

University of Groningen

## Predicting sample success for large-scale ancient DNA studies on marine mammals

Keighley, Xenia; Bro-Jørgensen, Maiken Hemme; Ahlgren, Hans; Szpak, Paul; Ciucani, Marta Maria; Sánchez Barreiro, Fátima; Howse, Lesley; Gotfredsen, Anne Birgitte; Glykou, Aikaterini ; Jordan, Peter

*Published in:*  
Molecular Ecology Resources

*DOI:*  
[10.1111/1755-0998.13331](https://doi.org/10.1111/1755-0998.13331)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2021

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Keighley, X., Bro-Jørgensen, M. H., Ahlgren, H., Szpak, P., Ciucani, M. M., Sánchez Barreiro, F., Howse, L., Gotfredsen, A. B., Glykou, A., Jordan, P., Lidén, K., & Olsen, M. T. (2021). Predicting sample success for large-scale ancient DNA studies on marine mammals. *Molecular Ecology Resources*, 21(4), 1149-1166. <https://doi.org/10.1111/1755-0998.13331>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).









The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Predicting sample success for large-scale ancient DNA studies on marine mammals

Xénia Keighley<sup>1,2</sup>  | Maiken Hemme Bro-Jørgensen<sup>1,3</sup>  |  
Hans Ahlgren<sup>3</sup> | Paul Szpak<sup>4</sup>  | Marta Maria Ciucani<sup>1</sup>  | Fátima Sánchez Barreiro<sup>1</sup>  |  
Lesley Howse<sup>5</sup>  | Anne Birgitte Gotfredsen<sup>6</sup> | Aikaterini Glykou<sup>3</sup> | Peter Jordan<sup>7</sup> |  
Kerstin Lidén<sup>3</sup>  | Morten Tange Olsen<sup>1</sup> 

<sup>1</sup>Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, Copenhagen K, Denmark

<sup>2</sup>Arctic Centre/Groningen Institute of Archaeology, Faculty of Arts, University of Groningen, AS Groningen, The Netherlands

<sup>3</sup>Archaeological Research Laboratory, Department of Archaeology and Classical Studies, Stockholm University, Stockholm, Sweden

<sup>4</sup>Department of Anthropology, Trent University, Peterborough, Ontario, Canada

<sup>5</sup>Archaeology Centre, University of Toronto, Toronto, Ontario, Canada

<sup>6</sup>Section for GeoGenetics, GLOBE Institute, University of Copenhagen, Copenhagen K, Denmark

<sup>7</sup>Department of Archaeology and Ancient History, Lund University, Lund, Sweden

## Correspondence

Xénia Keighley, Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, CSS Building 7, Øster Farimagsgade 5 DK-1353 Copenhagen K, Denmark.  
Email: xenia@palaeome.org

and

Morten Tange Olsen, Section for Evolutionary Genomics, GLOBE Institute, University of Copenhagen, CSS Building 7, Øster Farimagsgade 5 DK-1353 Copenhagen K, Denmark  
Email: morten.olsen@sund.ku.dk

## Funding information

Joint Proxies Research Group, financed by Research Council of Norway (RCN) Grant Number 246899 and Faculty for Humanities, Social Sciences and Education, UiT The Arctic University of Norway; Swedish Research Council, Sweden, Grant/Award Number: 2015-02151; European Union's Framework Programme for Research and Innovation Horizon 2020, Grant/Award Number: #676154

## Abstract

In recent years, nonhuman ancient DNA studies have begun to focus on larger sample sizes and whole genomes, offering the potential to reveal exciting and hitherto unknown answers to ongoing biological and archaeological questions. However, one major limitation to such studies is the substantial financial and time investments still required during sample screening, due to uncertainty regarding successful sample selection. This study investigates the effect of a wide range of sample properties including latitude, sample age, skeletal element, collagen preservation, and context on endogenous content and DNA damage profiles for 317 ancient and historic pinniped samples collected from across the North Atlantic and surrounding regions. Using generalised linear and mixed-effect models, we found that a range of factors affected DNA preservation within each of the species under consideration. The most important findings were that endogenous content varied significantly within species according to context, the type of skeletal element, the collagen content and collection year. There also appears to be an effect of the sample's geographic origin, with samples from the Arctic generally showing higher endogenous content and lower damage rates. Both latitude and sample age were found to have significant relationships with damage levels, but only for walrus samples. Sex, ontogenetic age and extraction material preparation were not found to have any significant relationship with DNA preservation. Overall, skeletal element and sample context were found to be the most

Maiken Hemme Bro-Jørgensen and Hans Ahlgren authors contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. Molecular Ecology Resources published by John Wiley & Sons Ltd.

influential factors and should therefore be considered when selecting samples for large-scale ancient genome studies.

**KEYWORDS**

aDNA, DNA damage, endogenous content, pinnipeds, sample age, seal, walrus, zooarchaeology

## 1 | INTRODUCTION

Ancient DNA (paleogenetics, or aDNA) has seen numerous developments over recent decades and is increasingly integrated into interdisciplinary studies (e.g., Cappellini et al., 2010; Lazaridis et al., 2016; Moss et al., 2016; Raghavan et al., 2014; Star et al., 2018). Although not without challenges, paleogenetics and more recently, paleogenomics (focusing on the entire mitochondrial or nuclear genome rather than smaller targeted regions) offer exciting and unique opportunities to understand an organism's past, including historic and ancient human-animal-environmental interactions. Contemporary genetic studies on modern organisms have allowed us to explore evolutionary processes, phylogenetic relationships, species' physiology, response to human activities and environmental change. However, paleogenomics allows us to answer many previously intractable questions by delving into the past with greater accuracy and detail than ever before.

Improvements in both laboratory and bioinformatic methodologies have enabled paleogenomic studies to focus on an increasingly diverse range of species and to focus on entire populations or species, rather than individuals. This has revealed, and will continue to reveal, a wealth of hitherto unknown information about numerous animals, plants and micro-organisms including: the distribution and structure of past populations (e.g. Brandt et al., 2018; Palkopoulou et al., 2018), the characteristics and fate of now extinct taxa (e.g. McLeod et al., 2014; Scheel et al., 2014), the timing of key demographic events (e.g. Cole et al., 2019; Markova et al., 2015), the details of past gene flow (e.g. Barlow et al., 2018; Cahill et al., 2018), as well as the origin of particular adaptations (Ramos-Madrigal et al., 2016; Sandoval-Castellanos et al., 2017). Additionally, palaeogenomic analyses on faunal remains offer new insights into human-animal interactions, trade networks and human cultural histories (e.g. Bro-Jørgensen et al., 2018; Keighley, Pálsson, et al., 2019; Larson et al., 2007; Star et al., 2017). Such information is often not discernible from modern material as genetic signatures of the past are increasingly likely to be lost with time, particularly when lineages go extinct (e.g. Keighley et al., 2019; McLeod et al., 2014; Palkopoulou et al., 2018), populations undergo bottlenecks (e.g. Alter et al., 2012; Palkopoulou et al., 2015) or are subject to selective sweeps (Foote et al., 2012; Leonardi et al., 2017). These processes all contribute uncertainty and a lack of resolution for analyses or projections based upon modern samples.

While the benefits of population or species-scale paleogenomic studies are realised, large palaeogenomic data sets remain relatively uncommon, particularly outside of human genetics. A major

ongoing challenge remains the unpredictability of sample preservation, which impacts on both the quantity and quality of target DNA yield through endogenous content and damage profiles respectively. It is common for paleogenomic studies to initially screen samples to select the most promising material for deeper, whole-genome sequencing. However, the investment in time and resources of such laboratory and bioinformatic analyses can be significant. Our inability to currently predict the suitability of samples for whole-genome sequencing also means that a greater number of precious samples are subject to destructive sampling. This need for screening becomes particularly problematic when dealing with a large sample size or very precious material.

To date, much of our understanding of the relationship between various sample or environmental properties, and DNA quantity or quality, has been limited to widely-held assumptions and a limited number of empirical studies (e.g., Götherström et al., 2002). It is generally accepted that genetic degradation occurs over time due to structural and chemical modification of the inorganic and organic components of bone or teeth. These modifications are more likely when there have been fluctuations in surrounding temperature or moisture, overall warmer conditions, direct exposure to sunlight, particularly acidic or alkaline soils, inherently more porous bone materials (spongy or cancellous bone), or higher levels of microorganism activity (Allentoft et al., 2012; Bollongino et al., 2008; Kendall et al., 2017; Lindahl, 1993; Nielsen-Marsh & Hedges, 2000; Pruvost et al., 2007; Sosa et al., 2013; Trueman & Tuross, 2002).

In this study we used a large data set ( $n = 317$ ) of ancient and historic pinniped bones to understand the real-world relationships between a range of sample properties and sequencing success (as measured by endogenous DNA and damage profiles). Specifically, we sought to identify if any of the following variables might serve as a predictive characteristic for suitable genomic sequencing: sampling latitude, sample context, sample age, bone element type, collagen preservation, date of collection, species, sex and ontogenetic age of the sample. Three species from across the North Atlantic were used: the Atlantic walrus (*Odobenus rosmarus rosmarus* Linnaeus, 1758) ( $n = 177$ ), grey seal (*Halichoerus grypus* Fabricius, 1791) ( $n = 53$ ) and harp seal (*Pagophilus groenlandicus* Erxleben, 1777) ( $n = 87$ ). These samples were all drilled, extracted and prepared for shotgun sequencing on Illumina platforms under strict clean-laboratory conditions. The results are discussed in light of generally accepted theoretical principles of DNA and sample degradation to highlight which real-world sample properties are most important to consider during sample selection.

## 2 | MATERIALS AND METHODS

### 2.1 | Sample selection and contextual information

Samples were selected based on geographic location, extent of contextual information and macroscopic preservation quality (Table 1, Figure 1). Samples that appeared very porous or that had dirt embedded in the material were avoided. Porosity is likely to impact negatively on the quality and quantity of DNA obtained as bones with higher inherent porosity are prone to dissolution of the mineral and organic phase, and hence more susceptible to DNA degradation (Hedges et al., 1995). Furthermore, bones showing porosity as a consequence of structural degradation, often from microbial activity, commonly show higher contaminant loads (Gilbert et al., 2005). Samples with visible dirt were excluded, to avoid extracting any non-target DNA originated from organisms found in soil, such as fungi, bacteria or algae. Small samples were also excluded to ensure material remained for future investigation. Sample properties including latitude, geographic region, archaeological site, context (e.g. cave site), and chronology [age in years Before Present BP] were taken from archaeological reports, personal correspondence and museum catalogues. Radiocarbon dated samples were preferentially chosen, followed by those excavated from a context with other directly dated terrestrial mammal remains (e.g., reindeer, *Rangifer tarandus*). Samples for which no radiocarbon data were available were instead

assigned to time periods reflecting the cultural period that the sample was excavated from (e.g., Pre-Dorset or Thule). The median of either the calibrated radiocarbon probability date distribution (see section 2.2.), or the cultural period range, were used to provide a single-point quantitative “year BP” for statistical analysis. In the rare cases where the cultural period could not be defined the sample was excluded or marked as “time unknown”.

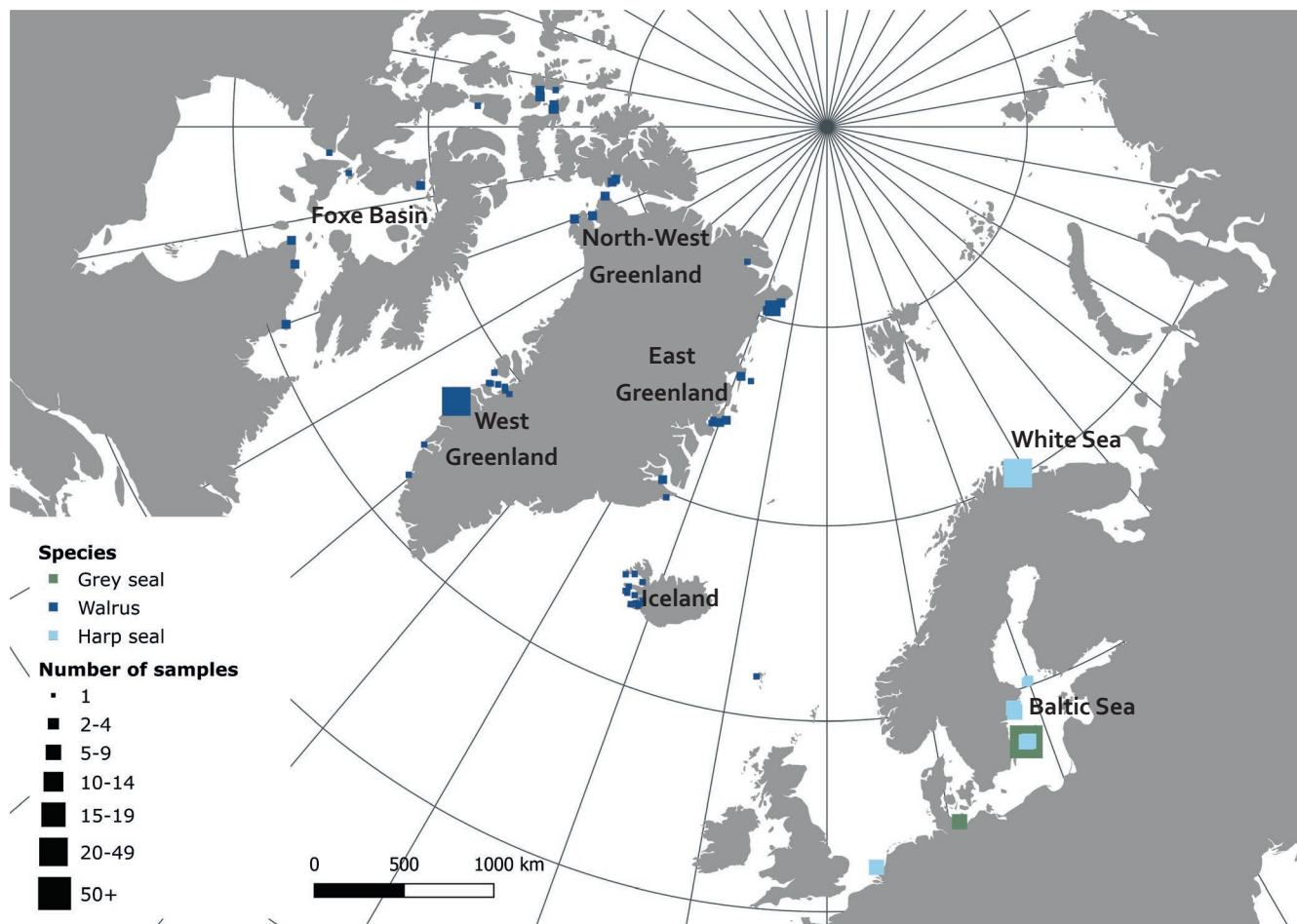
Any information on skeletal element, sex or ontogenetic age class (juvenile or adult) from zooarchaeological examination was recorded. When sample types were ambiguous or deemed “unidentifiable” they were classified as “fragmentary”. Across all samples, most had unknown sex or ontogenetic age. Additional sexing was performed using genetic methods (Bro-Jørgensen et al., 2019).

### 2.2 | DNA extraction and sequencing of walrus

Atlantic walrus samples derived from existing collections held at various institutes including the Icelandic Institute of Natural History, National Museum of Iceland, Natural History Museum of Denmark, National Museum of Denmark, Canadian Museum of History and Canadian Museum of Nature. Samples were selected based on geographic location, extent of contextual information and macroscopic preservation quality. All 177 walrus samples represent a collection amassed from 66 different localities (mostly archaeological

**TABLE 1** Characteristics of the successfully sequenced 285 (total 317) ancient and historical samples analysed for the three study species walrus, harp seal and grey seal collected across the Arctic and northern Europe

Species	Walrus	Harp seal	Grey seal
Number of individuals	154 successful (177 total)	78 successful (87 total)	53 successful (53 total)
Bone elements sampled	Baculum, digit, fragments, limb, mandible, rib, scapula, skull, tooth, tusk, vertebrae	Limb (42), skull (31), mandible (7), fragmentary (5), scapula (3)	Skull (auditory bulla), unknown (1)
Samples from cave site?	No	Yes (8)	Yes (52)
Age range (BP)	12–c. 15,000 mean = 1,691	1,800–4,950 mean = 4,492	4,950–9,450 mean = 9,034
Latitudinal range	60°56' N–82°11' N	52°16' N–70°10' N	57°17' N & 70°10' N
Longitudinal range	07°19' W–97°16' W	03°50' E–28°53' E	17°58' E & 28°53' E
Endogenous content (%) range	0–63.3 mean = 9.6%	0–44.7 mean = 2.8%	0.4–73.1 mean = 56.0%
N:Mt ratio range	0–14,273 mean = 777	62–11,095 mean = 1657	4–85 mean = 18
Damage G–A range	0–0.2823 mean = 0.07	0.1063–0.4865 mean = 0.24	0.1985–0.3755 mean = 0.32
Damage C–T range	0–0.2903 mean = 0.08	0.1188–0.4813 mean = 0.25	0.1975–0.3843 mean = 0.32
Extract concentration range (pM)	1–3998 mean = 294	Unknown	Unknown
Sequencing platform	Illumina HiSeq2500 (150PE), HiSeq4000 (80SR), MiSeq (100SR & 150PE)	HiSeq X, HiSeq 2500 (125PE), NovaSeq SI	Illumina HiSeq2500 (125PE) & NovaSeq SI
Extraction performed on chunk or powder	Chunk & powder	Powder	Powder



**FIGURE 1** Map of the distribution of the 285 successfully sequenced pinniped samples analysed. Symbol colour corresponds to species (dark blue for walrus, green for grey seal and light blue for harp seal). Sizes of symbols reflect the number of samples, clustered within sample size bins. Please note that symbols are placed at each site included in this study, but samples may have been taken from several different contexts within a single site. The exact details of this are included in the supplementary raw data

sites), across the Atlantic walrus' historic distribution including: the Canadian Arctic, Greenland, Iceland, Svalbard and northern Europe. Samples were taken from dated geological finds or archaeological contexts dating to broad cultural periods; Pre-Dorset (approximately 4950–2500 cal BP), Dorset (approximately 2500–650 cal BP), Thule (approximately 650–300 cal BP) or Historic (<300 cal BP) (Friesen & Mason, 2016). For samples that were directly dated, or those taken from contexts with contemporaneous dated terrestrial remains, calibrated radiocarbon probability distributions were taken from relevant publications and site reports. For unpublished dates, calibrations were made using INTCAL 13 (terrestrial) and Marine 13 (marine) (Reimer et al., 2013) using CALIB (Stuiver et al., 2019) or OxCAL (v4.1) (Bronk Ramsey, 2010) software. Where available local  $\delta R$  values were taken into account. Given the time periods under consideration here and the current lack of specific values for many geographic regions, the lack of  $\delta R$  for a small subset of samples should not have a major bearing upon the results of this study.

Walrus samples were prepared under strict clean laboratory conditions in a dedicated ancient DNA (aDNA) laboratory following expected guidelines (Cooper & Poinar, 2000; Gilbert et al., 2005), at the Animal Clean Laboratory, GLOBE Institute, University of

Copenhagen, Denmark. Specifically, all preamplification laboratory work was completed in a physically distinct locality and negative controls were maintained throughout the drilling, extraction, library build and amplification process (all were found to be free of any walrus DNA). DNA fragment length and damage patterns were analysed after sequencing, although as sequencing read length was limited in some cases by sequenced read length the reported value does not necessarily reflect the actual sample DNA fragment length. All equipment and materials used in the laboratory were sterilised with UV, bleach and ethanol. Drilling of 100–220 mg of bone into powder or small bone chunks was completed using a Dremel hand drill (Micro 8050) or an Osada dental drill (OS-40) and dental rosenbor drill bits (sizes 012–031). The surface was initially removed with a drill piece, the surrounding area wrapped in foil or parafilm, and then powder drilled using a new drill piece. Where possible, sampling was undertaken on cortical bone and the cementum of teeth/tusks. In some cases a small section was cut using a diamond cutter attached to the dremel tool. These chunks were cut into smaller pieces and typically weighed 200–300 mg. All drilling was completed at the lowest speed 2,000–5,000 rpm (depending on the tool), with frequent pauses to prevent the bone from overheating.

Powders and pieces of bone, tooth or tusk for all walrus samples were extracted following the protocol of Dabney et al. (2013). Bone chunks underwent a series of initial dilute bleach washes prior to extraction to increase endogenous DNA content (Boessenkool et al., 2016). Extraction concentrations were determined using a High Sensitivity TapeStation (Agilent Technologies), with yields converted to total Molarity (nM) and divided by material weight (mg) used in extraction. Illumina libraries were built following the BEST protocol (Carøe et al., 2018) and amplified following Barnett et al. (2018). Cycle number was determined based on extract yield and from the point of plateau using an Mx qPCR, where 1  $\mu$ l of water was substituted for 1  $\mu$ l of SYBRgreen fluorescent dye from the indexing reaction. Amplified libraries were purified and dual size selected using SPRI beads to remove adapter dimers and long fragment length non-target DNA when present (ratios 0.5x and 1.6x, targeting fragment lengths 60–600 base pairs [bp]). Samples for which amplified library was detected using a second High Sensitivity TapeStation were then pooled for sequencing. Low yielding samples were not sequenced when only primers and dimers were visible on the electrophoresis. A minimum of 12 libraries were pooled together with compatible 6 bp hexamer motifs single-indices with at least two mismatches. All samples sequenced on the HiSeq 4000 were dual-indexed with compatible 6 bp hexamer motif indices to avoid incorrect sample assignment due to index-hopping (van der Valk et al., 2019). Sequencing was conducted at the Danish National High-Throughput Sequencing Centre on a range of Illumina technologies MiSeq (100SR, 150PE), HiSeq 2500 (150PE) and HiSeq 4000 (80SR). To avoid any bias from laboratory preparation or sequencing technology samples were randomised across sites and ages for extraction, library build and sequencing. Walrus extracts were prepared by a single author; however, samples have been assigned into two “batches” for statistical analyses (see below), according to which of two authors built libraries.

### 2.3 | DNA extraction and sequencing of grey and harp seal

Grey seals were sampled from the collection of the Swedish Museum of Natural History. The majority of the grey seal samples were from the archaeological cave site Stora Förvar, near Gotland, Sweden. Radiocarbon dates for these samples have been previously published by Lindqvist and Possnert (1997). Harp seals were sampled from the collection of the Swedish Museum of Natural History, Ålands Landskapsregering Museibyån, Finland, and the University Museum of Bergen, Norway. The harp seal samples represent a total of 15 archaeological sites throughout the Baltic region, the North Sea and the Varanger Fjord, Norway. Information concerning radiocarbon dates, ontogenetic age class of samples and a detailed description of DNA laboratory work is included in Ahlgren et al. (2021).

Samples were prepared in the clean laboratory of the Archaeological Research Laboratory, Department of Archaeology

and Classical Studies, Stockholm University, Sweden. Between 130–150 mg of bone powder was used for extraction. Unlike the walrus samples, all grey and harp seals were extracted using a modified version (Ahlgren et al., 2021) of the protocol C of Yang et al. (1998). Double stranded DNA libraries were then prepared using the method described in Meyer and Kircher (2010). Size selection was performed using Ampure beads to 100–300 bp before samples were pooled and sent for sequencing at the SciLifeLab Uppsala, Sweden. All but one of the grey seal samples were sequenced on the Illumina HiSeq2500 platform (125PE). A selection of the harp seal samples were sequenced on Illumina HiSeqX (150PE), while the remaining samples were sequenced on NovaSeq S1 (150PE). Throughout the laboratory processing, grey seals were randomised according to age (they could not be further randomised given that all but one sample came from the same site and skeletal element). Harp seals were not fully randomised; however, samples prepared together during extraction, library and amplification have been assigned as a single “batch” for the purposes of statistical analyses to avoid any confounding effect driving the results (see below).

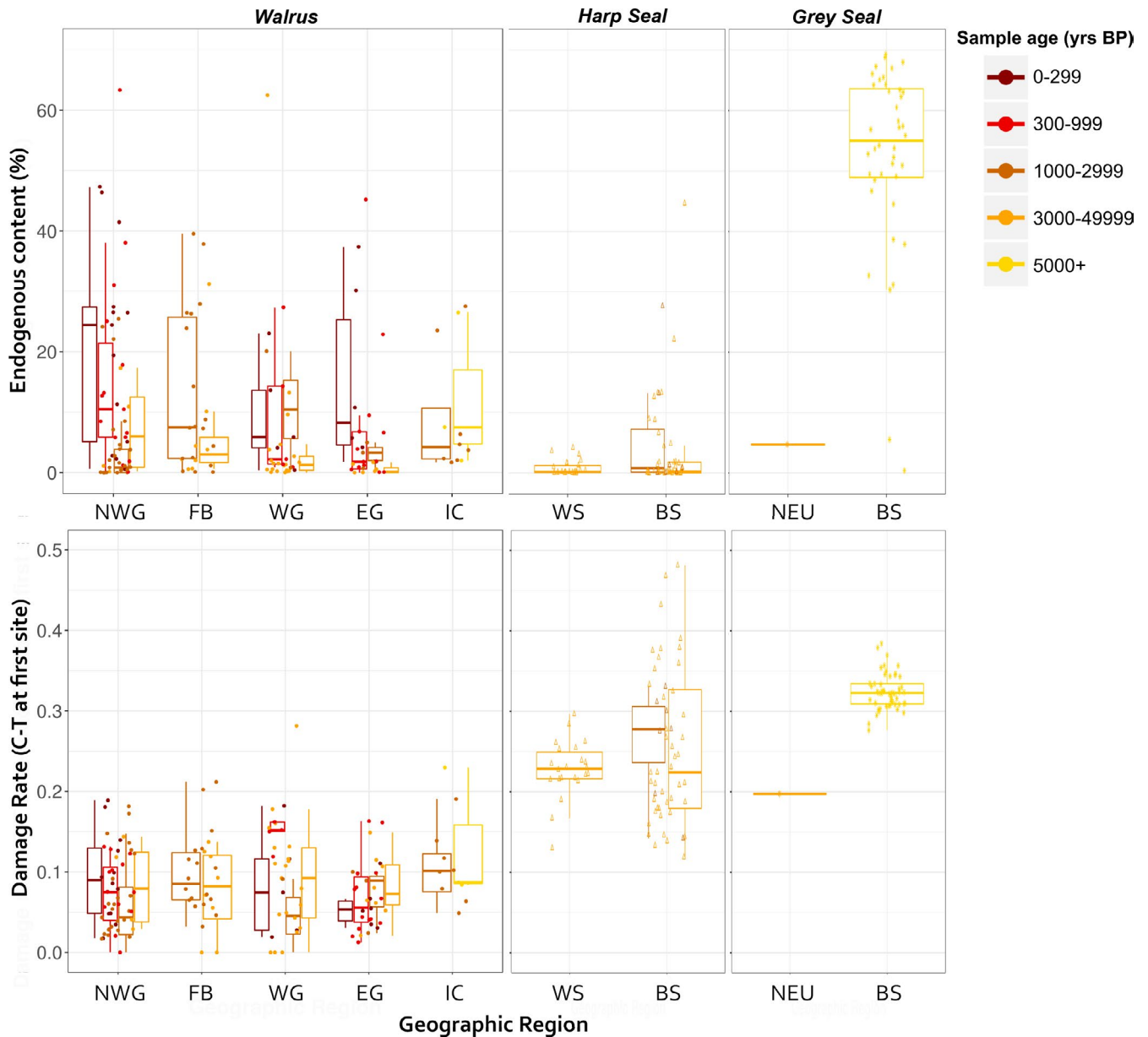
### 2.4 | Collagen extraction

Collagen was extracted from a subset of the walrus and harp seal specimens to provide an additional comparative measure of sample preservation. Bone chunks weighing ~200 mg were demineralized in 0.5 M HCl at 4°C. Samples which were dark in colour were rinsed to neutrality with type I water, then treated with 0.1 M NaOH for successive 30 min treatments with sonication to remove humic contaminants until the solution ran clear. Samples were again rinsed to neutrality with type I water, then the collagen residue was refluxed at 75°C in 0.01 M HCl (pH ~3). After heating, the solution containing the collagen was filtered using a 5–8  $\mu$ m filter to remove insoluble residues, and then with a Pall Microsep 30 kDa ultrafilter. The >30 kDa fraction was freeze dried and the collagen yield was calculated (extracted collagen/initial bone mass). Elemental compositions were determined using an IsoPrime continuous flow isotope-ratio mass spectrometer coupled to a Vario Micro elemental analyser with a glutamic acid standard (USGS40) used to calibrate the measurements.

### 2.5 | Data analysis

#### 2.5.1 | Bioinformatics

All DNA reads were trimmed, filtered and aligned to their closest respective reference genome using the PALEOMIX (v1.2.13) pipeline (Schubert et al., 2014). This utilised the following software: SAMTOOLS (v1.3.1) (Li et al., 2009), BWA (v0.7.15) (Li & Durbin, 2009), ADAPTERREMOVAL (v2.2.0) (Schubert et al., 2016), PICARD Tools (Broad Institute, n.d.) and MAPDAMAGE (V2.0.6) (Jónsson et al., 2013). The walrus reads were aligned to the walrus nuclear genome (NCBI



**FIGURE 2** Effect of the samples' geographic origin on endogenous DNA content (top row) and DNA damage rate (bottom row) represented as box-plots for each of the three study species. Geographic origin was summarised as one of seven regions: North West Greenland (NWG), Foxe Basin (FB), West Greenland (WG), East Greenland (EG), Iceland (IC), Neustadt (NEU), Varanger Fjord (WS) and Baltic Sea (BS). Sample points and boxes have been colour-coded to indicate the age of the sample (years BP) within one of five categories. Note that some geographic regions are represented by a single or limited number of samples

accession: GCA\_000321225.1), with the Atlantic walrus mitochondrial genome (NCBI accession: NC\_004029.2) as a "region of interest". Grey seal reads were aligned to the grey seal nuclear genome (Savriama et al. 2018) with a separate grey seal mitochondrial genome (NCBI accession: NC\_001602). Harp seal reads were aligned to the grey seal nuclear genome (Savriama et al., 2018), and the harp seal mitochondrial genome (NCBI accession: KP942581). PALEOMIX and mapDamage outputs provided information concerning summary statistics, endogenous content, clonality and damage level. Damage levels were estimated as the rate of cytosine to uracil (thymine) transitions at the first base of the 5' strand end. Finally, mitochondrial:nuclear ratios were calculated

as mitochondrial coverage divided by nuclear coverage following Hansen et al. (2017).

## 2.5.2 | Data

Successful DNA extraction, amplification and sequencing was possible for 154 of 177 walrus, 78 of 87 harp seals and all 53 grey seals, resulting in a total sample size of 285 animals (Table 1). The other 32 samples yielded insufficient amplified library for sequencing.

For the 285 samples that were successfully sequenced, we selected both endogenous content and damage levels to explore

**TABLE 2** *p*-values for multivariate analyses indicating the significance of the effect of each explanatory variable (by row) on the four measures of sequencing success (by column). Analyses were performed on all samples for eight explanatory variables and for walrus samples 10 explanatory variables according to data availability.

	Endogenous				Damage			
	All species	Walrus	Harp seal	Grey seal	All species	Walrus	Harp seal	Grey seal
Age BP	0.067	0.397	0.178	0.725	0.230	0.037*	0.593	0.581
Latitude	0.079	0.300	0.483		0.010**	0.001***	0.128	
Chunk	0.017*	0.150			0.982	0.092		
Species	0.022*				0.027*			
Element	0.032*	0.017*	0.142		0.316	0.068	0.130	
Cave	4.49E <sup>-05</sup> ***		2.11E <sup>-05</sup> ***		0.013**		0.014*	
Collection		0.006**				0.388		
Powder		0.146				0.114		
Extraction		0.648				0.337		
Mt:NuDNA	0.171	0.053	0.268	0.018*	0.821	2.20E <sup>-16</sup> ***	0.282	0.479
Damage	0.134	0.459	0.988	0.884				
Endogenous					0.898	0.737	0.310	0.505

Levels of significance are indicated with the use of \*(0.05 > *p* > 0.01), \*\*\*(0.01 > *p* > 0.001), \*\*\*(*p* < 0.001).

trends relating to sample age, geographic region, latitude, bone element type, ontogenetic age, sex, excavation date, context, as well as sample weight and bone element included in extraction.

### 2.5.3 | Single-factor statistical analyses

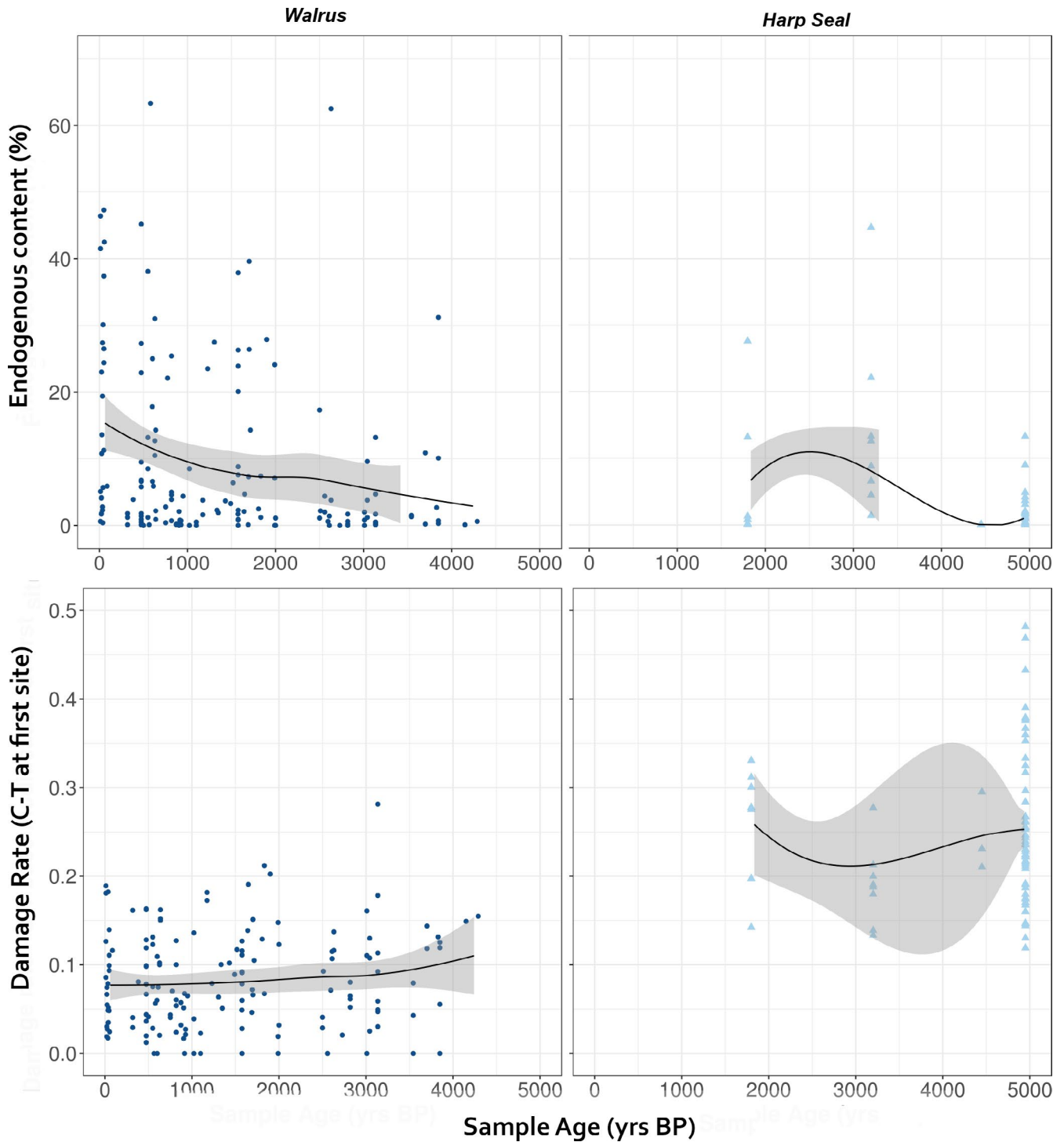
Initial relationships between explanatory variables and measures of sample success were examined visually. All graphs were created in R v.3.5.1. (R Core Team, 2018) using ggplot2 (Wickham, 2016) and lattice (Sarkar, 2008), and maps using QGIS v.3.4.2 (QGIS Development Team, 2018). In some cases, specimens with uncharacteristically high values for one of the four measures of sample success were omitted from the graphical representation as outliers, but are listed in Tables S1 and S2. Statistical analyses were completed separately according to species due to the lack of overlap for numerous variables across species. The main species of focus were walrus and harp seals due to the large effect of cave context for the grey seal samples (see below). Analyses with the total data set were also completed; however, their results should be interpreted with caution given the potential for confounding factors to be driving observed trends.

### 2.5.4 | Multifactorial statistical analyses

Generalised linear mixed-effect models (GLMM) were conducted for the combined three-species data, as well as the walrus and harp seal data separately, following Winter (2013). GLMMs were chosen to test the significance of any observed trends without the influence of potential confounding variables and to account for the combination of discrete, binary and continuous variables. GLMMs were

repeated for both endogenous content and damage levels. Each explanatory variable was included as a fixed effect. Random effects and random slopes were chosen to correspond with the sample success measure under consideration and to avoid overparameterising the model resulting in insufficient degrees of freedom (Table S3). When all three species were included there was a reduction in the number of fixed and random effects to account for missing data. The greater amount of information available for walrus samples allowed a greater number of fixed and random effects to be tested. However, not all explanatory variables were included as fixed effects when these were found to be nonindependent (e.g., latitude and longitude was reduced to the most important measure, latitude). In some cases data were only available for a very limited number of samples. Residual plots were examined visually prior to further analysis to ensure there were no obvious deviations from assumptions of normality or homoscedasticity. In some cases, these assumptions were found to be violated and the sample success measure was either log or square root transformed, thereby resolving these concerns (Table S3). *p*-values were obtained from likelihood ratio tests by comparing the likelihoods using ANOVA of the overall model and one which excluded the fixed effect under analysis. Grey seals were analysed using two generalised linear models (GLM) for both endogenous content and damage levels. This approach therefore differed from that used for walrus, harp seal and the combined three-species data outlined above. GLMs were used for grey seals instead of GLMMs as data concerning many of the fixed and random effects were not available, or there was no variation (i.e. all elements were auditory bulla and hence cannot be compared). The exact parameters of each GLM and GLMM can be found in Table S3. All statistical analyses were performed in R, using the lme4 package (Bates et al., 2015) and standard packages.





**FIGURE 3** Effect of sample age (years BP) on endogenous content (top row) and damage rate (bottom row). Plots have been separated and points coloured according to species. Smoothed trend lines with Standard Error (shaded area) are shown for each plot. Three walrus samples older than 5000 years BP have been excluded (see Figure S1 for the full version)

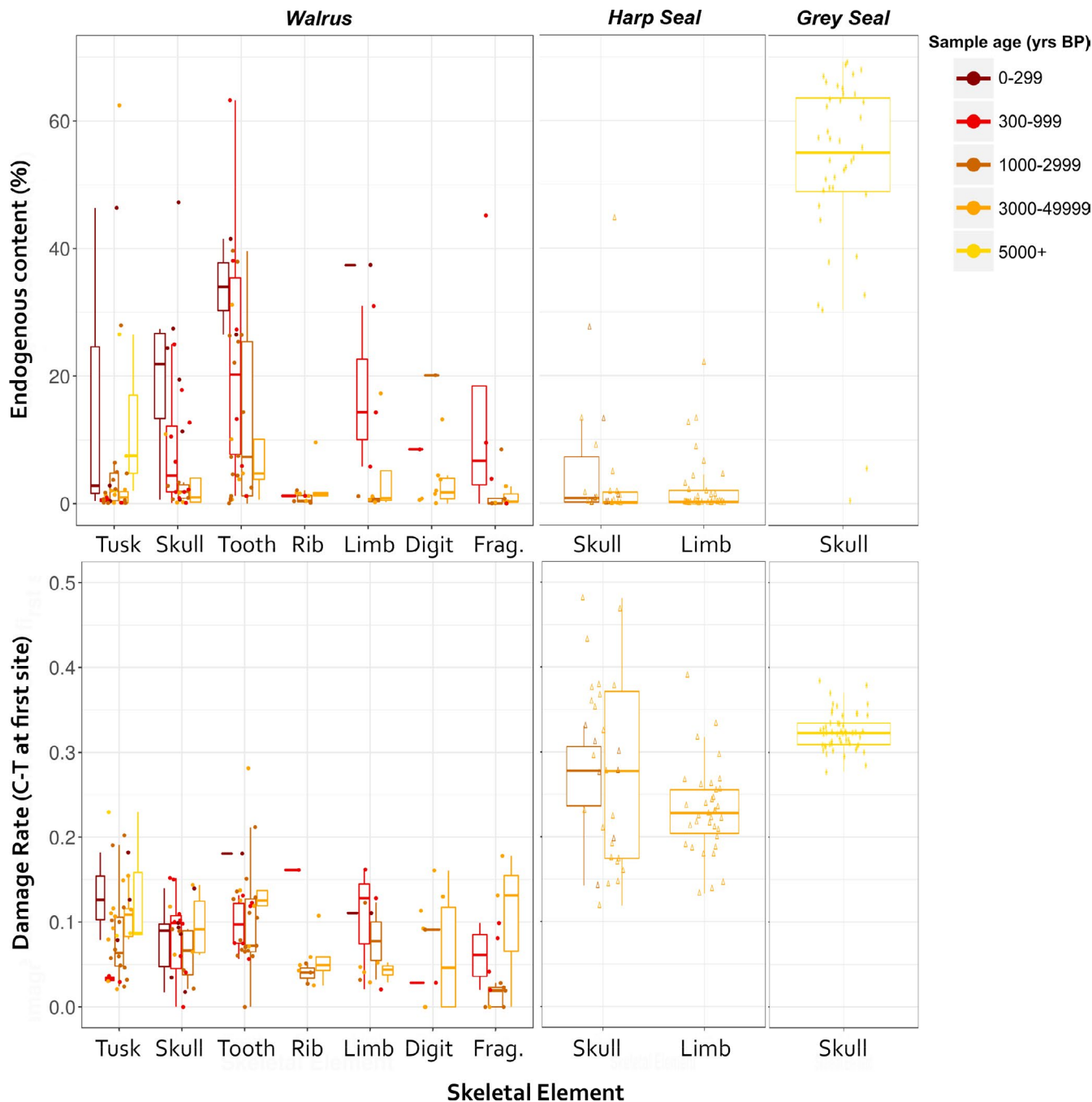
### 3 | RESULTS AND DISCUSSION

#### 3.1 | Effect of sample properties on sample success

##### 3.1.1 | Geographic region, latitude and context

The geographic region in which a sample is found has previously been linked to the preservation of organic material, and DNA more

specifically, most likely through a combination of climate, solar radiation, soil and depositional conditions (e.g., Allentoft et al., 2012; Bollongino et al., 2008; Higgins et al., 2015). When the walrus, harp seal and grey seal samples were grouped into broader geographic regions there was a general pattern for higher endogenous content and lower damage rates in samples collected from Arctic sites (both in Greenland and Canada), as compared with samples from sub-Arctic or temperate sites (e.g., Iceland and the Baltic Sea) (Figure 2, Table 2).

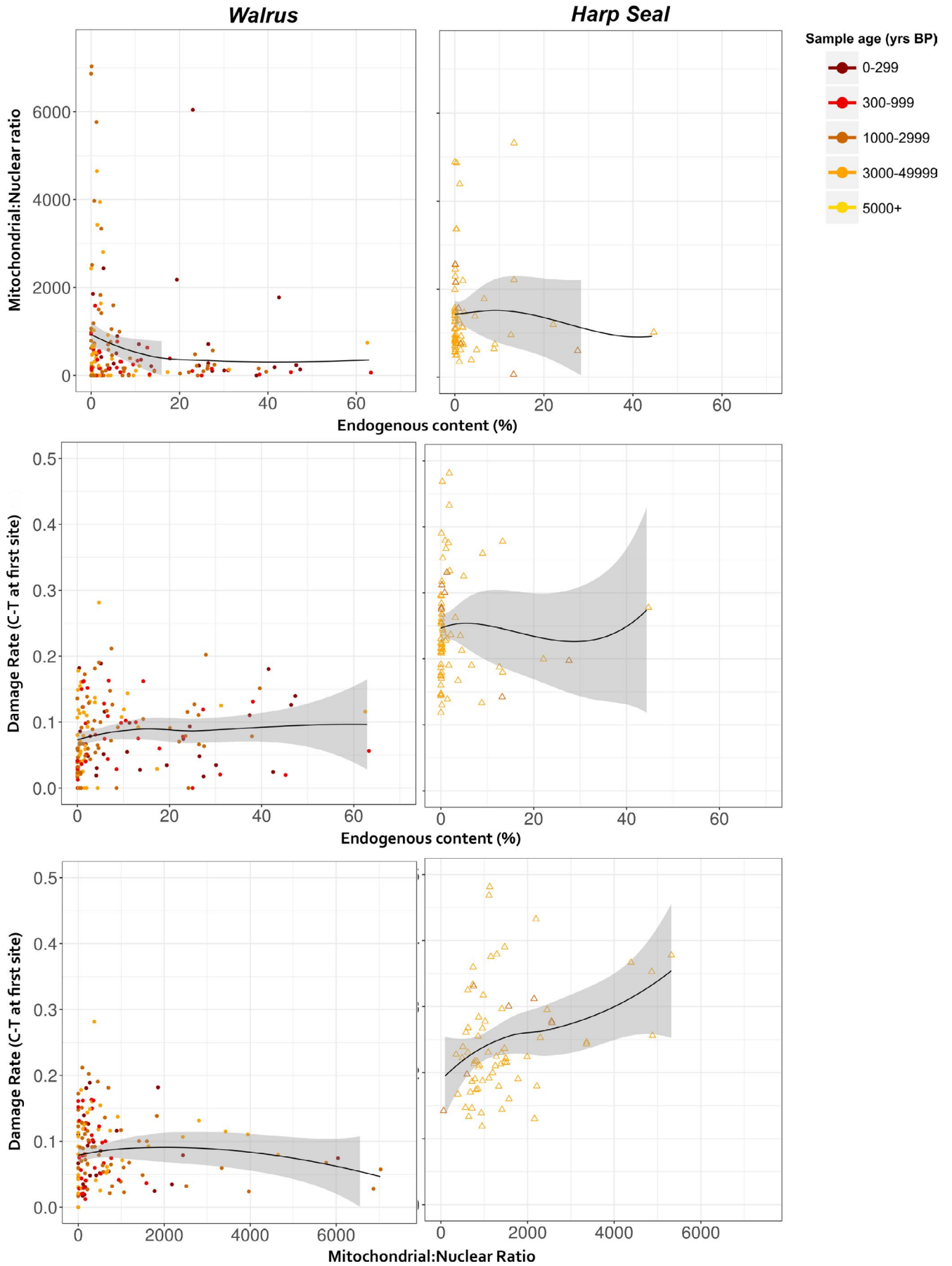


**FIGURE 4** Effect of skeletal element on endogenous DNA content (top row) and DNA damage rate (bottom row) represented as box-plots. Plots have been separated according to species, and skeletal elements grouped together into simple categories. Sample points and boxes have been colour-coded to indicate the age of the sample (years BP) within one of five categories. Element categories for which there were fewer than two time periods containing at least two samples were excluded

There was one notable exception to this pattern; unexpectedly high values of endogenous DNA content combined with high DNA damage levels for the oldest (>5,000 years BP) Baltic Sea seal samples from cave contexts (Figure 2). This was particularly so for grey seals from the cave site Stora Förvar, Sweden, which had striking 30%–65% endogenous contents. For harp seals, the effect of cave context on sample preservation was supported by the multivariate analyses with highly significant relationships for both endogenous ( $p < 0.05$ ,  $\chi^2(2) = 18.1$ ) and damage profiles ( $p = 0.014$ ,  $\chi^2(2) = 0.5$ ).

These findings agree with a number of previous studies which have found cave conditions to yield high quality samples (e.g., Bollongino et al., 2008; Höss et al., 1996; Nielsen-Marsh & Hedges, 2000). In contrast, the majority of walrus samples were predominately surface finds or from shallow terrestrial excavations (profiles often <15 cm).

To examine the effect of geography quantitatively, latitude was included in multivariate analyses (Table 2). Latitude was found to have a highly significant relationship with damage in walrus ( $p = 0.001$ ,  $\chi^2(2) = 10.8$ ). Additionally, there were mixed,



**FIGURE 5** Relationships between walrus and harp seal sample endogenous content, damage rate and mitochondrial:nuclear ratio. Dots have been colour-coded to indicate the age of the sample (years BP) within one of five categories. A smoothed trend line with a shaded area either side represents standard error. Three walrus samples older than 5000 years BP were excluded

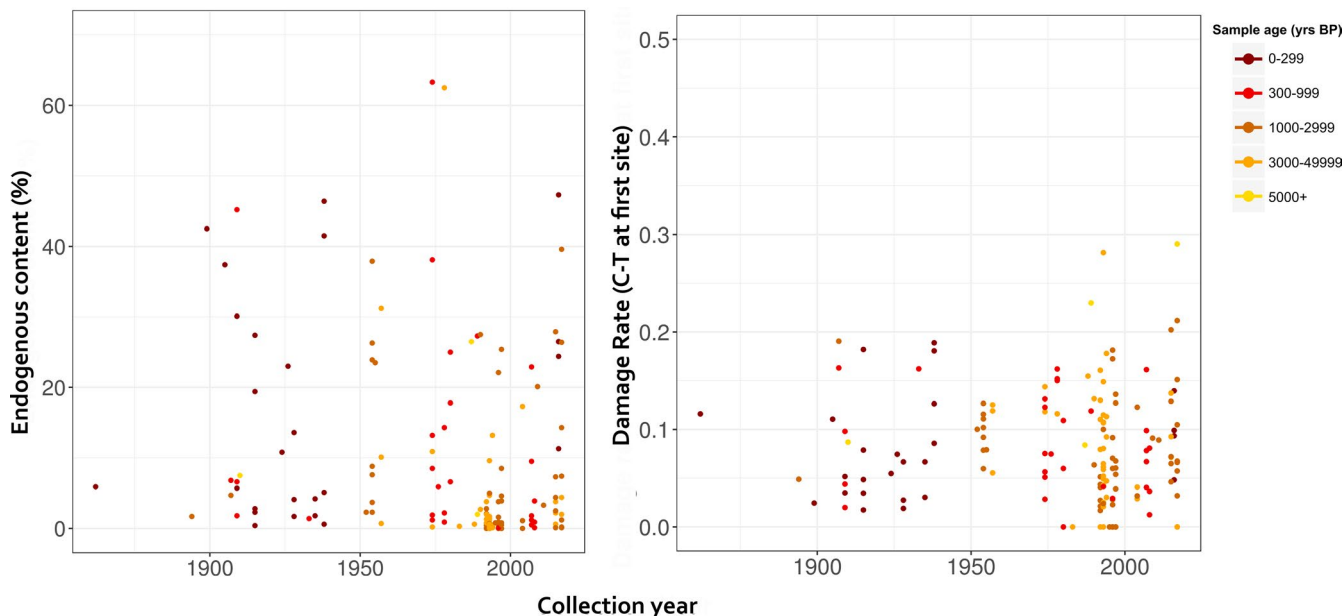
nonsignificant relationships between latitude and both endogenous DNA content and damage rate for the other species. For example, walrus samples that were <1,000 years old showed the expected decline in endogenous content with declining latitude; however, this trend was reversed or nonexistent for older samples. The absence of strong support for the expected relationship between latitude and endogenous content may be due to the more limited range of latitudes considered in this study. A link might only emerge across large geographical areas, such as between the Arctic and the tropics (Bollongino et al., 2008; Kendall et al., 2017; Smith et al., 2001, 2003; Sosa et al., 2013). Previous studies have also revealed a certain degree of overlap between amplifiable DNA from Arctic and temperate conditions (Pruvost et al., 2007). Furthermore, the exceptional preservation conditions described for the most northerly latitudes (e.g. Campos et al., 2010; Langeveld et al., 2017; Shapiro, 2004) may only be of benefit at much longer time-scales of tens of thousands of years (e.g., Orlando et al., 2013; Schwarz et al., 2009). For comparatively younger samples as in this study, the slow rate of soil deposition and annual freezing-thawing cycles may make them more vulnerable than initially expected (Kendall et al., 2017; Ping et al., 1998).

### 3.1.2 | Sample age

Sample age has been considered as a primary agent of DNA degradation despite the exact mechanisms remaining elusive (Allentoft

et al., 2012; Campos et al., 2012; Hansen et al., 2006; Sawyer et al., 2012). The process of nuclear DNA loss has been shown to occur exponentially in the initial phase post-deposition at approximately 2–2.5 times that of mitochondrial DNA loss. The greater preservation of mitochondrial DNA is attributed to the circular structure, double membrane and potential differences in the level of enzymatic activity (Allentoft et al., 2012; Higgins et al., 2015; Schwarz et al., 2009). Our study did not support a universal pattern of declining endogenous content with age (Figure 3, Table 2, Figure S1), with only walruses showing a statistically significant correlation between age and damage rates ( $p = 0.037$ ,  $\chi^2(2) = 4.3$  respectively). Interestingly, although not supported by the GLMMs, walruses also appear to show a negative relationship between age and endogenous content when comparing samples from within the same geographic region (Figure 2, Figure 3). The lack of a definitive and universal relationship between chronological age and DNA preservation may indicate the need for a more nuanced consideration of sample age. Indeed, Smith et al. (2003) combined the climatic conditions (or “thermal history”) of a sample’s context with the absolute chronological age of the sample to calculate “thermal age” as a more useful predictor of DNA survival or degradation.

Walrus samples included in this study suggest that age may be a driving factor within more homogenous sample sets, rather than as the clear, universal factor as is commonly thought. Indeed, focusing on radiocarbon dated samples of the extinct Moa (*Dinornithiformes*) Allentoft et al. (2012) were able to demonstrate that the importance of sample age on DNA preservation varied considerably according to



**FIGURE 6** Effect of walrus sample collection date on endogenous content (left) and damage rate (right). The points are colour-coded based on sample age separated into time periods

site. Similar site specific patterns may explain the lack of correlation between DNA yield and sample age reported in other studies (e.g., Campos et al., 2012; Hagelberg et al., 1991; Haynes et al., 2002). Also, the expected trend in declining structural and genetic integrity of skeletal materials through time may emerge over longer time periods than covered in our study (Gilbert et al., 2005).

### 3.1.3 | Bone element type

Bone element type was only found to have a significant relationship with endogenous content in walruses ( $p = 0.017$ ,  $\chi^2(2) = 24.5$ ). Within walruses, endogenous content was highest in teeth and skulls and bones such as ribs yielded poorly (Figure 4, Table 2). This agrees with expectations that certain elements such as porous ribs or thin-walled scapulae should be avoided if possible (Bollongino et al., 2008; O'Connor, 2008; Parker et al., 2020). The lack of correlation between bone element and DNA preservation in grey and harp seals may be a consequence of the limited diversity of bone elements available for sampling. Compared with walruses that are more easily distinguishable based on size, bones from the smaller seal species were included if they were diagnostic to species level. Therefore, certain bone elements, particularly fragmentary or less characteristic bones were not included. Additionally, where possible the same bone element was preferred to avoid resampling the same individual. In grey seals this resulted in only the pars mastoideus being sampled.

Although high endogenous content is commonly found in the dense pars petrosa (petrous bones) (Gamba et al., 2014; Pinhasi et al., 2015), such material was extremely rare to find from the sites included in this study, and in the few instances where it was available it was deliberately avoided as it constitutes an important diagnostic marker in comparative morphology.

In contrast to endogenous content, damage rates showed no clear pattern across skeletal elements when sample age, cave samples and species were taken into consideration (Figure 4). Grey and harp seals showed higher damage profiles than that for walruses, however as discussed below (see subsection on species), it is not possible with current data to determine whether this is a species difference or due to other confounding factors.

### 3.1.4 | Collagen

Samples producing higher collagen yields approaching those expected for modern cortical bone (c. 20% by weight) were characterized by higher endogenous DNA content in walruses, although this trend did not hold when collagen yield exceeded 20%. There was no clear or consistent relationship between collagen content and damage rate for either walruses or harp seals. Several other studies have attempted to examine the relationship between the preservation of DNA and protein in ancient bone (Götherström et al., 2002; Schwarz et al., 2009; Scorrano et al., 2015; Sosa et al., 2013).

The extent of aspartic acid racemization was initially believed to be a minimally destructive screening technique for DNA preservation of ancient bones (Poinar et al., 1996), but more recent research has demonstrated that the quaternary structure of the collagen helix impedes racemization to such an extent that only highly degraded collagen undergoes sufficient levels of racemization for these measurements to be useful (Collins et al., 2009). DNA forms complexes with collagen helices (Svintradze et al., 2008) and possesses a strong affinity for bioapatite (the inorganic component of bone) (Okazaki et al., 2001). These interactions increase the likelihood that DNA will survive in the burial environment (Salamon et al., 2005). Bioapatite and collagen confer stability onto one another within bone (Nielsen-Marsh & Hedges, 2000), which further helps to explain the correlation between collagen and DNA preservation. If a greater amount of collagen is present, the mineral component of the bone is less likely to be altered, and endogenous DNA is more likely to preserve.

There was no clear relationship between the atomic C:N ratio of the extracted collagen and DNA preservation (Figure S2). The elemental compositions (wt% C, wt %N, C:N ratio) of collagen are widely used to demonstrate that stable isotope measurements reflect the endogenous isotopic composition of the collagen (Ambrose, 1990; DeNiro, 1985; van Klinken, 1999). Degraded collagen is more likely to produce C:N ratios that deviate from the theoretical value of 3.23 for unaltered mammalian collagen (Szpak, 2011), with a range of 2.9–3.6 being frequently cited as indicative of unaltered or “well preserved” collagen (DeNiro, 1985). The lack of a clear relationship between the atomic C:N ratio and DNA preservation may relate to the fact that all of the samples considered here were from geologically young contexts in environments that tend to be more favourable to collagen preservation (Collins et al., 2002).

### 3.1.5 | Ontogenetic age class and sex

There was no consistent trend with either ontogenetic age class or sex for both endogenous content and damage rates (Figure S3). As information concerning ontogenetic age and sex was available for a small subset of samples from walrus and harp seals the discussion of these findings is provided in Supporting Information Text Material S1.

### 3.1.6 | Sample processing and amount

There was no significant relationship between endogenous content or DNA rates and whether extractions were taken from bone chunks or fine powder for individual species (Figure S5, Table 2). For samples that were powdered, there was no significant relationship between the amount (weight) of powder used in extraction and either endogenous content or damage rate (Figure S4, Table 2). The lack of relationship between sample success and amount of extraction material suggests that any advantage of increasing the quantity of starting material is offset by saturation or inhibition of extracts.

Thus, given that samples are often precious, and we did not find a positive effect of increasing sample quantity, we recommend using low quantities of material (max. 120–140 mg) in line with recent extraction protocols (e.g. Dabney & Meyer, 2019; Gamba et al., 2016; Korlević et al., 2015).

### 3.1.7 | Study species

Differences in species could not be tested in this study, due to a lack of overlap between numerous sample characteristics and differences in laboratory processing according to species. For example, walrus samples were not found in either the Baltic or Varanger Fjord and grey seals were only sampled from auditory bulla. There was a trend for higher endogenous content in grey seals, followed by walrus and then harp seals (Table 2). Although this may represent true biological variation according to species (such as physiological differences in bone structure) it may also reflect differences in depositional histories (such as cooking), sample context (cave vs. surface) or the differences in laboratory methods. Further research including numerous species found at the same site and subject to the same laboratory methods is required to test this.

## 3.2 | Relationship between measures of sample success

Overall, we found that samples with high endogenous contents exhibited relatively more abundant nuclear than mitochondrial DNA in grey seals ( $p = 0.018$ ) and a similar trend in walrus, albeit not statistically significant ( $p = 0.053$ ) (Figure 5; Table 2). This correlation agrees with standard assumptions of DNA degradation and is conflicts with recent findings for human aDNA by Furtwängler et al. (2018). DNA extraction yield and damage rates did not correlate with endogenous content for any of the species. The absence of any relationship between endogenous content and damage rates is surprising, as one would assume that better preserved samples would have both high endogenous content and lower damage rates. With respect to extraction yield, even if it might be a good indicator of amplification success or sequencing output, this does not guarantee high endogenous content and hence the quality of the resulting DNA sequence data. This is probably due to other nontarget DNA such as soil bacteria and fungi being measured in the extract (Gilbert et al., 2005).

## 3.3 | Other possible factors affecting DNA degradation

A major limiting factor to understanding the relationships discussed in this paper is that sample or site characteristics that can affect sample degradation were not available for all samples. This is also true for the majority of zooarchaeological remains currently held

in research and cultural institutions. Potential factors may include soil pH, moisture or micro-organism activity. Importantly, many of the Arctic zooarchaeological finds included in this study are surface finds or from comparatively shallow excavations (profile <15–25 cm) (e.g. Pre-Dorset and Dorset deposits of Foxe Basin; Howse et al., 2019). Depth of archaeological profile has previously been documented as affecting amplification success of mitochondrial DNA (Bollongino et al., 2008), and shallow profiles are likely to lead to greater surface exposure and reduced microclimatic buffering as compared with much deeper profiles (Campos et al., 2012; Kendall et al., 2017; Todisco & Monchot, 2008). The amount of organic matter content and hence microbial activity in soils is also likely to affect DNA preservation (Nielsen-Marsh & Hedges, 2000). For example, skeletal preservation can vary significantly within and between sites according to soil pH conditions (Ovchinnikov et al., 2001). In northern Norway, bone material typically shows better preservation when deposited in shell sands (Hodgetts, 1999). Finally, samples from archaeological sites can experience surface exposure during periods of site reoccupation when deposits become disturbed, as described for the Canadian Arctic (Dyke et al., 2018; Habu & Savelle, 1994; Savelle & Habu, 2004).

While this study has largely focused on sample and site properties relating to the phase of degradation following deposition in sediment, so-called diagenetic processes, two other critical phases of degradation cannot be so easily examined for many of the samples included in this study (especially from older excavations). These are the initial rapid degradation affected by any treatment prior to incorporation into the sediments (e.g., burning, cooking, animal predation; known as perthotoxic processes), as well as degradation following excavation, transport, handling and storage (curatorial or taphonomic factors) (O'Connor, 2008). There is a lack of information concerning the impact of such stages despite indications that they can have considerable influence on sample preservation. Indeed, Pruvost et al. (2007) found that the handling and treatment of bones following excavation (e.g., washing and climatic control), can result in degradation rates 70x faster than throughout burial, a halving of endogenous content and substantial differences in contaminant load. The potential importance of taphonomic factors can be seen in the significant difference in collection date of material and endogenous content for walrus, where more recent finds show decreased endogenous content ( $p = 0.049$ ,  $\chi^2(2) = 3.89$ ) (Figure 6, Table 2). This may indicate that changes in excavation and curatorial treatments have not led to the expected improvement in sample preservation. However, it may also be a result of more comprehensive collection of faunal remains, as in recent decades a greater proportion of zooarchaeological bones are now returned from an excavation site.

While sample selection for future aDNA studies can be guided by the relationships highlighted in this paper, there is a need for greater documentation of excavation and curatorial practices that may affect the viability of molecular analyses. The taphonomic pathways prior to deposition can be characterised to a certain extent through detailed taphonomic studies (i.e., assessing degree of weathering, burning, gnawing marks, butchery traces







- Hofreiter, M., Bradley, D. G., & Pinhasi, R. (2014). Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, 5, 52–57.
- Gilbert, M. T. P., Bandelt, H.-J., Hofreiter, M., & Barnes, I. (2005). Assessing ancient DNA studies. *Trends in Ecology & Evolution*, 20(10), 541–544.
- Götherström, A., Collins, M. J., Angerbjörn, A., & Liden, K. (2002). Bone preservation and DNA amplification. *Archaeometry*, 44(3), 395–404.
- Habu, J., & Savelle, J. M. (1994). Construction, use, and abandonment of a Thule whale bone house, Somerset Island, Arctic Canada. *The Quaternary Research (Daiyonki-Kenkyu)*, 33(1), 1–18.
- Hagelberg, E., Bell, L. S., Allen, T., Boyde, A., Jones, S. J., & Clegg, J. B. (1991). Analysis of ancient bone DNA: Techniques and Applications [and Discussion]. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 333(1268), 399–407.
- Hansen, A. J., Mitchell, D. L., Wiuf, C., Paniker, L., Brand, T. B., Binladen, J., Gilichinsky, D. A., Ronn, R., & Willerslev, E. (2006). Crosslinks rather than strand breaks determine access to ancient DNA sequences from frozen sediments. *Genetics*, 173(2), 1175–1179.
- Hansen, H. B., Damgaard, P. B., Margaryan, A., Stenderup, J., Lynnerup, N., Willerslev, E., & Allentoft, M. E. (2017). Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLoS One*, 12(1), e0170940.
- Haynes, S., Searle, J. B., Bretman, A., & Dobney, K. M. (2002). Bone preservation and ancient DNA: The application of screening methods for predicting DNA survival. *Journal of Archaeological Science*, 29(6), 585–592.
- Hedges, R. E. M., Millard, A. R., & Pike, A. W. G. (1995). Measurements and relationships of diagenetic alteration of bone from three archaeological sites. *Journal of Archaeological Science*, 22(2), 201–209.
- Higgins, D., Rohrlach, A. B., Kaidonis, J., Townsend, G., & Austin, J. J. (2015). Differential nuclear and mitochondrial DNA preservation in post-mortem teeth with implications for forensic and ancient DNA studies. *PLoS One*, 10(5), e0126935.
- Hodgetts, L. M. (1999). Animal bones and human society in the late Younger Stone Age of arctic Norway. <http://etheses.dur.ac.uk/4491/>
- Höss, M., Jaruga, P., Zastawny, T. H., Dizdaroğlu, M., & Pääbo, S. (1996). DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research*, 24(7), 1304–1307.
- Howse, L., Savelle, J. M., & Dyke, A. S. (2019). New Insights from the dorset type site at Alarniq, northern Foxe basin, arctic Canada: Beach level chronology and site use. *American Antiquity*, 84(3), 531–548.
- Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., & Orlando, L. (2013). mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 29(13), 1682–1684.
- Keighley, X., Tange Olsen, M., & Jordan, P. (2019). Integrating cultural and biological perspectives on long-term human-walrus (*Odobenus rosmarus rosmarus*) interactions across the North Atlantic. *Quaternary Research*, 1–21. <http://dx.doi.org/10.1017/qua.2018.150>
- Keighley, X., Pálsson, S., Einarsson, B. F., Petersen, A., Fernández-Coll, M., Jordan, P., Olsen, M. T., & Malmquist, H. J. (2019). Disappearance of Icelandic walrus coincided with Norse settlement. *Molecular Biology and Evolution*, 36(12), 2656–2667.
- Kendall, C., Høier Eriksen, A. M., Kontopoulos, I., Collins, M. J., & Turner-Walker, G. (2017). Diagenesis of archaeological bone and tooth. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 491, 21–37.
- Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., & Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *BioTechniques*, 59(2), 87–93.
- Langeveld, B. W., Mol, D., Zazula, G. D., Gravendeel, B., Eurlings, M., McMichael, C. N. H., Groenenberg, D., van Reenen, G. B. A., Palmeira, M., Vogel, J., & van Geel, B. (2017). A multidisciplinary study of a Late Pleistocene arctic ground squirrel (*Urocitellus parryii*) midden from Yukon, Canada. *Quaternary Research*, 89, 333–351.
- Larson, G., Albarella, U., Dobney, K., Rowley-Conwy, P., Schibler, J., Tresset, A., Vigne, J.-D., Edwards, C. J., Schlumbaum, A., Dinu, A., Balaşescu, A., Dolman, G., Tagliacozzo, A., Manaseryan, N., Miracle, P., Van Wijngaarden-Bakker, L., Masseti, M., Bradley, D. G., & Cooper, A. (2007). Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proceedings of the National Academy of Sciences*, 104(39), 15276–15281. <https://doi.org/10.1073/pnas.0703411104>
- Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D. C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., Connell, S., Stewardson, K., Harney, E., Fu, Q., Gonzalez-Forbes, G., Jones, E. R., Roodenberg, S. A., Lengyel, G., Bocquentin, F., ... Reich, D. (2016). Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), 419–424.
- Leonardi, M., Librado, P., Der Sarkissian, C., Schubert, M., Alfarhan, A. H., Alquraishi, S. A., Al-Rasheid, K. A. S., Gamba, C., Willerslev, E., & Orlando, L. (2017). Evolutionary patterns and processes: Lessons from ancient DNA. *Systematic Biology*, 66(1), e1–e29.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
- Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, 362(6422), 709–715.
- Lindqvist, C., & Possnert, G. (1997). The subsistence economy and diet at Jakobs/Ajvide, Eksta parish and other prehistoric dwelling and burial sites on Gotland in long-term perspective. *Remote Sensing*, 1, 29–90.
- Lyman, R. L. (2014). Bone density and bone attrition. In J. T. Pokines, & S. A. Symes (Eds.), *Manual of forensic taphonomy* (pp. 51–72). CRC Press.
- Markova, A. K., Puzachenko, A. Y., van Kolfschoten, T., Kosintsev, P. A., Kuznetsova, T. V., Tikhonov, A. N., Bachura, O. P., Ponomarev, D. V., van der Plicht, J., & Kuitert, M. (2015). Changes in the Eurasian distribution of the musk ox (*Ovibos moschatus*) and the extinct bison (*Bison priscus*) during the last 50 ka BP. *Quaternary International: the Journal of the International Union for Quaternary Research*, 378, 99–110.
- McLeod, B. A., Frasier, T. R., & Lucas, Z. (2014). Assessment of the extirpated Maritimes walrus using morphological and ancient DNA analysis. *PLoS One*, 9(6), e99569.
- Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6), p.db.prot5448.
- Moss, M. L., Rodrigues, A., Speller, C. F., & Yang, D. (2016). The historical ecology of Pacific herring: Tracing Alaska Native use of a forage fish. *Journal of Archaeological Science: Reports*, 8, 504–512.
- Nielsen-Marsh, C. M., & Hedges, R. E. M. (2000). Patterns of diagenesis in bone I: The effects of site environments. *Journal of Archaeological Science*, 27(12), 1139–1150.
- O'Connor, T. (2008). *The archaeology of animal bones*. Texas A&M University Press.
- Okazaki, M., Yoshida, Y., Yamaguchi, S., Kaneno, M., & Elliott, J. C. (2001). Affinity binding phenomena of DNA onto apatite crystals. *Biomaterials*, 22(18), 2459–2464.
- Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., Johnson, P. L. F., Fumagalli, M., Vilstrup, J. T., Raghavan, M., Korneliusen, T., Malaspina, A.-S., Vogt, J., Szklarczyk, D., Kelstrup, C. D., ... Willerslev, E. (2013). Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, 499, 74–78.

- Ovchinnikov, I. V., Götherström, O., Romanova, G. P., Kharitonov, V. M., Lidén, K., Goodwin, W. (2001). reply: Not just old but old and cold? *Nature*, 410(6830), 772–772.
- Palkopoulou, E., Lipson, M., Mallick, S., Nielsen, S., Rohland, N., Baleka, S., Karpinski, E., Ivancevic, A. M., To, T.-H., Kortschak, R. D., Raison, J. M., Qu, Z., Chin, T.-J., Alt, K. W., Claesson, S., Dalén, L., MacPhee, R. D. E., Meller, H., Roca, A. L., ... Reich, D. (2018). A comprehensive genomic history of extinct and living elephants. *Proceedings of the National Academy of Sciences of the United States of America*, 115(11), E2566–E2574.
- Palkopoulou, E., Mallick, S., Skoglund, P., Enk, J., Rohland, N., Li, H., Omrak, A., Vartanyan, S., Poinar, H., Gotherstrom, A., Reich, D., & Dalén, L. (2015). Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. *Current Biology*, 25(10), 1395–1400.
- Parker, C., Rohrlach, A. B., Friederich, S., Nagel, S., Meyer, N., Krause, J., Bos, K. I., & Haak, W. (2020). A systematic investigation of human DNA preservation in medieval skeletons. *Scientific Reports*, 10, 18225.
- Ping, C. L., Bockheim, J. G., Kimble, J. M., Michaelson, G. J., & Walker, D. A. (1998). Characteristics of cryogenic soils along a latitudinal transect in arctic Alaska. *Journal of Geophysical Research*, 103(D22), 28917–28928.
- Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., Anders, A., Pietrusewsky, M., Rollefson, G., Jovanovic, M., Trinhhoang, H., Bar-Oz, G., Oxenham, M., Matsumura, H., & Hofreiter, M. (2015). Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One*, 10(6), e0129102.
- Poinar, H. N., Höss, M., Bada, J. L., & Pääbo, S. (1996). Amino acid racemization and the preservation of ancient DNA. *Science*, 272(5263), 864–866.
- Pruvost, M., Schwarz, R., Correia, V. B., Champlot, S., Braguier, S., Morel, N., Fernandez-Jalvo, Y., Grange, T., & Geigl, E.-M. (2007). Freshly excavated fossil bones are best for amplification of ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 104(3), 739–744.
- QGIS Development Team. (2018). *QGIS Geographic Information System*. Retrieved from <http://qgis.osgeo.org>
- R Core Team. (2018). *R: A language and environment for statistical computing*. Retrieved from <https://www.R-project.org/>
- Raghavan, M., DeGiorgio, M., Albrechtsen, A., Moltke, I., Skoglund, P., Korneliusen, T. S., Gronnow, B., Appelt, M., Gullov, H. C., Friesen, T. M., Fitzhugh, W., Malmstrom, H., Rasmussen, S., Olsen, J., Melchior, L., Fuller, B. T., Fahrni, S. M., Stafford, T., Grimes, V., ... Willerslev, E. (2014). The genetic prehistory of the New World Arctic. *Science*, 345(6200), 1255832.
- Ramos-Madrigal, J., Smith, B. D., Moreno-Mayar, J. V., Gopalakrishnan, S., Ross-Ibarra, J., Gilbert, M. T. P., & Wales, N. (2016). Genome sequence of a 5,310-year-old maize cob provides insights into the early stages of maize domestication. *Current Biology*, 26(23), 3195–3201.
- Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Ramsey, C. B., Buck, C. E., Cheng, H., Edwards, R. L., Friedrich, M., Grootes, P. M., Guilderson, T. P., Hafliadason, H., Hajdas, I., Hatté, C., Heaton, T. J., Hoffmann, D. L., Hogg, A. G., Hughen, K. A., ... van der Plicht, J. (2013). IntCal13 and marine13 radiocarbon age calibration curves 0–50,000 Years cal BP. *Radiocarbon*, 55(4), 1869–1887.
- Salamon, M., Tuross, N., Arensburg, B., & Weiner, S. (2005). Relatively well preserved DNA is present in the crystal aggregates of fossil bones. *Proceedings of the National Academy of Sciences of the United States of America*, 102(39), 13783–13788.
- Sandoval-Castellanos, E., Wutke, S., Gonzalez-Salazar, C., & Ludwig, A. (2017). Coat colour adaptation of post-glacial horses to increasing forest vegetation. *Nature Ecology & Evolution*, 1(12), 1816–1819. <https://doi.org/10.1038/s41559-017-0358-5>
- Sarkar, D. (2008). *Lattice: Multivariate data visualization with R*. Springer Science & Business Media.
- Savelle, J. M., & Habu, J. (2004). A processual investigation of a thule whale bone house, Somerset Island. *Arctic Canada. Arctic Anthropology*, 41(2), 204–221.
- Savriama, Y., Valtonen, M., Kammonen, J. I., Rastas, P., Smolander, O.-P., Lyyski, A., Häkkinen, T. J., Corfe, I. J., Gerber, S., Salazar-Ciudad, I., Paulin, L., Holm, L., Löytynoja, A., Auvinen, P., & Jernvall, J. (2018). Bracketing phenogenotypic limits of mammalian hybridization. *Royal Society Open Science*, 5(11), 180903.
- Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., & Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One*, 7(3), e34131.
- Scheel, D.-M., Slater, G. J., Kolokotronis, S.-O., Potter, C. W., Rotstein, D. S., Tsangaras, K., Greenwood, A. D., & Helgen, K. M. (2014). Biogeography and taxonomy of extinct and endangered monk seals illuminated by ancient DNA and skull morphology. *ZooKeys*, 409, 1–33.
- Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M. D., Fernández, R., Kircher, M., McCue, M., Willerslev, E., & Orlando, L. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nature Protocols*, 9(5), 1056–1082.
- Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Research Notes*, 9, 88.
- Schwarz, C., Debruyne, R., Kuch, M., McNally, E., Schwarcz, H., Aubrey, A. D., Bada, J., & Poinar, H. (2009). New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains. *Nucleic Acids Research*, 37(10), 3215–3229.
- Scorrano, G., Valentini, F., Martínez-Labarga, C., Rolfo, M. F., Fiammenghi, A., Lo Vetro, D., Martini, F., Casoli, A., Ferraris, G., Palleschi, G., Palleschi, A., & Rickards, O. (2015). Methodological strategies to assess the degree of bone preservation for ancient DNA studies. *Annals of Human Biology*, 42(1), 10–19.
- Shapiro, B. (2004). Rise and fall of the beringian steppe bison. *Science*, 306(5701), 1561–1565. <https://doi.org/10.1126/science.1101074>
- Smith, C. I., Chamberlain, A. T., Riley, M. S., Cooper, A., Stringer, C. B., & Collins, M. J. (2001). Neanderthal DNA: Not just old but old and cold? *Nature*, 410(6830), 771–772.
- Smith, C. I., Chamberlain, A. T., Riley, M. S., Stringer, C., & Collins, M. J. (2003). The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution*, 45(3), 203–217.
- Sosa, C., Vispe, E., Núñez, C., Baeta, M., Casalod, Y., Bolea, M., Hedges, R. E. M., & Martínez-Jarreta, B. (2013). Association between ancient bone preservation and dna yield: A multidisciplinary approach. *American Journal of Physical Anthropology*, 151(1), 102–109.
- Star, B., Barrett, J. H., Gondek, A. T., & Boessenkool, S. (2018). Ancient DNA reveals the chronology of walrus ivory trade from Norse Greenland. *Proceedings of the Royal Society B: Biological Sciences*, 285(1884), 20180978–<https://doi.org/10.1098/rspb.2018.0978>
- Star, B., Boessenkool, S., Gondek, A. T., Nikulina, E. A., Hufthammer, A. K., Pampoulie, C., Knutsen, H., André, C., Nistelberger, H. M., Dierking, J., Peterleit, C., Heinrich, D., Jakobsen, K. S., Stenseth, N. C., Jentoft, S., & Barrett, J. H. (2017). Ancient DNA reveals the Arctic origin of Viking Age cod from Haithabu, Germany. *Proceedings of the National Academy of Sciences of the United States of America*, 114(34), 9152–9157.
- Stuiver, M., Reimer, P. J., & Reimer, R. W. (2019). *CALIB*. Retrieved from <http://calib.org>
- Svintradze, D. V., Mrevlishvili, G. M., Metreveli, N., Jariashvili, K., Namicheishvili, L., Skopinska, J., & Sionkowska, A. (2008). Collagen-DNA complex. *Biomacromolecules*, 9(1), 21–28.
- Szpak, P. (2011). Fish bone chemistry and ultrastructure: Implications for taphonomy and stable isotope analysis. *Journal of Archaeological Science*, 38(12), 3358–3372.

- Todisco, D., & Monchot, H. (2008). *Bone weathering in a periglacial environment: The Tayara site (KbFk-7), Qikirtaq Island, Nunavik (Canada)*. *Arctic*, 87–101.
- Trueman, C. N. G., Behrensmeyer, A. K., Tuross, N., & Weiner, S. (2004). Mineralogical and compositional changes in bones exposed on soil surfaces in Amboseli National Park, Kenya: Diagenetic mechanisms and the role of sediment pore fluids. *Journal of Archaeological Science*, 31(6), 721–739.
- Trueman, C. N., & Tuross, N. (2002). *Trace elements in recent and fossil bone apatite*. *Reviews in Mineralogy and Geochemistry*. Retrieved from <https://pubs.geoscienceworld.org/msa/ring/article-abstract/48/1/489/110622>
- van der Valk, T., Vezi, F., Ormestad, M., Dalén, L., & Guschanski, K. (2019). Index hopping on the Illumina HiSeqX platform and its consequences for ancient DNA studies. *Molecular Ecology Resources*, 20(5), 1171–1181. Available at, <https://doi.org/10.1111/1755-0998.13009>.
- van Klinken, G. J. (1999). Bone collagen quality indicators for palaeodietary and radiocarbon measurements. *Journal of Archaeological Science*, 26(6), 687–695.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.
- Winter, B. (2013). *Linear models and linear mixed effects models in R with linguistic applications*. arXiv [cs.CL]. Retrieved from <http://arxiv.org/abs/1308.5499>
- Yang, D. Y., Eng, B., Wayne, J. S., Dudar, J. C., & Saunders, S. R. (1998). Technical note: Improved DNA extraction from ancient bones using silica-based spin columns. *American Journal of Physical Anthropology*, 105(4), 539–543.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Keighley X, Bro-Jørgensen MH, Ahlgren H, et al. Predicting sample success for large-scale ancient DNA studies on marine mammals. *Mol Ecol Resour*. 2021;21:1149–1166. <https://doi.org/10.1111/1755-0998.13331>