
Article:
Biard, Vincent, Gol'din, Pavel, Vishnyakova, Karina et al. (6 more authors) (2017) Genomic and proteomic identification of Late Holocene remains: Setting baselines for Black Sea odontocetes. Journal of Archaeological Science Reports. ISSN 2352-409X

https://doi.org/10.1016/j.jasrep.2017.07.008

TITLE

Genomic and proteomic identification of Late Holocene remains: setting baselines for Black Sea odontocetes

AUTHORS

Vincent Biard[1*], Pavel Gol'din[2], Elena Gladilina[2, 3], Karina Vishnyakova[2,3], Krista McGrath[4], Filipe G. Vieira[1], Nathan Wales[1], Michael C. Fontaine[5], Camilla Speller[4], and Morten Tange Olsen[1*]

AFFILIATIONS

[1]Section for Evolutionary Genomics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen K, Denmark
[2]Schmalhausen Institute of Zoology, National Academy of Sciences of Ukraine, 15 Bogdan Khmelnytskyi Street, Kiev, 01601, Ukraine
[3]Ukrainian Scientific Centre of Ecology of the Sea, 89 Frantsuzsky Blvrd., Odessa, 65009, Ukraine
[4]BioArCh, Department of Archaeology, University of York, Environment Building, York, North Yorkshire YO10 5DD, UK
[5]Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, PO Box 11103 CC, Groningen, The Netherlands

CORRESPONDING AUTHORS

*Vincent Biard
4 Rue Roger Senturenne, D103
33140 Villenave d'Ornon, France
E-mail: vincentbiard33@gmail.com

*Morten Tange Olsen
Section for Evolutionary Genomics
Natural History Museum of Denmark
University of Copenhagen
Øster Voldgade 5-7
1350 Copenhagen K, Denmark
Fax: +45 353210102
E-mail: morten.olsen@snm.ku.dk

ABSTRACT

A critical challenge of the 21st century is to understand and minimise the effects of human activities on biodiversity. Cetaceans are a prime concern in biodiversity research, as many species still suffer from human impacts despite decades of management and conservation efforts. Zooarchaeology constitutes a valuable approach for informing conservation and management decisions by providing baseline information on the past distribution and human uses of species. However, traditional morphological species identification of mixed assemblage bones can be challenging, particularly in the case of cetaceans. To address this issue, we applied and evaluated the performance of three biomolecular approaches – Sanger sequencing, shotgun sequencing and collagen peptide fingerprinting (ZooMS) – for species identification in a mixed assemblage of 800 to 1600 years old odontocete (toothed whale) samples from the site of Chersonesus in Crimea, Ukraine. We found that ZooMS allowed for identification to the taxonomic level for 28 of our 30 samples (>90%), identifying them as either "porpoise" or "dolphin", and approximately half of those samples could be further identified to species level with the shotgun sequencing approach. In addition, shotgun sequencing produced several complete ancient odontocete mitogenomes and auxiliary nuclear genomic data for further exploration in a population genetic context. In contrast, both morphological identification and Sanger sequencing lacked taxonomic resolution and/or resulted in misclassification of samples. We found that the combination of ZooMS and shotgun sequencing provides a powerful tool in zooarchaeology, and here allowed for a deeper understanding of past marine resource use and its implication for current management and conservation of Black Sea odontocetes.

KEYWORDS

Species identification; mixed assemblage; odontocete; ZooMS; shotgun sequencing; Sanger sequencing; mitogenome.

INTRODUCTION

One of the critical challenges of the 21st century is to investigate and reduce the negative effects of human activities on the Earth's biodiversity. To address these issues, detailed baseline information on the ecology and evolution of plant and animal populations and species is required, including information on past abundance, distribution, species life history or phenology, as well as historic and prehistoric human resource use, context and impacts. While this is true for much of Earth's biota, cetaceans constitute an iconic group where many species and populations, despite decades of management and conservation efforts, still suffer from past and present human impacts, such as overexploitation, habitat alterations, bycatch, and ocean noise (Schipper *et al.* 2008; Pimm *et al.* 2014).

Zooarchaeology is particularly suited to provide such baseline information, as this approach documents a broad range of direct material evidences of animals, including their abundance, distribution, diversity, population structure and individual traits, as well as long-term changes of these parameters (Steadman 1995; Dietl *et al.* 2015; Hofman *et al.* 2015). However, in the case of cetaceans, accurate identification of species constitutes a special problem. On many archaeological sites, cetacean assemblages contain multiple species (Mulville 2002), which are represented by only incomplete fragments of postcranial bones and often are considered to be non-diagnosable by morphological methods (Mulville 2002; Amundsen *et al.* 2013). For example, isolated vertebrae of *Delphinus* sp. are hardly discriminated from *Stenella* sp., small *Tursiops sp.* or *Steno sp*. Furthermore, in an assemblage containing kitchen refuse of *Delphinus* sp., *Stenella* sp. and *Tursiops* sp., most of postcranial fragments were unidentified, unlike well informative cranial and, in particular, dental remains (Cooke *et al.* 2016). In

95 cetaceans, caudal vertebrae (comprising 35–50% of all vertebrae, depending on the species)
96 and vertebral centra with missing processes are especially difficult for diagnostics. This may
97 also be true for other mammalian groups, but many terrestrial mammalian taxa are often
98 represented by a single species in each assemblage or a few species of different size
99 categories and, thus, are more easily discriminated (Dunnell 1971; Grayson 1984).
100
101 The ability to extract and analyse ancient DNA (aDNA) from subfossil material have greatly
102 contributed to the resolution of this problem (Willerslev & Cooper 2005), allowing for
103 determining the species identity of e.g. historical whale remains from whaling stations at
104 South Georgia (Lindqvist *et al.* 2009; Sremba *et al.* 2010). Moreover, although driven
105 primarily by research in human (Briggs *et al.* 2009; Rasmussen *et al.* 2010; Allentoft *et al.*
106 2015) and terrestrial megafauna evolution (Gilbert *et al.* 2008; Lorenzen *et al.* 2011; Dabney
107 *et al.* 2013a), several studies have utilized aDNA to assess the evolutionary and demographic
108 history of cetaceans (see Foote *et al.* 2012a). For instance, McLeod *et al.* (2012) have
109 examined the demographic history of bowhead whales (*Balaena mysticetus*) in the Canadian
110 Arctic using mitochondrial DNA fragment, and have revealed a population expansion over
111 the past 30,000 years BP. Similar historic inference have been made from Eastern Atlantic
112 bowhead whales (Foote *et al.* 2013), as well as common bottlenose dolphin (*Tursiops*
113 *truncatus*) (Nichols *et al.* 2007), killer whales (*Orcinus orca*) (Foote *et al.* 2009) and North
114 Atlantic right whales (*Eubalaena glacialis*) (McLeod *et al.* 2010).
115
116 To date, most molecular approaches for ancient cetacean species determination have been
117 based on Sanger sequencing of short fragments of mitochondrial genes such as cytochrome-b
118 (CytB), ribosomal 12S or cytochrome oxidase-I (COI). However, for highly degraded

119 material, sufficient DNA template may not be preserved, or this approach might not yield
120 sufficient coverage and resolution for species distinction – in particular if these are from
121 closely related species (Goldstein & DeSalle 2011). Moreover, the focus on a single short
122 fragment often limits additional population genetic inference of evolutionary and
123 demographic trajectories, as well as associated historical anthropogenic and climatic impacts.
124 The recent development of peptide mass fingerprinting of bone collagen (ZooMS), where
125 species identification can be accomplished by comparing ancient peptide fingerprints to
126 modern reference databases, allows for analysis of degraded ancient material otherwise
127 suboptimal for DNA analysis (Buckley *et al.* 2009; Welker *et al.* 2015). However, due to the
128 relatively slow rate of evolution within the collagen chains, taxonomic resolution is often
129 limited to the genus or even family level. Studies combining analysis of aDNA Sanger
130 sequencing and collagen targeting have successfully been applied to a few diverse
131 assemblages of baleen whales (Buckley *et al.* 2014; Evans *et al.* 2016; Speller *et al.* 2016),
132 however for some species and materials, even the combined use of these methods may not
133 provide sufficient resolution for species identification. By generating massive amounts of data
134 across the genome, short-read DNA "shotgun" sequencing techniques may circumvent the
135 obstacles of aDNA fragmentation and taxonomic resolution, as well as provide additional data
136 for population genetic inference of demography and phylogeography (Leonardi *et al.* 2016).
137 Thus, the approach has a great potential to become the gold standard for species identification
138 of ancient material. However to date the applicability of shotgun sequencing relative to
139 Sanger sequencing and collagen profiling have not been evaluated..
140
141 All three odontocete populations inhabiting the Black Sea, the common bottlenose dolphin
142 (*Tursiops truncatus*), the short-beaked common dolphin (*Delphinus delphis*) and the harbour

143 porpoise (*Phocoena phocoena*), though of widespread species, are recognized as separate
144 subspecies and of special management and conservation concern. The IUCN has classified the
145 Black Sea harbour porpoise (*P. p. relicta*) and the Black Sea bottlenose dolphin (*T. t.*
146 *ponticus*) as endangered (Birkun Jr & Frantzis 2008; Birkun Jr 2012) and the Black Sea
147 common dolphin (*D. d. ponticus*) as vulnerable (Birkun Jr 2008). These have all been
148 severely depleted by decades of extensive hunting (Kleinenberg 1956; Birkun Jr 2002a), and
149 are now affected by fisheries bycatch and consequences of ctenophore *Mnemiopsis leidyi*
150 invasion, undermining their trophic base (Bushuev 2000; Vishnyakova & Gol'din 2015b). The
151 common bottlenose dolphin (Viaud-Martinez *et al.* 2008; Moura *et al.* 2013) and short-beaked
152 common dolphin (Amaral *et al.* 2007) both appear to have colonised the Black Sea a few
153 thousand years ago. Similarly, it has been suggested that a relict harbour porpoise population
154 in the Eastern Mediterranean founded the Black Sea population a few thousand years after the
155 reconnection of the two basins, probably tracking suitable habitats (Fontaine *et al.* 2012;
156 Fontaine *et al.* 2014; Fontaine 2016). Moreover, in contrast to the two other odontocete
157 species present in the basin, the Black Sea harbour porpoise is unique in its isolation to other
158 populations situated in the Atlantic Ocean, around 4,000 km apart. Its absence from the
159 Mediterranean Sea, already noted by Aristotle (350 BC) (Frantzis *et al.* 2001), might be the
160 consequence of changing oceanic conditions, notably warming temperatures and low
161 productivity of Mediterranean ecosystems (Thunell *et al.* 1977; Thunell 1979; Fontaine
162 2016). Thus isolated, the Black Sea harbour porpoise has evolved distinct morphological and
163 genetic characteristics (Viaud-Martínez *et al.* 2007; Galatius & Gol'din 2011), and additional
164 sub-structuring might even exist among the different water bodies of the Black Sea region
165 (Gol'din 2004; Gol'din & Vishnyakova 2016), although this question requires further
166 investigations.

Here we tested the applicability of different species identification methods on a mixed assemblage of odontocete (toothed whale) zooarchaeological remains excavated at the site of Chersonesus on the Black Sea coast of Crimea, Ukraine. Specifically, we i) compared and evaluated the performance of four approaches: morphology, Sanger sequencing, collagen peptide mass fingerprinting (ZooMS), and shotgun sequencing of short-read DNA; and ii) discuss our findings in the context of historic marine resource use, and in relation to setting baselines for contemporary conservation and management schemes in the region. To our knowledge this is one of the first studies to utilise the full extent of recent genomic and proteomic techniques to identify and analyse highly degraded ancient odontocete material. Importantly, although the present focus is on odontocetes, our findings should apply to other organisms and study systems.

MATERIALS & METHODS

Study site

Chersonesus (also known as Chersonesus Taurica or Tauric Chersonesus = Chersonese) is located on the Black Sea coast in the southern Crimea (Supplementary Figure S1). It was the greatest city, trade and cultural centre of the northern Black Sea region during the Hellenistic, Roman and Byzantine ages between 2400–600 years BP (Strabo 1929; Porphyrogenitus 1993; Carter & Mack 2003). Founded by Greeks from Asia Minor, it was populated by colonists and visited by traders from various Black Sea and Mediterranean localities, and thus, the archaeology of the city shows a great variety of regional economic and cultural practices (Kadeev 1970; Kadeev & Sorochan 1989). In particular, marine fisheries and seafood played an important role in economy and diet of the Chersonesus population, and diverse marine fauna was reported from zooarchaeological evidence, as well as from art representations and some descriptive sources (Semenov-Zuser 1947; Kadeev 1970; Højte 2005; Morales *et al.* 2007).

Zooarchaeological material

The zooarchaeological material comprised 259 samples of small toothed whales excavated in 2011-2013 from the Kruze basilica and adjoining area in Chersonesus (Ushakov 2011), consisting of partial skulls and skull fragments, mandibles, ear bones, isolated teeth, sternum bones, humeri, partial vertebrae, epiphyses of vertebral bodies, ribs and rib fragments and pelvic bones. Of these, 30 samples of vertebrae, teeth, mandibles, epiphyses and skull

fragments were used for comparison of species identification methods, as representing individuals that were identified as different specimens in terms of morphology, taphonomy and archaeological context (Table 1, Figure 1A). The samples dates ranged from 1600 to 800 years BP based on the archaeological context, as presented by age-specific ceramics and coins (Ushakov 2011, 2013; Ushakov *et al.* 2013); but most of specimens came from the layers dated as 1600-1500 years BP (400-530 CE).

Morphological species identification

Based on morphological characteristics, the majority of the samples were assigned to species level (Supplementary Table S1). Identification was conducted by comparing the samples with museum collection specimens (Supplementary Figure S2). Specifically, all three species are well distinguished from their teeth, skull vertices, facial skulls, ear bones, sternum, humerus and pelvic bones (Supplementary Table S2). In addition, bottlenose dolphins are significantly larger than harbour porpoises in all dimensions, independent of ontogenetic age. In contrast to the above distinctions, vertebrae and ribs are less indicative than other bones. Generally, vertebrae of bottlenose dolphins can be identified from their relative large size, and those of common dolphins by their narrow bodies and long spines. However, incomplete vertebrae can generally only be provisionally identified. Here, as for ribs, they were identified to species only with other associated bones.

Genetic species identification

DNA extraction

228

229 All DNA extractions were conducted in a designated clean laboratory for ancient DNA
230 analysis at the Centre for GeoGenetics, Natural History Museum of Denmark. First, the outer
231 layer of the bone was removed by drilling and bleaching to limit contamination from soil
232 composition. Then, 72 to 420mg of cleaned bone powder was obtained by drilling the bones
233 inner part or, when drilling was impossible, by grinding of the full samples. Drilling was
234 performed using a Dremel with a rounded drill head, running at <1000rpm to minimise
235 heating of the sample. DNA was extracted using a silica-in-solution method (Rohland &
236 Hofreiter 2007; Dabney *et al.* 2013a; Allentoft *et al.* 2015). Extraction blanks were included
237 in each batch to monitor potential contamination. DNA quality and quantity was assessed
238 using High Sensitivity D1000 ScreenTape for 2200 TapeStation (Agilent Technologies) and a
239 Qubit® 2.0 Fluorometer. The former estimates peak molarities of the short-fragment
240 component of the DNA extracts, and is hence expected to provide the most reliable
241 approximation of DNA concentration in our subfossil samples.
242

243 *Endogenous DNA content*

244

245 The DNA extracts were tested for the presence and amount of endogenous (odontocete) DNA
246 by real time qPCR using previously designed primers specifically developed for odontocete
247 species identification by targeting a 43bp region of the CytB gene (Foote *et al.* 2012b).
248 qPCRs were performed on a Stratagene Mx3000P (Agilent Technologies) and each 25µL
249 reaction contained 2µL of DNA, 1 x PCR buffer, 2.5mM $MgCl_2$, 0.8µg/µL BSA, 0.4µM of
250 each primer, 0.25µM mixed dNTPs, 1µL SYBR Green and 0.25µL AmpliTaq Gold® enzyme
251 (Applied Biosystems). Thermocycling was performed at 95°C for 10 min, 55 cycles of 95°C

for 30sec, 55°C for 30sec and 72° for 30sec, and concluded by a dissociation step at 95°C for 1min, 50°C for 30sec and 95°C for 30sec. DNA extracts were used at 1:5, 1:10 and 1:20 dilution. Furthermore, for each qPCR reaction, one PCR blank (containing ddH_2O instead of sample) was run in triplicate to control eventual contamination during the qPCR set up. The qPCRs also included a standard, composed of 1:1, 1:10, 1:100 and 1:1000 dilution of modern DNA from a harbour porpoise, which was run in triplicate to monitor the specificity of the amplification. The Ct values (cycle threshold) and amplification efficiency was determined using the Pfaffl method (Pfaffl 2001). The Ct value refers to the numbers of cycles needed to observe the amplification of the DNA template, as the fluorescence level is correlated to the amount of double stranded DNA that is synthetized. Such approach is therefore more specific targeting, theoretically, only endogenous DNA. The Ct values were determined in the exponential phase of the qPCR amplification and normalized to the standard for comparison between the different qPCR assays.

Furthermore, to investigate the eventual impact of the quantity of starting material in the DNA retrieval from the archaeological material, we compared the amounts of bone powder with extract peak molarities and the normalized Ct values from the qPCR assays.

Species identification by Sanger sequencing

To illustrate the traditional approach to species determination, the qPCR products were used as a template for PCR amplification and Sanger sequencing using the CytB primers mentioned above (Figure 1B). Each 25µL reaction contained 2µL of qPCR product, 1 x PCR buffer, 2.5mM $MgCl_2$, 0.4µg/µL BSA, 0.4µM of each primer, 0.25µM mixed dNTPs and

0.2µL AmpliTaq Gold® enzyme (Applied Biosystems) under thermocycling 95°C for 5min, 25 cycles of 95°C for 30sec, 55°C for 30sec and 72°C for 30sec, finishing by 72°C for 7min. The resulting PCR products were visualised by electrophoresis on 2% agarose gel and positive PCRs were purified and sent for Sanger sequencing at Macrogen®, Korea.

Species identification by shotgun sequencing

In order to test the applicability of shotgun sequencing of short-read sequence data for species determination (Figure 1C), DNA extracts were converted to Illumina indexed libraries using the NEBNext® DNA Library Prep Master Mix Set for 454™ (New England BioLabs®). The resulting libraries were tested by real time qPCR to determine the optimal number of cycle for the library PCR amplification prior to sequencing. qPCRs were performed on a Stratagene Mx3000P (Agilent Technologies) and each 100µL reaction contained 20µL of 1:20 diluted library, 1 x PCR buffer, 2.5mM $MgCl_2$, 0.6µg/µL BSA, 0.2µM of each primer (PE1.0 and Index), 0.15µM mixed dNTPs, 1µL SYBR Green and 2µL AmpliTaq Gold® enzyme (Applied Biosystems) under thermocycling 95°C for 12min, 40 cycles of 95°C for 20sec, 60°C for 30sec and 72° for 40sec, finishing by 72°C for 5min.

Libraries were amplified by PCR under the above-described conditions (excluding SYBR green). Amplified libraries were purified using the QIAquick® PCR Purification Kit following the manufacturer's instructions, and then tested for quality and quantity using High Sensitivity D1000 ScreenTape for 2200 TapeStation (Agilent Technologies). The final libraries were pooled into two groups (BS_01 to BS_17 and BS_18 to BS_30) to equimolarity based on the 2200 TapeStation profiles and single read sequenced on respectively one lane

300 and a partial lane of the Illumina HiSeq 2000 platform. Moreover, five samples (BS_07,
301 BS_10, BS_11, BS_12 and BS_15) were re-sequenced on a partial lane. The library of the
302 sample BS_16 did not reveal any insert based on the DNA profile and was consequently not
303 sequenced.
304
305 Given the lack of nuclear reference genomes for most cetaceans and general low yield of
306 endogenous nuclear DNA from zooarchaeological material, the genetic species determination
307 was done by use of mitochondrial genomic data (mitogenomes). First, sequencing adapters
308 were removed from the de-multiplexed reads and reads shorter than 30bp were discarded
309 using the program AdapterRemoval 1.5.4 (Lindgreen 2012). Then, all reads were mapped to
310 reference mitogenomes of the harbour porpoise (Accession number NC005280) and
311 bottlenose dolphin (Accession number KF570325). Full mitogenomes are not available for
312 short-beaked common dolphin so reference data for this species was created by compiling
313 available mtDNA sequence fragments in GenBank (Accession numbers KC312628,
314 AB481388, and EU685091), as well as by mapping the short-read data to the mitogenome of
315 a sister species, the long-beaked common dolphin (*Delphinus capensis*) (Accession number
316 NC012061). Mapping was done using the program BWA 0.7.5a-r405 (Li & Durbin 2009),
317 reads were quality filtered (Q>30) using the program samtools 1.2 (Li *et al.* 2009), and PCR
318 duplicates were removed using the program Picard (http://picard.sourceforge.net) before
319 being imported into Geneious (Kearse *et al.* 2012). Also, to aid in species identification by
320 phylogenetic matching, we constructed a phylogenetic tree, including all mitogenome
321 references and samples presenting high coverage mitogenomes, using the MrBayes 3.2.2.
322 (Huelsenbeck & Ronquist 2001) plugging in Geneious. The mitogenome reference of the
323 Ginkgo-toothed beaked whale (*Mesoplodon ginkgodens*) (Accession number NC027593) was

used as an out-group. Four independent chains for 1,000,000 generations under the General-time-reversible (GTR) model substitution were performed, with a burn-in length of 100,000 generations, and a sampling frequency of 5,000 generations.

Taxonomic identification by collagen peptide mass fingerprinting (ZooMS)

The identification of zooarchaeological remains through peptide mass fingerprinting of bone collagen (also known as ZooMS) has been applied in several contexts, including the identification of marine mammals (Buckley *et al.* 2014; Evans *et al.* 2016; Speller *et al.* 2016). We applied ZooMS to 29 of the odontocete samples after they had been processed for ancient DNA analysis (BS_16 was destroyed entirely during DNA analysis). The samples were processed in BioArCh, University of York, using protocols for ancient bone (Buckley *et al.* 2009). We immersed *ca.* 10-30mg of bone in 250µL of 0.6M hydrochloric acid at 4˚C until the samples were demineralised. The samples were centrifuged and the supernatant discarded, before being rinsed in 200µL of 0.1M NaOH to remove humics and other chromophoric compounds. The NaOH was removed, and the samples were rinsed three times in 200µL of 50mM ammonium bicarbonate solution (NH4HCO3) pH 8.0 (AmBic). A final 100µL of AmBic was then added to the samples, followed by incubation for one hour at 65°C to gelatinise the collagen; 50µL of this supernatant was incubated overnight at 37°C with 0.4µg of trypsin, then acidified to < pH 4 with 0.1% trifluoroacetic acid (TFA), and purified using a 100µl C_{18} resin ZipTip® pipette tip (EMD Millipore).

For mass spectrometry, we combined 1µL of extract with 1µL of matrix (α-cyano-hydroxycinnamic acid) and spotted them in triplicate along with calibration standards onto a

Bruker ground steel target plate and ran them on a Bruker ultraflex III MALDI TOF/TOF mass spectrometer, resulting in 90 individual spectra. Spectra were analyzed using the mMass software (Strohalm *et al.* 2008); spectra with low signal to noise ratios, or few to no discrete peaks were eliminated from the dataset. We averaged spectra from replicates of the same sample and compared them to the list of *m/z* markers for both marine and non-marine mammals (Buckley & Kansa 2011; Kirby *et al.* 2013; Buckley *et al.* 2014). We assigned taxonomic identifications at the most conservative level of identification (genus, or family level) based on the presence of unambiguous *m/z* markers (Figure 1D).

Population genetic analyses

Finally, we illustrate the additional use of short-read DNA sequence data from ancient odontocetes in a population genetic context by assembling and comparing it with publically available modern sequences from studies on same species and geographical regions. As an example, the ancient Black Sea harbour porpoise sequences were aligned with a 706bp fragment of the mitochondrial region of 64 modern harbour porpoise from the Black Sea and the Aegean Sea from Fontaine *et al.* (2012) and a parsimony-based TCS haplotype network was created using PopART ver. 1.7 (Leigh & Bryant 2015). This species was chosen based on the availability of modern reference data.

RESULTS

Morphological species determination

A total of 30 odontocete samples were analysed, comprising 16 vertebrae, four mandibles, four individual teeth, three epiphyses, two skulls and one occipital condyle (Table 1). Of these, 14 were morphologically identified as bottlenose dolphin remains, 12 were identified as harbour porpoises, one was identified as a short-beaked common dolphin, and three were unidentified.

Genetic species identification

Endogenous DNA content

The DNA concentrations were between 0.42ng/µL and 68.0ng/µL with an average of 4.96ng/µL, and peak molarities were ranging from 0.15nmol/L to 42.5nmol/L, with an average concentration of 10.12nmol/L (Supplementary Table S3). The average amplification efficiency of the qPCR assay was of 98.1±1.7%, and the melting curves were composed of single peaks for most of the samples, indicating the successful amplification of DNA from subfossil odontocetes. In a few samples, we also observed amplification in both extraction and qPCR blanks. However, these amplifications started after 40 cycles, more than 10 cycles after the amplification start of the latest amplified sample, and melting curves revealed qPCR products that were different from the standard (Supplementary Figure S3). We therefore attributed these amplifications to the formation of primer dimers and to the high number of

394 cycles performed compare to modern DNA template qPCR assays. A Pearson's correlation
395 was run to determine the relationship between the amounts of bone powder and the extract
396 peak molarities and no direct correlation was observed (r (28) = -0.042, P>0.5) (Figure 2 A).
397 Rather there seemed to be a weak but non-significant correlation with Ct values (r (28) = -
398 0.250, P>0.1) (Figure 2 B), indicating that amount of starting material determines endogenous
399 DNA yield, although it is not statistically significant.
400
401 *Species identification by Sanger sequencing*
402
403 In the Sanger sequencing approach, CytB sequence chromatograms were obtained from 28 of
404 the 30 samples. However, the quality of the data was poor and chromatograms difficult to
405 score, leaving uncertainties in the base calls. Moreover, even in the cases where quality was
406 sufficient for reliable base calling, the amplified CytB sequence fragment was found to
407 contain few diagnostic sites to reliably distinguish between bottlenose dolphin and short-
408 beaked common dolphin. Thus, Sanger sequencing data only allowed for species
409 identification of eight harbour porpoises, whereas the remaining 20 samples were classified as
410 Dolphin *sp*. (Table 1).
411
412 *Species identification by shotgun sequencing*
413
414 Fourteen of the 29 samples that were shotgun sequenced yielded sufficient short read data for
415 assembling partial and high coverage mitogenomes and species determination
416 (Supplementary Table S4). The identified species included seven harbour porpoises, four
417 short-beaked common dolphins and three bottlenose dolphins (Table 1, Supplementary Figure

418 S4). For these samples, the proportion of mitogenome recovery and depth of coverage ranged
419 from 72% to 100% (average = 90.2%) and 1.6 to 78.5 (average = 19.4) (Figure 3),
420 respectively, whereas the remaining 15 samples yielded a very low number of reads (<120) or
421 insufficient depth of coverage for species identification (Supplementary Table S4).
422
423 *Species identification by collagen peptide mass fingerprinting (ZooMS)*
424
425 We used ZooMS to assign broad level taxonomic identifications for all 29 samples tested for
426 both Sanger and shotgun sequencing. Twenty-eight of the samples were confirmed as
427 odontocetes, falling into two broad categories: 21 samples were identified as 'dolphins',
428 which includes bottlenose dolphin, striped dolphin (*Stenella coeruleoalba*), short-beaked
429 common dolphin, white-beaked dolphin (*Lagenorhynchus albirostris*); and 7 samples were
430 identified as 'porpoises', which includes Dall's porpoise (*Phocoenoides dalli*) and harbour
431 porpoise, but also white-sided dolphin (*Lagenorhynchus acutus/ obliquidens*) and killer whale
432 (*Orcinus orca*) (Table 1; Figure 1D; Supplementary Table S5). One sample was identified as
433 a carnivore based on a lack of peptide markers common to cetaceans (e.g. *m/z* value 1,079)
434 (Buckley *et al.* 2014), and the presence of peptides common to the order Carnivora (e.g. *m/z*
435 values 1,105; 1,453; 2853; 2869) (Buckley *et al.* 2009; Kirby *et al.* 2013). The spectra was
436 compared against available peptide markers for both terrestrial and marine carnivores (e.g.
437 seals) (Kirby *et al.* 2013; Buckley *et al.* 2014; Buckley *et al.* 2017); due to the presence of
438 peptide markers 2,131 (marker (D)) and particularly 1,576 (2t76), the spectra for BS_03
439 appears to be most consistent with the genus *Canis* (Buckley et al. 2017) (Supplementary
440 Figure S5).
441

442 **Comparison of species identification methods**

443

444 The three biomolecular methods for species identification – Sanger sequencing, shotgun
445 sequencing and ZooMS – only disagreed for two (BS_03 and BS_10) of the 28 samples
446 where data was obtained from at least two of the molecular methods (Table 1). Specifically,
447 the sample BS_10 – an occipital condyle, and morphologically determined to be Cetacean
448 species – was determined to be a dolphin species by Sanger sequencing, a harbour porpoise
449 by the shotgun sequencing method, and ZooMS identified this as a member of the 'porpoise'
450 group, thus agreeing with the shotgun species identification. In the case of BS_03, this sample
451 was morphologically identified as short-beaked common dolphin, but as a harbour porpoise
452 using the Cytb sequencing, while the ZooMS identification pointed to a carnivore. Shotgun
453 sequencing did not produce exploitable data for this sample, leaving its identification
454 uncertain.

455

456 For the remaining 26 samples, the three molecular approaches agreed, although only shotgun
457 sequencing provided identification at species level. The molecular approaches also revealed
458 multiple errors in identifications from morphology. For instance, species identification was
459 confirmed only for 6 of 11 specimens morphologically identified as harbour porpoises,
460 whereas the others were re-identified as dolphins (BS_08 was re-identified to species level as
461 a short-beaked common dolphin). Moreover, 2 of 13 specimens morphologically identified as
462 bottlenose dolphins, were re-identified as short-beaked common dolphins using biomolecular
463 approaches. Therefore, all the errors in morphology, which were tracked to the species level
464 by genetic analysis involved short-beaked common dolphins and were likely due to

misidentification or omission of this species, and there was no confusion in discrimination between harbour porpoises and bottlenose dolphins.

Phylogenetic analyses

Four ancient Black Sea harbour porpoises yielded sufficient data for comparison with the mitochondrial control region fragment available for modern data. One transition was observed between the four ancient sequences, segregating them into two distinct haplotypes: BS_09 and BS_14 to one, and BS_11 and BS_15 to the other. Comparison with modern Black Sea sequences reveals that both haplotypes are still present in the modern population, including the southern Crimean waters, with one being the most abundant and central in the star-like topology of network (Figure 4), indicating a recent population expansion. .

DISCUSSION

Species identification in zooarchaeology

Species identification of archaeological materials can be challenging. Morphological identifications can be subject to human error and is sometimes impossible, being limited by reference to a higher taxon. This is especially common in the case of cetaceans, which are often found in multispecies assemblages, consisting of numerous closely related taxa of similar size. Biomolecular approaches may provide an excellent alternative for confirming a morphological label or resolving uncertainties. Here, we applied and evaluated four approaches to identify material from a mixed assemblage of odontocetes; one morphology-based, one protein-based and two DNA-based. We found that the proteomic ZooMS approach and Sanger sequencing could assign taxonomic grouping ("porpoise" or "dolphin") to the majority of samples. In contrast, shotgun sequencing only yielded sufficient genetic information for half of the samples, but for those samples working well, it allowed for identification to species level and provided the first ancient mitogenomes from Black Sea odontocetes, allowing for species level identification and additional population genetic inference.

There are several advantages to a protein-based method such as ZooMS for species identification of subfossil material. First, collagen in particular, is a robust protein, and can generally be recovered from tropical or subtropical archaeological sites where DNA may not be preserved (Welker *et al.* 2015). Second, ZooMS can identify samples without prior taxonomic knowledge potentially identifying 'non-target' species that might fail PCR based

approaches such as Sanger sequencing due to a lack of primer specificity. The main limitation of ZooMS is its level of taxonomic resolution: although all odontocete samples were identified, they could only be assigned to broad taxonomic groups due to the lack of diversity within the collagen sequences of odontocetes and poor resolution within certain areas of the spectrum (Buckley *et al.* 2014). For instance, species such as killer whales and white-sided dolphin within the Delphinidae family have a peak at mass 1652 grouping them with porpoises, but differ from porpoises and matches other dolphins at masses in the 1500s area of the spectrum. However this area is often poorly resolved, making conclusive identification difficult. Thus, ZooMS classifications may not be correct in systems unlike ours, where we can rule out certain species based on their known taxonomic distributions, habitat preference and absence from the Black Sea. For example, we could quite confidently rule out killer whale, Dall's porpoise, white-beaked dolphin, and white-sided dolphin from the Black Sea, leaving the 'dolphin' group to include the genera *Tursiops*, *Delphinus* and, hypothetically, *Stenella*, and the 'porpoise' group to include only *Phocoena*. More precise taxonomic identifications can occasionally be obtained from the mass spectra, but this requires the successful recovery of multiple diagnostic peptides.

Sanger sequencing has previously been successfully applied for species identification in cetacean mixed assemblages (Lindqvist *et al.* 2009; Sremba *et al.* 2010; Evans *et al.* 2016). Although the primers used here have been successful in a previous study for identifying odontocetes (Foote *et al.* 2012b), we generally found the chromatograms difficult to interpret, the approach failed to identify samples beyond taxonomic level, and two samples (BS_03 and BS_10) were misclassified. The CytB fragment that we targeted contains 12 and 15 variations sites that differentiate between harbour porpoise and bottlenose dolphin or harbour porpoise

526 and common short-beaked dolphin respectively. The two dolphin species differ at 3 sites,
527 which in our case was insufficient for clear species identification. These obstacles may be
528 overcome by designing new primers and/or targeting larger mitochondrial fragments, as
529 reported for some baleen whale samples (e.g. Yang & Speller 2006; Evans *et al.* 2016).
530 However, for highly fragmented ancient DNA, amplification of large PCR products may be
531 problematic, so targeting of short regions are recommended. If such short fragment primers
532 can be designed and highly degraded ancient material can be avoided, Sanger sequencing may
533 provide species identification, but the obtained sequences will likely not provide much data
534 for population genetic inference.
535
536 In contrast, having previously been limited by costs and laborious methodology, shotgun
537 sequencing has emerged as an efficient and simple method for generating a large number of
538 short-read sequence data for species identification and population genetic inference from
539 degraded material (Blow *et al.* 2008). In particular, this method can by-pass difficulties
540 occurring using Sanger sequencing, including significant DNA fragmentation, which was an
541 issue in our study. In addition, the ability to recover larger regions of the mitochondrial
542 genome rather than just a short fragment can often provide a more robust taxonomic
543 identification. Still, shotgun sequencing is not without limitations. Specifically, although all
544 of the samples seemed to contain well-preserved collagen, only half of the samples yielded
545 sufficient amounts of sequence data for species identification. The geographical region from
546 which our samples originate is characterised by temperate to arid subtropical climate, and thus
547 typically not considered optimal for preservation of aDNA (Willerslev & Cooper 2005;
548 Dabney *et al.* 2013b). Moreover, the samples are likely kitchen refuse and some have
549 certainly been boiled or cooked, probably resulting in some DNA degradation already before

burying. As observed here, sequencing success could potentially be improved by increasing the amount of starting material (e.g. bone powder), and for small or valuable samples where material is limited, issues of poor sequence recovery may be solved by re-sequencing troublesome samples, applying capture-enrichment methods (Ávila-Arcos *et al.* 2011; Ávila-Arcos *et al.* 2015) or using bioinformatics software capable of handling low depth sequencing data (Korneliussen *et al.* 2014).

ZooMS has been suggested as a pre-screening method ahead of more intensive or destructive bimolecular approaches, including PCR-based and shotgun DNA analyses (Von Holstein *et al.* 2014; Brown *et al.* 2016; Speller *et al.* 2016), shotgun proteomics (Welker *et al.* 2016) or radiocarbon dating (Harvey *et al.* 2016). As ZooMS is typically less expensive and requires less bone powder than other biomolecular techniques, pre-screening with ZooMS can provide both taxonomic information, as well as an indication of overall biomolecular preservation. In this study, the ZooMS technique provided comparable taxonomic resolution as the Sanger sequencing approach, and could therefore act as a cost-effective pre-screening method ahead of shotgun sequencing. Thus, we recommend using a combination of ZooMS and shotgun sequencing for more effectively resolving uncertainties in species determination. Under such an approach, ZooMS will likely allow for the assignation to a taxonomic level of most or all species, as well as a pre-screen of suitable material that can be shotgun sequenced for further identification to species level and generation of mitogenomic and nuclear data for population genetic inference.

Marine resource use in ancient Chersonesus

574 The overall species composition of wild fauna on archaeological sites, and particularly in
575 ancient kitchen refuse, provides numerous insights for the reconstruction of ancient
576 environments, paleoclimatic conditions, catch techniques and seasonality of different hunting
577 and fisheries activities. For example, numerous anchovy (*Engraulis encrasicolus*) remains at
578 the Chersonesus sites are indicative of set net fisheries during the autumn and spring seasonal
579 migration; turbot (*Scophthalmus maximus*) is indicative for bottom gill net fisheries, which
580 are the most productive in spring; and tuna (*Thunnus sp*) is indicative for longline and
581 intensive boat tending activities in warm season. From these perspectives, the identify
582 odontocetes can be either a product of incidental catches in bottom gill nets set for turbot, as it
583 is in present times (Vishnyakova & Gol'din 2015a), specimens recovered from strandings, or
584 an object of direct hunt from boats in coastal waters, as it appeared in the early 20th century
585 (Kleinenberg 1956). Of these alternatives, the former idea is well supported by abundant
586 turbot remains in the Kruze basilica. Meanwhile, strandings of common dolphins are
587 relatively rare in Crimea, as well as bycatches in any modern fishery practices in the Black
588 Sea (Birkun Jr 2002b). On a global scale, bycatch is generally associated with modern types
589 of fisheries, such as pelagic trawling and big purse seine fisheries (Tregenza *et al.* 1997;
590 Bilgmann *et al.* 2008), rather than artisanal fisheries. Therefore, stable catches of common
591 odontocetes might have implied boat-based direct hunting. For instance, Black Sea fishermen
592 hunted common dolphins in early 20th century, using a few small boats setting a single purse
593 seine (*alaman*) (Kleinenberg 1956). Aelian described a similar net for tuna fisheries used in
594 the Roman-era southern Black Sea and mentioned dolphin catches in this context. In an
595 earlier evidence, Strabo (1929) mentioned ancient dolphin catches in the same area in
596 association with tuna and bonito fisheries. Also, Brito & Vieira (2009) found evidence for
597 catches of common dolphins and tuna in medieval Portugal with the same type of trap net

598 (*almadrava*). Additionally, Kadeev (1970) suggested that dolphins could be hunted by
599 Chersonesus fishermen with a harpoon. Our study showed a stable presence of common
600 dolphins in archaeological kitchen refuse assemblages, comprising more than half of the
601 specimens that were identify to "dolphin" by ZooMS and could be further identified to
602 species by shotgun sequencing. These were greatly underestimated by the morphological
603 analyses and, thus, support the idea of diverse odontocete catch practices in the ancient
604 Chersonesus. These practices seem to have affected all the local odontocete species and
605 possibly included, not only common fisheries bycatch products, but also, direct cetacean-
606 focused hunting.
607
608 **Implications for management and conservation of Black Sea odontocetes**
609
610 The geographically isolated Black Sea harbour porpoise, as well as other Black Sea
611 odontocetes, seem to have undergone dramatic population size reductions during the past
612 century as a result of substantial hunting and bycatch peaking between the 1930s and 1950s
613 and killing hundreds of thousands, if not a million, animals (Kleinenberg 1956; Birkun Jr
614 2002a; Fontaine *et al.* 2012; Fontaine *et al.* 2014). The extensive fisheries and other
615 anthropogenic disturbances also impacted fish abundance and resulted in an ecosystem
616 regime shift, intrusion of invasive species and economic collapse in the 1970s (Bushuev 2000;
617 Daskalov 2002, 2003; Daskalov *et al.* 2007; Llope *et al.* 2011). Consequently, the three
618 odontocete species inhabiting the Black Sea have been classified as species under special
619 management and conservation concern.
620

To supplement management policies, detailed baseline knowledge on past occurrence and human impacts is needed. Although human activities have clearly escalated in recent decades, our species identification of Late Holocene zooarchaeological material indicate that Black Sea odontocetes, similarly to other marine resources, have been impacted by hunting and fisheries bycatch for millennia. For instance, our population genetic analyses of ancient harbour porpoise samples supported previous studies based on modern material (Fontaine *et al.* 2012), suggesting a long term presence of harbour porpoise in the Black Sea. A more complete understanding of the relative roles of past and present human activities on odontocete demography will require additional zooarchaeological investigations, and upscaling current genetic efforts to include genomic data from both contemporary and ancient samples. In demonstrating the use of biomolecular approaches for species identification, and obtaining the first full mitogenomic and auxiliary nuclear data from ancient Black Sea odontocetes, we have taken the first step.

ACKNOWLEDGEMENTS

The authors wish to thanks the Dr Bøje Benzon Foundation and the Section for Evolutionary Genomics at the Natural History Museum of Denmark for their financial support. The authors also thank Sergey Ushakov, Alexey Ivanov and Larisa Sedikova for providing access to the Black Sea samples and the Danish National High-Throughput DNA Sequencing Centre for sequencing the samples.

REFERENCES

Aelian (1959) *On the Characteristics of Animals*, Trans. A.F. Schofield. 3 vols. TheLoeb Classical Library. London: Heinemann; Cambridge, MA: Harvard University Press.

Allentoft ME, Sikora M, Sjögren K-G*, et al.* (2015) Population genomics of Bronze Age Eurasia. *Nature* **522**, 167-172.

Amaral A, Sequeira M, Martínez-Cedeira J, Coelho M (2007) New insights on population genetic structure of Delphinus delphis from the northeast Atlantic and phylogenetic relationships within the genus inferred from two mitochondrial markers. *Marine Biology* **151**, 1967-1976.

Amundsen C, Perdikaris S, McGovern TH*, et al.* (2013) Fishing booths and fishing strategies in medieval Iceland: an Archaeofauna from the of Akurvík, North-West Iceland. *Environmental Archaeology* **10**, 127-142.

Andersen D, Kinze CC, Skov J (1992) The use of pelvic bones in the harbour forpoise Phocoena phocoena as an indication of sexual maturity. *Lutra* **35**, 105-112.

Ávila-Arcos MC, Cappellini E, Romero-Navarro JA*, et al.* (2011) Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. *Scientific reports* **1**, 74.

Ávila-Arcos MC, Sandoval‐Velasco M, Schroeder H*, et al.* (2015) Comparative performance of two whole‐genome capture methodologies on ancient DNA Illumina libraries. *Methods in Ecology and Evolution* **6**, 725-734.

Benke H (1993) Investigations on the osteology and the functional morphology of the flipper of whales and dolphins (Cetacea). *Investigations on Cetacea* **24**, 9-252.

Bilgmann K, Möller L, Harcourt RG, Gales R, Beheregaray LB (2008) Common dolphins subject to fisheries impacts in Southern Australia are genetically differentiated: implications for conservation. *Animal Conservation* **11**, 518-528.

Birkun Jr A (2002a) Cetacean direct killing and live capture in the Black Sea. *Cetaceans of the Mediterranean and Black Seas: State of Knowledge and Conservation Strategies. A report to the ACCOBAMS Secretariat, Monaco*.

Birkun Jr A (2002b) Interactions between cetaceans and fisheries in the Black Sea. *Cetaceans of the Mediterranean and Black Seas: State of knowledge and conservation strategies* **98**, 107.

Birkun Jr A (2008) *Delphinus delphis ssp. ponticus*. *The IUCN Red List of Threatened Species*.

Birkun Jr A (2012) *Tursiops truncatus ssp. ponticus*. *The IUCN Red List of Threatened Species*.

Birkun Jr A, Frantzis A (2008) *Phocoena phocoena ssp. relicta*. *The IUCN Red List of Threatened Species*.

680 Blow MJ, Zhang T, Woyke T*, et al.* (2008) Identification of ancient remains through genomic
681 sequencing. *Genome research* **18**, 1347-1353.
682 Briggs AW, Good JM, Green RE*, et al.* (2009) Targeted retrieval and analysis of five
683 Neandertal mtDNA genomes. *Science* **325**, 318-321.
684 Brito C, Vieira N (2009) Historical accounts about the occurrence and capture of common
685 dolphins in Portugal mainland. . *International Whaling Commission, Scientific*
686 *Committee, Document IWC - SC/61/SM17: 1-10.*
687 Brown S, Higham T, Slon V*, et al.* (2016) Identification of a new hominin bone from
688 Denisova Cave, Siberia using collagen fingerprinting and mitochondrial DNA
689 analysis. *Scientific reports* **6**.
690 Buckley M, Collins M, Thomas‐Oates J, Wilson JC (2009) Species identification by
691 analysis of bone collagen using matrix‐assisted laser desorption/ionisation time‐
692 of‐flight mass spectrometry. *Rapid communications in mass spectrometry* **23**, 3843-
693 3854.
694 Buckley M, Fraser S, Herman J*, et al.* (2014) Species identification of archaeological marine
695 mammals using collagen fingerprinting. *Journal of Archaeological Science* **41**, 631-
696 641.
697 Buckley M, Harvey VL, Chamberlain AT (2017) Species identification and decay assessment
698 of Late Pleistocene fragmentary vertebrate remains from Pin Hole Cave (Creswell
699 Crags, UK) using collagen fingerprinting. *Boreas*.
700 Buckley M, Kansa SW (2011) Collagen fingerprinting of archaeological bone and teeth
701 remains from Domuztepe, South Eastern Turkey. *Archaeological and Anthropological*
702 *Sciences* **3**, 271-280.
703 Bushuev SG (2000) Depletion of food resources as a factor limiting the number of the Black
704 Sea dolphins. In: *In Ecological Safety of Coastal and Shelf Zones and Comprehensive*
705 *Use of Shelf Resources*, pp. pp. 437–452 (Russian), Sevastopol.
706 Carter JC, Mack GR (2003) *Crimean Chersonesos: city, chora, museum, and environs*
707 University of Texas at Austin Inst of.
708 Cooke RG, Wake TA, Martínez-Polanco MF*, et al.* (2016) Exploitation of dolphins (Cetacea:
709 Delphinidae) at a 6000yr old Preceramic site in the Pearl Island archipelago, Panama.
710 *Journal of Archaeological Science: Reports* **6**, 733-756.
711 Dabney J, Knapp M, Glocke I*, et al.* (2013a) Complete mitochondrial genome sequence of a
712 Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments.
713 *Proceedings of the National Academy of Sciences* **110**, 15758-15763.
714 Dabney J, Meyer M, Pääbo S (2013b) Ancient DNA damage. *Cold Spring Harbor*
715 *perspectives in biology* **5**, a012567.
716 Daskalov GM (2002) Overfishing drives a trophic cascade in the Black Sea. *Marine Ecology*
717 *Progress Series* **225**, 53-63.
718 Daskalov GM (2003) Long-term changes in fish abundance and environmental indices in the
719 Black Sea. *Marine Ecology Progress Series* **255**, 259-270.
720 Daskalov GM, Grishin AN, Rodionov S, Mihneva V (2007) Trophic cascades triggered by
721 overfishing reveal possible mechanisms of ecosystem regime shifts. *Proceedings of*
722 *the National Academy of Sciences* **104**, 10518-10523.
723 Dietl GP, Kidwell SM, Brenner M*, et al.* (2015) Conservation paleobiology: leveraging
724 knowledge of the past to inform conservation and restoration. *Annual Review of Earth*
725 *and Planetary Sciences* **43**, 79.
726 Dunnell RC (1971) *Systematics in Prehistory.* McMillan, New York and London.

727 Evans S, i Godino IB, Álvarez M*, et al.* (2016) Using combined biomolecular methods to
728 explore whale exploitation and social aggregation in hunter–gatherer–fisher society in
729 Tierra del Fuego. *Journal of Archaeological Science: Reports* **6**, 757-767.
730 Fontaine M (2016) Harbour Porpoises, Phocoena phocoena, in the Mediterranean Sea and
731 Adjacent Regions: Biogeographic Relicts of the Last Glacial Period. *Advances in*
732 *Marine Biology* **75**, 333-358.
733 Fontaine MC, Roland K, Calves I*, et al.* (2014) Postglacial climate changes and rise of three
734 ecotypes of harbor porpoises, Phocoena phocoena, in western Palearctic waters.
735 *Molecular Ecology* **23**, 3306-3321.
736 Fontaine MC, Snirc A, Frantzis A*, et al.* (2012) History of expansion and anthropogenic
737 collapse in a top marine predator of the Black Sea estimated from genetic data.
738 *Proceedings of the National Academy of Sciences* **109**, E2569–E2576.
739 Foote AD, Hofreiter M, Morin PA (2012a) Ancient DNA from marine mammals: studying
740 long-lived species over ecological and evolutionary timescales. *Annals of Anatomy-*
741 *Anatomischer Anzeiger* **194**, 112-120.
742 Foote AD, Kaschner K, Schultze SE*, et al.* (2013) Ancient DNA reveals that bowhead whale
743 lineages survived Late Pleistocene climate change and habitat shifts. *Nature*
744 *communications* **4**, 1677.
745 Foote AD, Newton J, Piertney SB, Willerslev E, Gilbert MTP (2009) Ecological,
746 morphological and genetic divergence of sympatric North Atlantic killer whale
747 populations. *Molecular Ecology* **18**, 5207-5217.
748 Foote AD, Thomsen PF, Sveegaard S*, et al.* (2012b) Investigating the potential use of
749 environmental DNA (eDNA) for genetic monitoring of marine mammals. *PLoS ONE*
750 **7**, e41781.
751 Frantzis A, Gordon J, Hassidis G, Komnenou A (2001) The enigma of harbor porpoise
752 presence in the Mediterranean Sea. *Marine Mammal Science* **17**, 937-944.
753 Galatius A, Gol'din P (2011) Geographic variation of skeletal ontogeny and skull shape in the
754 harbour porpoise (Phocoena phocoena). *Canadian Journal of Zoology* **89**, 869-879.
755 Gilbert MTP, Drautz DI, Lesk AM*, et al.* (2008) Intraspecific phylogenetic analysis of
756 Siberian woolly mammoths using complete mitochondrial genomes. *Proceedings of*
757 *the National Academy of Sciences* **105**, 8327-8332.
758 Gol'din P (2004) Growth and body size of the harbour porpoise, Phocoena phocoena
759 (Cetacea, Phocoenidae), in the Sea of Azov and the Black Sea. *Vestnik Zoologii*.
760 Gol'din P, Vishnyakova K (2016) Habitat shapes skull profile of small cetaceans: evidence
761 from geographical variation in Black Sea harbour porpoises (Phocoena phocoena
762 relicta). *Zoomorphology* **135**, 387-393.
763 Goldstein PZ, DeSalle R (2011) Integrating DNA barcode data and taxonomic practice:
764 determination, discovery, and description. *Bioessays* **33**, 135-147.
765 Grayson DK (1984) *Quantitative Zooarchaeology: Topics in the analysis of archaeological*
766 *faunas* Elsevier.
767 Harvey VL, Egerton VM, Chamberlain AT, Manning PL, Buckley M (2016) Collagen
768 fingerprinting: a new screening technique for radiocarbon dating ancient bone. *PLoS*
769 *ONE* **11**, e0150650.
770 Hofman CA, Rick TC, Fleischer RC, Maldonado JE (2015) Conservation archaeogenomics:
771 ancient DNA and biodiversity in the Anthropocene. *Trends in Ecology & Evolution*
772 **30**, 540-549.
773 Højte JM (2005) The archaeological evidence for fish processing in the Black Sea region.
774 *Bekker-Nielsen* **2005**, 133-160.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755.
Kadeev VI (1970) Essays on the History and Economy of Chersonesos Tauricus in the I-IVth centuries AD. Izdatelstvo Kharkovskogo Universiteta. Kharkov. (Russian).
Kadeev VI, Sorochan SB (1989) Economical Ties of Antique Cities of the Northern Black Sea Region in the 1st cent. BC - 5th cent. AD. Izdatelstvo Kharkovskogo Universiteta. Kharkov. (Russian).
Kearse M, Moir R, Wilson A*, et al.* (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649.
Kirby DP, Buckley M, Promise E, Trauger SA, Holdcraft TR (2013) Identification of collagen-based materials in cultural heritage. *Analyst* **138**, 4849-4858.
Kleinenberg SE (1956) Mammals of the Black Sea and the Sea of Azov: an experience of biological and fisheries research. *Moscow: Izdatelstvo Akademii Nauk SSSR.*, (Russian).
Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: analysis of next generation sequencing data. *BMC bioinformatics* **15**, 356.
Leigh JW, Bryant D (2015) popart: full‐feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**, 1110-1116.
Leonardi M, Librado P, Der Sarkissian C*, et al.* (2016) Evolutionary patterns and processes: lessons from ancient DNA. *Systematic Biology*, syw059.
Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760.
Li H, Handsaker B, Wysoker A*, et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079.
Lindgreen S (2012) AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC research notes* **5**, 337.
Lindqvist C, Probst A, Martin AR, Wiig Ø, Bachmann L (2009) Molecular species identification of historical whale remains from South Georgia. *Marine Mammal Science* **25**, 229-238.
Llope M, Daskalov GM, Rouyer TA*, et al.* (2011) Overfishing of top predators eroded the resilience of the Black Sea system regardless of the climate and anthropogenic conditions. *Global Change Biology* **17**, 1251-1265.
Lorenzen ED, Nogués-Bravo D, Orlando L*, et al.* (2011) Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* **479**, 359-364.
McLeod BA, Brown MW, Frasier TR, White BN (2010) DNA profile of a sixteenth century western North Atlantic right whale (Eubalaena glacialis). *Conservation Genetics* **11**, 339-345.
McLeod BA, Frasier TR, Dyke AS, Savelle JM, White BN (2012) Examination of ten thousand years of mitochondrial DNA diversity and population demographics in bowhead whales (Balaena mysticetus) of the Central Canadian Arctic. *Marine Mammal Science* **28**, E426-E443.
Morales A, Antipina E, Roselló E, Antipina A (2007) An ichthyoarchaeological survey of the ancient fisheries from the Northern Black Sea. *Archaeofauna* **16**, 117-172.
Moura AE, Nielsen SC, Vilstrup JT*, et al.* (2013) Recent diversification of a marine genus (Tursiops spp.) tracks habitat preference and environmental change. *Systematic Biology* **62**, 865-877.

822 Mulville J (2002) The role of cetacea in prehistoric and historic Atlantic Scotland.
823 *international Journal of Osteoarchaeology* **12**, 34-48.
824 Nichols C, Herman J, Gaggiotti OE*, et al.* (2007) Genetic isolation of a now extinct
825 population of bottlenose dolphins (Tursiops truncatus). *Proceedings of the Royal*
826 *Society B: Biological Sciences* **274**, 1611-1616.
827 Pfaffl M (2001) A new mathematical model for relative quantification in real-time RT-PCR.
828 *Nucleic Acids Research* **29**, e45.
829 Picard. http://picard.sourceforge.net
830 Pimm SL, Jenkins CN, Abell R*, et al.* (2014) The biodiversity of species and their rates of
831 extinction, distribution, and protection. *Science* **344**, 1246752.
832 Porphyrogenitus CV (1993) *Constantine Porphyrogenitus, De administrando imperio.* Edited
833 by G Moravcsik and translated by RJH Jenkins, Dumbarton Oaks Center for
834 Byzantine Studies, Washigton D.C.
835 Rasmussen M, Li Y, Lindgreen S*, et al.* (2010) Ancient human genome sequence of an
836 extinct Palaeo-Eskimo. *Nature* **463**, 757-762.
837 Rohland N, Hofreiter M (2007) Ancient DNA extraction from bones and teeth. *Nature*
838 *protocols* **2**, 1756-1762.
839 Schipper J, Chanson JS, Chiozza F*, et al.* (2008) The status of the world's land and marine
840 mammals: diversity, threat, and knowledge. *Science* **322**, 225-230.
841 Semenov-Zuser SA (1947) Fish market in Chersonesus. *Vestnik Drevney Istorii* **2**, 237–246
842 (Russian).
843 Speller C, van den Hurk Y, Charpentier A*, et al.* (2016) Barcoding the largest animals on
844 Earth: ongoing challenges and molecular solutions in the taxonomic identification of
845 ancient cetaceans. *Phil. Trans. R. Soc. B* **371**, 20150332.
846 Sremba A, Martin A, Baker C (2010) Genetic approach to species identification of whale
847 bones from South Georgia Island whaling stations. *Report to Scientific Committee of*
848 *the International Whaling Commission SC/62/SH19*.
849 Steadman DW (1995) Prehistoric extinctions of Pacific island birds: biodiversity meets
850 zooarchaeology. *Science* **267**, 1123.
851 Strabo HLJ (1929) The Geography of Strabo.
852 Strohalm M, Hassman M, Košata B, Kodíček M (2008) mMass data miner: an open source
853 alternative for mass spectrometric data analysis. *Rapid communications in mass*
854 *spectrometry* **22**, 905-908.
855 Thunell RC (1979) Pliocene-Pleistocene paleotemperature and paleosalinity history of the
856 Mediterranean Sea: Results from DSDP Sites 125 and 132. *Marine Micropaleontology*
857 **4**, 173-187.
858 Thunell RC, Williams DF, Kennett JP (1977) Late Quaternary paleoclimatology, stratigraphy
859 and sapropel history in eastern Mediterranean deep-sea sediments. *Marine*
860 *Micropaleontology* **2**, 371-388.
861 Tregenza N, Berrow S, Hammond P, Leaper R (1997) Common dolphin, Delphinus delphis
862 L., bycatch in bottom set gillnets in the Celtic Sea. *Report of the International*
863 *Whaling Commission* **47**, 835-839.
864 Ushakov SV (2011) Kruze Basilica in Chersonesus: main results of studies. In: *Black Sea*
865 *Region in the Middle Ages.*, pp. 198–212 (Russian). Aleteia, St. Petersburg.
866 Ushakov SV (2013) A new fish salting cistern from Chersonesus (on the issue of time of
867 construction of the Kruze basilique). In: *The Ancient Black Sea Region*, pp. 590–596
868 (Russian). Fridman, A S, Odessa.

Ushakov SV, Lesnaya ES, Tyurin MI (2013) A new ceramic complex, 5th century BCE from the north-eastern part of Chersonesus. In: *The Black Sea Region in the Ancient and Early Middle Ages*, pp. 308–315 (Russian). Southern Scientific Centre of the Russian Academy of Sciences, Rostov-on-Don.

Viaud-Martinez KA, Brownell RL, Komnenou A, Bohonak AJ (2008) Genetic isolation and morphological divergence of Black Sea bottlenose dolphins. *Biological Conservation* **141**, 1600-1611.

Viaud-Martínez KA, Martínez Vergara M, Gol'din P*, et al.* (2007) Morphological and genetic differentiation of the Black Sea harbour porpoise *Phocoena phocoena*. *Marine Ecology Progress Series* **338**, 281-294.

Vishnyakova K, Gol'din P (2015a) Seasonality of strandings and bycatch of harbour porpoises in the Sea of Azov: the effects of fisheries, weather conditions, and life history. *ICES Journal of Marine Science: Journal du Conseil* **72**, 981-991.

Vishnyakova K, Gol'din PE (2015b) Differences in skull size of harbour porpoises, Phocoena phocoena (Cetacea), in the Sea of Azov and the Black Sea: evidence for different morphotypes and populations. . *Vestnik Zoologii* **49**, pp.171-180.

Von Holstein IC, Ashby SP, van Doorn NL*, et al.* (2014) Searching for Scandinavians in pre-Viking Scotland: molecular fingerprinting of Early Medieval combs. *Journal of Archaeological Science* **41**, 1-6.

Welker F, Collins MJ, Thomas JA*, et al.* (2015) Ancient proteins resolve the evolutionary history of Darwin/'s South American ungulates. *Nature* **522**, 81-84.

Welker F, Hajdinjak M, Talamo S*, et al.* (2016) Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proceedings of the National Academy of Sciences*, 201605834.

Willerslev E, Cooper A (2005) Review paper. ancient dna. *Proceedings of the Royal Society of London B: Biological Sciences* **272**, 3-16.

Yang DY, Speller C (2006) Co-amplification of cytochrome b and D-loop mtDNA fragments for the identification of degraded DNA samples. *Molecular Ecology Notes* **6**, 605-608.

Authors Contribution

V.B., P.G. and M.T.O. designed the research; E.G., K.V. and P.G. collected zooarchaeological samples and processed the morphological analysis; V.B. performed aDNA based laboratory work advised by N.W. and M.T.O; K.MG. and C.S. performed protein based laboratory work; V.B., K.MG., F.G.V., N.W., M.C.F., C.S. and M.T.O. analysed the molecular data; V.B., P.G., C.S. and M.T.O. wrote the manuscript; V.B. and M.T.O. provides funding; all authors commented on previous version of the manuscript and approved the final version.

Data Accessibility

DNA sequences: Genbank accession XXXXX (to be added upon accept of the manuscript)

Supporting Information

Table S1. Morphology of cetacean remains from the Kruze basilica in Chersonesus (400–530 CE), excavations 2011-13.

Table S2. Description of the morphological characteristics of bones structures used for the species identification.

Table S3. Description of the DNA extract characterisation per sample.

Table S4. Description of the sequencing results per samples. NS=Not Sequenced.

Table S5. Designated peptide markers used for taxonomic identification of the Black Sea samples.

Figure S1. Map showing the sampling locations of the Black Sea samples. *Inset* shows area of the main map.

Figure S2. Picture of caudal vertebrae of extant *Tursiops truncatus* (A) and *Delphinus delphis* (B) typically used for morphological species determination of subfossil material.

Figure S3. Melting curves for BS_29 qPCR amplifications. The gray line with dots, the green line with squares and the blue line with upside down triangles represent respectively the results for the 1:5, 1:10 and 1:20 dilutions. The other four yellow curves represent negative controls. Amplification can be observe for two of the negative controls, however the dissociated products are different form those obtained for the samples dilutions.

Figure S4. Bayesian phylogenetic tree of ancient mitogenomes infer with MrBayes algorithm. *Mesoplodon ginkgodens* was used as an out-group and numbers on nodes represent posterior probabilities.

Figure S5. Averaged MALDI-ToF mass spectra from sample BS_03. Rectangles indicate the peptides used to identify the sample as a probable carnivore.

FIGURES & TABLES

Figure 1. Summary of the different techniques for identification of subfossil odontocete remains. (A) Odontocete samples from Chersonesus showing the state of preservation and degree of fragmentation. (B) Chromatograms resulting from the Sanger sequencing of CytB gene fragments from odontocete samples. (C) Reads alignment to the *Phocoena phocoena* reference mitogenome resulting from the Shotgun sequencing of the sample BS_14. (D) MALDI-ToF mass spectra displaying peptide marker (P2) distinguishing the porpoise from the dolphin group.

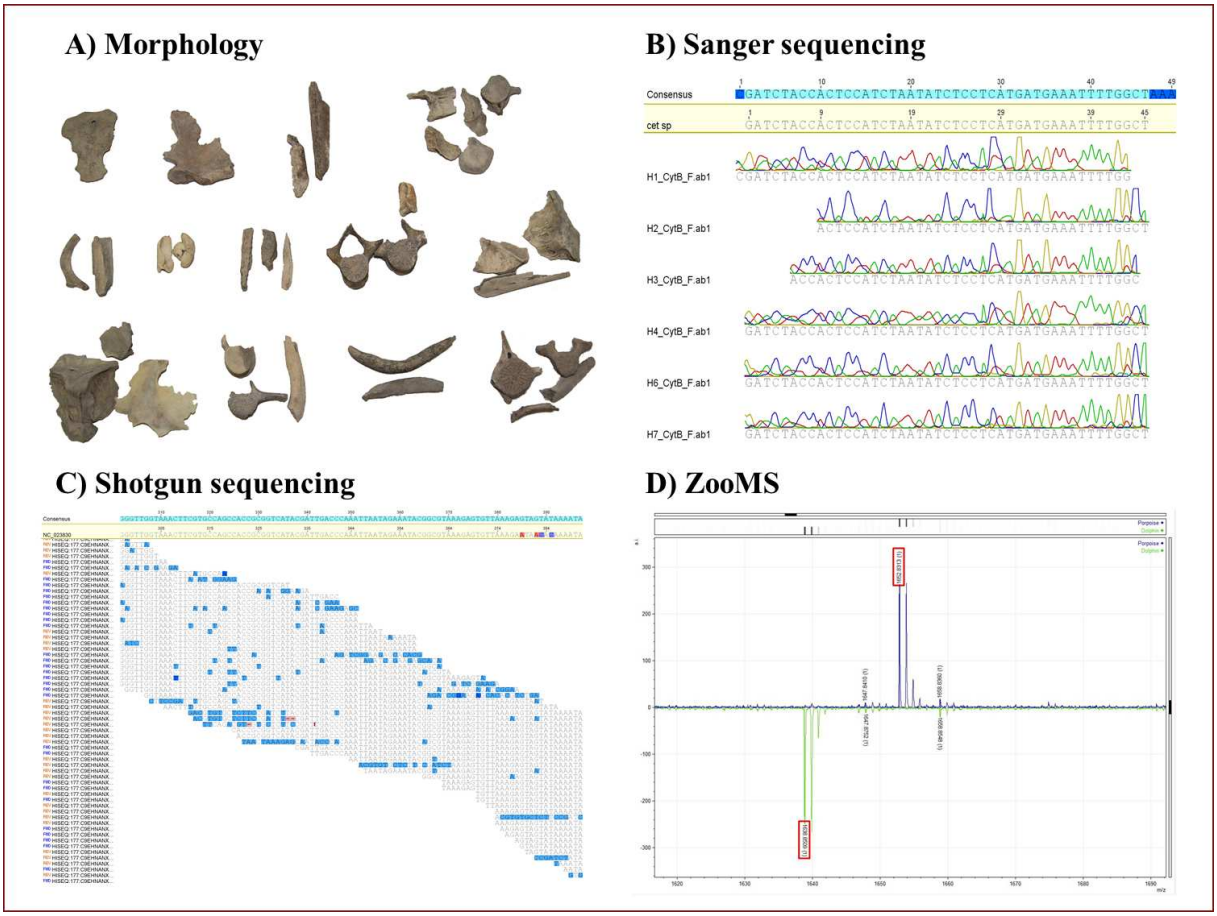

Figure 2. The correlation between amount of starting material (bone powder weight) and (A) peak molarity and (B) normalized qPCR Ct-values. Note that higher peak molarity and lower Ct-values typically signifies higher DNA content.

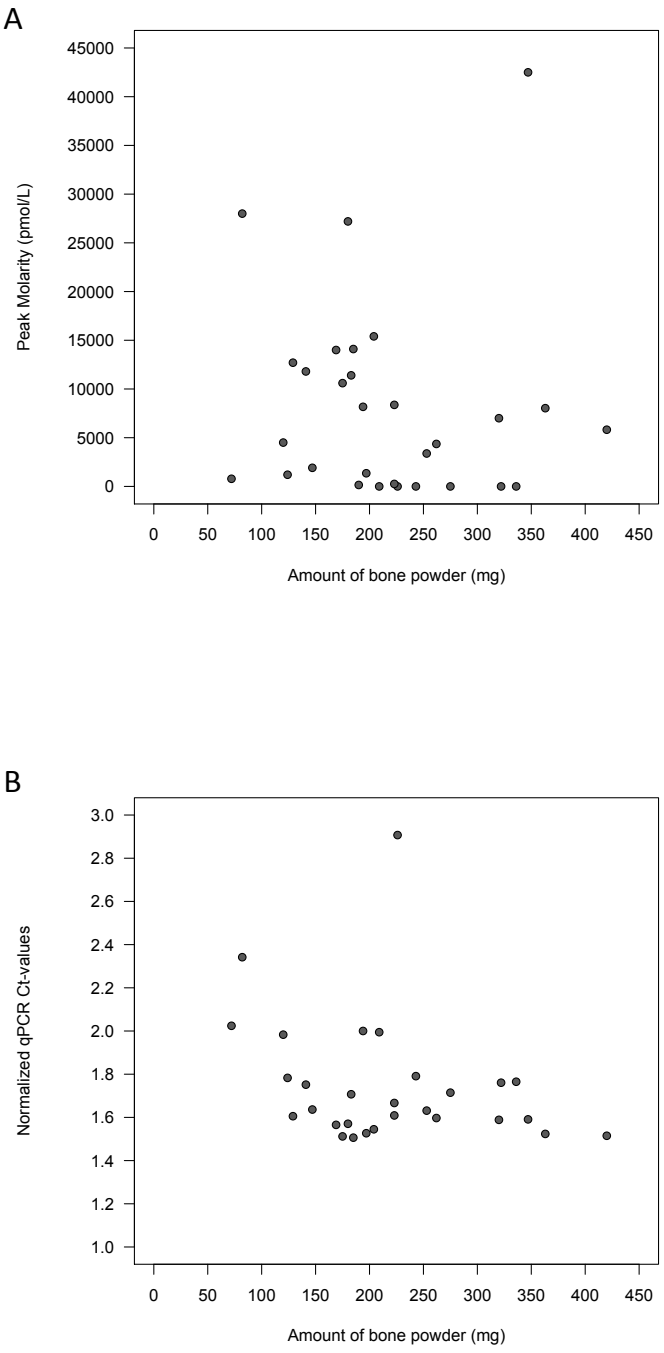

Figure 3. Mitogenome recovery and sequencing depth of coverage resulting from the shotgun sequencing.

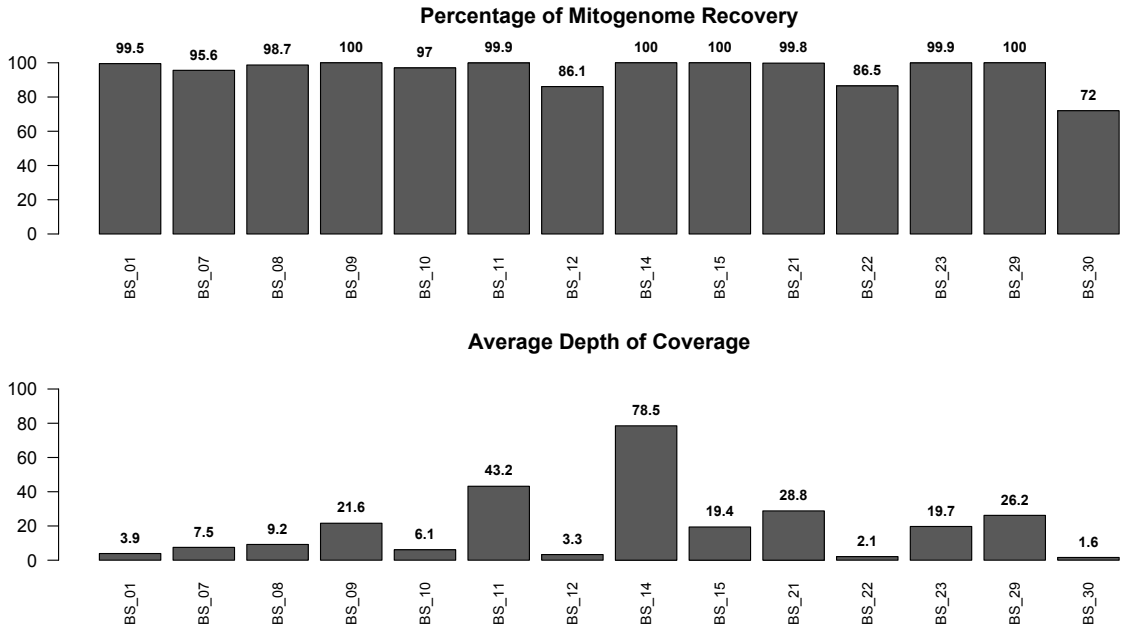

Figure 4. Haplotype network among Black Sea harbour porpoise based on the 706bp mitochondrial sequence.

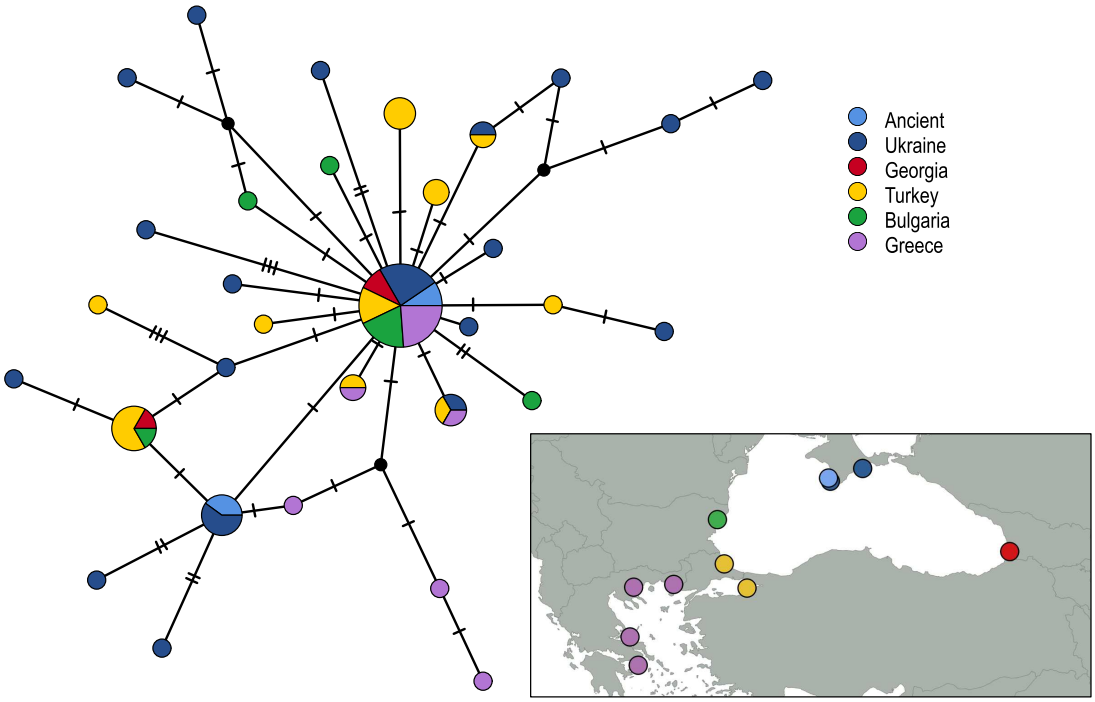

Table 1. Odontocete samples and results of the species determination by morphology, CytB Sanger sequencing, shotgun sequencing of mitogenomes and ZooMS. Notice that the three molecular approaches gave overall very similar results, with the exception of samples BS_3 and BS_10.

ID	Year (CE)	Bone	Morphology	Sanger	Shotgun	ZooMS[a]
BS_01	400-530	Mandible	Cetacea sp.	Dolphin sp.	*D. delphis*	'Dolphin'
BS_02	400-530	Mandible	Cetacea sp.	Dolphin sp.	Failed	'Dolphin'
BS_03	275-300	Mandible	*D. delphis*	*P. phocoena*	Failed	Carnivora
BS_04	400-530	Mandible	*P. phocoena*	Dolphin sp.	Failed	'Dolphin'
BS_05	400-530	Tooth	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_06	400-530	Vertebra	*P. phocoena*	Dolphin sp.	Failed	'Dolphin'
BS_07	400-530	Vertebra	*P. phocoena*	*P. phocoena*	*P. phocoena*	'Porpoise'
BS_08	400-530	Epiphysis	*P. phocoena*	Dolphin sp.	*D. delphis*	'Dolphin'
BS_09	400-530	Epiphysis	*P. phocoena*	*P. phocoena*	*P. phocoena*	'Porpoise'
BS_10	400-530	Occip. condyle	Cetacea sp.	Dolphin sp.	*P. phocoena*	'Porpoise'
BS_11	400-530	Vertebra	*P. phocoena*	*P. phocoena*	*P. phocoena*	'Porpoise'
BS_12	400-530	Vertebra	*P. phocoena*	*P. phocoena*	*P. phocoena*	'Porpoise'
BS_13	400-530	Vertebra	*P. phocoena*	Dolphin sp.	Failed	'Dolphin'
BS_14	400	Teeth	*P. phocoena*	*P. phocoena*	*P. phocoena*	'Porpoise'
BS_15	400	Skull	*P. phocoena*	*P. phocoena*	*P. phocoena*	'Porpoise'
BS_16	400	Skull	*P. phocoena*	*P. phocoena*	N/A	N/A
BS_17	1000-1200	Vertebra	*P. phocoena*	Dolphin sp.	Failed	'Dolphin'
BS_18	470-500	Tooth	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_19	470-500	Tooth	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_20	470-500	Vertebra	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_21	470-500	Epiphysis	*T. truncatus*	Dolphin sp.	*T. truncatus*	'Dolphin'
BS_22	470-500	Vertebra	*T. truncatus*	Failed	*D. delphis*	'Dolphin'
BS_23	470-500	Vertebra	*T. truncatus*	Dolphin sp.	*D. delphis*	'Dolphin'
BS_24	1000-1100	Vertebra	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_25	1000-1100	Vertebra	*T. truncatus*	Failed	Failed	'Dolphin'
BS_26	1000-1100	Vertebra	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_27	470-500	Vertebra	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_28	470-500	Vertebra	*T. truncatus*	Dolphin sp.	Failed	'Dolphin'
BS_29	470-500	Vertebra	*T. truncatus*	Dolphin sp.	*T. truncatus*	'Dolphin'
BS_30	470-500	Vertebra	*T. truncatus*	Dolphin sp.	*T. truncatus*	'Dolphin'

[a] 'Dolphin' includes '*Tursiops truncatus*, *Stenella coeruleoalba*, *Delphinus delphis*, *Lagenorhynchus albirostris*'; 'Porpoise' includes '*Orcinus orca*, *Phocoenoides dalli*, *Phocoena phocoena*, *Lagenorhynchus obliquidens*'; Note that sample BS_16 was not shotgun sequenced or tested using ZooMS.