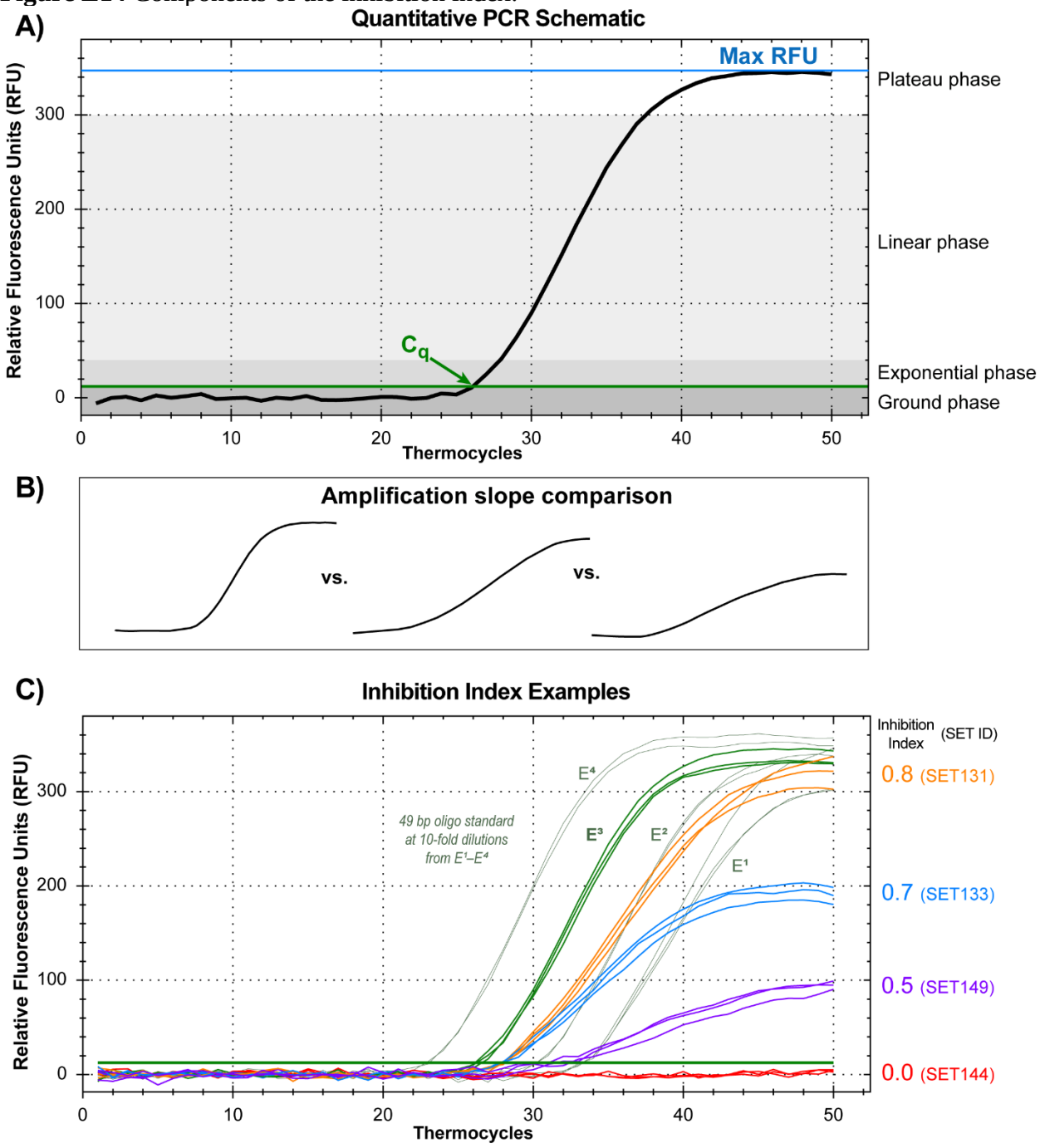
908
909 Minimum length = 24 bp, minimum mapping quality = 30.
910

Figure E13 Metagenomic comparison of extraction and library blanks, all reads (not map-filtered), absolute counts (non-normalized).



Values indicate total reads assigned to that taxon node.

916 **Figure E14** Components of the inhibition index.



917
918 **A)** A standard qPCR reaction showing C_q and max RFU. **B)** A comparison of various amplification slopes from a
919 typical reaction (left), towards increasingly inhibited reactions (right). **C)** Example inhibition indices derived from
920 averaging the C_q, max RFU, and by fitting a variable-slope sigmoidal dose-response curve to the raw fluorescence
921 data (using GraphPad Prism v. 7.04) based on King et al.(2009) for each PCR replicate by sample against the spiked
922 E^3 standard. Inhibition index values <0.5 tend to occur when individual PCR replicates fail in a triplicate series;
923 blanks and standard serial dilutions E^2 and E^1 tend to have inhibition indices >0.9 despite their 10- and 100-fold
924 reduction in starting DNA causing a 3 or 6 cycle C_q shift. QPCR standard curve: E = 94.2%, R^2 = 0.997, slope =
925 −3.469. See Table S7 for PCR assay specifications.