Paleogenomics in a Temperate Environment: Shotgun Sequencing from an Extinct Mediterranean Caprine

Oscar Ramírez[1], Elena Gigli[1], Pere Bover[2], Josep Antoni Alcover[2], Jaume Bertranpetit[1], Jose Castresana[1], Carles Lalueza-Fox[1]*

1 Institute of Evolutionary Biology (UPF-CSIC), Barcelona, Spain, 2 Institut Mediterrani d'Estudis Avançats (CSIC-UIB), Esporles, Mallorca, Spain

Abstract

Background: Numerous endemic mammals, including dwarf elephants, goats, hippos and deers, evolved in isolation in the Mediterranean islands during the Pliocene and Pleistocene. Most of them subsequently became extinct during the Holocene. Recently developed high-throughput sequencing technologies could provide a unique tool for retrieving genomic data from these extinct species, making it possible to study their evolutionary history and the genetic bases underlying their particular, sometimes unique, adaptations.

Methodology/Principals Findings: A DNA extraction of a ~6,000 year-old bone sample from an extinct caprine (*Myotragus balearicus*) from the Balearic Islands in the Western Mediterranean, has been subjected to shotgun sequencing with the GS FLX 454 platform. Only 0.27% of the resulting sequences, identified from alignments with the cow genome and comprising 15,832 nucleotides, with an average length of 60 nucleotides, proved to be endogenous.

Conclusions: A phylogenetic tree generated with *Myotragus* sequences and those from other artiodactyls displays an identical topology to that generated from mitochondrial DNA data. Despite being in an unfavourable thermal environment, which explains the low yield of endogenous sequences, our study demonstrates that it is possible to obtain genomic data from extinct species from temperate regions.

Introduction

Paleogenomics, the study of genomes from extinct organisms, is an emerging scientific field that has been fuelled by recently developed technologies in high-throughput DNA sequencing [1,2]. In the first of such approaches to be undertaken [3], about 27,000 base pairs (bp) of cave bear (*Ursus spelaeus*) genomic sequences were obtained with cloning vectors from 42,000 and 44,000 years-old cave bear samples, respectively. Using sequencing-by-synthesis (SBS) technology, as applied through the Roche/454 Life Sciences GS20 and FLX pyrosequencing platforms [4], 13 million bp of the woolly mammoth (*Mammuthus primigenius*) genome were generated from a 28,000-year-old permafrost mammoth bone [5], complete mitochondrial genomes from mammoth and thylacine hairs [6,7] and, finally, about 80% of the nuclear genome from ~20,000-year-old mammoth hairs [8]. The same approach has been applied to other ancient bones, including Neanderthal samples that provided around 1 million bp of its genome [9], as well as other Pleistocene mammals from Denisova cave in Siberia [10]. However, the efficiency of these metagenomic analyses is notably variable: while in the mammoth bone it was possible to identify from 45.4% (mainly in bone samples) to 90.45% (in hair samples) of the sequences as endogenous, this fraction was significantly

reduced in cave bear (between 1.1 and 5.8%), Neanderthal (6%, although a significant fraction of contamination was posteriorly estimated to be present in this particular extract [11]), ancient horse (0.7%), ancient wolf (1.8%) and cattle (1.1%) [3,5,9,10]. In addition, due to the low genomic coverage, the degradation of the template DNA, and the innate error rate of the sequencing platforms [12], the paleogenomic data contains a significant number of sequencing errors, resulting in an excess of C to T substitutions due to cytosine deaminations as compared to the corresponding reference genome [3,5,9,13]. Thus, it is likely that in the future, specific loci in regions with low shotgun coverage would need to be verified by targeted approaches, such as the polymerase chain reaction (PCR). Several studies in mammoths and Neanderthals have already focussed on the specific retrieval of nuclear genes and the problems of distinguishing endogenous variants from DNA damage [14–16].

Paleogenomic data can be useful for understanding the rate and nature of some evolutionary processes, because it allows us to investigate the genetic basis of adaptive traits in extinct species [14]. At present, however, it is not clear what the limits of these new technical approaches are, in terms of efficiency (ratio of endogenous versus exogenous DNA retrieved), age of the sample, geographic location and/or thermal history. For instance, some mammal species,
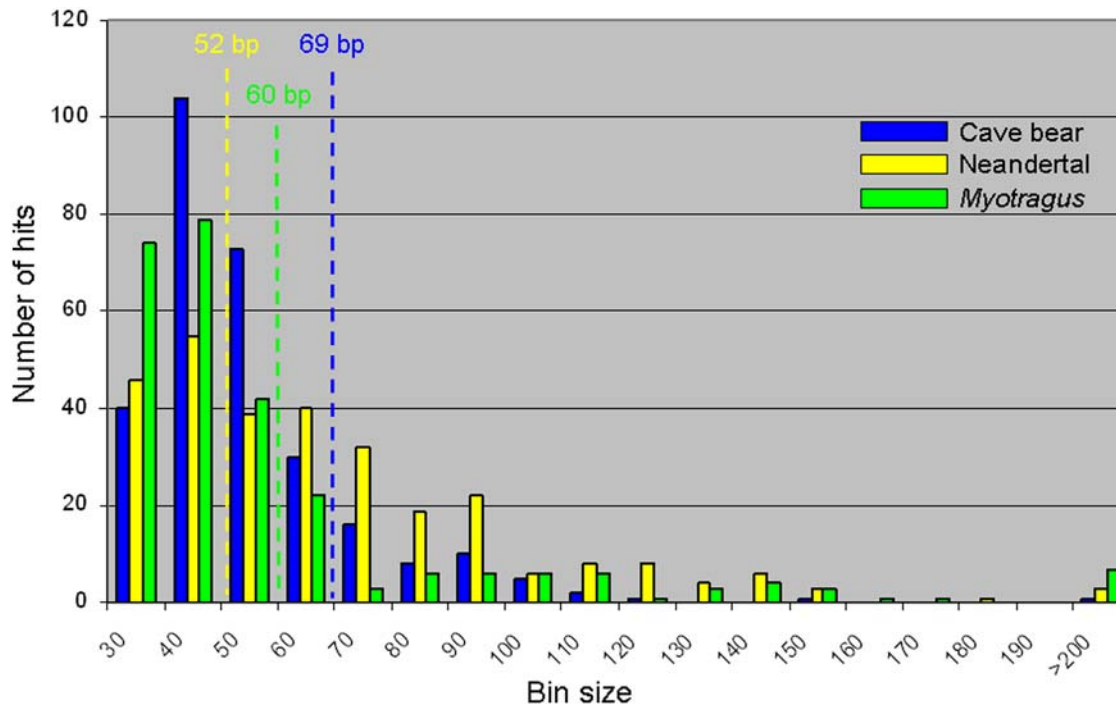
Figure 1. Size distribution, plotted in 10 bp bins, of Neandertal [10]**, cave bear** [3] **and** *Myotragus* **sequences obtained from metagenomic analyses.** The average hit size in each case is indicated by a dotted line.
doi:10.1371/journal.pone.0005670.g001

including goats, cervids, elephants and hippos, have gone extinct in the last few thousands years in the Mediterranean islands [17 and references therein], a temperate area which is clearly not favourable to DNA preservation. The possibility of having access to the genome of these species is therefore of great interest for exploring unique insular evolutionary patterns.

In previous studies [18–20], we have retrieved by PCR mitochondrial genes (Cytochrome *b*, 12S rRNA) and a multi-copy nuclear gene (28S rRNA) from one of these species, *Myotragus balearicus*, an extinct goat from the Western Balearic Islands (Western Mediterranean). *Myotragus* is an extremely modified caprine [21] that evolved in insularity conditions since the end of the Messinian crisis (5.35 million years ago) in the islands of Mallorca and Menorca [22,23]. It became extinct between 3,700 and 2,040 years BC, probably after the arrival of modern humans to the Balearic Islands [24], that took place between 2,350 and 2,150 years BC [25]. The unclear taxonomic position of this caprine is related to its amazing morphological peculiarities, which include extreme size reduction (250–500 mm shoulder height), a single, ever-growing rodent-like lower incisor, shortened distal limb bones, frontal eyes, and reduced brain size [26–31]. Although ancient mitochondrial DNA (mtDNA) data have provided statistical support for a *Myotragus* clade with *Ovis* [20], the general phylogeny of the caprine group is not yet fully established [20]. Here, we demonstrate that it is possible to undertake shotgun sequencing approaches from ancient bones from the thermally unfavourable Mediterranean area. Additionally, the paleogenomic data obtained from *Myotragus* support phylogenetic relationships previously generated with mtDNA sequences.

Results

A total of 96,357 singleton GS-FLX sequence reads were obtained and analysed by means of database searches. No

significant identity was found for 98.49% of the sequences, a figure higher than that found in the ancient wolf, horse and cattle (86.8% on average), Neanderthal (79%) and mammoth shotgun (5.53%, 18.4% and 24.92%, depending on the study) [5,8,9,10]. A fraction of those sequences could be endogenous, but remain unidentified due to the incompleteness of the cow and specially of the sheep genome. Alternatively, the high fraction of sequences without any match may reflect a lack of environmental DNA studies in the Mediterranean area. The remaining 1.51% of the sequences were taxonomically classified by the highest identity found in the database.

Only 0.27% of the sequences, comprising 15,832 nucleotides, gave the best hit to the cow genome, with an average percentage identity of 94.95%. This figure seems to be in agreement with divergence times of about 12–14.3 million years for the cattle-caprine lineages, as suggested from genetic and morphological data [32]. In addition, 0.35% of the sequences gave the best hit to the human genome, with almost 100% of identity, indicative of exogenous contamination. The most represented taxonomic group, however, was bacteria (0.69%), followed in decreasing order by invertebrates (0.12%), plants (0.05%), fungi (0.02%) environmental sequences (0.02%) and others (0.02%). The average length of these sequences was 59.97 nucleotides, and they ranged from 30 bp (determined by the length cut-off in the analysis) to 245 bp (limited by the GS-FLX technology) (Figure 1). The length was similar to those of putatively endogenous sequences found in the Neanderthal and the cave bear metagenomic library (52 and 69 nucleotides, respectively) [3,33]. The presence of two sequences deriving from the *Bos* Y chromosome indicated that the *Myotragus* specimen analysed was likely male.

The human contaminant sequences were significantly longer on average than the *Myotragus* ones (81.57 and 59.97 respectively, *P*<0.0001), suggesting that they were more recent and therefore, less degraded. The longest (>200 bp) *Myotragus* sequences did not

have higher identities to the cow sequences than the shorter ones (85.6% versus 94.95%, respectively), indicating that they are unlikely to derive from recent cow-mediated DNA contamination. Furthermore, no remains of cow were found inside the Cova Estreta cave.

To additionally confirm that the bovid-like DNA fragments were endogenous, we designed five primer pairs from the shotgun sequences that matched *Myotragus* specific substitutions in their 3′ ends and represented unambiguous (those that did produce only one match to the *Bos* genome) BLAST hits. These nuclear fragments, varying between 80 and 112 nucleotides in length, were co-amplified with a previously known 113 bp fragment of the 12S mtDNA gene [20]. In the PCR, we used, to overcome inhibitors present in ancient extracts, rat serum albumin (RSA) [34] instead of the usual bovine serum albumin (BSA) to avoid possible cow contamination in the BSA. One nuclear fragment, along with the mtDNA gene, showed an amplification product and was subsequently cloned and sequenced. The nuclear sequence was identical to that obtained in the shotgun sequencing except for two nucleotide changes that could be related to DNA damage, both in the shotgun and in the PCR-generated sequence (Figure S1).

The plotting of the *Myotragus* sequences along the *Bos* chromosomes showed an excess of sequences in chromosomes 3, 16 and 23, although they were not statistically significant after applying a Bonferroni correction (Figure 2). This pattern could correspond to chromosomal duplications unique to the *Myotragus* lineage or shared by all the Caprinae species, although more sequences and the completion of the *Ovis* genome are needed to explore in the future this possibility. Most of the identified *Myotragus* sequences correspond to unannotated genomic regions of the cow, with only 3.42% of the sequences and 3.90% of the nucleotides belonging to coding regions (Figure S2). The predicted *Myotragus* genes are listed in Table S1.

To explore the phylogenetic signal of the *Myotragus* sequences, we further searched for orthologous sequences in three Bovidae species (*Bos taurus*, *Ovis aries* and *Capra hircus*) and one Cervidae species (*Muntiacus*) in GenBank. However, we noticed a greater genomic coverage of the *Bos* genome that generated an excess of matches due to the presence of multiple paralogs. These sequences might remain undetected in the other genomes due to their more limited coverage. Therefore, we created a sub-dataset of 80 sequences (accounting for a total of 1,987 nucleotides after removing gaps and missing data) that included only those sequences that did not produce multiple matches in none of the genomes. With these sequences, we generated a maximum-likelihood phylogenetic tree that showed the topology previously established from mtDNA data for these species [20] (Figure 3), in which *Myotragus* grouped first with *Ovis*. However, the bootstrap support for this tree was low (64%). The same topology was found with Bayesian trees with a probability of 0.97 for the *Myotragus-Ovis* group. The overall congruence of this partial genomic phylogeny and the mtDNA tree further supports the authenticity of the *Myotragus* sequences.

The large branch found for *Myotragus* in the phylogenetic tree (Figure 3) could be attributed to sequence changes due to DNA damage or to an accelerated evolution of the *Myotragus* genome. To test these possibilities, we characterised the nucleotide changes exclusively present in the aligned *Myotragus* sequences (and different to those from *Ovis* and *Bos*) and found a statiscally significant ($P<0.05$) bias towards higher C to T/G to A ratios, as compared to the T to C/A to G (Figure 4, Table S1). This feature has been previously described as damage-derived lesions due to cytosine deaminations [35]. However, the removal of these substitutions from the alignment only barely shortened the *Myotragus* branch in subsequently generated trees. Specifically, the *Myotragus* branch was 3.1 times longer than the *Ovis* branch in the original alignment and it was still 2.8 times longer than *Ovis* after removing putatively
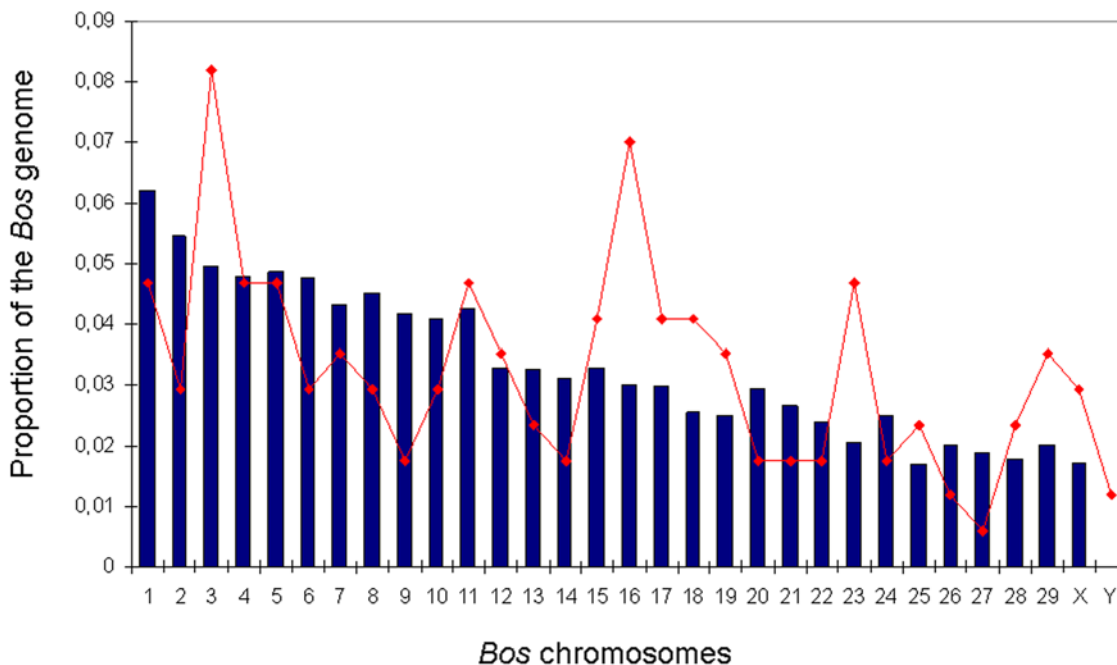

Figure 2. The proportion of the *Bos* genome contained on each chromosome (blue bars) is shown with the proportion of *Myotragus* sequences (red line) aligning to each *Bos* chromosome with exactly 1 hit with e-value<1e-3 BLAST. The observed distribution is not significantly different from the expected one when we compare all the chromosomes together ($P = 0.081$) or when we tested each one independently and correct for multiple testing.
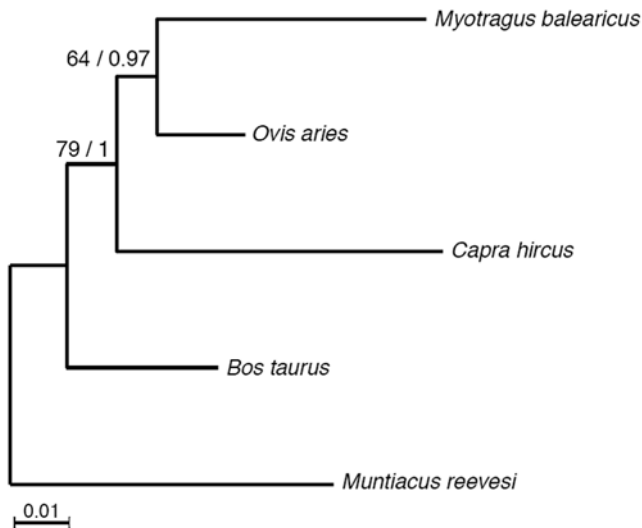doi:10.1371/journal.pone.0005670.g002

Figure 3. Maximum-likelihood phylogenetic tree of *Myotragus balearicus* and other artiodactyls. The tree was rooted in the cervid *Muntiacus reevesi*. Numbers along the branches indicate bootstrap support of the maximum-likelihood analyses (first number) and Bayesian support of an independent Bayesian analysis (second number). The scale bar represents 0.01 substitutions/site.
doi:10.1371/journal.pone.0005670.g003

damaged positions. Thus, an important contribution of accelerated evolution in the *Myotragus* genome cannot be discarded. However, a similar acceleration in *Capra* indicates that this phenomenon is not specific to the Balearic lineage.

Discussion

Molecular studies, mainly based on mtDNA data, have failed until now to fully resolve the caprine phylogeny, probably due to the explosive radiation of this group [20]. The phylogenetic analysis of the present paleogenomic data supports the previous caprine relationships established from mtDNA, but also indicates the potential of this approach for testing evolutionary hypothesis and establishing robust phylogenies.

Despite being excavated in a region with a mean annual temperature of 14°C and below 40 degrees North latitude, we have been able to successfully retrieve nuclear genome sequences from a ~6,000 years old *Myotragus balearicus* bone. The extremely low efficiency of the paleogenomic retrieval is striking, as is the fact that the level of human contaminant sequences is higher than that of the endogenous ones (0.34 vs 0.27). In contrast, the ratio of endogenous to contaminant human sequences among the colder preserved Denisova mammalian samples was 49:1 [10], and the human sequences accounted for less than 0.015% [10]. In a mammoth sample from the Artic Circle [5], this ratio was 32:1, and the human contaminants up to 1.4% of the total sequences. The *Myotragus* sample was retrieved with no special precautions against contamination. However, the histological structure of the bone also correlates with contamination levels [36], and the cortical tissue in *Myotragus* limb bones is thinner than in other, larger extinct mammals, such as mammoths and Neanderthals. Thus, it is not clear if this figure can be taken as an estimate of potential human contamination in ancient bone specimens stored in museums.

The efficiency ratio of retrieval of endogenous *Myotragus* sequences is the lowest among those observed in some other bone-based metagenomic studies, obtained from samples at higher latitudes: 47.4°N [3], 74°N [5], 45.5°N [8] and 51.23°N [10].
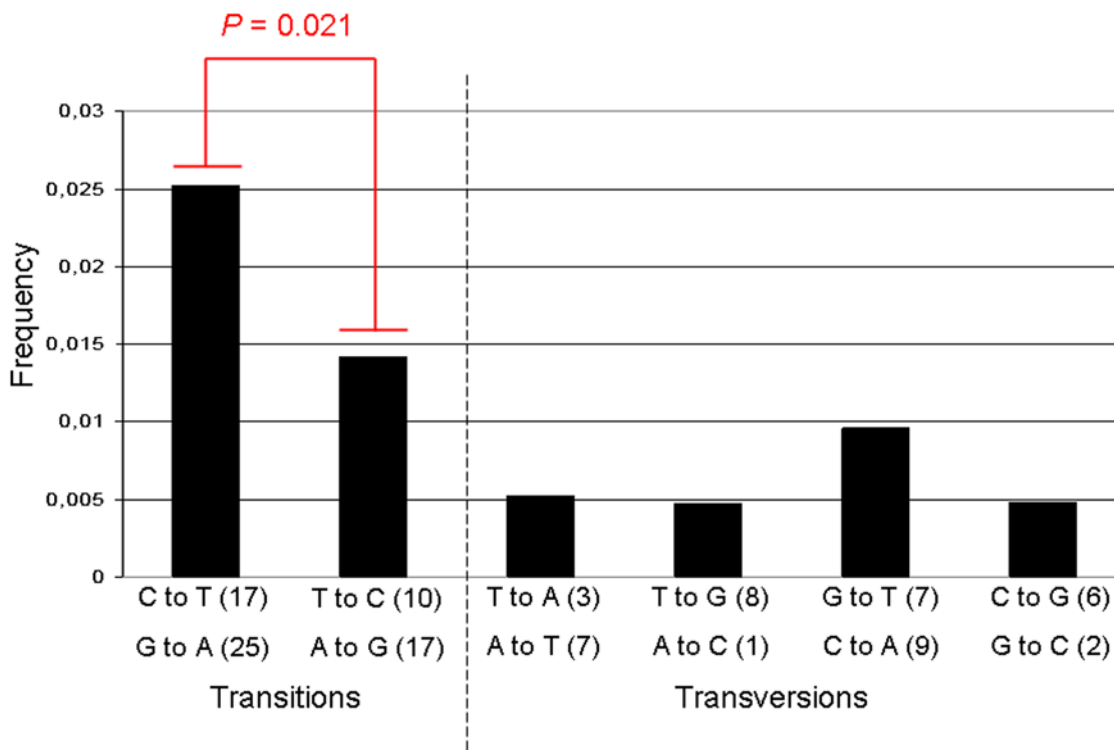
Figure 4. Frequency distribution of 113 *Myotragus*-specific substitutions observed in 3,602 bp of aligned *Ovis*, *Myotragus* and *Bos* genomic sequences. Complementary substitutions (such as C to T and G to A) are considered equivalent events. Fisher's exact test was used to calculate the excess of *Myotragus*-specific C to T and G to A transitions. The total number of each substitution is in parentheses.
doi:10.1371/journal.pone.0005670.g004

Despite the low efficiency values, the mean fragment length and the range value of the *Myotragus* sequences are similar to those found in samples with higher efficiencies. In addition, the frequency of damage-derived lesions in the *Myotragus* sequences is 4.2 times lower than those found in Neanderthal sequences [33]. Even so, some inconsistencies have been found between the shotgun and the PCR-based sequence, indicating the need for targeted approaches in genomic regions with low coverage. A previous study has estimated that a 12-fold coverage would be needed to have an error rate of 1 in 10,000 nucleotides [37], something extremely expensive to achieve in highly degraded ancient samples.

These somehow contradictory results between low retrieval efficiency and low DNA damage can be due to a combination of factors. On one hand, the temperate climatic conditions of the Mediterranean islands are highly unfavourable to paleogenomic preservation, although the cave where the bones were found has maintained a rather constant temperature inside. On the other hand, the *Myotragus* sample used is much more recent than those from wolf, horse, cattle, cave bear, Neanderthal and mammoth, all of them dated between 20,000 and 69,000 years ago [3,5,8,9,10]. However, the estimated thermal age [38] for this bone at the excavation is 26,206 years at 10°C (David Harker, personal communication). This age is older than that estimated for the 100,000 years-old Scladina Neanderthal [38], which is in agreement with the low efficiency of DNA retrieval found in the present study.

Our findings imply that we are working at the very limits of the current paleogenomic approaches, but still they are more efficient than PCR-based strategies, which are problematic for genomic studies on similarly preserved samples. In fact, under these unfavourable environmental conditions, only paleogenomic approaches can provide the amount of sequence data generated here. In the future, with greater genomic coverage, paleogenomic approaches could provide further data to study other aspects of this Balearic endemism, such as evidences for selective sweeps in the *Myotragus* genome related to its particular adaptations. Also, our results suggest that livestock domestication events that took place in the Fertile Crescent could be approachable from paleogenomics.

Materials and Methods

A left *Myotragus* radius bone (IMEDEA 43619) from Cova Estreta (Pollença, Mallorca) was chosen for analysis because of its excellent macroscopic preservation. Previous analyzed bones were excavated in a different site, Cova des Gorgs (Escorca, Mallorca) [20]. Cova Estreta is a deep and narrow cave discovered in 1996 that acted as a natural trap for *Myotragus* [31]. Radiocarbon dates from bones obtained from the same stratigraphical unit [UtC-5175, 6,357±44 BP (5469–5225 calBC)] and [UtC-5171, 5,720±60 (4716–4449 calBC)] allow us to establish a narrow chronological age for the studied material of ~6,000 years.

A sample of 3 g of cortical tissue was powdered, digested with proteinase K and extracted with phenol-chloroform, following a protocol described elsewhere [5]. Previous metagenomic studies have described an overwhelming fraction of environmental DNA found in ancient bones. Following a previously published procedure, the bone powder was incubated with bleach for five

minutes, prior to extraction [39]. It was assumed that this could remove part of the pervasive exogenous DNA and thus increase the efficiency of the endogenous DNA retrieval. The fact that so much contamination is still seen afterwards is intriguing. Further studies could help clarify the efficiency of the bleach treatment prior to GS-FLX 454 sequencing.

One hundred microliters of extract were subjected to the GS-FLX 454 sequencing platform. The nebulization and Ampur purification steps were omited for the library building process, following, except for this, the manufacturer's guidelines (Roche Diagnostics). The amount of DNA in the libraries was estimated by Quantitative PCR (qPCR) [40] and found to be too low for successful sequencing. Therefore, libraries were amplified with the emulsion primers prior to the emulsion PCR (ePCR) to increase the amount of DNA. This procedure generated redundant sequences that were posteriorly identified and eliminated. Subsequently half of a full sequencing run was performed on the commercial Cogenics Genome Express FLX platform (Grenoble, France). To confirm the authenticity and accuracy of the GS-FLX generated data, a small number of mtDNA and nuclear DNA sequences were targeted using conventional PCR protocols, following a two-steps protocol [41] and 50 degrees of annealing temperature. Amplification products were cloned using the TOPO TA cloning kit (Invitrogen), and sequenced using an ABI3730 capillary sequencer (Applied Biosystems).

Obtained sequences were identified with BLAST searches [42] (using the megaBLAST program with an e-value threshold of 0.001) using the cow and human genomes, the environmental sample sequences database in the GenBank *env*, and the general nucleotide sequences *nt*. Sequences of other bovids (*Bos taurus*, *Ovis aries* and *Capra hircus*) as well as one cervid (*Muntiacus reevesi*) were aligned to those of *Myotragus* with Multialin [43]. Discrepancies in homopolymeric tracts were not considered, as 454 technology is known to have problems dealing with these regions [33]. Best match to target sequences in the blast didn't include the edge nucleotides, since these are known to accumulate postmortem damages associated to the breakage DNA process.

Phylogenetic trees were constructed by maximum likelihood with the Phyml program, version 2.4.4 [44]. A general time reversible (GTR) model with four rate categories and a proportion of invariable sites was used, with parameters estimated from the data. A bootstrap analysis with 100 replicas was also performed. In addition, a Bayesian tree was calculated with MrBayes 3.1 [45] using a GTR model with invariable sites and rate heterogeneity. Two runs of four chains of 5,000,000 trees were generated, sampling every 100 trees, with burning completed by the 20,000th tree.

Author Contributions

Conceived and designed the experiments: CLF. Performed the experiments: OR EG CLF. Analyzed the data: OR EG JB JC CLF. Contributed reagents/materials/analysis tools: JB. Wrote the paper: OR CLF. Provided samples and paleontological information: PB JAA.

References

1. Hofreiter M (2008) Paleogenomics. C R Palevol 7: 113–124.
2. Gilbert MTP (2008) Paleogenomics using the 454 sequencing platform. In: Janitz M, ed. Next Generation Genome Sequencing: Towards Personalized Medicine Wiley-VCH Verlag GmbH & KgaA, Weinheim. pp 183–199.
3. Noonan JP, Hofreiter M, Smith D, Priest JR, Rohland N, et al. (2005) Genomic sequencing of Pleistocene cave bears. Science 309(5734): 597–599.
4. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437(7057): 376–380.

5. Poinar HN, Schwarz C, Qi J, Shapiro B, Macphee RD, et al. (2006) Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. Science 311(5759): 392–394.

6. Gilbert MT, Tomsho LP, Rendulic S, Packard M, Drautz DI, et al. (2007) Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. Science 317(5846): 1927–1930.

7. Miller W, Drautz DI, Janecka JE, Lesk AM, Ratan A, et al. (2009) The mitochondrial genome sequence of the Tasmanian tiger (Thylacinus cynocephalus). Genome Res 19(2): 213–220.

8. Miller W, Drautz DI, Ratan A, Pusey B, Qi J, et al. (2008) Sequencing the nuclear genome of the extinct woolly mammoth. Nature 456: 387–390.

9. Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, et al. (2006) Analysis of one million base pairs of Neanderthal DNA. Nature 444(7117): 330–336.

10. Blow MJ, Zhang T, Woyke T, Spelle CF, Krivoshapkin A, et al. (2008) Identification of the source of ancient remains through genomic sequencing. Gen Res 18(8): 1347–1353.

11. Wall JD, Kim SJ (2007) Inconsistencies in Neanderthal genomic DNA sequences. PLoS Genet 3: 1862–1866.

12. Gilbert MT, Binladen J, Miller W, Wiuf C, Willerslev E, et al. (2007) Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. Nucleic Acids Res 35(1): 1–10.

13. Stiller M, Green RE, Ronan M, Simons JF, Du L, et al. (2006) Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. Proc Natl Acad Sci USA 103: 13578–13584.

14. Römpler H, Rohland N, Lalueza-Fox C, Willerslev E, Kuznetsova T, et al. (2006) Nuclear gene indicates coat-color polymorphism in mammoths. Science 313(5783): 62.

15. Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, et al. (2007) The derived FOXP2 variant of modern humans was shared with Neanderthals. Curr Biol 17(21): 1908–1912.

16. Lalueza-Fox C, Römpler H, Caramelli D, Stäubert C, Catalano G, et al. (2007) A melanocortin 1 receptor allele suggests variiying pigmentation among Neanderthals. Science 318: 1453–1455.

17. Ramis D, Alcover JA (2005) Holocene extinction of endemic mammals of the Mediterranean Islands: some methodological questions and an update. In: Alcover JA, Bover P, eds. Proceedings of the International Symposium "Insular Vertebrate Evolution: the Palaeontological Approach". Palma de Mallorca: SHNB, IMEDEA. pp 309–318.

18. Lalueza-Fox C, Bertranpetit J, Alcover JA, Shailer N, Hagelberg E (2008) Mitochondrial DNA from Myotragus balearicus, an extinct bovid from the Balearic Islands. J Exp Zool (Mol Dev Evol) 288: 56–62.

19. Lalueza-Fox C, Shapiro B, Bover P, Alcover JA, Bertranpetit J (2002) Molecular phylogeny and evolution of the extinct bovid Myotragus balearicus. Mol Phylogenet Evol 25: 501–510.

20. Lalueza-Fox C, Castresana J, Sampietro L, Marqués-Bonet T, Alcover JA, et al. (2005) Molecular dating of caprines using ancient DNA sequences of Myotragus balearicus, en extinct endemic Balearic mammal. BMC Evol Biol 5: 70.

21. Bate DMA (1909) A new artiodactyle from Majorca. Geol Mag NS 6: 385–388.

22. Alcover JA, Moyà-Solà S, Pons-Moyà J (1981) Les quimeres del passat. Els vertebrats fòssils del Plio-Pleistocè de les Balears i Pitiusses. Palma de Mallorca: Moll.

23. Bover P, Quintana J, Alcover JA (2008) Three islands, three worlds: paleogeography and evolution of the vertebrate fauna from the Balearic Islands. Quat Int 182: 135–144.

24. Bover P, Alcover JA (2003) Understanding Late Quaternary Extinctions: the case of Myotragus balearicus Bate 1909. J Biogeograph 30: 771–782.

25. Alcover JA (2008) The first Mallorcans: prehistoric colonization in the Western Mediterranean. J World Prehist 21: 19–84.

26. Andrews CW (1915) A description of the skull and skeleton of a peculiarly modified rupicaprine antelope (Myotragus balearicus Bate 1909) with a notice on a new variety Myotragus balearicus var. major. Phil Trans Roy Soc London B 206: 281–305.

27. Spoor CF (1988) The body proportions of Myotragus balearicus Bate, 1909. Proc Konink Ned Akad Wetens B 91(3): 285–293.

28. Spoor CF (1988) The limb bones of Myotragus balearicus Bate, 1909. Proc Konink Ned Akad Wetens B 91(3): 295–308.

29. Bover P, Alcover JA (1999) The evolution and ontogeny of the dentition of Myotragus balearicus Bate, 1909 (Artiodactyla, Caprinae): evidence from new fossil data. Biol J Linnean Soc 68: 401–428.

30. Encinas JA, Alcover JA (1997) El jaciment fossilífer de la cova Estreta (Pollença). Endins 21: 83–92.

31. Kohler M, Moyà-Solà S (2004) Reduction of brain size and sense organs in the fossil insular bovid Myotragus. Brain Behav Evol 63: 125–140.

32. Hassanin A, Douzery EJP (1999) The Tribal Radiation of the Family Bovidae (Artipdactyla) and the Evolution of the Mitochondrial Cytochrome b Gene. Mol Phylogenet Evol 13(2): 227–243.

33. Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, et al. (2006) Sequencing and Analysis of Neanderthal Genomic DNA. Science 314: 1113–1118.

34. Beja-Pereira A, Caramelli D, Lalueza-Fox C, Vernesi C, Ferrand N, et al. (2006) The origin of European cattle: Evidence from modern and ancient DNA. Proc Natl Acad Sci USA 103: 8113–8118.

35. Hofreiter M, Jaenicke V, Serre D, Haeseler Av A, Pääbo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res 29(23): 4793–4799.

36. Gilbert MTP, Rudbeck L, Willerslev E, Hansen AJ, Smith C, et al. (2005) Biochemical and physical correlates of DNA contamination in archaeological human bones and teeth excavated at Matera, Italy. J Archaeol Sci 32: 783–795.

37. Green RE, Malaspinas AS, Krause J, Briggs A, Johnson PLF, et al. (2008) A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. Cell 134: 416–426.

38. Smith CI, Chamberlain AT, Riley MS, Stringer C, Collins M (2003) The thermal history of human fossils and the likelihood of successful DNA amplification. J Hum Evol 45: 203–217.

39. Malmström H, Svensson EM, Gilbert MTP, Willerslev E, Götherström A, Holmlund G (2007) More on contamination: the use of asymmetric molecular behavior to identify authentic ancient human DNA. Mol Biol Evol 24(4): 998–1004.

40. Meyer M, Stenzel U, Myles S, Prüfer K, Hofreiter M (2007) Targeted high-throughput sequencing of tagged nucleic acid samples. Nucleic Acids Res 35: e97.

41. Krause J, Dear PH, Pollack JL, Slatkin M, Spriggs H, et al. (2006) Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. Nature 439(7077): 724–727.

42. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410.

43. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16: 10881–10890.

44. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.

45. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.