Biodiversity lost: The phylogenetic relationships of a complete mitochondrial DNA genome sequenced from the extinct wolf population of Sicily

Stefano Reale[1], Ettore Randi[2], Floriana Bonanno[1], Valentina Cumbo[1], Ignazio Sammarco[1], Antonio Spinnato[1], Salvatore Seminara[1]

[1]Istituto Zooprofilattico Sperimentale della Sicilia "A. Mirri", Via Gino Marinuzzi 3, 90129 Palermo, Italy; e-mails: *stefano.reale@izssicilia.it; floriana.bonanno5@gmail.com, netsamma@gmail.com, valentina.cumbo@gmail.com, antoniospinnato@libero.it, salvatore.seminara@izssicilia.it*

[2]Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy; and Faculty of Engineering and Science, Department of Chemistry and Bioscience, University of Aalborg, Aalborg, Denmark

Corresponding author:

Ettore Randi; BIGEA, University of Bologna, Via Selmi, 3, 40126 Bologna, Italy

E-mail: *ettore.randi@unibo.it*

Abstract

Using next-generation sequencing, we obtained for the first time a complete mitochondrial DNA genome from a museum specimen of the extinct wolf (*Canis lupus*) population of the island of Sicily (Italy). Phylogenetic analyses showed that this genome, which was aligned with a number of historical and extant complete wolf and dog mtDNAs sampled worldwide, was closely related to an Italian wolf mtDNA genome (TN93 and *p*-distances = 0.0012), five to seven times shorter than divergence among Sicilian and any other known wolf mtDNA genomes (distance range = 0.0050 – 0.0070). Sicilian and Italian haplotypes joined a basal clade belonging to the mtDNA haplogroup-2 of ancient western European wolf populations (Pilot et al. 2010). Bayesian calibration of divergence times indicated that this clade coalesced at MRCA = 13.400 years (with 95% HPD = 4000 – 21.230 years). These mtDNA findings suggest that wolves probably colonized Sicily from southern Italy towards the end of the last Pleistocene glacial maximum, when the Strait of Messina was almost totally dry. Additional mtDNA and genomic data will further clarify the origin and population dynamics before the extinction of wolves in Sicily.

Key words: Sicilian wolf; *Canis lupus*; complete mtDNA genome; next-generation sequencing; island extinctions.

Introduction

During the last few centuries biodiversity has been dramatically destroyed worldwide (WWF 2018). With a few exceptions, range, abundance and genetic diversity of many animal populations declined pressed by the impacts of human population expansion and degradation of natural habitats (Li et al. 2016). In particular, large vertebrates and top predators paid the price of environmental anthropogenic changes. Wolf (*Canis lupus*) is one of the few large predators that managed to survive safely the Pleistocene faunal turnover (Loog et al. 2018). Nevertheless, wolf populations fluctuated widely in the Old and New World through the Pleistocene and more recently, during the last few centuries, in consequence of deforestation, prey overhunting and direct persecution (Leonard et al. 2005; Randi 2011). In the recent past, wolves disappeared from southern and central regions of North America and from most of central European countries. Although wolves are currently recovering, aided by legal protection, controlled hunting and active conservation, the recent demographic declines led to a number of local extinctions, in particular small and isolated populations (Linnell et al. 2008). For instance, the last Honsu wolf (*C. l. hodophilax*), the endemic dwarf subspecies living in three of the main islands of Japan, was killed in 1905 (Ishiguro et al. 2009). The Ezo wolf (*C. l. hattai*), endemic of the Hokkaido, was eradicated from the island by the end of the 1800s (Ishiguro et al. 2010). Recently, also the very small isolated wolf population in the Sierra Morena of central Spain has been definitely eradicated (Gómez-Sánchez et al. 2018).

Fossil remains indicate that wolves were present in Europe at least by the end of the Middle Pleistocene at about 0.5–0.3 million years ago (Sotnikova & Rook 2010). However, those ancestral wolf populations, which showed distinct ecology, morphology and genetics, were completely substituted by contemporary wolves that rapidly spread across all Europe *c*. 25.000 – 20.000 year ago (Loog et al. 2018). The extant wolves of the Italian peninsula(*C. l. italicus*) are genetically divergent from all the other wolf populations in Europe, likely due to their long-lasting isolation south of the Alps, to historical and recent anthropogenic bottlenecks (Lucchini et al. 2004; Pilot et al. 2014). Wolves in peninsular Italy hardly survived at the end of the 1970' in the southernmost parts of the Apennine, distributed in two small isolated subpopulations counting less than 100 individuals in total (Boitani 1984). In the last two centuries wolves in the Alps and Apennine harboured much more mtDNA diversity that presently (Dufresnes et al. 2018). Since

then, the recovery of the Italian wolves has been spectacular, but at that time wolves living in the island of Sicily already went extinct (Angelici et al. 2018).

Information on the Sicilian wolf population is scanty: its phylogeographic origin, historical distribution and abundance in the island are largely unknown and only partially reconstructed anecdotally. Most probably wolves were already rare in first half of the 1800', likely due to habitat and natural prey loss, and direct persecution. Crossbreeding with free-ranging dogs was already a threaten (Minà Palumbo 1868). However, purpose hunting has likely been the main cause of the Sicilian wolf population decline and final extinction. Documented reports mentioned the killing of seven wolves in 1891 in San Fratello (Messina) and one wolf in 1902 in San Pietro (Caltagirone) (Pratesi 1978). Although reports mentioned the alleged presence of wolves until 1959, the last documented wolf was killed in 1935 in the woodlands in Ficuzza (Giovanni Giardina and Andrea Milazzo; *pers. com.*). Morphologic and genetic analyses showed that the skin of a canid shoot in 1924 in Bellolampo (Palermo; preserved at the Regional Museum of Terrasini), belongs to a domestic dog or a hybrid (Angelici et al. 2018). A few other wolves killed before 1935 are preserved in museums in Sicily or in peninsular Italy (Angelici et al. 2018).

Wolves in Sicily have been cut of their mainland populations in consequence of Mediterranean sea-level fluctuations and flooding of land-bridge connections. Paleogeographic reconstructions and paleontological data documented periods of intense African-Sicily faunal interchange through the strait of Sicily during the Messinian land-bridge, *c*. 5.3 million years ago (Stock et al. 2008). More recently, concurrently with the Pleistocene glacial maxima, the north-eastern coasts of the island were in connection with the southern tips of peninsular Italy (Calabria) (Antonioli et al. 2014). The Messina strait, currently 3.2 km wide and 80-120 m deep, has been repeatedly dried at glacial maxima, and Pleistocene temporary land-bridges have been used by a number of animal populations to colonize the island (Antonioli et al. 2014). Wolf remains have been identified in the fossil record of peninsular Italy since the late Middle Pleistocene, *c*. 340,000 years ago (Anzidei et al. 2012), but modern wolf populations (*Canis lupus subspp*.) colonized the peninsula much more recently, towards the end of the last glaciations (Bertè and Pandolfi 2014). Although morphological traits and preliminary molecular identifications of a few Sicilian wolves specimens and body remains preserved in museum collections were recently described (Angelici et al. 2018), the origin and phylogenetic relationships of the Sicilian wolves are still largely unknown.

In this study we present a new sequence of a complete mtDNA genome obtained from a stuffed Sicilian wolf specimen, aiming at: 1) evaluating the position of this Sicilian wolf mtDNA genomes within the phylogenetic framework of extant and historical wolves in Italy and worldwide; 2) obtaining reliable estimates of haplotype divergence times using complete mtDNA genomes and not limited to very short mtDNA sequences; 3) identifying the likely origin of the sequenced mtDNA genome, either if from a wolf or a dog ancestral maternal population via hybridization; and 4) contributing to reconstruct a plausible scenario of wolf colonization of the island of Sicily.

Materials and Methods

DNA extraction

A single tissue sample (the inner part of a nail) was collected from a stuffed wolf specimen preserved at the Civic Museum "Baldassare Romano", Termini Imerese (Palermo; Italy). Although some authors assert that the wolf (Fig. 1) was probably killed on Monte San Calogero near Termini Imerese (Palermo) in the last years of the 19th century (Angelici and Rossi 2018), we didn't found truthful historical data regarding the its origins and the years of acquisition by the museum. The nail surface was decontaminated by UV radiation for 30 minutes. We collected inner dry tissue remains by drilling the nail, and the powder was stored in sterile UV-decontaminated test tubes. DNA was extracted twice using a specific forensic kit and procedure (ChargeSwitch® Forensic DNA Purification Kit, Thermo Fisher). In short: the sample was lysed overnight under agitation at 50°C; the DNA was bind to magnetic beads and then cleaned through three washing step employing a magnetic tube support; finally the DNA was recovered washing the magnetic beads with 50 µl of elution buffer. The DNA quality and concentration was checked and quantified by spectrophotometer analysis using Nanodrop (Thermo Fisher Scientific, Waltham, Massachusetts, USA), Qubit (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and TapeStation (Agilent Technologies Inc. Santa Clara, California) equipments. To minimize risk of contaminations by exogenous DNAs, all sample manipulations were done in a DNA-free area that was never used before for DNA extractions. All bench-tops and equipments were flamed or

cleaned with bleach, and ethanol and UV irradiated for 60 minutes before and after their use. We used pipette tips with aerosol filters. The DNA library was prepared in a DNA-free area using decontaminated equipments only.

NGS library preparation and DNA sequencing

We used the Illumina Truseq DNA Nano kit (Illumina Inc., San Diego, CA, USA) for library preparation according to the manufacturer instructions with some modifications. We did not perform the fragmentation step due to the small size of the input DNA as shown by TapeStation spectrophotometer analysis. We performed end repair to blunt the DNA fragments on a thermal cycler at 30°C for 30 minutes using the End-Repair Mix2. Appropriate library size for sequencing was selected removing the shortest DNA fragments with 250 µl of undiluted Sample Purification Beads reagent (SPB). A single adenine nucleotide was added to the 3' ends of the blunt fragments to prevent them from ligating to each other and to provide complementary bases for adapters. In detail: 12.5 µl of A-Tailing Mix were added and the following thermal cycler program was used: 37°C for 30 minutes, 70°C for 5 minutes, 4°C for 5 minutes, hold at 4°C. Adapter ligation was performed adding 2.5 µl Resuspension buffer, 2.5 µl LIGation Mix2, 2.5 µl DNA adapter solution, and incubating at 30°C for 10 minutes. The reaction was stopped adding 5 µl of Stop Ligation Buffer. Adapter dimers were removed from the library cleaning the ligated fragments by SPB. The library was PCR-amplified with 5 µl of PCR Primer Cocktail that anneals to the ends of the adapters and 20 µl of Enhanced PCR Mix at 95°C for 3 minutes, 8 cycles of: 98°C for 20 seconds, 60°C for 15 seconds, 72°C for 30 seconds, 72°C for 5 minutes, hold at 4°C. The amplified DNA was cleaned with SPB. Accurate quantification of DNA libraries, a critical step to produce optimal cluster densities across every lane of the flow cell and achieve the highest quality sequencing data, was assayed using a fluorometric method based on dsDNA binding dyes. The distribution of the amplified DNA fragments size was centered at the 295 bp. To obtain the input DNA solution for the sequencing reaction on Illumina Cartridge, the selected size of DNA was at first normalized at 4nM and then diluted at 12.5 pM. The sequencing step was performed with a Illumina MiSeq sequencer using a SBS MiSeq Reagent Kit v2 (Illumina). PhiX Control library (v2; Illumina) was added to the library. The libraries were sequenced with a 150 Paired End MiSeq run. Image analysis, base calling and data quality assessment were performed on the MiSeq and the BaseSpace cloud software was used to generate the final FASTQ files.

Bioinformatic analyses and mtDNA genome assemblage

The output files from the Illumina sequencer were pre-processed removing adapter sequences, low quality and short read sequences using AdapterRemoval (Schubert et al. 2016) and Cutadapt (Martin 2011). Edited paired-end reads were merged with their overlapping regions using PEAR (Zhang et al. 2014). The processed reads were aligned using an Italian wolf complete mtDNA reference genome (GenBank accession number KF661048) and the BWA-MEM aligner algorithm (Li and Durbin 2010). Mapped reads in BAM format were filtered using MapDamage 2.0 (Jónsson et al. 2013) and quality scores of C->T or G->A transitions, potentially due to post-mortem DNA damage, were recalculated according to the position in reads and damage patterns. Filtered reads were extracted from BAM file using SAMTools 1.4 (Li et al. 2009) and then assembled into a complete consensus sequence with Spades 3.11 (Bankevich et al. 2012). The quality of assembled mtDNA genome was evaluated using Quast tool (Gurevich et al. 2013). Male-specific DNA sequences from the sex-determining region Y protein (SRY gene) were extracted from the reads of the Sicilian wolf genomes by means of Bowtie 2.3.4.1 (Langmead et al. 2012), using a dog sequence (AF107021.1) as reference. To confirm our sequences, we blasted (Altschul et al. 1990) the extracted reads, which had an average length of 120-130 bp, against GenBank database.

Phylogenetic analyses and estimates of divergence times

We used ClustalW (Higgins et al. 1994) in MEGA X (Kumar et al. 2018) to align the new Sicilian wolf mtDNA concatenated genome with the following mtDNA genomes downloaded from GenBank:

- Set#1: the complete wolf mtDNA genomes (including the control-region; CR) used by Koblmuller et al. (2016), including three historical wolf samples, 14 dog (*C. l. familiaris*) genomes and five Himalayan wolves (named *C. l. laniger* or *C. l. chanco*) used as outgroups;
- Set#2: the modern wolf mtDNA genomes (CR excluded) used by Thalmann et al. (2013), including four coyotes (*C. latrans*) used as outgroups;
- Set#3: a subset of wolf mtDNA genomes (CR excluded) used by Panget al. (2009) and Matsumura et al. (2014), including two historical Ezo wolf (*C. l. hattai*) and five historical Honsu wolf (*C. l. hodophilax*); four coyote mtDNA genomes were used as outgroups.

The alignments were manually checked and adjusted. The control-regions, which were incompletely sequenced in some genomes, were extracted and aligned separately with partial sequences by Dufresnes et al. (2018) obtained from museum specimens originating from the historical wolf populations which lived in the Alps, Italian peninsula and Sicily. Moreover, the mtDNA CR were blasted in GenBank to search for eventual matching with dogs or other wolf CRs. The Sicilian wolf mtDNA protein coding genes were extracted, concatenated and aligned with the homologous coding gene sequences of the three Italian wolf mtDNA genomes in GenBank: KF661048.1 (Thalmann et al. 2013); KU696389.1 and KU644662.1 (Koblmuller et al. 2016). Short overlapping DNA segments, eventual incomplete stop codons and mutations at the three codon positions were identified using MEGAX. We analysed the Set#1 and Set2# alignments using neighbor-joining procedure (NJ; Saitou and Nei 1987) implemented in MEGAX, with the TN93 model (Tamura and Nei 1993), assuming heterogeneous gamma distribution of mutation rates (α = 0.5) and heterogeneous lineages evolution. All positions containing gaps, including a variable copy number repeat unit in the CR (about 511 bp from nucleotide 16,040 to 16,550 in the reference dog sequence NC002008), and missing data were pairwise-deleted in the analyses. Support to the phylogenetic tree internodes were determined by 1000 interior-branch length test of minimum evolution trees (ME; Nei and Kumar 2000) and by 10,000 bootstrap samplings of NJ trees (Tamura et al. 2013). Four coyote mtDNA sequences: DQ480509, DQ480510 and DQ480511 (Björnerfeldt et al. 2006); KF661096 1 (Thalmann et al. 2013) were used as outgroups. Bayesian phylogenetic trees were obtained using BEAST 2.5.1 (Drummond et al. 2012), with the HKY+G model (Hasegawa et al. 1985). Markov chain Monte Carlo (MCMC) samples were drawn every 1000 generations from a total of 1,000,000 generations, following a discarded burn-in of 100,000 generations. BEAST was used for age estimates of the nodes of phylogenetic trees. The HKY+G model was assumed for the nucleotide substitution. Both the uncorrelated log-linear model (Drummond et al. 2006) and the strict clock model were tested for the molecular clock, and both the Bayesian skygrid model (Gill et al. 2013) and a constant model for the population size. The convergence and performance of different models were assessed using Tracer 1.7.1 (Drummond and Rambaut 2007). A 10% of the MCMC generations were discarded as a burn-in. BEAST was used to infer the age of the most recent common ancestor (MRCA) of a clade joining the Italian and Sicilian wolf mtDNA, which was calibrated using the divergence times among Japanese wolves (MRCA = 46,800; 95% highest

probability density HPD = 37,500–58,000 years) as estimated by Matsumura et al. (2014). The trees were visualized with FigTree 1.4.4, or with TreeAnnotator in BEAST.

Results

We successfully extracted good-quality DNA from the inner nail tissue of an ancient Sicilian wolf museum specimen. The quantity and quality of the extracted DNA was good enough to obtain reliable sequences, which allowed reconstructing a complete mtDNA genome by next-generation sequencing procedures on Illumina platform. DNA quality was preliminary assessed to ensure successful NGS results. A Nanodrop value of the ratio 260/280 nm = 1.9 indicated a low presence of inhibitors; a Q-bit spectrophotometer analysis showed a total double-strand DNA concentration of 1.6 ng/µl. A TapeStation capillary electrophoresis led to visualize a normal distribution of the DNA fragments, ranging from 40 to 1040 bp and centered at 250 bp, with a concentration of 0.6 ng/µl (Fig. 2a). These quality-controls suggested to skip an initial fragmentation step of input DNA for library preparation, and the DNA was directly PCR-enriched with fragments spanning from 165 to 655 bp, and centered at 295 bp, as showed by TapeStation results (Fig. 2b). The reads obtained from the input DNA, which concentration was 2.49 ng/µl, were assembled into a complete Sicilian wolf mtDNA. The length of this mtDNA genome (GeneBank accession number: MH891616.1) was 16 678 bp.

This sequence aligned with reference dog and Italian wolf homologous sequences. All the tRNA, rRNA and protein-coding genes were correctly identified and mapped; these sequences did not show any anomalous stop codon and translated into the expected RNAs or proteins. Thus, we assumed that this mtDNA genome was authentic. In comparison with the known Italian wolf mtDNA genomes, the Sicilian wolf exhibited 14 silent transition substitutions and only one first-position G-A mutation that changed a V into an M aminoacid residue at codon 21 of the ATP6 gene in the Sicilian wolf. The mtDNA CR of the Sicilian wolf was 1219 bp, that is 50 bp shorter than the corresponding sequence of the Italian wolf CR due to 10 missing copies of a CGGTACACGT repeat (Kim et al. 1998). The *p*-distance between the complete mtDNA genome of Sicilian and Italian wolves was $D = 0.0012$ (almost identical to the TN93 distance), that is five to seven times shorter than *Ds* among the Sicilian and any other wolf mtDNA genomes ($D = 0.0050 – 0.0070$). The *p*-

distance between Sicilian and Italian wolves was *D* = 0.0181; hence, as expected, the CR evolved about 10 times faster than the RNA and protein-coding mtDNA sequences.

We analysed the Set#1 (Fig. 3) and Set#2 (Fig. 4) mtDNA genomes by the NJ and ME procedures in MEGAX. The mtDNA genome of the Sicilian wolf always joined a basal phylogenetic clade that included the Italian wolf and two mtDNA genomes respectively sequenced from a wolf sampled in Poland (KF661045.1) and in Belarus (KU696390.1). This clade (hereafter named the Italian clade) was basal to all the other modern wolf and dog genomes, including the Japanese wolf clade (*C. l. hodophilax*; Matsumura et al. 2014). The Sicilian wolf genome was basal to the peninsular Italian wolf mtDNAs. Bootstrap and interior-branch length tests showed that the Italian clade was 99% - 100% supported.

The Bayesian consensus phylogenetic trees (Fig. 5) obtained analysing the Set#3 mtDNA genomes fully supported the MEGAX results. The Japanese and Italian wolf clades were basal to all the other modern wolf and dog genomes. The average coalescence time of the Italian clade, as estimated in BEAST following Matsumura et al. (2014), was MRCA = 13,400 years (with 95% highest probability density HPD = 4000 – 21,230; Fig. 6). The divergence times of the Italian wolf and Japanese wolf from the other wolf clades were similar (*c*. 100,000 years), further highlighting the ancient origins of wolves in peninsular Italy and Sicily in comparison to the other wolves and dogs worldwide. The Italian and Sicilian wolf mtDNA haplotypes belong to the wolf haplogroup-2, that includes all the ancient wolves sampled in western Europe dating from between 44,000 and 1200 years ago (Pilot et al. 2010).

The mtDNA CR of the Sicilian wolf was blasted in GenBank to search for eventual matching with domestic dog CRs. The CR of the Sicilian wolf did not match with any of the dog sequences known so far, thus supporting its origin in a wild wolf populations. The sequence was identical to a partial mtDNA CR sequence found by Dufresnes et al. (2018) in a different wolf sample from Sicily (their haplotype H3 from sample AN855; Museo di Zoologia P. Doderlein, Palermo). Thus, this mtDNA genomes was apparently unique for the extinct wolf population of Sicily.

We analyzed the stored genomic DNA reads to search for specific sex markers. We identified SRY sequences matching with the homologous *Canis lupus* chromosome Y genomic sequences present in GenBank (AF107021.1), thus indicating unequivocally that the studied specimen was a male.

Discussion

In this study we obtained for the first time a complete mtDNA genome of a Sicilian wolf by NGS technologies. This wolf was likely killed on Mt. San Calogero, near the city of Termini Imerese, in the last years of the nineteenth century, very near to the extinction of the island population, for the last documented wolf was killed in 1935. The control-region of this mtDNA is identical to a partial CR sequenced from a different Sicilian wolf sample (Dufresnes et al. 2018), and is apparently very closely related to a partial CR sequence obtained from another Sicilian wolf specimen, as mentioned by Angelici et al. (2018; although at the moment this sequence is not stored in GenBank). These results suggest that, during the last few decades before the extinction, the wolf population of Sicily showed (at least) two distinct but very closely related mtDNA haplotypes. Phylogenetic analyses also indicate that the mtDNA genome of the Sicilian wolf is closely related to the predominant mtDNA genome of the past and extant wolf population in peninsular Italy (Dufresnes et al. 2018; Randi et al. 2014). The Sicilian and Italian wolf mtDNAs join in a strongly supported clade (the Italian clade) which includes also two mtDNAs sequenced from two wolves sampled in Poland and in Belarus, respectively. The Italian clade is basal to all the other modern wolf and dog haplogroups sequenced so far, with the exception of most of the ancestral sequences obtained from historical wolves (Thalmann et al. 2016; Koblmuller et al. 2016), and from the now extinct Japanese wolves (*C. l. hodophilax*; Matsumura et al. 2014). The origin and fate of the Japanese wolves has been described Matsumura et al. (2014). Both the Japanese and Italian wolf clades, which apparently split *c*. 130,000 – 100,000 years from all the other modern wolf haplogroups worldwide, belong to the mtDNA haplogroup-2 (Pilot et al. 2010). This haplogroup has been detected in the ancient western European wolf population that were largely substituted by the recent spread of modern wolves, which showed the more recent mtDNA haplogroup-1. However, the mtDNA genomes clearly indicate that both extant wolves in Italy and extinct wolves in Sicily are by far more recent than Himalayan wolves, formerly considered a subspecies of *C. lupus* and named *C. l. lanigero*r *C. l. chanco*, but now ranked as distinct species *C. himalayensis* (Aggarwal et al. 2007). They diverged *c*. 550,000 (95% HDP = 495,100–605,600) years ago (Matsumura et al. 2014), and predated the evolutionary radiation of Eurasian and New World *C. lupus*.

In this study we used Matsumura et al. (2014) estimates of Japanese wolf mtDNA divergence time to compute a MRCA = 13,400 years (95% HPD = 4000 – 21,230) of the Italian wolf

clade. The mtDNA genome of the Sicilian wolf is basal to the Italian wolf clade, thus a MRCA = 13,400 years can be considered as an approximate estimate of island-mainland DNA divergence. Although phylogenetic relationships and divergence time estimates obtained from complete mtDNA genomes should, in principle, outperform estimates obtained using only shorter sequences, they should anyway be used with caution. First, the mtDNA is a maternal haploid genome informative only on single-gene relationships and not on population-species phylogeny.Then, the sample size used in our and other studies (e.g., Angelici et al. 2018) are by far too small to exclude uncertainty. The wolf population in Sicily is extinct, the available museum specimens are few and perhaps not always suitable for genomic studies, thus the sample size of the target population could not be much expanded. However, we believe that the addition of complete mtDNA genomes from other haplogroup-2 wolf populations could improve the phylogenetic structure and connections of the Italian wolf clade, allowing more reliable estimates of divergence times. Moreover, sequences form chromosomal genes could contribute to better describe the phylogeographic history of the wolf population in Sicily.

The mtDNA CRs of Sicilian wolves are distinct from homologous CR sequences of historical Italian wolves obtained by Dufresnes et al. (2018). Wolves in peninsular Italy were certainly abundant a few centuries ago, but the museum specimens suitable to DNA sequencing are too few to conclude that the Sicilian haplotypes were absent in the historical mainland population. Hence, we cannot exclude that the Sicilian mtDNA haplotypes evolved in peninsular Italy. However, based on the available data, the most parsimonious hypothesis is that those haplotypes evolved in Sicily following wolf island colonization. The divergence time of 13,400 years (95% HPD = 4000 – 21,230) of the Sicilian wolf mtDNA is compatible with the age of the last land bridge between the island and the south-western tip of Italy, that is 21,500 – 20,000 (Antonioli et al 2012). A late Pleistocene colonization of peninsular wolves before the Messina strait was definitely flooded does not exclude earlier colonization waves, which, seems, nevertheless undocumented by the available Sicilian wolf specimens.

During late Holocene a number of species, and in particular ungulates (red deer, fallow deer, roe deer, wildboar), the natural prey of wolves, went extinct in Sicily like due to anthropogenic pressures (La Mantia and Cannella, 2008). The concomitant consequences of habitat transformations, ungulate decline and overhunting most probably pushed the wolf population of Sicily to decline and finally disappearing. The few available stuffed specimens evidence smaller body size and paler coat colours of the last wolves in Sicily in comparison to the

Italian wolves. Moreover wolves in Sicily did not show the darker fur strip on the forearms, a peculiar morphological trait of the peninsular Italian wolf population (Altobello 1921; Ciucci and Boitani 2003). Dwarfism and local phenotypic adaptations are typical of some island vertebrate populations. Moreover, we cannot exclude that during the final population bottleneck wolves in Sicily crossbred and hybridized with free-ranging dogs, perhaps accelerating the speed of the extinction vortex (see: Gómez-Sánchez et al. 2018). Future genomic data set and analyses could perhaps shed more light on the extent of homozygosity and eventual domestic dog introgression in the lost populations of wolves in Sicily.

REFERENCES

Aggarwal R. K., Kivisild T., Ramadevi J., Singh L., 2007. Mitochondrial DNA coding region sequences support the phylogenetic distinction of two Indian wolf species. J. Zool. Syst. Evol. Res. 45: 163–172.

Altobello G.,1921. *Mammiferi. IV. I Carnivori (Carnivora). Fauna dell'Abruzzo e del Molise*. Colitti, Campobasso, 61 pp.

Altschul S. F., Gish W., Miller W., Myer, E. W., Lipman D. J., 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

Angelici F. M., Rossi L., 2018. A new subspecies of gray wolf, recently extinct, from Sicily, Italy (Carnivora, *Canidae*). BioRxiv preprint first posted online May, 11, 2018; doi: http://dx.doi.org/10.1101/320655.

Angelici F. M., Ciucani M. M., Angelini S., Annesi F., Caniglia R, Castiglia R., Fabbri E., Galaverni M., Palumbo D., Ravegnini G., Rossi L., Siracusa A. M., Cilli E., 2018. The Sicilian wolf: Genetic identity of a recently extinct insular population. Bioxriv preprint not peer-reviewed; doi: https://doi.org/10.1101/453365.

Antonioli F., Lo Presti V., Morticelli M. G., Bonfiglio L., Mannino M. A., Palombo M. R., Tonielli R., 2014. Timing of the emergence of the Europe–Sicily bridge (40–17 cal ka BP) and its implications for the spread of modern humans. Geological Society, London, Special Publications 411(1): 111–144; https://doi.org/10.1144/sp411.1.

Anzidei A. P., Bulgarelli G. M., Catalano P., Cerilli E., Gallotti R., Lemorini C., Milli S., Palombo M. R., Pantano W., Santucci E., 2012. Ongoing research at the late Middle Pleistocene site of La Polledrara di

Cecanibbio (central Italy), with emphasis on human–elephant relationships. Quaternary International 255: 171–187.

Bankevich A., Nurk S., Antipov D., Gurevich A. A., Dvorkin M., Kulikov A. S., Lesin V. M., Nikolenko S. I., Pham S., Prjibelski A. D., Pyshkin A. V., Sirotkin A. V., Vyahhi N., Tesler G., Alekseyev M. A., Pevzner P. A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput. Biol. 19(5): 455-77.

Bertè F., Pandolfi L., 2014. *Canis lupus* (Mammalia, *Canidae*) from the late pleistocene deposit of Avetrana (Taranto, Southern Italy). Rivista Italiana di Paleontologia e Stratigrafia 120(3); DOI: 10.13130/2039-4942/6079.

Bjornerfeldt S., Webster M. T., Vilà C., 2006. Relaxation of selective constraint on dog mitochondrial DNA following domestication. Genome Research 16: 990–994.

Boitani L., 1984. Genetic considerations on wolf conservation in Italy. Boll. Zool. 51: 367-373.

Ciucci P., Boitani L., 2003. *Canis lupus*, Linnaeus, 1758. *In*: Boitani, L., Lovari, S. & Vigna Taglianti, A. (Eds.). *Fauna d'Italia. Mammalia III. Carnivora – Artiodactyla*. Calderini, Bologna, pp. 20-47.

Drummond A. J., Rambaut A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7: 214.

Drummond A. J., Ho S. Y., Phillips M. J., Rambaut A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4: e88.

Drummond A. J., Suchard M. A., Xie D., Rambaut A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29: 1969–1973.

Dufresnes C., Miquel C., Remollino N., Biollaz F., Salamin N., Taberlet P., Fumagalli L., 2018. Howling from the past: historical phylogeography and diversity losses in European grey wolves. Proc. Biol. Sci. 285(1884). doi: 10.1098/rspb.2018.1148.

Gill M. S., Lemey P., Nuno R., Faria N. R., Rambaut A., Shapiro B., Suchard M. A., 2013. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol. Biol. Evol. 30: 713–724.

Gómez-Sánchez D., Olalde I., Sastre N., Enseñat C., Carrasco R., Marques-Bonet T., Lalueza-Fox C., Leonard J. A., Vilà C., Ramírez O., 2018. On the path to extinction: inbreeding and admixture in a declining grey wolf population. Molecular Ecology 27: 3599–3612; DOI: 10.1111/mec.14824.

Gurevich A., Saveliev V., Vyahhi N., Tesler G., 2013. QUAST: quality assessment tool for genome assemblies, Bioinformatics 29 (8): 1072-1075; doi: 10.1093/bioinformatics/btt086.

Hasegawa M., Kishino H., Yano T., 1985. Dating of human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22 (2): 160–174.

Higgins D., Thompson J., Gibson T. Thompson J. D., Higgins D. G., Gibson T. J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

Ishiguro N., Inoshim, Y., Shigehara N., 2009. Mitochondrial DNA analysis of the Japanese wolf (*Canis lupus hodophilax* Temminck, 1839) and comparison with representative wolf and domestic dog haplotypes. Zool. Sci. 26: 765–770; https://doi.org/10.2108/zsj.26.765.

Jónsson H., Ginolhac A., Schubert M., Johnson P., Orlando L., 2013. MapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29 (13): 1682–1684; https://doi.org/10.1093/bioinformatics/btt193.

Kim K. S., Lee S. E., Jeong H. W., Ha J. H., 1998. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. Mol. Phylogenet. Evol. 10: 210–220.

Koblmuller S., Vilà C., Lorente-Galdos B., Dabad M., Ramirez O., Marques-Bonet T., Wayne R. K., Leonard J. A., 2016. Whole mitochondrial genomes illuminate ancient intercontinental dispersals of grey wolves (*Canis lupus*) Journal of Biogeography 43: 1728–1738; doi:10.1111/jbi.12765.

Kumar S., Stecher G., Li M., Knyaz C., Tamura K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35: 1547-1549.

La Mantia T., Cannella Z., 2008. Presenza storica dei grossi Mammiferi in Sicilia. (pp. 87-112). In: *Atlante della Biodiversità della Sicilia: Vertebrati terrestri*. Studi e Ricerche 6, Arpa Sicilia, Palermo, 536 pp.

Langmead B, Salzberg S., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357-359.

Leonard J., Vilà C., Wayne R.K., 2005. Legacy lost: genetic variability and population size of extirpated US grey wolves (*Canis lupus*) Molecular Ecology 14 (1): 9-17. https://doi.org/10.1111/j.1365-294X.2004.02389.x.

Li H., Durbin R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26 (5): 589–595; https://doi.org/10.1093/bioinformatics/btp698.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., 1000 Genome Project Data Processing Subgroup (2009). The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25: 2078-9. [PMID: 19505943].

Minà Palumbo F., 1858. Storia naturale delle Madonie. Catalogo dei Mammiferi. *La Scienza e la Letteratura*, A. I, 3: 154-170.

Li H., Xiang-Yub J., Daib G., Gub Z., Ming C., Yang Z., Ryder O. A., Li W.-H., Fu Y.-X., Zhang Y.-P., 2016. Large numbers of vertebrates began rapid population decline in the late 19th century. PNAS 113 (49): 14079–14084; doi/10.1073/pnas.1616804113.

Linnell J., Salvatori V., Boitani L., 2008. Guidelines for population level management plans for large carnivores in Europe. A Large Carnivore Initiative for Europe report prepared for the European Commission (contract 070501/2005/424162/MAR/B2).

Loog L., Thalmann O., Mikkel-Holger S., Sinding M. H. S., Schuenemann V. J., Perri A., Germonpré M., Bocherens H., Witt K. E., Samaniego Castruita J. A., Velasco M. S., Lundstrøm I. K. C., Wales N., Sonet G., Frantz L., Schroeder H. , Budd J., Jimenez E. L., Fedorov S., Gasparyan B., W. Kandel A. W., Lázničková-Galetová M., Napierala H., Uerpmann H. P., Nikolskiy P. A., Pavlova E. Y., Pitulko V. V., Herzig K. H., Malhi R. S., Willerslev E., Hansen A. J., Dobney K., Gilbert M. T. P., Krause J., Larson G., Eriksson A., Manica A., 2018. Modern wolves trace their origin to a late Pleistocene expansion from Beringia. BioRxiv Posted July 18, 2018 doi: https://doi.org/10.1101/370122. This article is a preprint and has not been peer-reviewed.

Lucchini A., Galov E., Randi E., 2004. Evidence of genetic distinction and long-term population decline in wolves (*Canis lupus*) in the Italian Apennines. Molecular Ecology 13 (3): 523-536; https://doi.org/10.1046/j.1365-294X.2004.02077.x.

499 Martin M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMB
500 net. journal 17: 10–12.

501

502 Matsumura S., Inoshima Y., Ishiguro N., 2014. Reconstructing the colonization history of lost wolf
503 lineages by the analysis of the mitochondrial genome Molecular Phylogenetics and Evolution 80: 105–
504 112; http://dx.doi.org/10.1016/j.ympev.2014.08.004.

505

506 Minà Palumbo, F., 1858. Storia naturale delle Madonie. Osservazioni sopra i Mammiferi. *La Scienza e la*
507 *Letteratura*, A. I, 4: 5-14.

508

509 Minà Palumbo, F., 1868. Catalogo dei mammiferi della Sicilia. *Annali di Agricoltura Siciliana*, 12 (serie 2):
510 3-123.
511

512 Nei M., Kumar S. 2000. *Molecular Evolution and Phylogenetics.* New York. Oxford University Press.

513

514 Pang J., Kluetsch C., Zou X. J., Zhang A. B., Luo L. Y., Angleby H., Ardalan A., Ekstrom C., Skollermo A.,
515 Lundeberg J., Matsumura S., Leitner T., Zhang Y. P., Savolainen P., 2009. MtDNA data indicate a single
516 origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. Mol. Biol.
517 Evol. 26(12): 2849-64; doi: 10.1093/molbev/msp195.

518

519 Pilot M., Greco C., vonHoldt B. M., Jędrzejewska B., Randi E., Jędrzejewski W., Sidorovich V. E.,
520 Ostrander E. A., Wayne R. K., 2014. Genome-wide signatures of population bottlenecks and diversifying
521 selection in European wolves. Heredity 112(4): 428–442; doi: 10.1038/hdy.2013.122.

522

523 Pilot M., Branicki W., Jedrzejewsk W., Goszczynski J., Jedrzejewska B., Dykyy I., Shkvyrya M., Tsingarska
524 E., 2010. Phylogeographic history of grey wolves in Europe. BMC Evol. Biol. 10, 104;
525 doi.org/10.1186/1471-2148-10-104.

526

527 Pratesi F., 1978. Esclusi dall'arca: animali estinti e in via di estinzione in Italia. Arnoldo Mondadori
528 Editore Spa, Milano. 32-43.

529

530 Randi E., 2011. Genetics and conservation of wolves *Canis lupus* in Europe. Mammal Rev. **41**(2)**:** 99-111;
531 https://doi.org/10.1111/j.1365-2907.2010.00176.x.
532

Randi E., Hulva P., Fabbri E., Galaverni M., Galov A., Kusak J., Bigi D., Černá Bolfíková B., Smetanová M., Caniglia R., 2014. Multilocus Detection of Wolf *x* Dog Hybridization in Italy, and Guidelines for Marker Selection. PlosOne January 22, 2014; https://doi.org/10.1371/journal.pone.0086409.

Saitou N., Nei M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406–425.

Schubert M., Lindgreen S., Orlando L., 2016 AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Research Notes 9(1):88; doi: 10.1186/s13104-016-1900-2.

Sotnikova M., L. Rook., 2010. Dispersal of the Canini (Mammalia, *Canidae: Caninae*) across Eurasia during the Late Miocene to Early Pleistocene. Quaternary International 212(2): 86-97.

Stöck M., Sicilia A., Belfiore N. M., Buckley D., Lo Brutto S., Lo Valvo M., Arculeo M., 2008. Post-Messinian evolutionary relationships across the Sicilian channel: Mitochondrial and nuclear markers link a new green toad from Sicily to African relatives. BMC Evolutionary Biology 8:56; doi: 10.1186/1471-2148-8-562008.

Tamura K., Nei M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10: 512–526.

Thalmann O. B., Shapiro B., Cui P., Schuenemann V. J., Sawyer S. K., Greenfield D. L., Germonprè M. B., Sablin M. V., Lopez-Giraldez F., Domingo-Roura X., Napierala H., Uerpmann H.-P., Loponte DM., Acosta A. A., Giemsch L., Schmitz R. W., Worthington B., Buikstra J. E., Druzhkova A., Graphodatsky A. S., Ovodov N. D., Wahlberg N., Freedman A. H., Schweizer R. M., Koepfli K.-P., Leonad J. A., Meyer M., Krause J., Paabo S., Green R. E., Wayne R. K., 2013. Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. Science 342: 871–874.

WWF 2018. Living Planet Report. Grooten M., Almond R. E. A. (Eds). WWF, Gland, Switzerland.

Zhang J., Kobert K., Flouri T., Stamatakis A., 2014. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30: 614–620.

Acknowledgements

We gratefully thank the Civic Museum "Baldassare Romano" (Termini Imerese) and dr. Fabio Lo Bono, curator of Mammals.

Figure Legends

Figure 1. A picture of the sampled and genotyped Sicilian wolf from the Civic Museum "Baldassare Romano", Termini Imerese (Palermo; Italy).

Figure 2. (a) TapeStation spectrophotometer analysis of the extracted DNA from the Sicilian wolf sample. On the left the ladder and the sample run on gel visualization. On the right the fragment size composition of the extracted DNA, spanning from 40 to 1040 bp, with a 287 bp average size at 600 pg/μl concentration. (b) TapeStation analysis of the Sicilian wolf DNA after PCR-amplification. On the left the ladder and the sample run on gel visualization. On the right the fragment size composition of the amplified DNA spanning from 156 to 655 bp, with a 295 bp average size at 2490 pg/μl concentration.

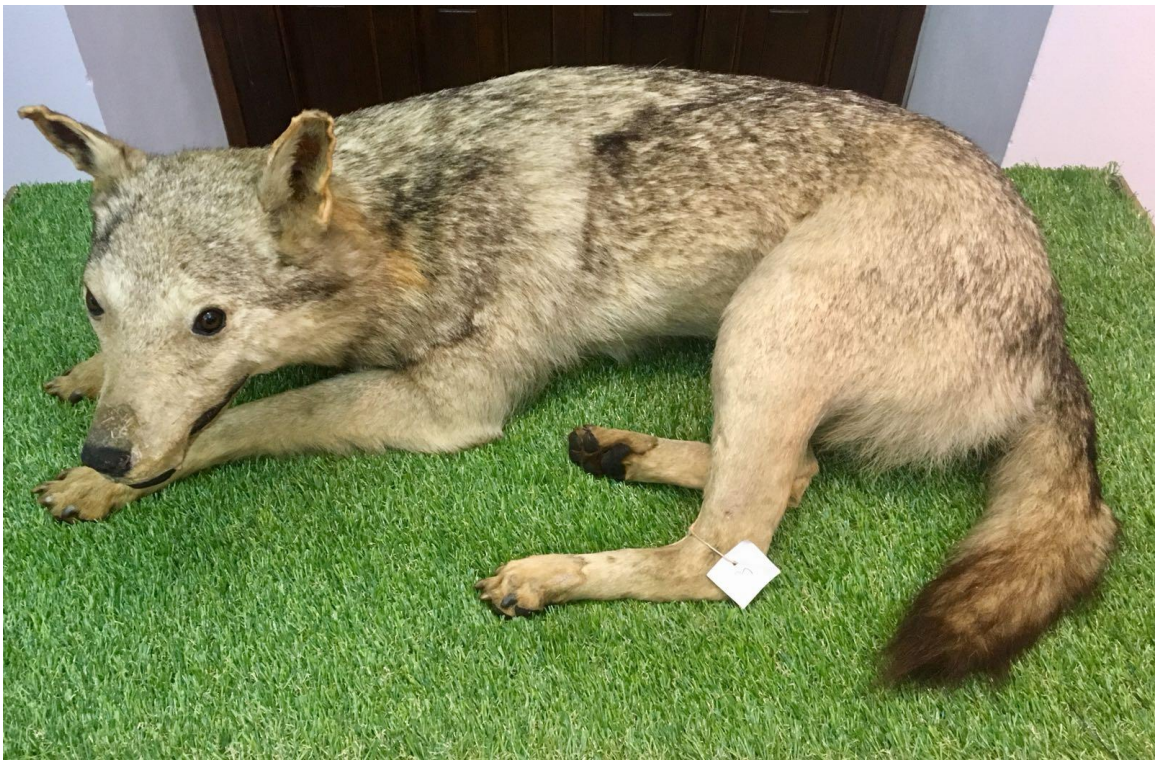
Figure 3. Neighbor-joining tree of complete wolf mtDNA genomes used by Koblmuller et al. (2016), including three historical wolf samples, 14 dog (*C. l. familiaris*) genomes and the new Sicilian wolf mtDNA genome. Five Himalayan wolf (here named *C. l. laniger* or *C. l. chanco*) mtDNAs were used as outgroups. The Italian clade is indicated. Bootstrap values at the internodes.

Figure 4. Neighbor-joining tree of modern wolf mtDNA genomes (control-region excluded) used by Thalmann et al. (2013), and the new Sicilian wolf mtDNA genome. Four coyote (*C. latrans*) mtDNAs were used as outgroups. The Italian clade is indicated. Bootstrap values at the internodes.

Figure 5. Consensus Bayesian phylogenetic tree computed by BEAST 2.5.1 (Drummond et al. 2012) with the HKY+G model (Hasegawa et al. 1985). We used a subset of wolf mtDNA genomes (control-region excluded) published by Pang et al. (2009) and Matsumura et al. (2014), including two historical Ezo wolf (*C. l. hattai*), five historical Honsu wolf (*C. l. hodophilax*) and the new Sicilian wolf mtDNA genome. Four coyote mtDNAs were used as outgroups. The Markov chain Monte Carlo samples were drawn every 1000 generations from a total of 1.000.000 generations, following a discarded burn-in of 100.000 generations.

Figure 6. Age estimates (indicated by bar lengths) of the nodes of the consensus phylogenetic trees computed by BEAST 2.5.1 with the HKY+G nucleotide substitution model.

606 Figure 1.

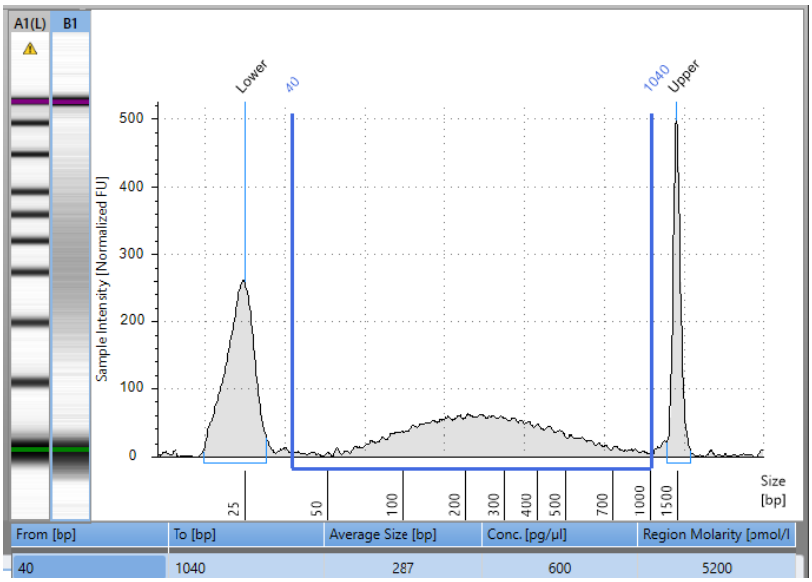


607

608

609

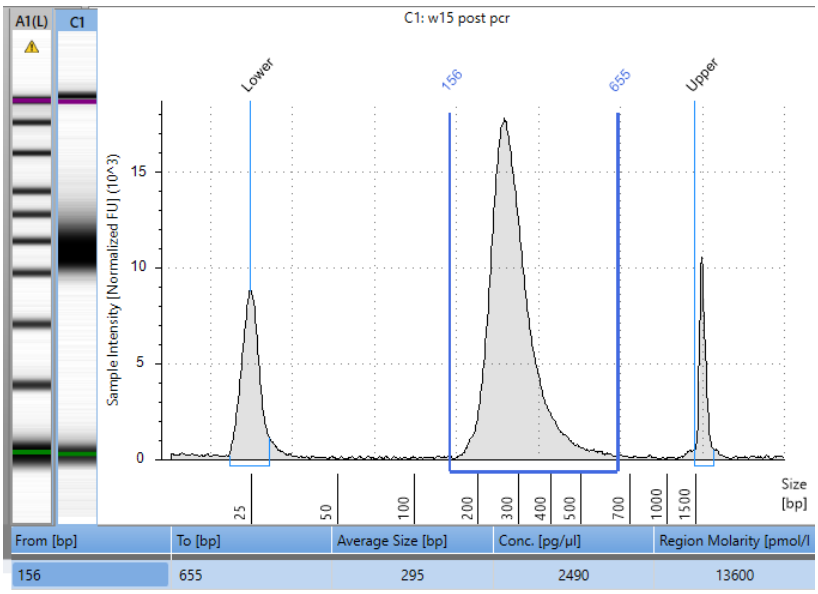
610 Figure 2a.



611

612

613 Figure 2b.



614

615

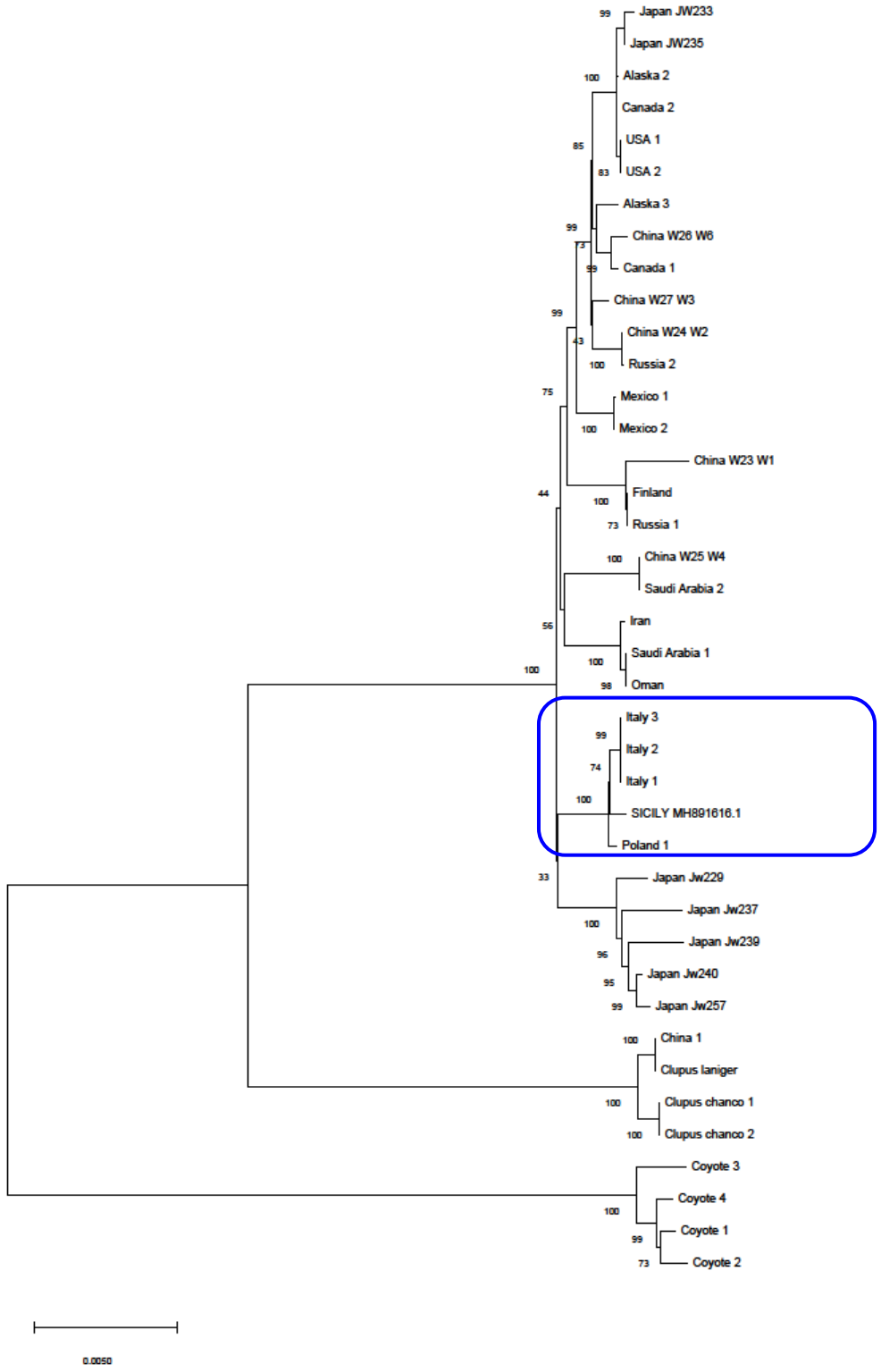
616

617 Figure 3.

618

619

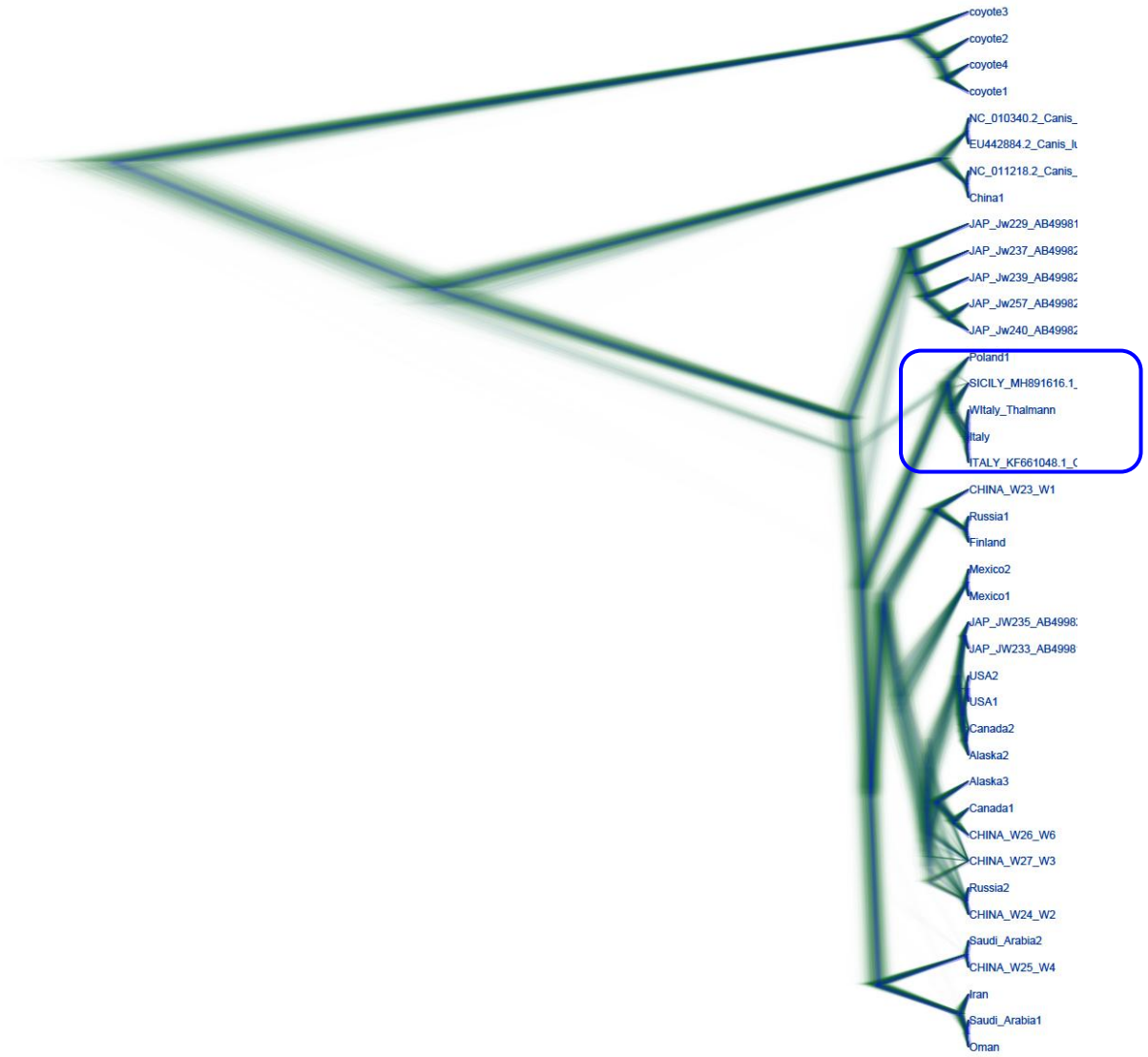
620 Figure 4.



621

622

623 Figure 5.

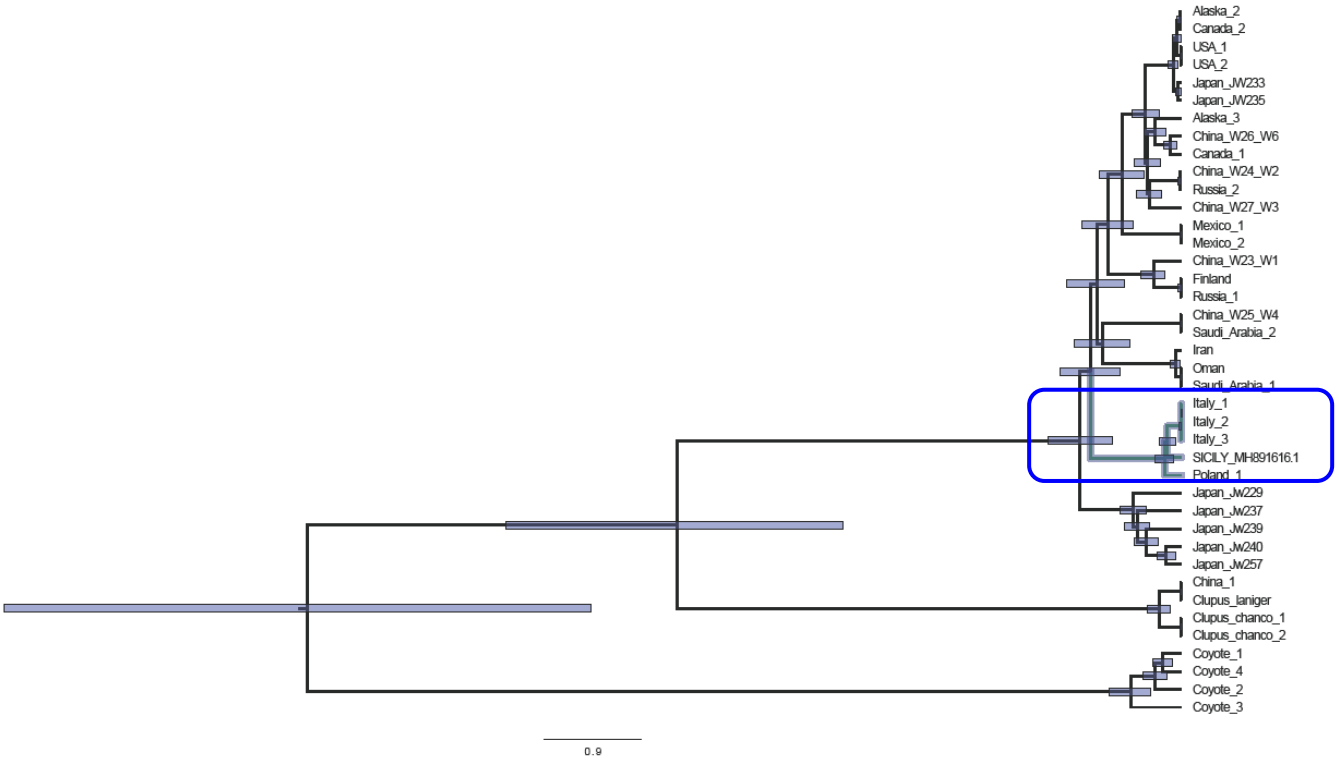


624

625

626

627 Figure 6.



628
629