Resource Article: Permanent Genetic Resources

Combined Hybridization Capture and Shotgun Sequencing for Ancient DNA Analysis of Extinct Wild and Domestic Dromedary Camel

Elmira Mohandesan [1,2*], Camilla F. Speller [3], Joris Peters [4], Hans-Peter Uerpmann [5], Margarethe Uerpmann [5,] Bea De Cupere [6], Michael Hofreiter [3,7], Pamela A. Burger [1*]

[1]Research Institute of Wildlife Ecology, Vetmeduni Vienna, Savoyenstraße 1, 1160 Vienna, Austria

[2]Institute of Population Genetics, Vetmeduni Vienna, Veterinärplatz 1, 1210 Vienna, Austria

[3]BioArCh, Department of Archaeology, University of York, Wentworth Way, York, YO10 5DD, UK

[4]Institute of Palaeoanatomy, Domestication Research and the History of Veterinary Medicine, LMU Munich and SNSB, Bavarian State Collection of Anthropology and Palaeoanatomy, Munich, Germany

[5]Institut für Naturwissenschaftliche Archäologie, Abteilung Archäozoologie, Eberhard-Karls-Universität Tübingen, Rümelinstrasse 23, 7207 Tübingen, Germarny

[6]Royal Belgian Institute of Natural Sciences, Vautierstraat 29, B-1000 Brussels, Belgium

[7]Evolutionary and Adaptive Genomics, Institute for Biochemistry and Biology, Department of Mathematics and Natural Sciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, Potsdam, 14476, Germany

[*]Corresponding authors:

Elmira Mohandesan

Institute of Wildlife Ecology, Vetmeduni Vienna
Savoyenstraße 1
1160 Wien, Austria
Tel: +43-1-25077-7141
Fax: +43-1-25077-7941
Email: elmira.mohandesan@vetmeduni.ac.at
Pamela Burger
Institute of Wildlife Ecology, Vetmeduni Vienna
Savoyenstraße 1
1160 Wien, Austria
Tel: +43-1-25077-7141
Fax: +43-1-25077-7941
Email: pamela.burger@vetmeduni.ac.at

Running title: Ancient dromedary mitogenomes

Key words: *Camelus dromedarius*, ancient DNA (aDNA), degraded DNA, mitochondrial genome (mtDNA), capture enrichment, next generation sequencing (NGS)

Abstract

The performance of hybridization capture combined with next generation sequencing (NGS) has seen limited investigation with samples from hot and arid regions until now. We applied hybridization capture and shotgun sequencing to recover DNA sequences from bone specimens of ancient-domestic dromedary (*Camelus dromedarius*) and its extinct ancestor, the wild dromedary from Jordan, Syria, Turkey and the Arabian Peninsula, respectively. Our

results show that hybridization capture increased the percentage of mitochondrial DNA (mtDNA) recovery by an average 187-fold and in some cases yielded virtually complete mitochondrial (mt) genomes at multi-fold coverage in a single capture experiment. Furthermore we tested the effect of hybridization temperature and time by using a touchdown approach on a limited number of samples. We observed no significant difference in the number of unique dromedary mtDNA reads retrieved with the standard capture compared to the touchdown method. In total, we obtained 14 partial mitochondrial genomes from ancient-domestic dromedaries with 17 - 95% length coverage and 1.27 – 47.1-fold read depths for the covered regions. Using whole genome shotgun sequencing, we successfully recovered endogenous dromedary nuclear DNA (nuDNA) from domestic and wild dromedary specimens with 1 – 1.06-fold read depths for covered regions. Our results highlight that despite recent methodological advances, obtaining ancient DNA (aDNA) from specimens recovered from hot, arid environments is still problematic. Hybridization protocols require specific optimization, and samples at the limit of DNA preservation need multiple replications of DNA extraction and hybridization capture as has been shown previously for Middle Pleistocene specimens.

Introduction

The pioneering world of next generation sequencing (NGS) (Margulies *et al*. 2005; Millar *et al.* 2008; Shendure & Ji 2008) has advanced the field of aDNA tremendously, from sequencing short fragments of mtDNA (Higuchi *et al.* 1984) to generating datasets of genome scale from extant and extinct species (Green *et al.* 2010; Reich *et al.* 2010; Orlando *et al.* 2011; Meyer *et al.* 2012; Orlando *et al.* 2013; Prüfer *et al.* 2014; Rasmussen *et al.* 2014). Although whole ancient genomes are becoming more readily accessible, mitochondrial

genomes (mitogenomes) are still the marker of choice in aDNA studies dealing with samples with very poor DNA preservation (Dabney *et al.* 2013; Meyer *et al.* 2014), or when comparing mitochondrial diversity between ancient and modern populations (Zhang *et al.* 2013; Thalmann *et al.* 2013; Almathen *et al.* 2016). Despite recent methodological progress, aDNA research is still fraught with technical complications, such as low template quantities, high fragmentation, miscoding lesions (Stiller *et al.* 2006; Briggs *et al.* 2007; Brotherton *et al.* 2007; Briggs *et al.* 2010; Sawyer *et al*. 2012), and contamination with modern DNA (Green *et al.* 2006; Surakka *et al.* 2010; Rasmussen *et al.* 2011). Only in few cases, such as permafrost samples (Palkopoulou *et al.* 2015), rare cave findings (Reich *et al.* 2010; Prüfer *et al.* 2014) or when sampling the petrous bone of the cranium (Gamba *et al.* 2014; Pinhasi *et al.* 2015) a high ratio of endogenous DNA (4 - 85%) *versus* environmental and contaminant DNA has been reported. Moreover, the rate of DNA integrity is negatively correlated to the ambient temperature to which the samples were exposed (Smith *et al.* 2001; Allentoft *et al.* 2012; Hofreiter *et al.* 2015). While poor DNA preservation from palaeontological samples collected in arid regions poses significant technical challenges (Paijmans *et al.* 2013), aDNA sequences have occasionally been reported from arid regions, and contributed significantly to understanding prehistoric events (*e.g.*, Orlando *et al.* 2006; Meiri *et al.* 2013; Bollongino *et al.* 2013; Fernández *et al*. 2014; Almathen *et al.* 2016). In this study, we focused on archaeological samples from wild and domestic dromedaries, a species typically associated with hot and arid regions.

The single-humped dromedary (*C. dromedarius*) is the most numerous and widespread domestic camel species inhabiting northern and eastern Africa, the Arabian Peninsula and southwest Asia; a large feral population exists in Australia (Köhler-Rollefson 1991; Spencer & Woolnough 2010). Dromedaries are bred for multiple purposes including meat, milk, wool,

transportation and sport (Bulliet 1990; Grigson 2012). They are particularly well adapted to hot, desert conditions and show a variety of biological and physiological characteristics of evolutionary, economic and medical importance (Wu *et al.* 2014). Zooarchaeological research suggests that the domestication of dromedaries (*C. dromedarius*) occurred around 2000 - 1000 BCE (before the common era) on the Southeast coast of the Arabian Peninsula (Rowley-Conwy 1988; Uerpmann & Uerpmann 2002; Iamoni 2009; Grigson 2012; Uerpmann & Uerpmann 2012; Magee 2015). This has recently been confirmed by phylogenetic and phylogeographic analyses of modern global dromedary populations, including aDNA analysis of wild dromedaries (Almathen *et al.* 2016), which likely became extinct in the early first millennium BCE (Uerpmann & Uerpmann 2002; von den Driesch & Obermaier 2007; Uerpmann & Uerpmann 2012; Grigson 2014).

The remains of a single large-sized Late Pleistocene camel individual recovered from the site 1040 near Wadi Halfa were first evaluated by Gautier (1966), who assigned them to *Camelus thomasi*, the giant North African camel. Based on a limited number of comparative specimens and few metrical data, the author at that time concluded that the Site 1040 specimen exhibited close relationship to the two-humped domestic camel *C. bactrianus*. Following this study, Peters (1998) revisited the same assemblage by using a much larger set of comparative specimens and drawing on the work of Steiger (1990). This revision concluded that all specimens available for re-study, *i.e.* distal humerus, distal radius-ulna, distal tibia and calcaneus exhibited features that are characteristic not of the two-humped but of the one-humped camel *C. dromedarius.* Towards the end of the Pleistocene, *C. thomasi* likely disappeared from Africa, given its absence in archaeological sites, natural deposits and rock art dating to the Holocene. The proximity of Northeast Africa and the Arabian Peninsula opens up the possibility that either *C. thomasi* or a closely related form survived in Southwest

121 Asia, giving rise in to the wild ancestor of domestic population at the transition of the Late
122 Bronze to the Iron Age.
123 The study of aDNA thus presents a unique opportunity to explore the genetic make-up and
124 variation in a wild progenitor population prior to the species' domestication. In other
125 livestock species, an increasing number of genetic studies have taken advantage of ancient
126 and historical samples from both extant and extinct species (Elbaum *et al.* 2006; Amaral *et al.*
127 2011; Cai *et al.* 2011; Kimura *et al.* 2011; Zhang *et al.* 2013; Girdland *et al.* 2014; Schubert *et*
128 *al.* 2014) to investigate the historical domestication process. However, no genetic data from
129 archaeological dromedary specimens have been available until recently (Almathen *et al.*
130 2016). This could be due to the general rarity of *C. dromedarius* specimens in archaeological
131 contexts, even within the current and historical geographical distributions of dromedaries, and
132 the challenging task of obtaining DNA from archaeological remains in desert regions.
133 In this study we explore two methodological strategies to recover mitochondrial genomes
134 from ancient dromedary specimens: 1) double- or single-stranded DNA library (DSL or SSL)
135 preparation (Meyer & Kircher 2010; Gansauge & Meyer 2013; Fortes & Paijmans 2015)
136 followed by hybridization enrichment (Briggs *et al.* 2009; Maricic *et al.* 2010; Fu *et al.* 2013)
137 and NGS sequencing; and 2) DSL preparation followed by whole genome shotgun-
138 sequencing. We describe the efficiency of the enrichment method, when applied to aDNA
139 libraries with variable levels of endogenous DNA. We also compare the effect of
140 hybridization condition on recovering the captured targets after the hybridization step in two
141 different enrichment methods. This study highlights one of the few successful recoveries of
142 DNA sequences from specimens excavated in hot and arid environments.
143

144 **Materials and Methods**

Ancient-domestic and wild dromedary samples

We analysed 54 ancient-domestic dromedary samples (100 BCE - 1870 CE) from excavation sites in Sagalassos, Turkey (Early Byzantine: 450-700 CE); Apamea, Syria (Early Byzantine: 400-600 CE); Palmyra, Syria (100 BCE- 300 CE) and Aqaba, Jordan (Ottoman: 1456-1870 CE, Mamluk: 1260 -1456 CE). We also analysed 22 wild dromedary specimens (5000 – 1130 BCE) from archaeological sites of Al Sufouh-2 (Wadi Suq Middle Bronze Age *ca.* 2000-1600 BCE); Tell Abraq (Late Bronze – Iron Age: 1260-500 BCE); Muweilah (older than 1000-586 BCE); Umm an-Nar (Early Bronze Age: 3000-2000 BCE) and Al-Buhais 18 (5000-4000 BCE) in the United Arab Emirates (UAE). In addition, we analysed one Upper Palaeolithic wild giant camel sample (*C. thomasi*) found below sediments dated to *ca.* 20,000 BCE and collected during the Combined Nubian Prehistory and Geological Campaign in the early 1960s at Site 1040, located in the northern Sudanese Nile valley close to Wadi Halfa, near the boundary with Egypt. The description of the samples and their geographical location are detailed in Table S1 and Fig 1.

Holocene climate change in regions of sample collection

After the initial warming at the end of the Ice Age (around 10,000 BCE) the climate in the Middle East began to change from cooler and moister (~ 4000 BCE) to warmer and more arid (~ 3000 BCE), reaching today's condition only at the very beginning of the Iron Age (~1200 BCE) (Preston *et al.* 2015; Hume *et al*. 2016), which according to present data coincides with the early domestication stages of the dromedary. Nevertheless, there is no evidence that the aridification caused the domestication of camels in this region. It may, however, have increased the value of tamed camels, which would have become more useful during times of drought. Although the climatic and environmental conditions from where the samples were

collected varied to some extent during the Holocene, they allowed for the existence of dromedaries in all the respective areas.

Ancient DNA extraction

The bone samples were prepared in a dedicated and highly contained aDNA laboratory at the Palaeogenetic Core Facility of the ArchaeoBioCenter at the LMU Munich, Germany, with appropriate contamination precautions in place (Knapp *et al.* 2012). For each sample, approximately 200 - 250 mg of bone powder were used for DNA extraction. Two independent DNA extractions in the presence of extraction blanks (one blank per six extractions) were conducted following a silica-based extraction protocol (Rohland & Hofreiter 2007; Rohland *et al.* 2010). DNA was eluted in 50 µL TET buffer and stored at -20°C. In addition, we extracted DNA from a subset of wild dromedaries (six samples) and one ancient giant camel (*C. thomasi*) in the presence of one extraction blank, using the Dabney *et al.* (2013) DNA extraction protocol. In this method, we used approximately 120 - 125 mg of bone powder and the final DNA extracts were eluted in 25 µL TET. The DNA extracts obtained by applying the Rohland *et al.* (2010) protocol were used for double-stranded DNA library preparation (DSL) (Meyer & Kircher 2010), while the DNA extracts following Dabney *et al.* (2013) were used for single-stranded library (SSL) preparation (Gansauge & Meyer 2013). To recover greater quantities of short DNA fragments we combined Dabney *et al.* (2013) DNA extraction and SSL methods (Gansauge & Meyer 2013), as both methods have been proposed for highly degraded samples.

Illumina sequencing library preparation

192 The quality of DNA extraction in each batch (12 bone samples and 2 blanks per batch) was
193 evaluated by amplification of an 80 bp (base pair) fragment (including primers) of the
194 dromedary mtDNA d-loop (see supplementary information). Only a subset of ancient-
195 domestic samples with successful PCR amplification (44 out of 54 samples) was further used
196 for library construction and NGS sequencing, while all 22 wild dromedary DNA extracts
197 regardless of positive / negative PCR results were included in further analyses (Fig 2). The
198 Illumina DSLs were built directly from the DNA extracts as well as extraction blanks and
199 negative controls (library blanks), following the Fortes and Paijmans (2015) protocol. This
200 protocol is based on the original Illumina library construction method by Meyer and Kircher
201 (2010) with specific optimizations for samples with degraded DNA. Purification steps
202 throughout the library construction protocol were performed with MinElute purification
203 columns (Qiagen) according to the manufacturer's instructions. The libraries were constructed
204 using an 8 bp barcode on the 3' end of the P5 adapter (directly adjacent to the 5' end of the
205 aDNA template), which served as an additional means to assign sequences to samples (Fortes
206 and Paijman 2015). In addition, it provided extra information to filter chimeric reads (or
207 jumping PCR) from the dataset, and thus increased the confidence in assigning the reads to a
208 particular library. This barcoding method did not require an additional sequence read; the 8 bp
209 P5 barcode was retrieved as part of the R1 forward reads. The 8 bp P5 barcode for each
210 sample was identical to its P7 index; sequences of the indices and the modified Illumina
211 adapters are listed in Tables S1 and S2, respectively.
212 Following library construction and pre-indexing amplification, we performed parallel
213 indexing PCRs (to apply the P5 barcode) to maintain more complexity of each library during
214 amplification (see supplementary information). As endogenous DNA in ancient samples is
215 usually present in low quantity, amplification of the library can introduce biases by

amplifying certain fragments. We reduced this loss of complexity by amplifying each library in six parallel indexing PCR (to apply the P5 barcode) reactions, each containing a unique subset of the original library as starting templates (see supplementary information; library preparation and indexing PCR to apply the P5 barcode). The PCR products were pooled in equimolar ratios, purified through a single Qiagen MinElute spin column, and eluted in 20 μL elution buffer (EB) following 10 min incubation at room temperature. The DSL preparation was performed in a dedicated aDNA laboratory at the University of York, UK, following standard contamination precautions (Knapp *et al.* 2012). In addition, we constructed seven single-stranded libraries (SSL) (Gansauge & Meyer 2013) from six wild dromedaries and one giant one-humped camel (*C. thomasi*) in the presence of one extraction and one library blank (Table S1). The SSL preparations were conducted in a dedicated aDNA laboratory at the University of Copenhagen, Denmark.

In-solution hybridization capture and sequencing

Dromedary complete mtDNA was enriched in indexed DSLs (domestic and wild) by in-solution hybridization capture (Table S3), using MYcroarray's MYbaits kit according to the manufacturer's instructions. We also performed the alternative 'MYbaits-touchdown' (TD) method (Li *et al.* 2013) on DSLs from four domestic and four wild dromedary samples (see supplementary information; Table S3; Fig 2). The hybridization conditions for MYbaits capture were 65°C for 36 hours, versus 48 hours for the MYbaits-touchdown method with the temperature decreasing from 65°C to 50°C. Following the capture enrichment, 2-4 μL of the indexed libraries were quantified on an Agilent Bioanalyzer 2100 (software version 1.03). The indexing PCRs (to apply the P5 barcode), in-solution hybridization enrichment and post-capture amplification were performed in a molecular laboratory at the University of York.

The TD hybridization method and the respective post-capture amplification were performed at the Vetmeduni in Vienna, Austria. Among the 66 prepared indexed DSLs, the expected product size of 150 – 300 bp for three libraries (two ancient-domestic and one wild) were not detected on 1.5% agarose gel, therefore these samples were excluded from further analysis (Fig 2).

Initially, 63 enriched indexed libraries and two library blanks were pooled in equimolar concentrations and single-end (SE) sequenced (read length 100 bp) on one lane of the HiSeq2000 Illumina platform (National High-throughput DNA Sequencing Centre, University of Copenhagen, Denmark). In another attempt, only indexed libraries from wild samples (21 libraries) were paired-end (PE) shotgun sequenced (read length 100 bp) on 1/16 of an Illumina platform lane (Beijing Genomic Institute, China). We also SE sequenced a set of 25 indexed libraries (15 shotgun and 8 TD enriched) on another 1/16 of an Illumina platform lane (Beijing Genomic Institute, China).

Data processing and mapping

The raw reads obtained from the sequenced libraries were trimmed for adapter and index /barcode sequences using the software *cutadapt* v1.2.1 (Martin 2011). During index/barcode trimming, one error in the index sequence was allowed (parameter –e 0.125). The reads were filtered to a minimum phred-scaled quality score of 20. The individual read collections were then mapped to the dromedary mtDNA reference (GenBank accession no. NC_009849.1), using the Burrows-Wheeler Alignment v.0.7.3a (Li & Durbin 2009) with the following parameters (-l 1024 -i 0 -o 2 -n 0.03 -t 6) as optimized for aDNA in Schubert *et al*. (2012). Shotgun sequences were additionally mapped to the dromedary reference genome (Wu *et al.* 2014) (GenBank accession no. GCA_000767585.1), using the same parameters as described.

PCR duplicates were removed using Picard MarkDuplicates (http://www.picard.sourceforge.net) to avoid the effect of clonality (PCR duplicates) on downstream analysis. In each sample, the consensus and the polymorphic sites were called with agreement threshold of 50% using Samtools package v.0.1.19 (Li *et al*. 2009). The assembly was then checked by eye at each informative polymorphic site to identify sequencing reads conflicting with the reference sequence. Only those sites covered by three unique reads with different start and end positions were accepted as true polymorphism.

To authenticate the sequences obtained as endogenous dromedary mtDNA, we ran mapDamage2.0 (Ginolhac *et al*. 2011; Jónsson *et al*. 2013) to identify DNA damage patterns typical for ancient or degraded DNA. The program uses misincorporation patterns, particularly deamination of cytosine to uracil within a Bayesian framework (Briggs *et al*. 2007; Brotherton *et al*. 2007; Krause *et al*. 2010; Sawyer *et al*. 2012). Nucleotide misincorporations, observed as elevated C to T substitution towards sequencing starts (and complementary increased G to A rates towards the end) are considered as indicative of genuine (endogenous) aDNA. Similarly, an excess of purines at the first nucleotide position of the reference preceding the sequencing reads (and complementary, excess of pyrimidines at the first sequence position following the end of the read) is considered as a typical breakage pattern for aDNA. In order to estimate the performance of different methods (In solution capture / TD capture, and shotgun-sequencing) in terms of the percentage of uniquely mapped reads obtained we performed the Wilcoxon signed rank test.

Summary statistics and phylogenetic analysis of modern and ancient-domestic dromedary mtDNA sequences

Analysis of the ancient-domestic mtDNA sequences, including the number of variable sites and mitochondrial genetic diversity summary statistics as number of segregating sites (s), number of haplotypes (h), haplotype diversity (H_d), nucleotide diversity (π), average number of pairwise nucleotide differences (k), Tajima's D, Fu and Li's F test, as well as a mismatch distribution based on the number pairwise nucleotide differences was completed with the software DnaSP V.5 (Librado *et al.* 2009). For comparisons with modern dromedary mitochondrial diversity we aligned the ancient mtDNA sequences to nine recently sequenced mitochondrial genomes (Mohandesan *et al*. personal communication; GenBank accession numbers are listed in data accessibility section) as well as to the dromedary mitochondrial reference genome (GenBank accession no. NC_009849.1) and estimated the same diversity parameters from the modern sequences only. For the phylogenetic study of modern and ancient-domestic dromedary sequences we performed a median-joining network (MJN) analysis with NETWORK 5.0 (Bandelt *et al.* 1999) with default parameters, displaying the parsimonious (shortest) consensus tree. The program MODELTEST implemented in MEGA6 (Tamura *et al.* 2013) was used to identify the appropriate substitution model for the mtDNA sequences. A maximum likelihood tree with HKY nucleotide substitution model as best-fitting model based on Bayesian Information Criterion (BIC) was reconstructed from 16,401 bp of mitochondrial sequences from seven ancient-domestic dromedary and the available reference sequences from domestic Old World camels (*C. dromedarius:* GenBank accession no: NC_009849.1, *C. bactrianus:* NC_009628.2, and *C. ferus*: NC_009629.2), using MEGA6. Gaps and missing data were treated with partial deletion and the 95% site coverage cut-off was used as default. To obtain statistical support for each node we used the bootstrap resampling procedure with 100 replications.

Results

DNA sequencing

In this study, we investigated the success rate of obtaining DNA sequences from ancient dromedary specimens from prehistoric and historic archaeological sites in Turkey, Syria, Jordan, and the UAE. We extracted DNA from 54 ancient-domestic and 22 wild dromedary bone samples, from which we successfully built 63 DSLs, which were enriched for camel mtDNA using the MYbaits kit. Among these libraries we recovered reads uniquely mapped to dromedary mtDNA for 58 libraries; four libraries (one ancient-domestic and three wild samples) produced no camel reads (Table S3, Fig 2). In addition, we applied TD enrichment to eight out of 63 DSLs (four ancient-domestic and four wild samples) and obtained camel mtDNA reads in all of them (Table S3, Fig 2).

Furthermore, we SE / PE shotgun sequenced 15 (10 ancient-domestic and five wild) and 21 (wild) DSLs, respectively (Table S3, Fig 2). Although in SE shotgun sequencing, 10 samples failed to produce endogenous mtDNA camel reads (six domestic, four wild) (Fig 2), we successfully recovered nuDNA from these libraries. Using PE shotgun sequencing we recovered both mt/nuDNA from all libraries.

Endogenous mtDNA content

Sequencing DSLs using both post-capture and shotgun NGS revealed an extremely low endogenous content of mtDNA ranging from 0.0001% - 0.34% and 0.0001% - 0.004%, respectively (Table 1 and S3). From all successfully sequenced libraries, we obtained a total of 261,961,806 reads of which 25,721 unique sequence reads were mapped to the dromedary mtDNA reference genome (Table S3). The proportions of raw, trimmed and uniquely mapped

reads to dromedary mtDNA for a few samples using MYbaits /-TD and shotgun-sequencing approaches are shown in Fig S1-3.

The post-capture mtDNA reads of the ancient-domestic samples exhibited DNA damage patterns typical of post-mortem depurination and cytosine deamination, indicating that the sequence data truly originated from ancient DNA templates (Fig S4). The damage pattern was not investigated in wild samples due to the fact that too few reads (2 - 60 reads) could uniquely be mapped to dromedary mtDNA (Table S3). Overall, we recovered 2,850 – 15,843 bp (17-95%) of the mitochondrial genome from the 14 domestic-ancient dromedaries, with average read depths of 1.27 – 47.1-fold for covered regions over the entire genome (Table 1). We obtained short sequence reads (20-100 bp) from ancient-domestic enriched libraries with mean fragment length of 65 bp (Table S4, Fig S5-6).

Endogenous nuclear DNA content

To exhaustively investigate the endogenous DNA preservation and endogenous DNA in domestic and wild samples, we mapped the shotgun sequences (SE and PE) to the dromedary whole genome sequences (WGS; Wu *et al*. 2014) (Table S5). From all 36 shotgun-sequenced libraries, we obtained a total of 107,007,621 reads of which 3,735,270 unique sequence reads (3.53%) were mapped to dromedary WGS with average read depths of 1 – 1.06-fold for covered regions over the entire dromedary genome (Table S5). These results show that despite the low amount of total endogenous mtDNA (0.00056 %) recovered from these samples in shotgun-sequencing experiment, there is a greater quantity of nuclear DNA (3.53%) preserved (Table S3-S5).

Enrichment performance on DSL

To evaluate the performance of the in-solution enrichment method (MYbaits), we computed the percentage of the unique reads that were mapped to the dromedary mtDNA reference sequence. We observed a significant increase in the percentage of on-target mapped reads in ancient-domestic camels in the captured libraries (range 0.0017 - 0.1230, mean 0.0785) compared to shotgun-sequenced libraries (range 0 - 0.0042, mean 0.0007; Wilcoxon signed rank *P*-value = 0.01563). For example, in the sample AQ40 the percentage of the uniquely mapped reads increased by three orders of magnitude post-capture (0.00039% to 0.34%; Table S3). Overall, the capture method increased the percentage of on-target mapped reads an average of 187-fold in our dataset of seven samples (ancient-domestic and wild) for which we performed both shotgun and capture approaches (Table 1). In addition, we observed an increase of average 400-fold enrichment considering only domestic samples (Table 1). It should be noted that this result is based on only three samples, since seven of the 10 domestic samples did not yield a single camel mtDNA read using shotgun sequencing, despite successful recovery of up to 73% of the mitochondrial genome in the capture approach. Overall, our observed enrichment ranges and averages are similar to those detected in other comparative studies (Avila-Arcos *et al.* 2015; Paijmans *et al.* 2015).

Effect of temperature and hybridization time

We explored the effects of temperature and hybridization time by comparing the number of uniquely mapped reads in the MYbaits capture (65°C, 36 hours) and the alternative MYbaits-TD (65-50°C, 48 hours) in four ancient-domestic and four wild individuals. In three domestic samples (AP3, AQ30 and Palm152), we observed a decrease in the percentage of unique mapped reads from the total number of mapped reads in the MYbaits-TD method. For example in AP3, we recovered 0.29% unique mapped reads with the capture method, while in

382 the TD method the percentage decreased to 0.17%. However in the wild sample (Tel622) and
383 one domestic sample (SAG2) we observed a slight increase in the percentage of the mapped
384 reads with the TD method (Table S3). For these five samples, however, differences in the
385 percentage of endogenous DNA recovered using the TD method are not significant (Wilcoxon
386 signed rank test *P*-value = 0.4375). An increase in the percentage of PCR duplicate reads
387 (measured as the fraction of the total mapped reads that are PCR duplicates) was observed for
388 80% of the samples used in the TD experiment (Table S6).
389
390 *Mitochondrial genetic diversity of modern and ancient-domestic dromedaries*
391 We obtained 14 partial mitogenomes from ancient-domestic dromedaries (GenBank accession
392 numbers are listed in data accessibility) with 2,850 – 15,843 bp covered and a mean read
393 depth of 1.27 – 47.1-fold (Table 1). Aligning seven ancient-domestic mtDNA genomes with
394 higher length coverage (59-95%), we obtained 6,694 aligned nucleotide sites. These seven
395 ancient samples showed 61 segregating sites with 5 haplotypes, H_d of 0.857 and π of 0.00263.
396 In comparison, the 10 modern dromedary sequences (accession numbers for nine genomes are
397 listed in data accessibility) aligned to the same 6,694 bp displayed 59 segregating sites, 7
398 haplotypes, H_d = 0.867 and π = 0.00185 (Table S7). From the ancient-domestic and modern
399 dromedary mtDNA, we obtained negative values of Tajima's D (-1.69635; *P*-value < 0.05
400 and -2.03913; *P*-value < 0.01) and Fu's and Li's F test (-1.96090; *P*-value < 0.02 and -
401 2.60322; *P*-value < 0.02), respectively (Table S7). As a test of recent population expansion,
402 we applied mismatch distribution analysis and calculated the observed and expected number
403 of pairwise nucleotide differences in 6,694 bp mtDNA from seven ancient-domestic and 10
404 modern dromedaries (Fig S8). The MJN including modern and ancient-domestic sequences
405 revealed two haplogroups separated by 50 fixed polymorphic sites, and one haplotype in

higher frequency (7/17 samples) and shared between modern and ancient-domestic samples (Fig 3). A phylogenetic tree displaying the relationship of the ancient-domestic mitogenomes to the reference sequences from domestic Old World camels is presented in Fig S7. The ancient-domestic dromedaries and modern dromedary (*C. dromedarius*: GenBank accession no. NC_009849.1) cluster together, while the domestic Bactrian camels (*C. bactrianus:* NC_009628.2) and the only remaining wild two-humped camels (*C. ferus:* NC_009629.2) form a separate sister group.

Discussion

The ancient-domestic samples (100 BCE - 1870 CE) used in this study were recovered from sites located in semi-arid to arid environments whereas the wild population samples (5000 - 1400 BCE) originated from hot and partly very humid habitats characterizing the Southeast coast of the Arabian Peninsula. Taking into account their archaeological age and the conditions of preservation, we observed a better recovery of endogenous mtDNA from ancient-domestic dromedary samples in comparison to the wild ones. This is consistent with the observation that arid conditions may be relatively less damaging to DNA than humid conditions even in hot climates (Poinar *et al.* 2003; Haile *et al*. 2009). However, this difference was not observed in the recovery of endogenous nuDNA in the shotgun experiment.

Effect of temperature and hybridization time on enrichment performance

Despite the use of various target-enrichment methods in aDNA research, the efficiency and effectiveness of different hybridization techniques have not yet been fully understood. Paijmans *et al.* (2015) investigated the impact of a key parameter, *i.e.* hybridization

temperature, on the recovery of mitogenomes from different types of samples (fresh, archival and ancient). They observed better sequence recovery with a constant hybridization temperature of 65ºC in degraded samples, while the touchdown method (65ºC down to 50ºC) yielded the best results for fresh samples. In our study, with a limited sample size (four ancient-domestic and one wild) we observed no significant effect on the recovery of uniquely mapped reads comparing regular capture and the TD method.

The factors like hybridization time and binding temperature did not dramatically affect the efficiency of the capture; however, the number of PCR duplicates (clones) increased using the TD method. To obtain adequate amounts of DNA for NGS sequencing, all libraries were amplified 20 cycles during library construction, 10 cycles for indexing and 10-20 cycles post capture (see supplementary information). Although the initial DNA concentration used for both capture protocols was the same (>300 ng), the MYbaits-TD method required an additional 10 cycles of post-capture PCR to generate optimal DNA concentrations for sequencing (Table S6). These additional post-capture PCR cycles may account for the greater sequence clonality observed in the majority of the MYbaits-TD libraries. At this stage, the reasons underlying the observed differences in capture success are not clear and more datasets and systematic experimental studies are needed to be able to understand the effect of different parameters on capture success.

Enrichment capture versus shotgun sequencing in ancient-domestic samples

We noted a greater recovery (approximately 400-fold) of endogenous DNA with the capture method for the presumably better preserved ancient-domestic samples in comparison with shotgun sequencing. This is demonstrated by the recovery of virtually complete mitogenomes from a few ancient-domestic samples using capture enrichment on just a single sequencing

454 library. This pattern has been observed in other studies where an increase in enrichment of 20
455 – 2488-fold (Paijmans *et al*. 2015) and 6–159-fold (Carpenter *et al.* 2013) of on-target content
456 in comparison to shotgun libraries were observed. In addition, the same pattern has been
457 observed by Dabney *et al.* (2013); using shallow shotgun sequencing on a subset of libraries
458 obtained from a Middle Pleistocene cave bear did not recover a single sequence read that
459 aligned with the published Late Pleistocene cave bear mitochondrial genome (Krause *et al*.
460 2008) while hybridization capture successfully enriched the libraries, aligning with ~4% of
461 the capture reads.
462 One alternative and cost effective approach to enrichment through hybridization is a highly
463 targeted amplicon sequencing technology. Amplicon sequencing allows specifically targeting
464 and deep sequencing multiple regions of interest containing informative genetic variations.
465 This approach reduces the costs and turnaround time where sequencing a large number of
466 samples with high coverage is required. However, in case of highly degraded samples most of
467 the fragments are too small for amplification, leaving enrichment through hybridization as
468 method of choice in many studies.
469
470 *Enrichment capture versus shotgun sequencing in wild samples*
471 Our results demonstrate that neither capture nor shotgun methods are efficient in the recovery
472 of mtDNA from wild dromedary samples, whose bones lingered for thousands of years in
473 soils, and which were subjected to varying degrees of humidity and salinity due to
474 fluctuations of the groundwater table. In samples with such low concentration of endogenous
475 DNA, it would be necessary to construct more libraries per sample and to run fewer samples
476 per sequencing lane (cf. Dabney *et al*. 2013; Meyer *et al*. 2014). While this strategy would

increase the percentage of endogenous reads, the financial resources in many laboratories preclude this approach.

Endogenous nuDNA content in ancient-domestic and wild samples

Mapping the sequence reads obtained from 36 shotgun-sequenced libraries to the published dromedary genome (Wu *et al*. 2014), we noted a greater recovery of nuDNA (3.53%) in comparison to mtDNA (~ 0.00056 %). We observed that due to the size difference between dromedary mitochondrial (16 Kb) and nuclear genome (2.27 Gb) (Wu *et al*. 2014; Fitak *et al.* 2015), the nuDNA sequence reads outnumber the mtDNA in shotgun sequences. Nevertheless, the data indicate that mt/nuDNA is preserved in our wild samples, and possibly with more DNA extraction and much deeper sequencing for each sample we would be able to recover more nuDNA from this extinct species.

Enrichment capture on SSLs in wild samples

Recently, optimized protocols for DNA extraction (Dabney *et al*. 2013) and library preparation (Gansauge & Meyer 2013) have been proposed for highly degraded samples. In particular, the silica-spin column method proposed in Dabney *et al.* (2013) seems to recover a greater quantity of short DNA fragments, which could significantly enhance the amount of endogenous DNA recovered from archaeological specimens collected in hot environments. The mean fragment length recovered from our ancient-domestic samples was 65 bp (Table S4, Fig S5-6), significantly higher than the fragment length pattern observed in the Sima de los Huesos samples from Spain (Dabney *et al.* 2013). Additional optimization may be obtained using a SSL preparation method (Gansauge & Meyer 2013). Although this method is

more costly and time-consuming, refinements to the SSL construction method may make it more accessible in the future (Bennett *et al*. 2014).

We tested the Dabney *et al.* (2013) DNA extraction and SSL methods followed by the in-solution target enrichment on seven wild dromedary camel specimens. However, these methods did not improve the number of obtained DNA sequence reads. This lack of success may be the result of combining these two methods with the capture enrichment. Although the silica-spin column DNA extraction methods and single-stranded library protocol are recommended for recovering greater quantities of short DNA fragments, the capture enrichment is generally more efficient on longer fragments. More systematic comparisons of extractions techniques, library building protocols and hybridisation capture methodologies will be required in order to optimize the recovery of short ancient DNA templates.

Mitogenome diversity and demography in ancient-domestic and modern dromedaries

During the process of domestication, population growth or dispersion of domestic animals across a wider geographic range can be inferred from molecular signals of sudden expansion (Bruford *et al.* 2003). From the mitogenomes of ancient-domestic and modern dromedaries we received negative values of Tajima's *D* and Fu and Li's F test (Table S7), respectively, which can indicate demographic expansion assuming absence of selection. In the MJN (Fig 3) we observed two haplogroups separated by 50 fixed polymorphisms and a star-shaped radiation starting from one haplotype in higher frequency, a typical pattern of recent population expansion. Although the mismatch distribution calculated on the number of pairwise differences showed a multimodal distribution related to the two haplogroups, the beginning of the curve is smooth indicative of an expanding population (Fig S8). Two major haplogroups (H_A and H_B) and signals of population growth in the context of domestication

524 have also been detected in a global sample set of modern dromedary populations (Almathen *et*
525 *al.* 2016). Comparing mitogenome diversity between ancient-domestic and modern
526 dromedaries, we observed higher pairwise nucleotide diversity but a slightly lower number of
527 haplotypes and haplotype diversity in the ancient-domestic dromedary sequences (Table S7).
528 This result can be interpreted as higher retained ancestral diversity in the early-domestic
529 (ancient) dromedary samples (Troy *et al*. 2001); while in the modern population new
530 haplotypes emerged with only one to two mutational steps (Fig 3). Evidence for dromedary
531 domestication was found in the Southeast coast of the Arabian Peninsula, with a mode of an
532 initial domestication followed by introgression from wild, now-extinct individuals (Almathen
533 *et al.* 2016).
534
535 **Conclusion**
536 The low amount of endogenous sequences in ancient dromedary specimens is an example of
537 the extreme DNA degradation in bone samples from hot and arid environments. Despite the
538 availability of a number of optimized protocols, the recovery of aDNA from poorly preserved
539 samples is still an unresolved issue and hybridization protocols require specific optimization
540 for such specimens. Much deeper sequencing would be necessary; however this would come
541 at very high costs. This study highlights one of the few successful recoveries of genetic
542 materials from specimens collected from prehistoric and historic archaeological sites located
543 in hot and (hyper)arid environments and reports the first nearly complete mitogenome
544 recovery from ancient-domestic dromedaries. We also highlight the first recovery of nuDNA
545 from ancient-domestic and extinct wild dromedary camels.
546

Acknowledgements

We are very grateful to A. Schmidt-Collinet, R. Saleh and G. Forstenpointer for the support in sample collection and anatomical analyses. We acknowledge general support from C. Schlötterer. We are thankful to MTP. Gilbert and his laboratory members for helpful discussions and support in constructing SSLs, and J. Paijmans for technical assistance in capture enrichment experiments. EM was supported through the Austrian Science Fund (FWF): P24706-B25 to PAB, recipient of an APART fellowship (11506) of the Austrian Academy of Sciences. The project was supported by MH, recipient of the ERC consolidator grant: 310763 GeneFlow. CFS was supported through ORCA FP7-PEOPLE-2011-IOF 299075.

References

Allentoft ME, Collins M, Harker D, *et al.* (2012) The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc Biol Sci,* **279:**4724-4733

Almathen F, Charruau P, Mohandesan E, *et al*. (2016) Dynamics of domestication and trading revealed by the global genetic diversity of the dromedary. *PNAS* doi:10.1073/pnas.1519508113

Amaral AJ, Ferretti L, Megens HJ, *et al.* (2011) Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. *PLoS ONE*, **6,** e14782.

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.

Bennett EA, Massilani D, Lizzo G, *et al.* (2014) Library construction for ancient genomics: single strand or double strand? *Biotechniques,* **56**, 289-290.

Bollongino R, Nehlich O, Richards MP*, et al.* (2013) 2000 years of parallel societies in Stone Age Central Europe. *Science*, **342**, 479-481.

Bandelt H, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, **16,** 37-48.

583 Briggs AW, Stenzel U, Meyer M, *et al.* (2010) Removal of deaminated cytosines and detection
584 of in vivo methylation in ancient DNA. *Nucleic Acids Res*, **38**: e87.
585
586 Briggs AW, Good JM, Green RE, *et al.* (2009) Targeted retrieval and analysis of five Neandertal
587 mtDNA genomes. *Science*, **325**, 318-321.
588
589 Briggs AW, Stenzel U, Johnson PL, *et al* (2007) Patterns of damage in genomic DNA sequences
590 from a Neandertal. *Proc Natl Acad Sci U S A*, **104**, 14616-14621.
591
592 Brotherton P, Endicott P, Sanchez JJ, *et al.* (2007) Novel high-resolution characterization of
593 ancient DNA reveals C > U-type base modification events as the sole cause of post mortem
594 miscoding lesions. *Nucleic Acids Res*, **35**, 5717-5728.
595
596 Bruford M, Bradley D, Luikart G (2003) Genetic analysis reveals complexity of livestock
597 domestication. *Nat Rev Genet*, **4**:900-910.
598
599 Bulliet R (1990) *The Camel and the Wheel. Columbia University Press*, New York.
600
601 Carpenter ML, Buenrostro JD, Valdiosera C, *et al.* (2013) Pulling out the 1%: whole-genome
602 capture for the targeted enrichment of ancient DNA sequencing libraries. *Am J Hum Genet.*
603 **93**, 852-864.
604
605 Cai D, Tang Z, Yu H, *et al.* (2011) Early history of Chinese domestic sheep indicated by ancient
606 DNA analysis of Bronze Age individuals. *J Arch Sci,* **38**, 896-902.
607
608 Dabney J, Knapp M, Glocke I, *et al.* (2013) Complete mitochondrial genome sequence of a
609 Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad*
610 *Sci U S A*, **110**, 15758-15763.
611
612 Elbaum R, Melamed-Bessudo C, Boaretto E, *et al.* (2006) Ancient olive DNA in pits:
613 preservation, amplification and sequence analysis. *J Arch Sci*, **33**, 77-88.
614
615 Fernández E, Pérez-Pérez A, Gamba C, *et al* (2014) Ancient DNA analysis of 8000 B.C. near
616 eastern farmers supports an early neolithic pioneer maritime colonization of Mainland Europe
617 through Cyprus and the Aegean Islands. *PLoS Genet*, **10**, e1004401
618
619 Fitak RR, Mohandesan E, Corander J, *et al.* (2015) The de novo genome assembly and
620 annotation of a female domestic dromedary of North African origin. *Molecular Ecololy*
621 *Resources*, **16**, 314-324.
622
623 Fortes GG, Paijmans JL (2015) Analysis of Whole Mitogenomes from Ancient Samples.
624 *Methods Mol Biol*, **1347**,179-195.
625
626 Fu Q, Meyer M, Gao X, *et al.* (2013) DNA analysis of an early modern human from Tianyuan
627 Cave, China. *Proc Natl Acad Sci U S A,* **110**, 2223-2227.
628
629 Gamba C, Jones ER, Teasdale MD, *et al.* (2014) Genome flux and stasis in a five millennium
630 transect of European prehistory. *Nat Commun*, **5**, 5257.

631
632 Gansauge MT, Meyer M (2013) Single-stranded DNA library preparation for the sequencing of
633 ancient or damaged DNA. *Nat Protoc*, **8**, 737-748.
634
635 Gautier A (1966) *Camelus thomasi* from the northern Sudan and its bearing on the relationship
636 *C. thomasi-C. bactrianus*. *Journal of Paleontology*, **40**, 1368-1372.
637
638 Ginolhac A, Rasmussen M, Gilbert MT, *et al*. (2011) mapDamage: testing for damage patterns
639 in ancient DNA sequences. *Bioinformatics.* **27**, 2153-2155.
640
641 Girdland Flink L, Allen R, Barnett R, *et al.* (2014) Establishing the validity of domestication
642 genes using DNA from ancient chickens. *Proc Natl Acad Sci U S A*, **111**, 6184-6189.
643
644 Green RE, Krause J, Ptak SE, *et al.* (2006) Analysis of one million base pairs of Neanderthal
645 DNA. *Nature*, **444,** 330-336.
646
647 Green RE, Krause J, Briggs AW, *et al.* (2010) A Draft Sequence of the Neandertal Genome.
648 *Science*, **328**, 710-722.
649
650 Grigson C (2014) The history of the camel bone dating project. *Anthropozoologica*, **49**, 225-235.
651
652 Grigson C (2012) Camels, copper and donkeys in the Early Iron Age of the Southern Levant:
653 Timna revisited. *Levant***, 44**, 82–100.
654
655 Haile J, Froese DG, Macphee RDE, *et al.* (2009) Ancient DNA reveals late survival of
656 mammoth and horse in interior Alaska. *Proc. Natl. Acad. Sci. U. S. A*, **106**, 22352–22357.
657
658 Higuchi R, Bowman B, Freiberger M, *et al.* (1984) DNA sequences from the quagga, an extinct
659 member of the horse family. *Nature,* **312**, 282-284.
660
661 Hofreiter M, Paijmans JL, Goodchild H, *et al.* (2015) The future of ancient DNA: Technical
662 advances and conceptual shifts. *Bioessays.* **37**, 284-293.
663
664 Hume BC, Voolstra CR, Arif C, *et al.* (2016) Ancestral genetic diversity associated with the
665 rapid spread of stress-tolerant coral symbionts in response to Holocene climate change. *Proc*
666 *Natl Acad Sci U S A*. **113**: 4416-4421.
667
668 Iamoni *M (2009) The Iron Age ceramic tradition in the Gulf: a re-evaluation from the Omani*
669 *perspective. Proceedings of the Seminar for Arabian Studies,* **39**, 223-236.
670
671 Jónsson H, Ginolhac A, Schubert M, *et al*. (2013) mapDamage2.0: fast approximate Bayesian
672 estimates of ancient DNA damage parameters. *Bioinformatics*, **29**, 1682-1684.
673
674 Kimura B, Marshall FB, Chen S, *et al.* (2011) Ancient DNA from Nubian and Somali wild ass
675 provides insights into donkey ancestry and domestication. *P Biol Sci*, **278**, 50-57.
676
677 Knapp M, Clarke AC, Horsburgh KA, *et al*. (2012) Setting the stage—building and working in
678 an ancient DNA laboratory. *Ann. Anat*, **194**, 3–6.

679
680 Köhler-Rollefson IU (1991) *Camelus dromedarius*. *Mammalian Species*, **375**, 1-8.
681
682 Krause J, Briggs AW, Kircher M, *et al.* (2010) A complete mtDNA genome of an early modern
683 human from Kostenki, Russia. *Curr Biol*, **20**, 231-236.
684
685 Krause J, Unger T, Noçon A, *et al.* (2008) Mitochondrial genomes reveal an explosive radiation
686 of extinct and extant bears near the Miocene-Pliocene boundary. *BMC Evol Biol*, **8**:220.
687
688 Li C, Hofreiter M, Straube N, *et al.* (2013) Capturing protein-coding genes across highly
689 divergent species. *Biotechniques*, **54**, 321-326.
690
691 Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform.
692 *Bioinformatics*, **25**, 1754-1760.
693
694 Li H, Handsaker B, Wysoker A, *et al.* (2009) The Sequence Alignment/Map format and
695 SAMtools. *Bioinformatics,* **25,** 2078-2079.
696
697 Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA
698 polymorphism data. *Bioinformatics,* **25**, 1451-1452.
699
700 Magee P (2015) When was the dromedary domesticated in the ancient Near East? *ZOrA*, **8**.
701
702 Margulies M, Egholm M, Altman WE, *et al.* (2005) Genome sequencing in microfabricated
703 high-density picolitre reactors. *Nature*, **437**, 376-380.
704
705 Maricic T, Whitten M, Pääbo S (2010) Multiplex DNA sequence capture of mitochondrial
706 genomes using PCR products. *PLoS One*, **5**, e14004.
707
708 Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads.
709 *EMBnet J*, **17**:10-12.
710
711 Meiri M, Huchon D, Bar-Oz G, *et al*. (2013) Ancient DNA and population turnover in southern
712 levantine pigs--signature of the sea peoples migration? *Sci Rep,* **3**, 3035.
713
714 Meyer M, Fu Q, Aximu-Petri A, *et al*. (2014) A mitochondrial genome sequence of a hominin
715 from Sima de los Huesos. *Nature*, **505**, 403-406.
716
717 Meyer M, Kircher M, Gansauge MT, *et al*. (2012) A High-Coverage Genome Sequence from an
718 Archaic Denisovan Individual. *Science*, **338**, 222-226.
719
720 Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed
721 target capture and sequencing. *Cold Spring Harb Protoc,* (6): pdb.prot5448. doi:
722 10.1101/pdb.prot5448.
723
724 Millar CD, Huynen L, Subramanian S, *et al.* (2008) New developments in ancient genomics.
725 *Trends Ecol Evol*, **23,** 386-393.
726

727 Orlando L, Ginolhac A, Zhang G, *et al*. (2013) Recalibrating *Equus* evolution using the genome
728 sequence of an early Middle Pleistocene horse. *Nature,* **499**, 74-78.
729
730 Orlando L, Ginolhac A, Raghavan M, *et al.* (2011) True single-molecule DNA sequencing of a
731 pleistocene horse bone. *Genome Research*, **21**, 1705-1719.
732
733 Orlando L, Mashkour M, Burke A, *et al.* (2006) Geographic distribution of an extinct equid
734 (*Equus hydruntinus*: Mammalia, Equidae) revealed by morphological and genetical analyses
735 of fossils. *Mol Ecol*, **15**, 2083-2093.
736
737 Paijmans JL, Fickel J, Courtiol A, *et al.* (2015) Impact of enrichment conditions on cross-species
738 capture of fresh and degraded DNA. *Mol Ecol Resour*, doi: 10.1111/1755-0998.12420.
739
740 Peters J. (1998). *Camelus thomasi* Pomel, 1893, a possible ancestor of the one-humped camel?
741 *Int. J. Mamm. Biol.* (Z. Säugetierk.) **63**: 372-376.
742
743 Preston GW, Thomas DS, Goudie AS, *et al.* (2015) A multi-proxy analysis of the Holocene
744 humid phase from the United Arab Emirates and its implications for southeast Arabia's
745 Neolithic populations. *Quaternary International*, **382,** 277-292.
746
747 Pinhasi R, Fernandes D, Sirak K, *et al*. (2015) Optimal Ancient DNA Yields from the Inner Ear
748 Part of the Human Petrous Bone. *PLoS One*, **10**, e0129102.
749
750 Palkopoulou E, Mallick S, Skoglund P, *et al.* (2015) Complete genomes reveal signatures of
751 demographic and genetic declines in the woolly mammoth. *Curr Biol,* **25,** 1395-1400.
752
753 Poinar H, Kuch M, McDonald G, *et al*. (2003) Nuclear gene sequences from a late Pleistocene
754 sloth coprolite. *Curr. Biol,* **12**, 1150–1152.
755
756 Prüfer K, Racimo F, Patterson N, *et al.* (2014) The complete genome sequence of a Neanderthal
757 from the Altai Mountains. *Nature*, **505**, 43-49.
758
759 Rasmussen M, Guo X, Wang Y, *et al*. (2011) An Aboriginal Australian genome reveals separate
760 human dispersals into Asia. *Science,* **334**, 94-98.
761
762 Rasmussen M, Anzick SL, Waters MR, *et al.* (2014) The genome of a Late Pleistocene human
763 from a Clovis burial site in western Montana. *Nature,* **506**, 225-229.
764
765 Reich D, Green RE, Kircher M, *et al.* (2010) Genetic history of an archaic hominin group from
766 Denisova Cave in Siberia. *Nature*, **468**, 1053-1060.
767
768 Rohland N, Siedel H, Hofreiter M (2010) A rapid column-based ancient DNA extraction method
769 for increased sample throughput. *Mol Ecol Resour*, **10**, 677-683.
770
771 Rohland N, Hofreiter M (2007) Comparison and optimization of ancient DNA extraction.
772 *Biotechniques*, **42**, 343-352.
773
774 Rowley-Conwy P (1988) The camel in the Nile Valley: new radiocarbon accelerator (AMS)

775 dates from Qasr Ibrim. *The Journal of Egyptian Archaeology*, **1**, 245-248.
776
777 Sawyer S, Krause J, Guschanski K, *et al.* (2012) Temporal patterns of nucleotide
778 misincorporations and DNA fragmentation in ancient DNA. *PLoS ONE*, **7**, e34131.
779
780 Schubert M, Jónsson H, Chang D, *et al.* (2014) Prehistoric genomes reveal the genetic
781 foundation and cost of horse domestication. *Proc Natl Acad Sci U S A*, **111,** 5661-5669.
782
783 Schubert M, Ginolhac A, Lindgreen S (2012) Improving ancient DNA read mapping against
784 modern reference genomes. *BMC Genomics*, **13**, 178.
785
786 Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol,* **26**, 1135-1145.
787
788 Smith CI, Chamberlain AT, Riley MS, *et al.* (2001) Neanderthal DNA: not just old but old and
789 cold? *Nature*, **10**, 771-772.
790
791 Spencer PBS, Woolnough AP (2010) Assessment and genetic characterisation of Australian
792 camels using microsatellite polymorphisms. *Livestock Science*, **129,** 241-245.
793
794 Steiger C. (1991). Vergleichend morphologische Untersuchungen an Einzelknochen des
795 postkranialen Skelettes der Altweltkamele. Diss. med. vet. München.
796
797 Stiller M, Green RE, Ronan M, *et al*. (2006) Patterns of nucleotide misincorporations during
798 enzymatic amplification and direct large-scale sequencing of ancient DNA. *Proc Natl Acad*
799 *Sci U S A,* **103**, 13578-13584.
800
801 Surakka I, Kristiansson K, Anttila V, *et al*. (2010) Founder population-specific HapMap panel
802 increases power in GWA studies through improved imputation accuracy and CNV tagging.
803 *Genome Res,* **20,** 1344-51.
804
805 Tamura K, Stecher G, Peterson D, *et al.* (2013) MEGA6: Molecular Evolutionary Genetics
806 Analysis Version 6.0. Mol Biol Evol, **30**, 2725-2729.
807
808 Thalmann O, Shapiro B, Cui P, *et al.* (2013) Complete Mitochondrial Genomes of Ancient
809 Canids Suggest a European Origin of Domestic Dogs. *Science*, **342**, 871-874.
810
811 Troy CS, MacHugh DE, Bailey JF, *et al.* (2001) Genetic evidence for Near-Eastern origins of
812 European cattle. *Nature.* 410, 1088-1091.
813
814 Uerpmann M, Uerpman HP (2012) Archeozoology of camels in South-Eastern Arabia. *Camels*
815 *in Asia and North Africa*. Interdisciplinary perspectives on their significance in past and
816 present, eds Knoll E & Burger P (Academy of Science Press, Vienna, Austria), 109-122.
817
818 Uerpmann M, Uerpmann HP (2002) The appearance of the domestic camel in SE-Arabia.
819 *Journal of Oman Studies*, **12**, 235-260.
820
821 von den Driesch A, Brückner H, Obermaier H, *et al*. (2008) The hunt for wild dromedaries at the
822 United Arab Emirates coast during the 3rd and 2nd millennia BC. Camel bones from the

excavations at Al Sufouh 2, Dubai, UAE. In: E. Vila, L. Gourichon, A. M. Choyke & H. Buitenhuis (Eds.), Archaeozoology of the Near East VIII. Proceedings of the eighth international Symposium on the Archaeozoology of southwestern Asia and adjacent areas. Travaux de la Maison de l'Orient et de la Méditerranée, Lyon, **49**, 487-497.

von den Driesch A, Obermaier H (2007) The hunt for wild dromedaries during the 3rd and 2nd millennia BC on the United Arab Emirates coast. Camel bone finds from the excavations at Al Sufouh 2 Dubai, UAE. Skeletal series and their socio-economic context, Documenta Archaeobiologiae, eds Grupe G & Peters J (Verlag Marie Leidorf GmbH, Rahden, Germany), 133–167.

Wu H, Guang X, Al-Fageeh MB, *et al.* (2014) Camelid genomes reveal evolution and adaptation to desert environments. *Nat Commun*, **5,** 5188.

Zhang H, Paijmans JL, Chang F, *et al*. (2013) Morphological and genetic evidence for early Holocene cattle management in northeastern China. *Nat Commun*, **4**, 2755.

Data Accessibility

The partial mitochondrial genome assemblies from ancient dromedary are archived in GenBank with accession numbers listed below: AP2: KU605058, AP3: KU605059, AQ5: KU605067, AQ24: KU605060, AQ30: KU605061, AQ34: KU605062, AQ40: KU605063, AQ46: KU605064, AQ48: KU605065, AQ49: KU605066, Palm152: KU605068, Palm157: KU605069, Palm171: KU605070, SAG2: KU605071.

The complete modern dromedary mitochondrial genomes used for genetic diversity analysis are deposited in GenBank with accession numbers listed below: Drom439 (Qatar, Jordan border): KU605072, Drom795 (Saudi Arabia): KU605073, Drom796 (Saudi Arabia): KU605074, Drom797 (Saudi Arabia): KU605075, Drom801A (Austria): KU605076, Drom802 (UAE, Dubai): KU605077, Drom806 (Kenya): KU605078, Drom816 (Sudan): KU605079, Drom820 (Pakistan): KU605080.

In addition, the raw sequence reads from all the libraries sequenced in this study are deposited in Sequence Read Archive under SRA accession: SRP073444 at the National Center for Biotechnology Information (NCBI).

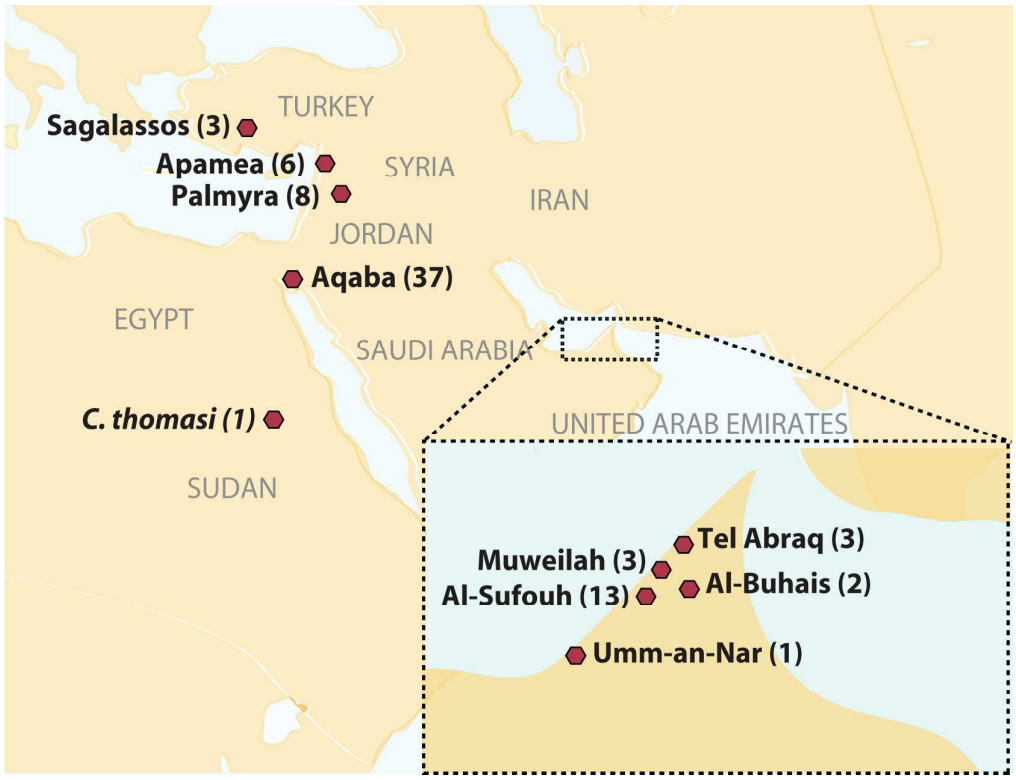
855 **Author Contributions**

856 EM wrote the paper and performed laboratory work and bioinformatic analyses. CFS
857 performed laboratory work and revised the manuscript. JP and BDC provided the samples and
858 revised the manuscript. MU and HPU provided the samples. MH supported part of the
859 laboratory work and revised the manuscript. PAB managed the project, and revised the
860 manuscript.

Table 1: Sample details and the sequencing scheme used for each sample. All the libraries were built using the double-stranded library (DSL) method, and subjected to sequencing both pre- and post-capture using MYbaits. The samples with an asterisk were only sequenced post-capture. The percentage and average coverage of the unique reads mapped to the dromedary mitochondrial genome and the total length of the recovered mtDNA for each sample is shown. For the wild samples, the length of the genome is not calculated, as a result of low numbers of reads mapped to the reference genome.

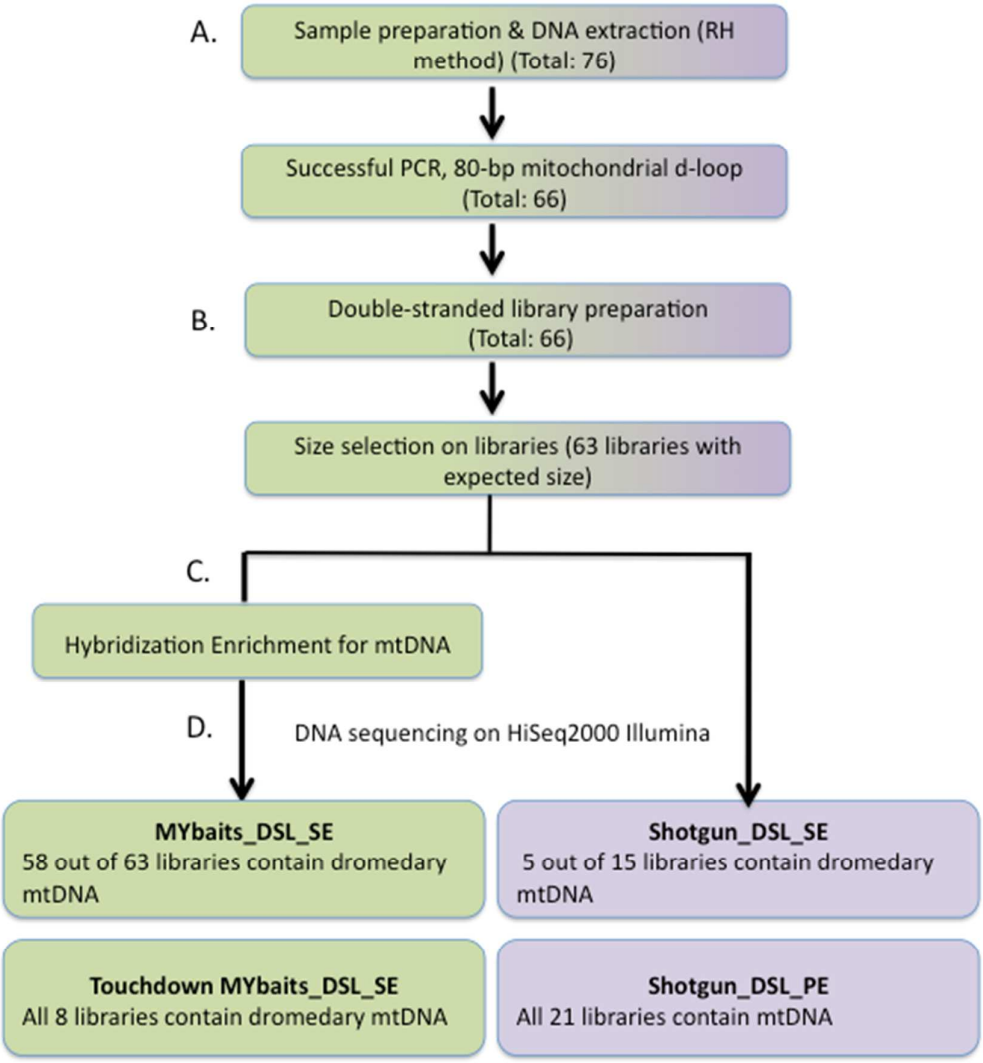
	% Unique mapped reads to *C. dromedarius* mtgenome			Mtgenome length (bp)	%Mtgenome recovered	Average read depth	GenBank accession no.
Sample ID	MYbaits Capture	MYbaits-TD Capture	Shotgun				
AP2	0.123		0.0008	9,943	59.7	2.45	KU605058
AP3	0.294	0.175		15,315	92.0	10.63	KU605059
AQ5	0.013			4,083	24.5	2.75	KU605067
AQ24	0.011		0.004	5,516	33.1	3.56	KU605060
AQ30	0.241	0.088		15,843	95.1	47.10	KU605061
AQ34	0.058		0	12,162	73.0	8.87	KU605062
AQ40	0.346		0.0003	12,422	74.6	19.33	KU605063
AQ46	0.006		0	4,143	24.8	1.44	KU605064
AQ48	0.002		0	3,829	23.0	1.56	KU605065
AQ49	0.001		0	2,850	17.1	1.62	KU605066
Palm152	0.005	0.001		5,149	30.9	1.27	KU605068
Palm157*	0.010			10,890	65.4	2.26	KU605069
Palm171*	0.011			7,402	44.4	1.82	KU605070
SAG2	0.028	0.046		14,514	87.2	8.48	KU605071
Tel622	0.0001	0.0006	0.0005				
Tel623	0.0002		0.0009				
Also1	0.0003		0.0008				
Also10	0.0007		0.0008				

868 **Figure 1:** Geographical locations of the ancient-domestic dromedary, its extinct
869 ancestor the wild dromedary and the giant camel (*C.thomasi*) used in this study.
870



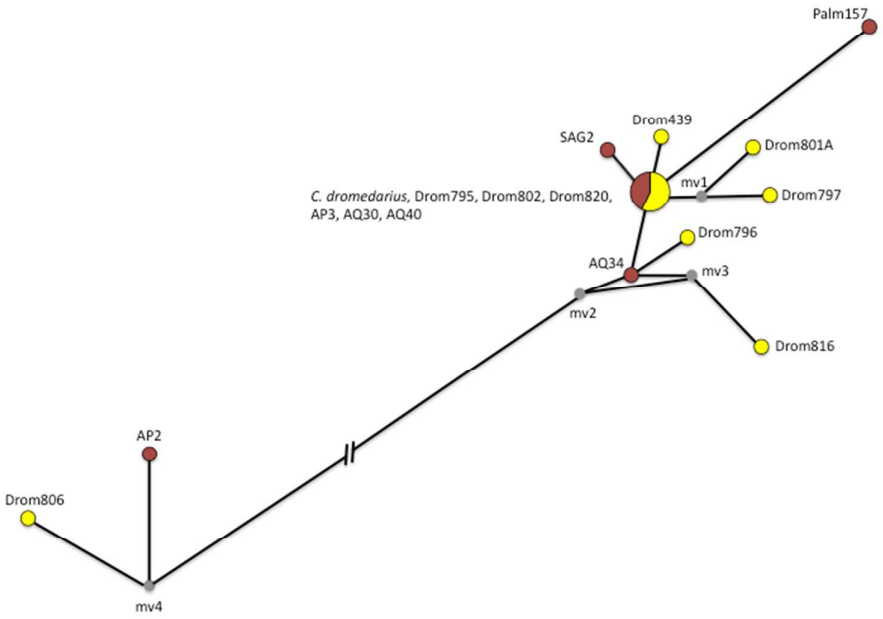
871
872

873 **Figure 2:** Basic workflow illustrating different steps prior to Illumina sequencing.
874 Summary of the results for enrichment hybridization and shotgun sequencing is shown.
875

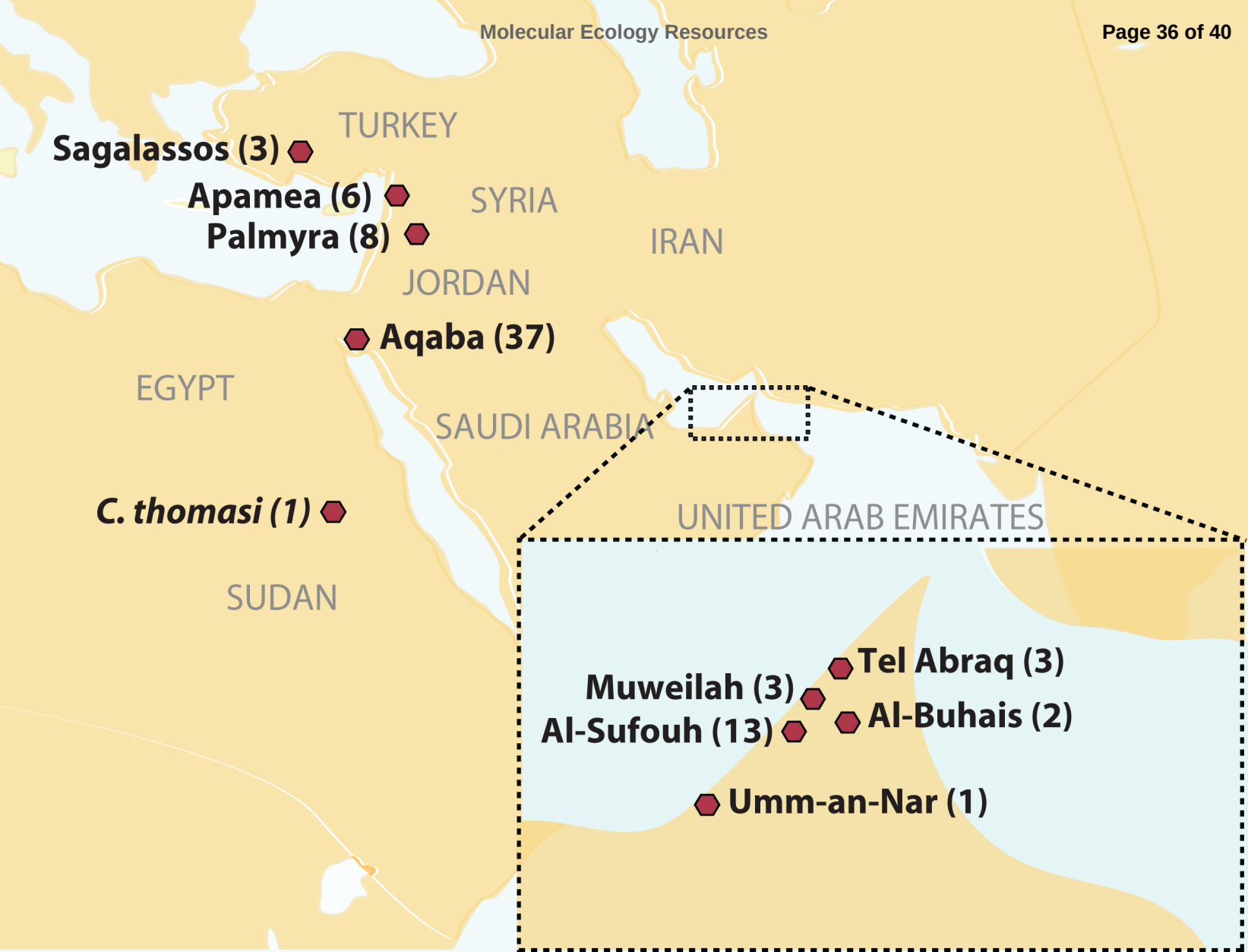


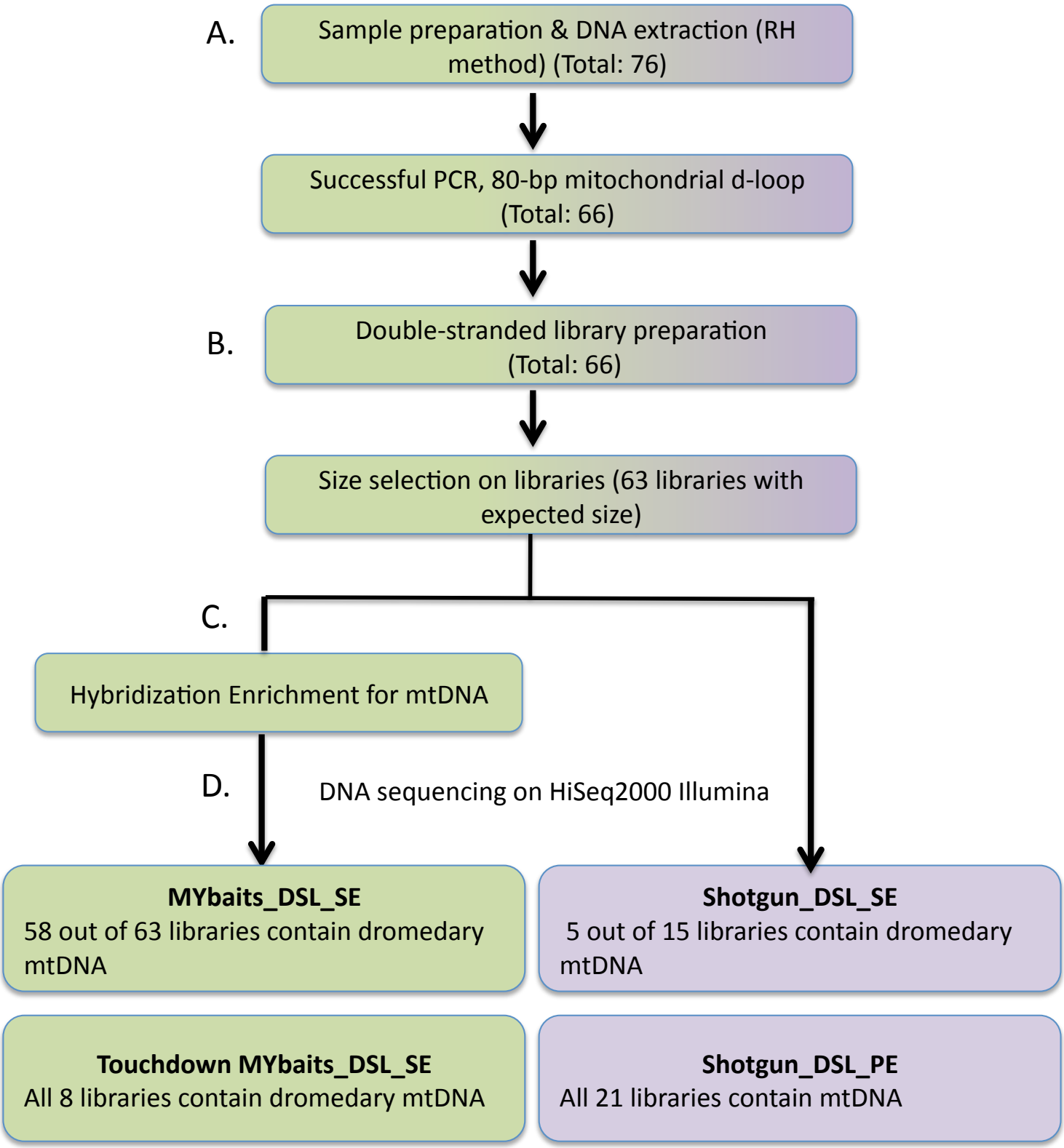
876
877

878 **Figure 3:** Representation of the mitochondrial haplotypes (6,694 bp) retrieved from 10
879 modern (yellow) and seven ancient (red) domestic dromedaries. Circles are proportional
880 to the sample size. Small grey circles represent median vectors corresponding to
881 missing haplotypes. The genetic distance of 50 fixed polymorphic sites between two
882 haplogroups in not displayed in the graph and is shown with a discontinuous line.



883


TURKEY
Sagalassos (3)
Apamea (6)
SYRIA
Palmyra (8)
IRAN
JORDAN
Aqaba (37)
EGYPT
SAUDI ARABIA
C. thomasi (1)
UNITED ARAB EMIRATES
SUDAN
Tel Abraq (3)
Muweilah (3)
Al-Buhais (2)
Al-Sufouh (13)
Umm-an-Nar (1)


A.
Sample preparation & DNA extraction (RH method) (Total: 76)
Successful PCR, 80-bp mitochondrial d-loop (Total: 66)
B.
Double-stranded library preparation (Total: 66)
Size selection on libraries (63 libraries with expected size)
C.
Hybridization Enrichment for mtDNA
D.
DNA sequencing on HiSeq2000 Illumina
MYbaits_DSL_SE
58 out of 63 libraries contain dromedary mtDNA
Shotgun_DSL_SE
5 out of 15 libraries contain dromedary mtDNA
Touchdown MYbaits_DSL_SE
All 8 libraries contain dromedary mtDNA
Shotgun_DSL_PE
All 21 libraries contain mtDNA

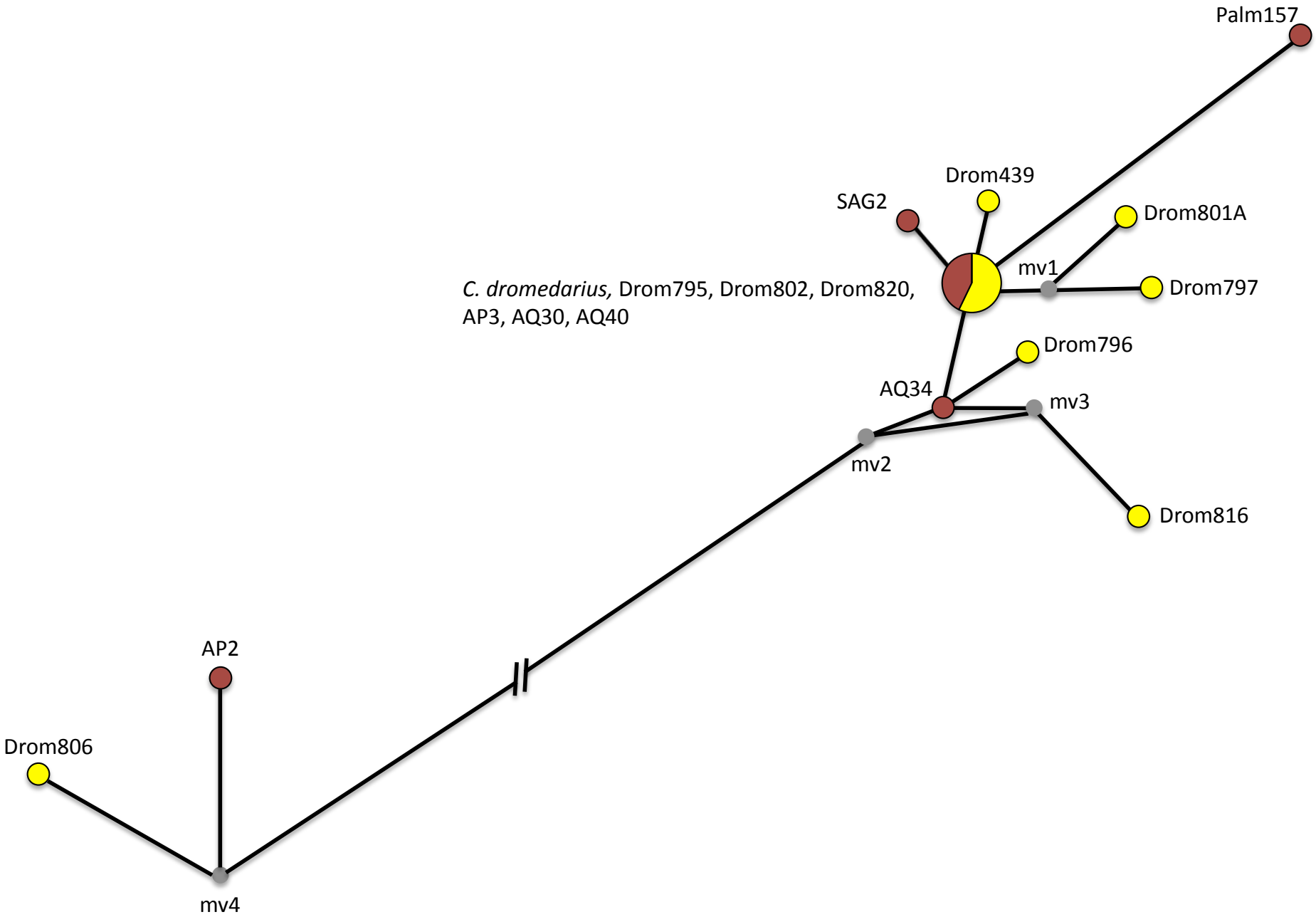

Palm157
Drom439
SAG2
Drom801A
mv1
C. dromedarius, Drom795, Drom802, Drom820,
AP3, AQ30, AQ40
Drom797
Drom796
AQ34
mv3
mv2
Drom816
AP2
Drom806
mv4

Table 1: Sample details and the sequencing scheme used for each sample. All the libraries were built using the double-stranded library (DSL) method, and subjected to sequencing both pre- and post-capture using MYbaits. The samples with an asterisk were only sequenced post-capture. The percentage and average coverage of the unique reads mapped to the dromedary mitochondrial genome and the total length of the recovered mtDNA for each sample is shown. For the wild samples, the length of the genome is not calculated, as a result of low numbers of reads mapped to the reference genome.

	% Unique mapped reads to *C. dromedarius* mtgenome			Mtgenome length (bp)	%Mtgenome recovered	Average read depth	GenBank accession no.
Sample ID	MYbaits Capture	MYbaits-TD Capture	Shotgun				
AP2	0.123		0.0008	9,943	59.7	2.45	KU605058
AP3	0.294	0.175		15,315	92.0	10.63	KU605059
AQ5	0.013			4,083	24.5	2.75	KU605067
AQ24	0.011		0.004	5,516	33.1	3.56	KU605060
AQ30	0.241	0.088		15,843	95.1	47.10	KU605061
AQ34	0.058		0	12,162	73.0	8.87	KU605062
AQ40	0.346		0.0003	12,422	74.6	19.33	KU605063
AQ46	0.006		0	4,143	24.8	1.44	KU605064
AQ48	0.002		0	3,829	23.0	1.56	KU605065
AQ49	0.001		0	2,850	17.1	1.62	KU605066
Palm152	0.005	0.001		5,149	30.9	1.27	KU605068
Palm157*	0.010			10,890	65.4	2.26	KU605069
Palm171*	0.011			7,402	44.4	1.82	KU605070
SAG2	0.028	0.046		14,514	87.2	8.48	KU605071
Tel622	0.0001	0.0006	0.0005				
Tel623	0.0002		0.0009				
Also1	0.0003		0.0008				
Also10	0.0007		0.0008				