Molecular Biology and Evolution, 2025, **42**, 1–17
https://doi.org/10.1093/molbev/msaf031
Advance access publication 5 February 2025
Discoveries

MBE OXFORD

Red Deer Resequencing Reveals the Importance of Sex Chromosomes for Reconstructing Late Quaternary Events

Menno J. de Jong [1,2,3]*, **Gabriel Anaya** [4], **Aidin Niamir** [1], **Javier Pérez-González** [5], **Camilla Broggini** [4], **Alberto Membrillo del Pozo** [4], **Marcel Nebenfuehr** [1,2,3], **Eva de la Peña** [4,6], **Jordi Ruiz-Olmo** [4], **Jose Manuel Seoane** [4], **Giovanni Vedel** [4], **Aurelie Barboiron** [7], **Luděk Bartoš** [8,9], **Elena Buzan** [10,11], **Ruth F. Carden** [12], **Giorgi Darchiashvili** [13], **Alain C. Frantz** [14], **Dragan Gačić** [15], **Adrien Gérard** [16], **Araceli Gort-Esteve** [17], **Etienne Guillaumat** [16], **Anja Hantschmann** [18], **Mahmoud-Reza Hemami** [19], **Jacob Höglund** [20], **Joost F. de Jong** [21], **Nikoleta Karaiskou** [22], **Niko Kerdikoshvili** [23], **Christian Kern** [18], **Dean Konjevic** [24], **Petr Koubek** [25], **Jarmila Krojerová-Prokešová** [25,26], **Allan D. McDevitt** [27], **Stefan Merker** [28], **Maryline Pellerin** [7], **Markus Pfenninger** [1,3], **Knut H. Røed** [29], **Christine Saint-Andrieux** [30], **Fatih Sarigol** [31], **Maciej Sykut** [32,33,34], **Alexandros Triantafyllidis** [22], **Josephine Pemberton** [35], **Urmas Saarma** [36], **Laura Iacolina** [10,37,38], **Magdalena Niedziałkowska** [32], **Frank E. Zachos** [39,40,41,42], **Juan Carranza** [4,†], **Axel Janke** [1,2,3,†]

[1]Biodiversity and Climate Research Center, Senckenberg Institute, Frankfurt am Main 60325, Germany
[2]Institute for Ecology, Evolution and Diversity, Goethe University, Frankfurt am Main, Germany
[3]LOEWE-Centre for Translational Biodiversity Genomics (TBG), Senckenberg Nature Research Society, Frankfurt am Main, Germany
[4]Wildlife Research Unit (UIRCP), University of Cordoba, Cordoba, Spain
[5]Biology and Ethology Unit, Faculty of Veterinary, University of Extremadura, Caceres, Spain
[6]Instituto de Investigación en Recursos Cinegéticos, IREC (CSIC, UCLM, JCCM), Ciudad Real, Spain
[7]Office Français de la Biodiversité, Direction de la Recherche et de l'Appui Scientifique, Service Conservation et Gestion Durable des Espèces Exploités, France
[8]Department of Ethology, Institute of Animal Science, Czech University of Life Sciences, 10400 Prague 10-Uhrineves, Czechia
[9]Department of Game Management and Wildlife Biology, Faculty of Forestry and Wood Sciences, University of Life Sciences, 16500 Prague 6-Suchdol, Czechia
[10]Department of Biodiversity, Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper 6000, Slovenia
[11]Faculty of Environmental Protection, Velenje 3320, Slovenia
[12]School of Archaeology, University College Dublin, Dublin 4, Ireland
[13]Tbilisi Zoo, Tbilisi 0171, Georgia
[14]Scientific Research Center, Musee National d'Histoire Naturelle, Luxembourg L-2160, Luxembourg
[15]Faculty of Forestry, University of Belgrade, Belgrade 11000, Serbia
[16]Domaine National de Chambord, Chambord, France
[17]Department of Animal and Food Science, Universitat Autònoma de Barcelona, Barcelona 08193, Spain
[18]Tierpark Berlin-Friedrichsfelde GmbH, Berlin 10319, Germany
[19]Department of Natural Resources, Isfahan University of Technology, Isfahan, Iran
[20]Dept. of Ecology and Genetics, Uppsala University, Uppsala SE-75236, Sweden
[21]Wildlife Ecology & Conservation Group, Wageningen University, Wageningen PB 6708, The Netherlands
[22]Department of Genetics, Developmental and Molecular Biology, School of Biology, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece
[23]Institute of Ecology, Ilia State University, Tbilisi 0162, Georgia
[24]The Faculty of Veterinary Medicine, University of Zagreb, Zagreb, Croatia
[25]Institute of Vertebrate Biology, Czech Academy of Sciences, Brno 603 00, Czech Republic
[26]Department of Zoology, Fisheries, Hydrobiology and Apiculture, Faculty of AgriSciences, Mendel University in Brno, Brno 613 00, Czech Republic
[27]Marine and Freshwater Research Centre, Department of Natural Resources and the Environment, Atlantic Technological University, Galway, Ireland
[28]Department of Zoology, State Museum of Natural History Stuttgart, Stuttgart 70191, Germany

Received: May 2, 2024. **Revised:** November 21, 2024. **Accepted:** January 13, 2025

[29]Department of Preclinical Sciences and Pathology, Norwegian University of Life Sciences, P.O. Box 5003, Ås 1432, Norway
[30]Office Français de la Biodiversité, Direction de la Recherche et de l'Appui Scientifique, Service Anthropisation et fonctionnement des écosystèmes terrestres, France
[31]Max Perutz Labs, Vienna BioCentre, Vienna 1030, Austria
[32]Mammal Research Institute, Polish Academy of Sciences, Białowieża 17-230, Poland
[33]Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), Department of Biology, Aarhus University, Aarhus C 8000, Denmark
[34]Department of Archaeology and Heritage Studies, Aarhus University, Højbjerg DK-8270, Denmark
[35]Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, Edinburgh EH93FL, UK
[36]Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Tartu 50409, Estonia
[37]Department of Veterinary Medicine, University of Sassari, Sassari 07100, Italy
[38]Department of Chemistry and Bioscience, Aalborg University, Aalborg 9220, Denmark
[39]Department of Mammal Collection, Natural History Museum Vienna, Vienna, Austria
[40]Department of Genetics, University of the Free State, Bloemfontein, South Africa
[41]Department of Evolutionary Biology, University of Vienna, Vienna, Austria
[42]Research Institute for the Environment and Livelihoods, Charles Darwin University, Casuarina, NT, Australia

[†]Same contribution as group leadership.
***Corresponding author:** E-mail: menno.de-jong@senckenberg.de.
Associate editor: Xuming Zhou

Abstract

Sex chromosomes differ in their inheritance properties from autosomes and hence may encode complementary information about past demographic events. We compiled and analyzed a range-wide resequencing data set of the red deer (*Cervus elaphus*), one of the few Eurasian herbivores of the Late Pleistocene megafauna still found throughout much of its historic range. Our analyses of 144 whole genomes reveal striking discrepancies between the population clusters suggested by autosomal and X-chromosomal data. We postulate that the genetic legacy of Late Glacial population structure is better captured and preserved by the X chromosome than by autosomes, for two reasons. First, X chromosomes have a lower N_e and hence lose genetic variation faster during isolation in glacial refugia, causing increased population differentiation. Second, following postglacial recolonization and secondary contact, immigrant males pass on their X chromosomes to female offspring only, which effectively halves the migration rate when gene flow is male mediated. Our study illustrates how a comparison between autosomal and sex chromosomal phylogeographic signals unravels past demographic processes that otherwise would remain hidden.

Keywords: population genomics, phylogeography, postglacial recolonization

Introduction

It is widely accepted that the present-day genetic structure of temperate Palearctic species has been shaped by climate-driven range contractions and expansions (Hewitt 1996; Bhagwat and Willis 2008; Shi and Chen 2012; Bertl et al. 2018). While this legacy of Quaternary Ice Ages is for many species indeed evident from mitochondrial genomes (Hewitt 1999), population-genetic theory predicts, for two reasons, that this legacy is less obvious from nuclear genomes. First, the effective population size (N_e) of nuclear loci is up to four times higher than that of the mitogenome, which restrains population stratification during isolation in glacial refugia (Zink 2010). Second, following postglacial recolonization and subsequent secondary contact, the inheritance properties of nuclear loci facilitate the introgression of nonnative DNA through gene flow and recombination. As a result, the genetic information stored in nuclear genomes will eventually reflect a new migration–drift equilibrium, obscuring former population structures (Eckert and Carstens 2008; de Jong et al. 2023).

For few species, the phylogeographic effects of Quaternary glacial cycles have been studied in such detail as for the red deer (*Cervus elaphus*). Large-scale mtDNA profiling has produced detailed maps depicting the geographical distribution of mtDNA haplotypes through time and space (Ludt et al. 2004; Sommer and Nadachowski 2006; Sommer et al. 2008; Pérez-Espona et al. 2009; Skog et al. 2009; Sommer and Zachos 2009; Niedziałkowska et al. 2011; Carden et al. 2012; Meiri et al. 2017; Doan et al. 2022; Mackiewicz et al. 2022). Meanwhile, the construction of a rich, calibrated fossil database uncovered a time-lapse picture of oscillating range boundaries throughout the Late Pleistocene and Holocene (Sommer et al. 2008; Carden et al. 2012; Niedziałkowska et al. 2021). These findings are corroborated by morphological and behavioral studies of present-day phenotypic variation (Frey et al. 2012; Passilongo et al. 2013; Della Libera et al. 2015; Volodin et al. 2019). Through the synthesis of the fossil, genetic, and phenotypic evidence, well-supported inferences can be drawn regarding putative postglacial migration routes.

At present, a prominent feature of red deer population structure is a chain of mtDNA discontinuities which bisects the species range into a western European and an eastern European clade. This elongated contact zone meanders from the Alps through Austria and Czechia into Poland, Belarus, and the Baltic States (Ludt et al. 2004; Skog et al. 2009; Krojerová-Prokešová et al. 2015; Zachos et al. 2016; Doan et al. 2022). A side branch, which splits off in the direction of the Black Sea, demarcates a further subdivision between a northeastern and southeastern subclade (Doan et al. 2022; Valnisty et al. 2024). Studies of ancient-mtDNA indicate that the observed discontinuities represent secondary contact zones which arose following postglacial recolonization from the south (Doan et al. 2022). This northward expansion culminated into reestablished contact between populations which had survived the Last Glacial Maximum (LGM) and subsequent stadials, most recently the Younger Dryas, in isolated refugia surrounding the Mediterranean Sea, Black Sea and Caspian Sea (Doan et al. 2017; Niedziałkowska et al. 2021; Doan et al. 2022).

Until now, these range-wide genetic inferences rely mostly on analyses of mtDNA data. To reassess red deer population structure using nuclear data, we compiled a data set of 144 nuclear genomes, carefully selected to faithfully represent native local populations and the various mtDNA haplogroups. Unlike previous multilocus studies of red deer (Hajji et al. 2007; Zachos et al. 2009; Fickel et al. 2012; Krojerová-

Prokešová et al. 2015; Carranza et al. 2016; Zachos et al. 2016; Queirós et al. 2019; de Jong et al. 2020; Carranza et al. 2024), our data set spans the entire species range. Moreover, unlike SNP data, the whole-genome data allow us to compare signals from different inheritance units, including sex chromosomes.

We set out to investigate to what extent the genetic legacy of population fragmentation in glacial refugia can still be found in the nuclear genomes of extant red deer. On the premises that X-chromosomal N_e is at most three-fourth of autosomal N_e and that gene flow is primarily male mediated, we hypothesized that the glacial population structure should be more evident from X chromosomes than from autosomes (Webster and Wilson Sayres 2016; de Jong et al. 2023). One part of the underlying reasoning is that the lower N_e of the X chromosome allows it to differentiate faster during isolation in glacial refugia (i.e. more genetic drift and hence less incomplete lineage sorting). The second part of the argument, which concerns secondary contact following postglacial recolonization, is that immigrant males pass on their X chromosomes to female offspring only, which effectively halves the migration rate in case of male-mediated gene flow. We evaluated this hypothesis, and the relative importance of the two underlying arguments, by comparing the genetic structure of red deer according to autosomes, sex chromosomes, and the mitogenome.

Materials and Methods

Sample Collection

This study focuses on the European red deer (*C. elaphus*), and not on the closely related species of Central Asian red deer (*Cervus hanglu*) or wapiti (*Cervus canadensis*). Throughout, we use the term "red deer" to refer to *C. elaphus* specifically.

We selected from our collections a total of 142 red deer samples for whole-genome resequencing, among them 3 zoo individuals and 139 free-ranging individuals (Fig. 1a; supplementary table S1, Supplementary Material online). From NCBI's SRA repository, we downloaded short-read data for two additional red deer individuals, as well as one Central Asian red deer (namely Yarkand deer, *Cervus hanglu yarkandensis*) and five North American wapiti (supplementary table S1, Supplementary Material online). We additionally sequenced two Scottish sika deer (*Cervus nippon*), selected based on their pureness as determined by a previous study (McFarlane et al. 2020).

Red deer samples were carefully chosen to represent native local populations and the various mtDNA haplogroups, using insights from previous studies. Red deer occurring in Sardinia, Corsica, and Barbary Coast (North Africa) are thought to have been introduced from the European mainland by Holocene seafarers (Vigne 1992; Doan et al. 2017, 2022). They can be considered "museum populations" as they represent the mtDNA lineage B of now extinct deer that inhabited the Italian peninsula in the Late Glacial (Doan et al. 2017).

Individuals from the Mesola Wood reserve in the Po Valley of northeastern Italy are thought to be the only surviving native population of the Italian mainland. Their mtDNA haplotype D used to be common in southeastern Europe (Doan et al. 2022), but is at present not found elsewhere, except sparsely in the Carpathian Mountains (Borowski et al. 2016). Throughout, we will use the term Mesola deer to refer to this particular population in northeastern Italy. In contrast, we use the term Italian deer to refer to the combined group of Mesola deer (*Cervus elaphus italicus*, haplogroup D), Sardinian deer (*Cervus elaphus corsicanus*, haplogroup B), and Barbary deer (*Cervus elaphus barbarus*, haplogroup B).

As confirmed by our study findings, red deer in the eastern Pyrenees were assumed to derive from a recent human-mediated reintroduction of red deer stemming from Chambord Castle in central France. Two individuals from the western Pyrenees clustered separately and hence were treated as a distinct, local population ("Pyrenees"). These individuals likely are related, owing to another translocation event, to deer in Cabañeros in central Spain, here referred to as "SpainMisc" ("MO" and "SM2" in Carranza et al. 2024). Based on assignment to the two major mtDNA haplogroups (Carranza et al. 2016), all other Iberian samples were divided in a western clade and a central clade ("SP1" and "SM1," respectively, in Carranza et al. 2024; supplementary table S1, Supplementary Material online).

No animals were killed specifically for this study. The red deer is listed as a species of Least Concern in the IUCN Red List and is not CITES listed, except for Barbary deer. Hair samples of female individuals of this subspecies, as well as a tissue sample of a deceased male, were obtained from Tierpark Berlin and sequenced at a sequencing facility in Germany.

DNA was isolated using standard extraction protocols, either from tissue or from hair samples. We sequenced all samples at minimum 10× coverage (supplementary fig. S1 and table S2, Supplementary Material online) using an Illumina NovaSeq 6000 (Novogene Cambridge and Novogene Munich) or a BGI DNBseq machine (BGI Hong Kong).

Read Mapping and Genotype Calling

Read quality check was performed using the software FastQC and MultiQC (Ewels et al. 2016). Reads with high adapter content were removed using the software AdapterRemoval v2 (Schubert et al. 2016). Reads were mapped, using the software BWA (Li and Durbin 2009), against a male Rocky Mountain wapiti mt-genome (CM033226.1) and nuclear genome with chromosome-level resolution (SAMN15814149) (Masonbrink et al. 2021). We opted against using a red deer genome assembly of similar quality (SAMEA7523520) because this assembly was generated from a female individual and hence lacked the Y chromosome. Samtools (Li et al. 2009) was used to remove reads with a mapping quality below 20 and/or alignment scores below 100, as well as reads that mapped discordantly. Read duplicates were removed using the software tool Picard MarkDuplicates (DePristo et al. 2011), with the option "REMOVE_DUPLICATES" set to TRUE.

Genotype likelihoods and calls were generated using the bcftools mpileup and call pipeline (Li 2011a). Genotype likelihoods were inferred (bcftools mpileup) only considering bases with a base quality of at least 13 (i.e. -min-BQ 13) and with a minimum mapping quality of 20 (–min-MQ 20) after correcting for excessive mismatches using the option -C50. When calling genotypes from genotype likelihoods (bcftools calls), the ploidy level was set according to genomic region (i.e. mtDNA, autosomal, and sex chromosome) and sample sex. For samples with missing sex information, the sex was inferred from Y-chromosomal mapping rates. Pseudoautosomal regions were identified based on deviations of sequencing depth relative to chromosome-wide means and subsequently excluded from the X-chromosomal and Y-chromosomal data sets. Indels were normalized and realigned using "bcftools norm-check-ref w."

When calling genotypes from genotype likelihoods (bcftools call), we used the "group-samples' option to assign each individual to its unique group, such that genotype calls were not

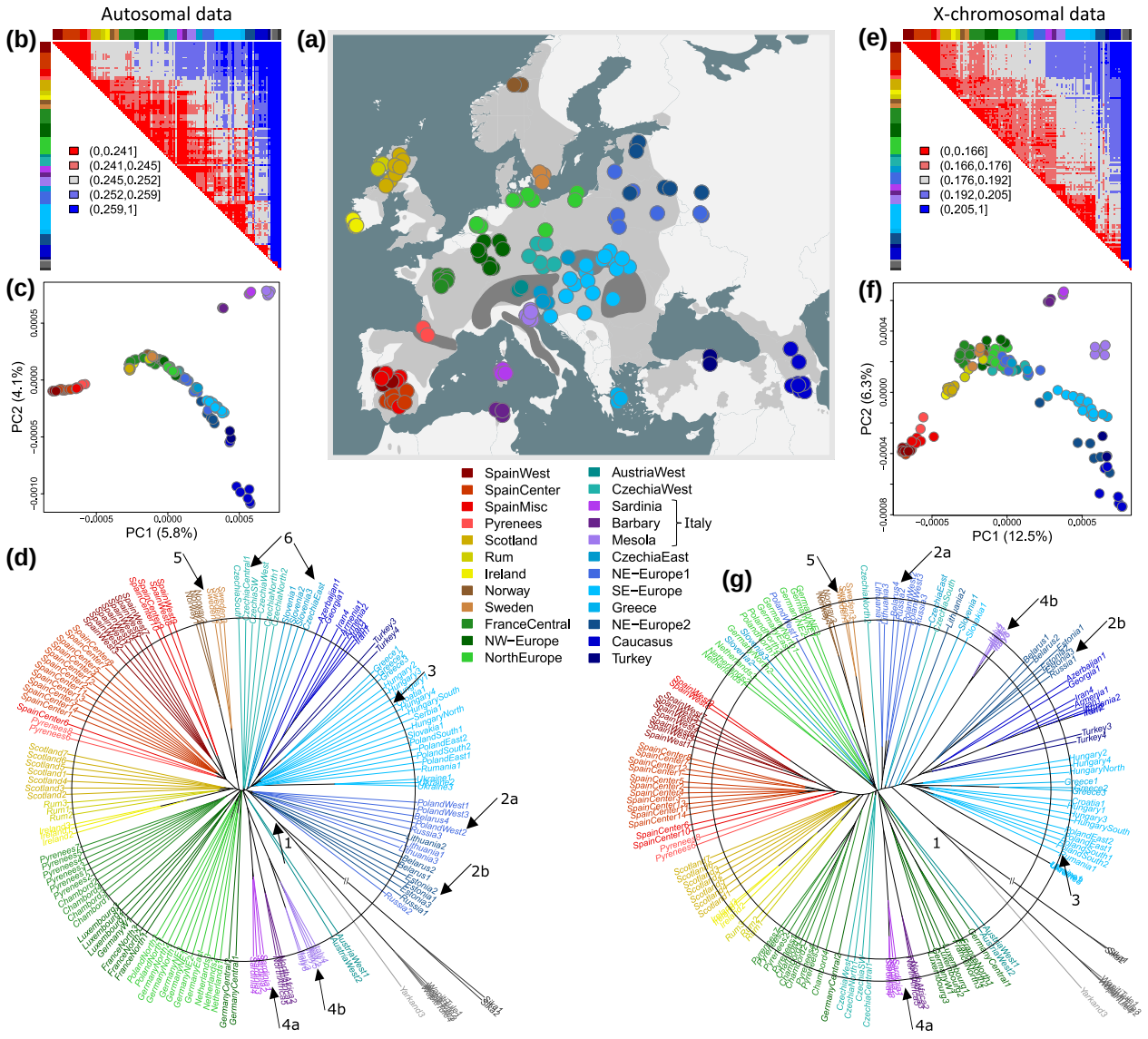

Fig. 1. Find the difference: autosomal (left) versus X-chromosomal (right) population structure. a) Sample distribution, with the present-day species range suggested by IUCN data in light gray and in dark grey the four major European mountain ranges (Pyrenees, Alps, Apennines, and Carpathians). Sample coordinates have been jittered, in order to reduce overlap. b) Heatmap depicting autosomal sequence dissimilarity (d_{auto}) between individuals. c) PCoA scatterplot inferred from d_{auto}. d) BioNJ dendrogram inferred from d_{auto}, with sika deer as outgroup. The circle has been added to aid interpretation of branch lengths. They emphasize the distinctness of Spanish, Anatolian ("Turkey"), and Caspian ("Caucasus," "NE-Europe2") red deer. The numbered arrows highlight inconsistencies between X-chromosomal and autosomal population structure (see main text for more details). Note that the branch lengths of individual-level multilocus tree are not proportional to population split time, but instead to mean coalescence time between haplotype pairs. e to g) Same as b to d), but for X-chromosomal sequence dissimilarity (d_X). Populations "CzechiaWest" and "CzechiaEast" cluster paraphyletically, and hence, arrow 6 is omitted. Apart from the annotated differences, also note the sharper boundaries between genetic clusters in the distance matrix, the higher explained variances of the PCoA axes, and the more oval shape of the dendrogram, which all indicate that X-chromosomal differences between populations are more pronounced than autosomal differences. The intermediate clustering of central populations in the dendrogram suggests a complex history of repeated fission and fusion, which affected all populations except those in the peripheral regions of the Middle East, Iberia, and, to a lesser extent, the British Isles.

influenced in any way by the sequencing data of other individuals. Admittedly, our approach results in a higher genotype call error rate in case a data set contains multiple individuals from a single, panmictic populations (Liu et al. 2022). However, we established experimentally, through the comparison of heterozygosity scores, that our approach yields the most unbiased results for data sets that contain uneven numbers of samples from highly differentiated populations (de Jong et al. 2023). In addition, because our approach does not make a priori assumptions about population assignment, the results are reproducible regardless of the sample set.

The bcftools filter pipeline was used to mask sites with a read depth below three, after establishing experimentally this provided a balance between disposal of useful data and incorrect heterozygosity estimation (supplementary fig. S2, Supplementary Material online) (de Jong et al. 2023). For the autosomal data set, we retained sites with a total read depth between 1,150 and 2,150 for all individuals combined (supplementary fig. S3a, Supplementary Material online) and with maximum 5% missing data. For the X-chromosomal and Y-chromosomal data sets, the depth range was set to 600-1,500 and 500-800, respectively, thereby accounting for ploidy differences

(supplementary fig. S3b to d, Supplementary Material online). The genotype call error rate was estimated as the proportion of homozygous alternative calls in the reference genome sample ("WapitiRocky1"). After filtering, we observed for this individual 6,971 homozygous alternative calls out of 88,678,215 retained sites on the first two chromosomes, suggesting a genotype call error rate of less than 0.008%, or less than 1 in 12,500 sites. For the Y chromosome, we observed 229 alternative calls out of 871,176 retained sites in total, suggesting an error rate of 0.026%, or over 1 in 4,000 sites.

Genome-Wide Statistics (D_{xy}, π, F_{ST}, *He*, F_{ROH})

The total number of homozygous and heterozygous sites per sample was counted in nonoverlapping windows with a fixed size of 20 kb. This counting was performed using the "Darwindow" pipeline (de Jong et al. 2023), which depends on the software Tabix (Li 2011b) for the extraction of genomic regions and which subsequently converts the count data into estimates of observed heterozygosity (*He*) and run-of-homozygosity content (F_{ROH}) using R functions (R Core Team 2019). Based on visual examination of the sensitivity of ROH analyses to various settings, ROHs were defined as continuous regions of at least 200 kb (i.e. ≥10 adjacent windows of 20 kb) with an average *He* value below 0.05%.

A custom-built Unix script was used to estimate pairwise sequence dissimilarity (p) (Xia et al. 2009) for each pair of individuals, from each of the full data sets (autosomes, X chromosome, Y chromosome, and mt-genome), containing polymorphic as well as monomorphic sites. When comparing two diploid individuals, p can be estimated without data phasing, using this simple set of rules: $p = 0$ for AA/AA; $p = 0.5$ for AA/AT or AT/AT; $p = 0.75$ for AT/CT; $p = 1$ for AA/TT or AC/GT. For haplodiploid data, three rules apply: $p = 0$ for A/AA; $p = 0.5$ for A/AT; $p = 1$ for A/TT. For haploid data, two rules apply: $p = 0$ for A/A; $p = 1$ for A/T. This calculation produces the same outcome regardless of whether individuals or regions are haploid, diploid, or haplodiploid.

The resulting matrices of uncorrected genetic distances between individuals were used as input for nonhierarchical cluster analyses (principal coordinate analyses, PCoA) as well as hierarchical cluster analyses (biological neighbor joining, bioNJ), by running the functions "pcoa" and "bionj" of the R package ape-5.3 (Paradis and Schliep 2018). X-chromosomal uncorrected genetic distances were used to infer genetic clusters, using threshold values of 0.166% or 0.176%.

For each population, nucleotide diversity (π) values were derived from the sample pairwise p values, namely as the mean of all possible sample comparisons within a population. π values are robust to sample size variation and, in panmictic populations, equal genome-wide heterozygosity (de Jong et al. 2024). Population pairwise D_{xy} values were derived as the mean of all possible sample pairwise comparisons between two populations. Relative distances between populations were estimated using the formula: $F_{ST} = (D_{xy} - \pi_{xy})/D_{xy}$ (Hudson et al. 1992; de Jong et al. 2024).

D_{xy} is a function of the time since population split (T), ancestral nucleotide diversity (π_{anc}), and the mutation rate (μ), as described by $D_{xy} = \pi_{anc} + 2\mu T$. In contrast, nucleotide diversity is described by $\pi = \pi_{anc}(1 - 1/(2N_e))^T$ (de Jong et al. 2024). Hence, elevated F_{ST} values may result either from loss of genetic variation in small populations (i.e. low π) or alternatively from the accumulation of novel mutations following deep population splits (i.e. high D_{xy}). We evaluated the underlying cause by investigating whether F_{ST} correlates with D_{xy} or instead with π_{xy} (i.e. mean nucleotide diversity in a pair of populations).

Y Chromosome and Mitogenome Phylogeny

VCF files containing genotype calls for the Y chromosome and the mitogenome, for monomorphic and polymorphic sites alike, were converted into phylip files using the software vcf2phylip (Ortiz 2019). Character-based maximum likelihood phylogenies were generated from these input files with the software IQ-TREE (with default settings) (Minh et al. 2020) and linearized using the function "chronoMPL" of the R package "ape," which implements the mean path length method (Britton et al. 2002). Genetic distance measures were converted into split time estimates (which may serve as upper limits of population splits) assuming a mutation rate of 2×10^{-9} per site per year for the Y chromosome and 2.0×10^{-8} per site per year for mtDNA (Nabholz et al. 2008).

Admixture Analyses

To leverage all available information, we strived to perform analyses on the full, genome-wide data set, comprising all monomorphic and polymorphic sites. An exception was made for admixture analyses, for which SNP data sets were generated by selecting biallelic sites with at maximum 5% missing data and by subsequently thinning the data using vcftools (Danecek et al. 2011). For autosomal data, we selected one SNP for every 40 kb, and for X-chromosomal data, we selected one SNP every 2 kb, accepting this could imply selecting multiple SNPs per coalescent locus.

Plink version 1.90b.20 (Purcell et al. 2007) was used to convert the SNP data from VCF format into PED/RAW and MAP/BIM, using the flags make-bed, recode A, chr-set 95, and allow-extra-chr. SNP data management and analyses were performed in R-4.2.0, using wrapper functions of the R package SambaR (de Jong et al. 2021). The data sets were imported into R and stored in a genlight object using the function "read.PLINK" of the R package adegenet-2.1.1 (Jombart 2008). The autosomal data set was filtered using the function "filterdata" of the R package SambaR, with indmiss = 0.25, snpmiss = 0.2, min_mac = 2, and dohefilter = TRUE.

We used two methods to detect admixture. First, ancestry coefficients were calculated from the autosomal SNP data set using the R package LEA-2.8.0 (functions "snmf" and "Q," with alpha set to 10, tolerance to 0.00001, and number of iterations to 200) (Frichot and François 2015). The optimal number of clusters (K) was determined using the elbow method on cross-entropy scores generated for $K = 2$ to $K = 12$ (with 50 independent runs each), with the assumption that the optimal K coincides with the starting point of a plateau.

Second, *f3* statistics were generated with the software Admixtools (Reich et al. 2009; Patterson et al. 2012). The *f3* statistic for population triplet (A;B,C) is defined as $(a - b) \times (a - c)$, in which a, b, and c represent vectors with the allele frequencies in the putatively admixed population A and the two putative donor populations B and C, respectively. If, and only if, a is intermediate between b and c will the product "$(a - b) \times (a - c)$" be negative. While the *f3* score has been designed to detect admixture, in theory, a negative *f3* score may also indicate a scenario in which a large source population buds of two smaller sister populations, as genetic drift can cause allele frequencies to randomly drift in opposite directions. Another

downside of the *f3* statistic is that the signal decays over time, with the *f3* statistic increasing each generation by $1/(4N_e)$, assuming equal and constant effective population sizes (de Jong et al. 2023). Therefore, for populations with low genetic diversity, we evaluated which triplets returned the lowest *f3* score, even if highly positive.

Climate Suitability Modeling

To estimate the relationship between species occurrences and their climate characteristics, we employed the maximum entropy species distribution model (MaxEnt) (Phillips et al. 2006). We trained the models through 5,000 iterations of random sampling from the historic distribution of red deer (Geptner 1988), and 19 bioclimatic variables obtained from Chelsa Climate (Karger et al. 2017). In each run, the sampled occurrences were split; 50% were used to train the models and the remaining 50% were used to evaluate the performance of the models with their discriminatory power. To prevent overfitting, we selected models with an AUC score of at maximum 0.8.

We transferred the climate suitability models through time using three PMIP3 generalized circulation models: CCSM4, MIROC-ESM, and MPI-ESM-P (Karger et al. 2017). Selected time periods were the LGM (here set at ~21 kya), Heinrich Stadial (Oldest Dryas, 17 to 14.7 kya), Bølling-Allerød interstadials (14.7 to 12.9 kya, thus including the Older Dryas ~14 kya), Younger Dryas stadial (12.9 to 11.7 kya), Early Holocene Greenlandian (11.7 to 8.3 kya), mid-Holocene Northgrippian (8.3 to 4.2 kya), and Late Holocene Meghalayan (4.2 to 0.3 kya), all obtained from PaleoClim.org (Fordham et al. 2017). We performed all the geospatial analyses in R version 4.1.1 (R Core Team 2019), using raster version 3.4-13 (Hijmans 2019), rdgal version 1.5-23 (Bivand et al. 2021), and dismo 1.3-3 (Hijmans et al. 2020).

Results

Autosomal Versus X-Chromosomal Population Structure

For autosomal as well as X-chromosomal data, the first PCoA axis accentuates the difference between Spanish deer and eastern European populations (the latter including Italian deer), while the second PCoA axis accentuates the difference between eastern European populations and Italian deer (Fig. 1c and f). The observed correlation between sample longitude and sample loadings with the first two PCoA axes (supplementary fig. S4, Supplementary Material online) implies that longitude predicts whether individuals are genetically more similar (i) to either eastern European deer or to Spanish deer and (ii) to either eastern European deer or to Italian deer. More specifically, relative to eastern European deer, individuals in western Europe resemble Spanish as well as Italian deer—even though Spanish deer and Italian deer are genetically very different from each other.

The autosomal and X-chromosomal dendrograms are generally consistent. They both indicate that the largest genetic distances are observed between the peripheral populations of Spanish deer and Caspian deer ("Caucasus", "Turkey"). (Fig. 1d and g; supplementary fig. S5a, Supplementary Material online). Other populations cluster intermediate, suggesting a longitudinal isolation-by-distance trend without pronounced genetic discontinuities (Fig. 1; supplementary fig. S5b, Supplementary Material online). Relative to autosomal data, the X-chromosomal genetic distance between Spanish deer and Caspian deer is more pronounced, as reflected by the more oval shape of the X-chromosomal dendrogram (Fig. 1d and g), as well as by the higher explained variance of the first axis of the X-chromosomal PCoA plot (12.5% vs. 5.8%, Fig. 1c and f). Other inconsistencies, which have been annotated in the dendrograms with numbered arrows (Fig. 1d and g), are as follows:

1. X-Chromosomal data suggest that all red deer populations are roughly equidistant to the sister species (wapiti and Yarkand deer), with the exception of Iberian deer. In contrast, autosomal data suggest that wapiti and Yarkand deer are genetically slightly closer to eastern red deer populations, particularly those in northeastern Europe.
2. Unlike autosomal data, X-chromosomal data suggest the existence of two genetic clusters in northeastern Europe: one cluster ("NE-Europe2," 2b) groups with individuals from the Caucasus and Turkey, and the other ("NE-Europe1," 2a) with individuals from northwestern Europe.
3. X-Chromosomal distances of southeastern European individuals ("SE-Europe") to western European populations are larger than the corresponding autosomal distances.
4. Autosomal data group Mesola deer (4b) with Sardinian and Barbary deer (4a), while X-chromosomal data indicate that Mesola deer are more similar to southeastern European deer.
5. Based on X-chromosomal data, Scandinavian individuals ("Norway" and "Sweden") cluster with individuals of northern Germany. In contrast, based on autosomal data, Scandinavian individuals cluster as a distinct unit, closer to Spanish, British, and Irish deer.
6. While based on autosomal data, populations in central Europe (i.e. "CzechiaWest", "CzechiaEast") cluster as distinct monophyletic clades, X-chromosomal data reveal paraphyletic clustering of these populations, with a fan-like pattern suggestive of ongoing admixture.

For all study populations, X-chromosomal π values are below their autosomal π values (0.03% to 0.14% vs. 0.09% to 0.23%, Fig. 2b), consistent with the theoretical notion that X-chromosomal N_e is at maximum three quarters relative to autosomal N_e (Laan et al. 2005; Nussey et al. 2005; Holsinger and Weir 2009). Genetic diversity in ancestral populations sets the base line of genetic distances between its offspring populations (i.e. $D_{xy} = \pi_{anc} + 2\mu t$), and hence, X-chromosomal D_{xy} values are also below autosomal D_{xy} values (Fig. 2a). Because of the lower π_{anc}, a relatively higher proportion of D_{xy} can be attributed to novel mutations, and hence, X-chromosomal F_{ST} values are above autosomal F_{ST} values (Fig. 2c).

High F_{ST} Values Are Caused by Drift, Not by Novel Mutations

As indicated by the short internal branches of the autosomal dendrogram, sequence dissimilarity between populations often barely exceeds sequence dissimilarity within populations (i.e. $D_{xy} \approx \pi_{xy}$), resulting in generally low F_{ST} values (Fig. 2). There are exceptions: going by their long internal branches and their high relative genetic distances (i.e. high F_{ST} values), populations from Ireland, Scandinavia ("Norway" and "Sweden"), Italy (including Barbary deer), and the Caucasus and Turkey stand out as genetically distant (Fig. 1; supplementary fig. S6a, Supplementary Material online). However, none of these populations differ strongly from other populations in terms of

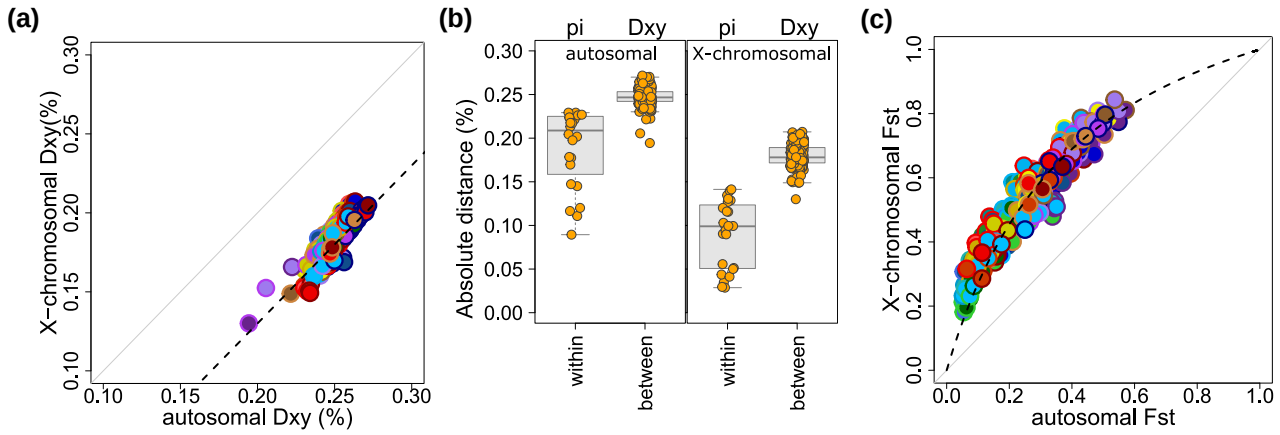

Fig. 2. The X chromosome has a relatively low N_e, and thus higher sensitivity for detecting population structure, as shown by its higher F_{ST} values. a) Scatterplot depicting correlation between autosomal and X-chromosomal population pairwise D_{xy} estimates. Each data point represents a pair of populations and hence consists of two colors. The dashed line is predicted by the formula $D_{xy} = \pi + 2\mu T$, assuming $\pi_{autosome} = 4N_e\mu$ and $\pi_X = 0.3(4N_e\mu)$. The close fit suggests that the differences between X-chromosomal and autosomal D_{xy} values are not due to differences in mutation rates, but rather to the lower N_e of the X chromosome. b) Boxplots depicting mean autosomal and X-chromosomal distances within (π) and between (D_{xy}) populations. The marginal difference between autosomal π and D_{xy} is reflected by the small inner branch lengths of the bioNJ dendrograms (see Fig. 1). c) Idem as a), but for pairwise Hudson F_{ST} estimates: $F_{ST} = (D_{xy} - \pi_{xy})/D_{xy}$. Note again the fit between predicted (dashed line) and observed values.

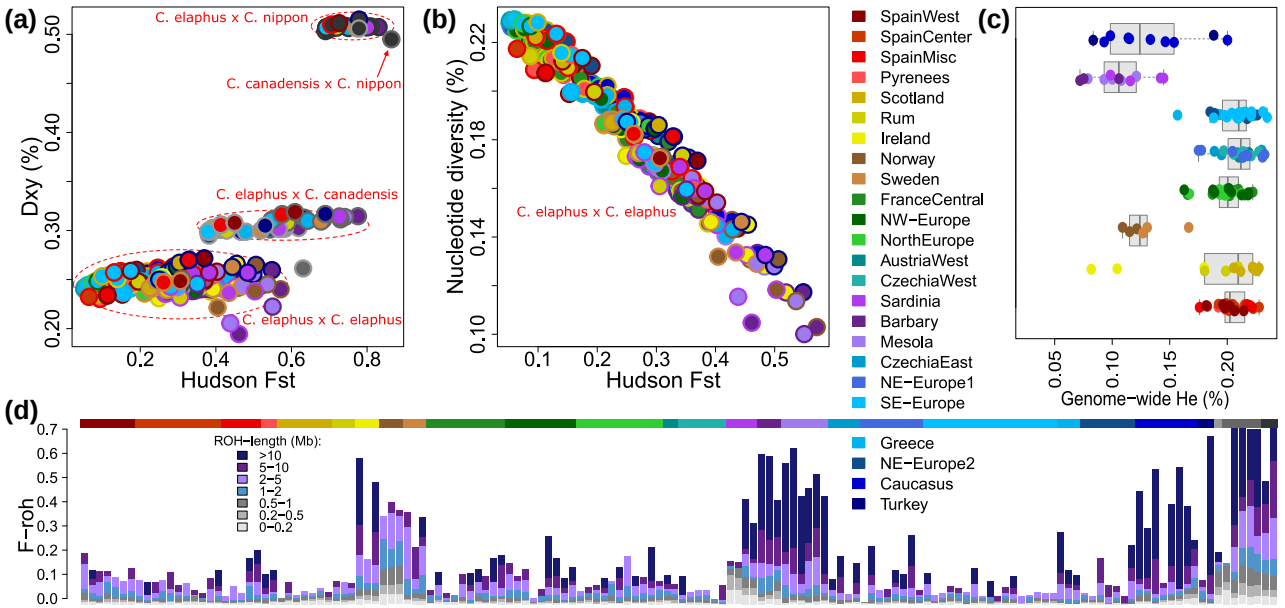

Fig. 3. High intraspecific F_{ST} values are caused by genetic drift, not by novel mutations. a) Scatterplot depicting population pairwise absolute distances versus relative genetic distances (i.e. D_{xy} vs. F_{ST}). Each data point represents a pair of populations and hence consists of two colors. The correlation between D_{xy} and F_{ST} is much stronger when considering between-species comparisons only. b) Scatterplot depicting mean nucleotide diversity versus population pairwise genetic distances (i.e. π_{xy} vs. F_{ST}) within *C. elaphus* only. The strong correlation indicates that high F_{ST} values are not caused by novel mutations (i.e. deep population splits), but by loss of genetic variation due to recent population bottlenecks. c) Boxplots overlayed with strip charts, depicting genome-wide *He* estimates for each individual, grouped per geographic region. d) Stacked bar plots depicting the genome-wide proportion of runs of homozygosity (ROH) per individual, with ROHs divided into length bins. Included are the sister species of wapiti, Yarkand, and sika deer.

absolute genetic distances (i.e. D_{xy} values, Fig. 3a; supplementary fig. S6b, Supplementary Material online), and for intraspecific comparisons, no relationship is observed between absolute and relative genetic distance (Fig. 3a). This indicates that the elevated intraspecific F_{ST} values are caused by loss of genetic variation within small populations (i.e. decrease of π) and not by the accumulation of private mutations following deep population splits (i.e. not by an increase of D_{xy}).

Indeed, all populations with high F_{ST} values are characterized by low genetic diversity and by high inbreeding values (Fig. 3b to d; supplementary fig. S6c, Supplementary Material online). The presence of long runs of homozygosity (ROHs), of over 5 Mb length, in Irish deer, Mesola deer, and Barbary deer suggests that the reduction of genetic variation in these populations occurred recently, consistent with documented demographic events during the 20th century (Hajji et al. 2007; Zachos et al. 2009; Carden et al. 2012). In contrast, the presence of mainly shorter ROHs (<5 Mb) in Scandinavian individuals (Fig. 3d) indicates that the loss of genetic diversity in these populations occurred further back in time (Haanes et al. 2011). For the Swedish deer, it is known that at the start of the 20th century, less than 50 individuals remained (Höglund et al. 2013).

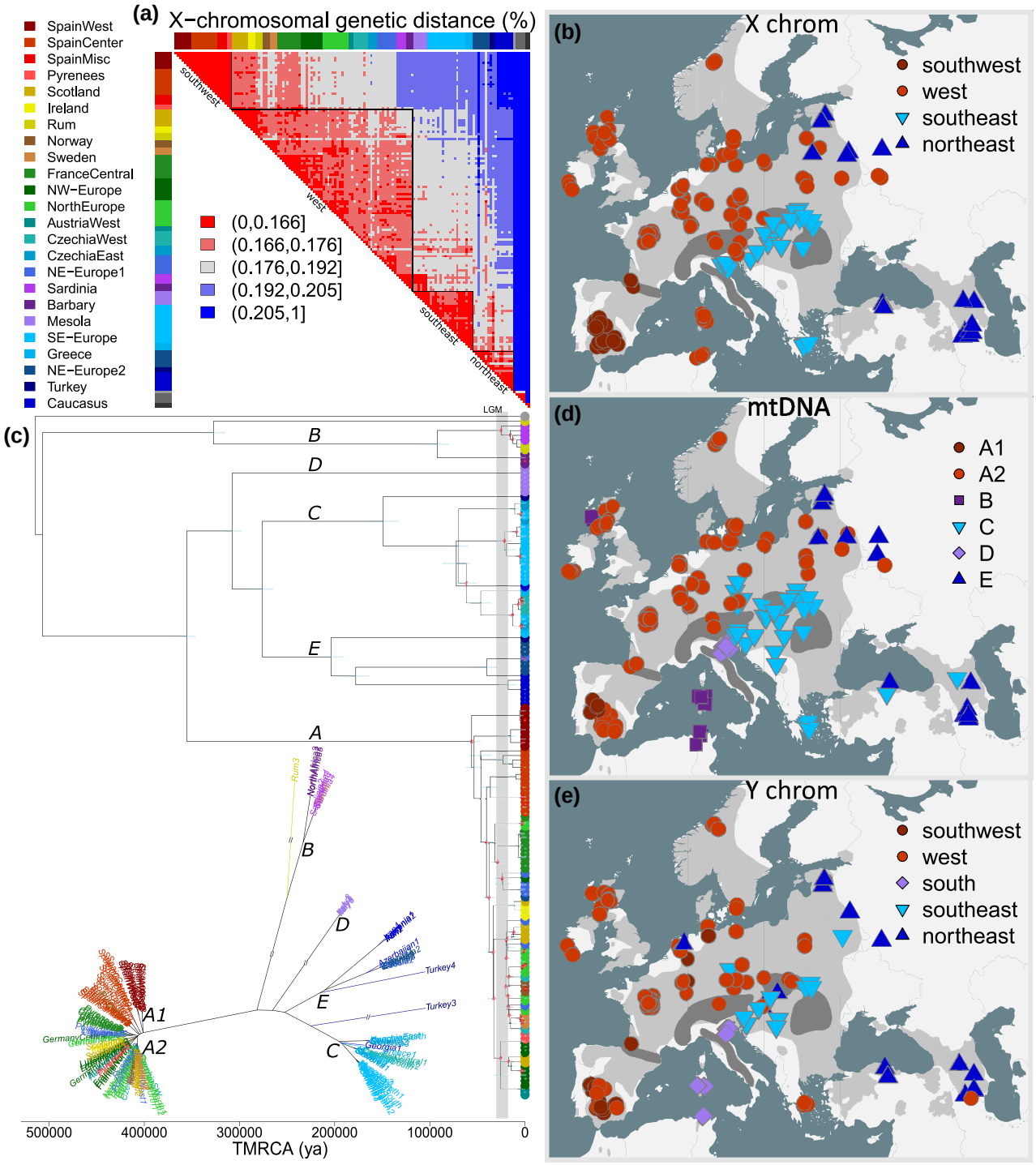

Fig. 4. The genetic legacy of postglacial recolonization. a) Heatmap depicting X-chromosomal sequence dissimilarity estimates between individuals, in percentages. Genetic clusters are inferred using arbitrary threshold values of 0.166% and 0.176%. b) Geographical distribution of X-chromosomal clusters (see a). Sample coordinates have been jittered, in order to reduce overlap. Note the general consistency with the geographical distribution of mtDNA clades (see d) and Y-chromosomal clades (see e). Also note that Spanish deer form a distinct cluster, suggesting that they did not contribute to the recolonization of northwestern Europe. c) MtDNA maximum likelihood phylogenies in unrooted and phylogram format (the latter linearized), generated with the software IQ-TREE, assuming a mutation rate of 2.0×10^{-8} per site per year. Gray vertical bar indicates the LGM (25 to 18 kya). d) Geographical distribution of mtDNA haplogroups (see c). Sample coordinates have been jittered. e) Geographical distribution of Y-chromosomal haplogroups, corresponding to the maximum likelihood phylogeny in Fig. 5. Sample coordinates have been jittered.

X-Chromosomal Data: Distinct Genetic Clusters

When grouping pairs of individuals into distance bins based on their X-chromosomal genetic distances, several clusters emerge, which are distributed parapatrically over different parts of the European continent (Fig. 4a and b). One cluster is restricted to southwestern Europe and comprises Spanish individuals only, while a second cluster comprises individuals across western Europe. A third cluster is restricted to southeastern Europe, and a fourth cluster contains individuals from the Caucasus and Turkey as well as a few individuals from northeastern Europe (Fig. 4a).

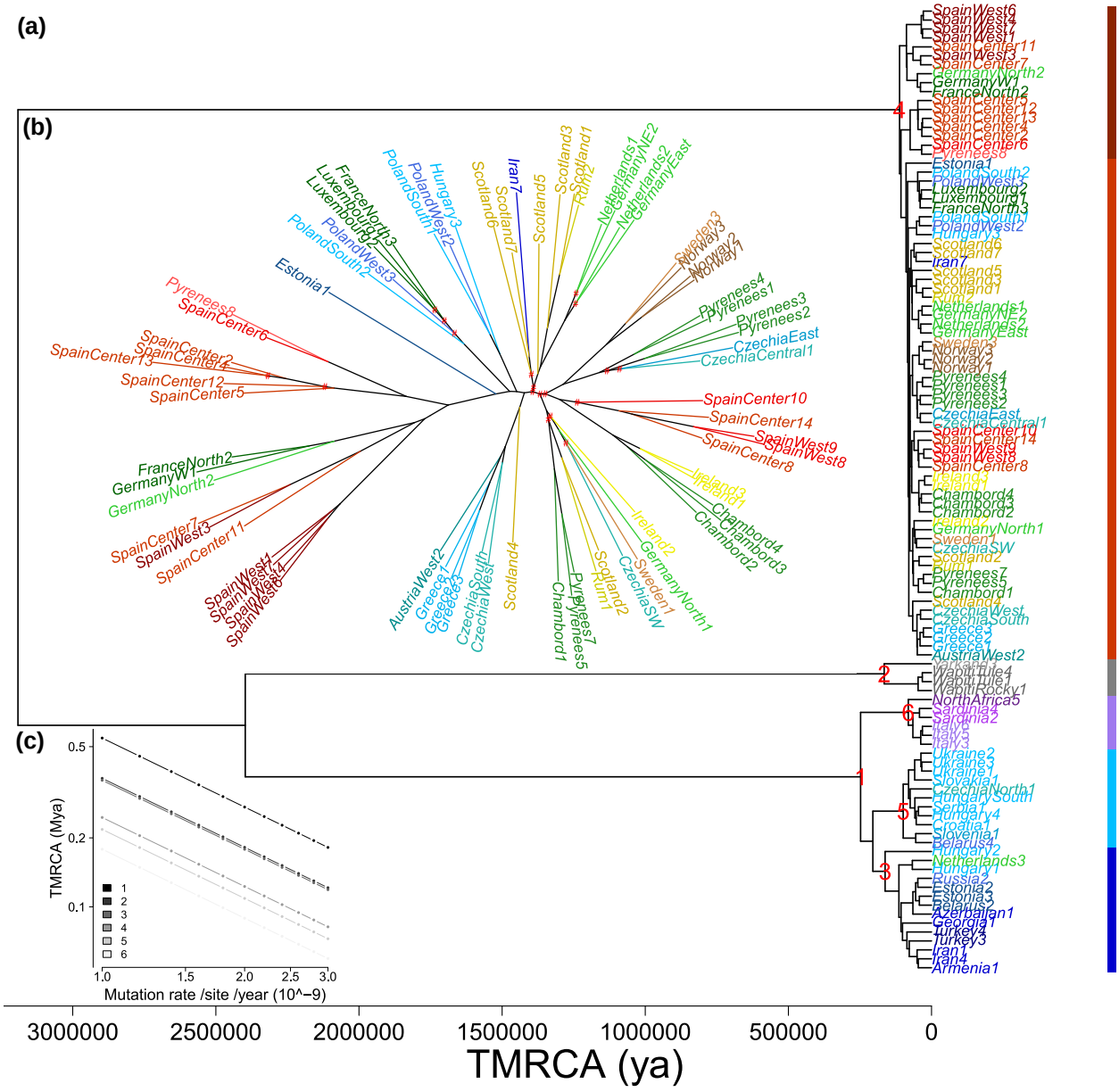

Fig. 5. Y-Chromosomal phylogeny. a) Linearized Y-chromosomal maximum likelihood phylogeny, generated with the software IQ-TREE, assuming a mutation rate of 2.0×10^{-9} per site per year. The right-hand color bar corresponds to the clades in Fig. 4e. Note that eastern European deer cluster with sister species (i.e. Yarkand and wapiti), which given the deep coalescent time likely reflects incomplete lineage sorting rather than gene flow. Sequence dissimilarity in western Europe is on average lower than in eastern Europe. b) Same as a), but for a subset of individuals (western clade), and here presented as an unrooted phylogeny. Nodes with low bootstrap support (<80%) have been marked with a hashtag. Due to the relatively recent coalescence times, implying few informative sites relative to genotype call errors, the observed topology is not robust. c) Sensitivity of node age estimates to mutation rate. Numbers refer to nodes highlighted in a).

The geographical range of the western cluster is not clear-cut. Red deer of Britain and Ireland cluster ambiguously and can be grouped either within the widespread western cluster or within the southwestern cluster, together with Spanish deer (Fig. 4a). Likewise, Mesola deer are genetically similar to Sardinian/Barbary deer, but also to southeastern European individuals (Fig. 4a).

mtDNA Clades Are Largely Consistent with X-Chromosomal Clusters

Consistent with previous studies (Ludt et al. 2004; Skog et al. 2009; Niedziałkowska et al. 2011; Queirós et al. 2019; Doan et al. 2022; Mackiewicz et al. 2022; Valnisty et al. 2024), our mtDNA-based phylogenetic inferences suggest that red deer individuals can be assigned to five major clades (Fig. 4c and d; supplementary fig. S6d, Supplementary Material online). Haplogroup A occurs throughout western Europe, including Iberia, Scandinavia, Britain, and Ireland. Haplogroup B is endemic to Barbary Coast and the Tyrrhenian islands (Corsica and Sardinia), as well as the Isle of Rum, where it has been supposedly recently introduced (Nussey et al. 2006; Skog et al. 2009; Doan et al. 2017, 2022). Haplogroup C occurs in southeastern and central Europe. Haplogroup D is found in the Po Valley in northeastern Italy (i.e. Mesola deer). Haplogroup E is restricted to the Caucasus, Turkey and northeastern Europe.

The geographical distribution of these mtDNA clades largely corresponds with the geographical distribution of

X-chromosomal clusters (Fig. 4a and b), albeit with a few discrepancies. The most prominent inconsistency concerns Sardinian and Barbary deer, which according to X-chromosomal data belong to the western cluster, but constitute a distinct mtDNA lineage (clade B, Fig. 4). Another apparent inconsistency involves Spanish deer. According to X-chromosomal data, Spanish deer are a distinct, monophyletic lineage. In contrast, in the mtDNA phylogeny, Spanish deer are paraphyletic with respect to other western red deer of haplogroup A (Fig. 4c) (Carranza et al. 2016).

Y-Chromosomal Clades Are Also Largely Consistent with X-Chromosomal Clusters

Assuming an arbitrary threshold of 0.06% sequence divergence, Y-chromosomal data reveal four main clusters (Figs. 4e and 5). As observed for mtDNA clades, the geographic distribution of these four Y-chromosomal clades largely corresponds with the geographic distribution of X-chromosomal clusters, except for identifying Sardinian and Barbary deer as a distinct lineage. Unlike mtDNA and X-chromosomal data, but in line with autosomal data, Y-chromosomal data suggest that this distinct lineage includes Mesola deer (Figs. 4e and 5). The aberrant clustering of individuals "Netherlands3," "Iran7," "Greece1," "Greece2," and "Greece3" (Figs. 4e and 5) might concur with documented human-made translocation events (Karaiskou et al. 2014; de Jong et al. 2020), although this is not supported by autosomal data.

Y-Chromosomal haplotypes found in eastern red deer populations differ less from the haplotypes of the sister species of wapiti and Yarkand deer than from those found in western red deer populations (Fig. 5). However, coalescence times between the Y-chromosomal haplotypes of eastern European red deer and those of these sister species are deep (>2 Mya, Fig. 5), which suggests that this instance of gene tree discordance reflects incomplete lineage sorting rather than gene flow.

Even within the western and eastern clades itself, the observed coalescence times of Y-chromosomal haplotypes are generally older than those of mtDNA haplotypes. For instance, the coalescence times of western Y-chromosomal haplotypes are estimated to be approximately 100 kya (Fig. 5), double those of the corresponding west European mtDNA haplogroup A (Fig. 4c). Because the mitogenome and the Y chromosome have in theory the same effective population size (i.e. 1 copy per mating, thus $\frac{1}{2}N_e$), this difference is unexpected. However, it should be noted that these coalescence times are highly sensitive to the mutation rate setting (Fig. 5c). Furthermore, the huge difference in mean sequencing depth between mtDNA (~6,000) and the Y chromosome (~10) is expected to translate to a relatively higher genotype call error rate for Y-chromosomal data.

Admixture Analyses

When assuming three ancestral demes ($K = 3$), LEA ancestry analyses suggest that red deer in northwestern Europe, including Scandinavian, British, and Irish deer, share most of their genetic variation with Sardinian and Barbary deer, not with Spanish deer (Fig. 6a). With the exception of Scandinavian, British, and Irish populations, red deer in western Europe also share some genetic variation with eastern European red deer populations, with the relative proportions depending on their proximity to the mtDNA contact zone in central and northeastern Europe (Fig. 6a).

The *f3* analyses indicate that populations that occur near the mtDNA phylogeographic break in central and northeastern Europe are admixed between western and eastern lineages (Fig. 6b and c). In addition, and consistent with LEA admixture analyses, they indicate that populations in western Europe have allele frequencies intermediate between Spanish deer and Barbary/Sardinian deer (Fig. 6b), potentially suggesting a hybridization event between these two lineages that occurred before or during the recolonization of northern latitudes. This would also explain why the path lengths of the autosomal bioNJ dendrogram underrepresent the genetic distance between Spanish deer and Barbary/Sardinian deer (supplementary fig. S5a, Supplementary Material online).

Owing to a severe population bottleneck (Zachos et al. 2009), all population triplets in which Mesola deer are the hypothetical admixed population, yield highly positive *f3* scores (supplementary fig. S7, Supplementary Material online). However, among those triplets, the lowest *f3* scores are observed when the hypothetical donors are pairs containing one southeastern population ("SE-Europe," "Greece," or "Turkey") and one mtDNA lineage B population (Sardinian or Barbary deer) (supplementary table S3, Supplementary Material online). An admixture event between these lineages would explain why X-chromosomal loci and mtDNA indicate shared ancestry between Mesola deer and southeastern European deer (Figs. 1 and 4c) (Doan et al. 2022), while Y-chromosomal and autosomal data suggest that Mesola deer instead share ancestry with Sardinian and Barbary deer (Fig. 5).

Climate Suitability Projections

Out of the 19 bioclimatic variables, three variables each explain 10% to 20% of the variation in the data, namely, in order of decreasing importance, minimum temperature of coldest month (BIO6), precipitation of coldest quarter (BIO19), and precipitation of driest month (BIO14) (supplementary fig. S8, Supplementary Material online).

Projections of this selected model to past climate conditions suggest that for over 20 millennia red deer climate suitability zones have been restricted to Europe and isolated pockets in the Middle East and thus barely extended into eastern Eurasia (Fig. 7). Within this already restricted range in western Eurasia, red deer frequently contracted to even smaller, fragmented range pockets at southern latitudes, which were distributed over the Middle East, the Mediterranean basin, and southwestern France, roughly consistent with outcomes of previous studies (Sommer et al. 2008; Queirós et al. 2019; Niedziałkowska et al. 2021). These range contractions did not only occur during the LGM but also, and in fact particularly, during subsequent stadials, most notably the Younger Dryas (Fig. 7). In the Early Holocene, suitable climatic conditions existed along the coastal region of northwestern Europe.

Discussion

Autosomal Versus X-Chromosomal Population Structure

The data set of 144 red deer genomes provides a comprehensive overview of the range-wide genetic structure of the species. In addition, it allows to reconstruct the historical demographic processes that shaped the present-day structure. This reconstruction of events is fairly straightforward if lineages remain indefinitely and fully isolated after their initial split, but the task becomes more challenging in the face of

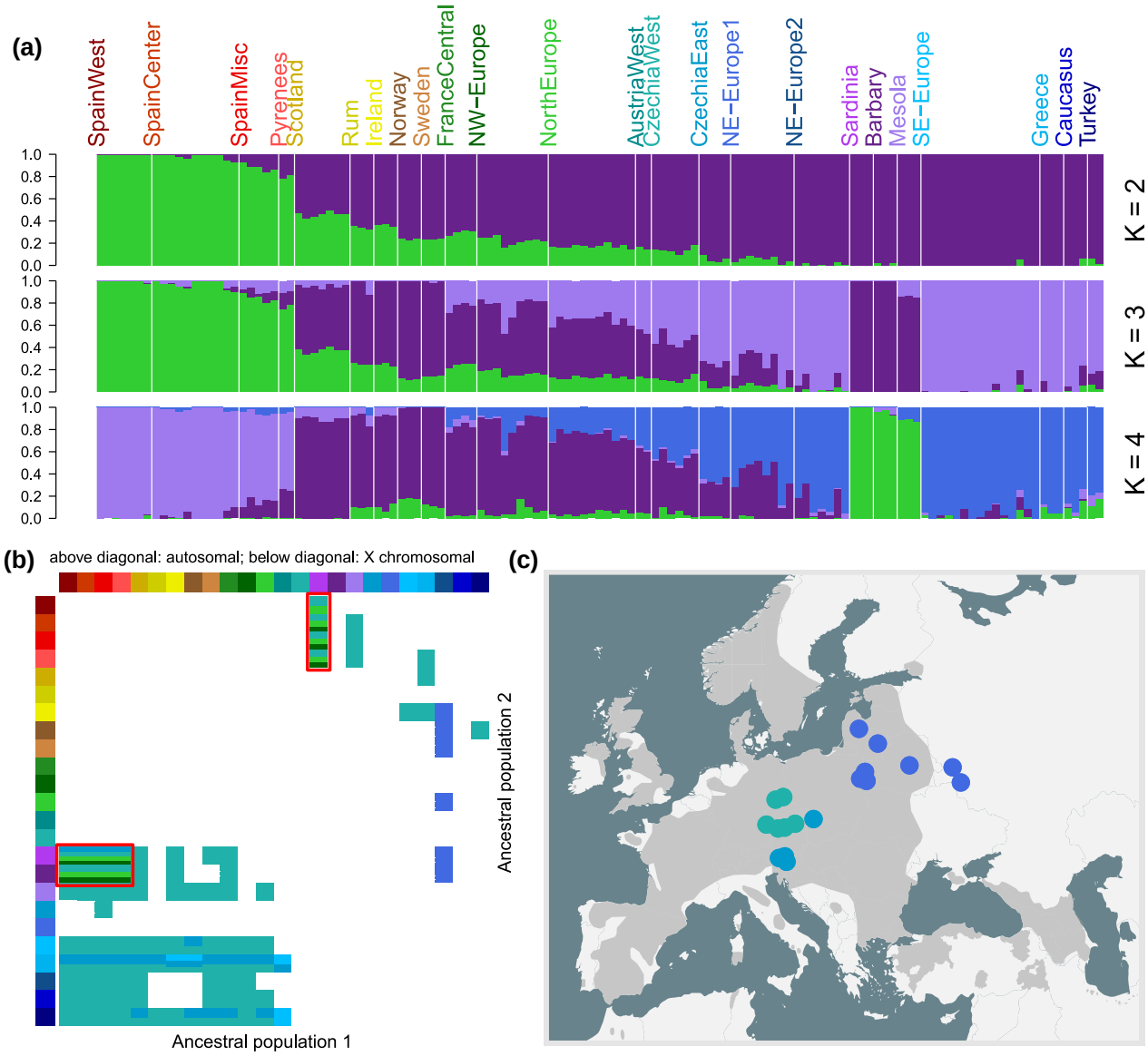

Fig. 6. Admixture analyses reveal recent hybridization at the secondary contact zone in central Europe. a) LEA admixture plots depicting inferred ancestry coefficients assuming two to four ancestral populations. For $K=2$, note the west-to-east gradient of admixture proportions. For $K=3$, note that populations in northwestern Europe (including Scandinavia, Ireland, and the British Isles) share more of their genetic ancestry with Sardinian, Barbary, and Mesola deer than with Spanish populations. b) Matrix highlighting population triplets (*Y;X,Z*) with negative or near-negative autosomal (above diagonal, <0.002) and X-chromosomal (below diagonal, <0) *f3* scores. The *f3* score is calculated as $(y-x)\times(y-z)$, with *x*, *y*, and *z* denoting the allele frequencies in populations *X*, *Y*, and *Z*, respectively. A negative *f3* score indicates that allele frequencies in population *Y* are intermediate between those of populations *X* and *Z*. The row and column colors represent populations *X* and *Z*, while field colors denote admixed population *Y*. The frames highlight that northwestern European deer are admixed between Spanish deer and Sardinian/Barbary deer. c) Geographical locations of individuals belonging to the populations that are marked most consistently as admixed by *f3* analyses. These individuals occur at or near the genetic discontinuity in central Europe (Fig. 4), supporting the hypothesis of a Holocene secondary contact zone that arose following postglacial recolonization.

admixture. While admixture itself can be accurately identified with available gene flow detection methods (such as *f3* and *D* statistics), establishing the original population structure prior to gene flow events remains difficult.

Fortunately, because gene flow in many mammals is predominantly male mediated (Nussey et al. 2005; Frantz et al. 2008; Fickel et al. 2012), the X chromosome can provide a work-around solution. X-Chromosomal DNA from immigrant males is inherited by female offspring only, which reduces the migration rate for this particular chromosome by up to a half, and therefore allows the X chromosome to retain the signal of a former population structure for longer (de Jong et al. 2023). Furthermore, considering that X-chromosomal N_e is approximately three-fourth of autosomal N_e, higher X-chromosomal F_{ST} values are expected even prior to secondary contact. With these assumptions, the puzzle pieces fall into place, and the observed inconsistencies between autosomal and X-chromosomal data become meaningful.

Consider, for instance, the genomic profiles of red deer in Poland, Belarus, the Baltic States, and the bordering regions of Russia. There, the western lineage (mtDNA haplogroup A) bulges into northeastern Europe to meet with the northeastern lineage originating from the Caucasus (mtDNA haplogroup E) (Fig. 4d). The autosomal data set suggests that all individuals in this region constitute a monophyletic clade, distant from both western and Caspian red deer ("Caucasus," "Turkey")

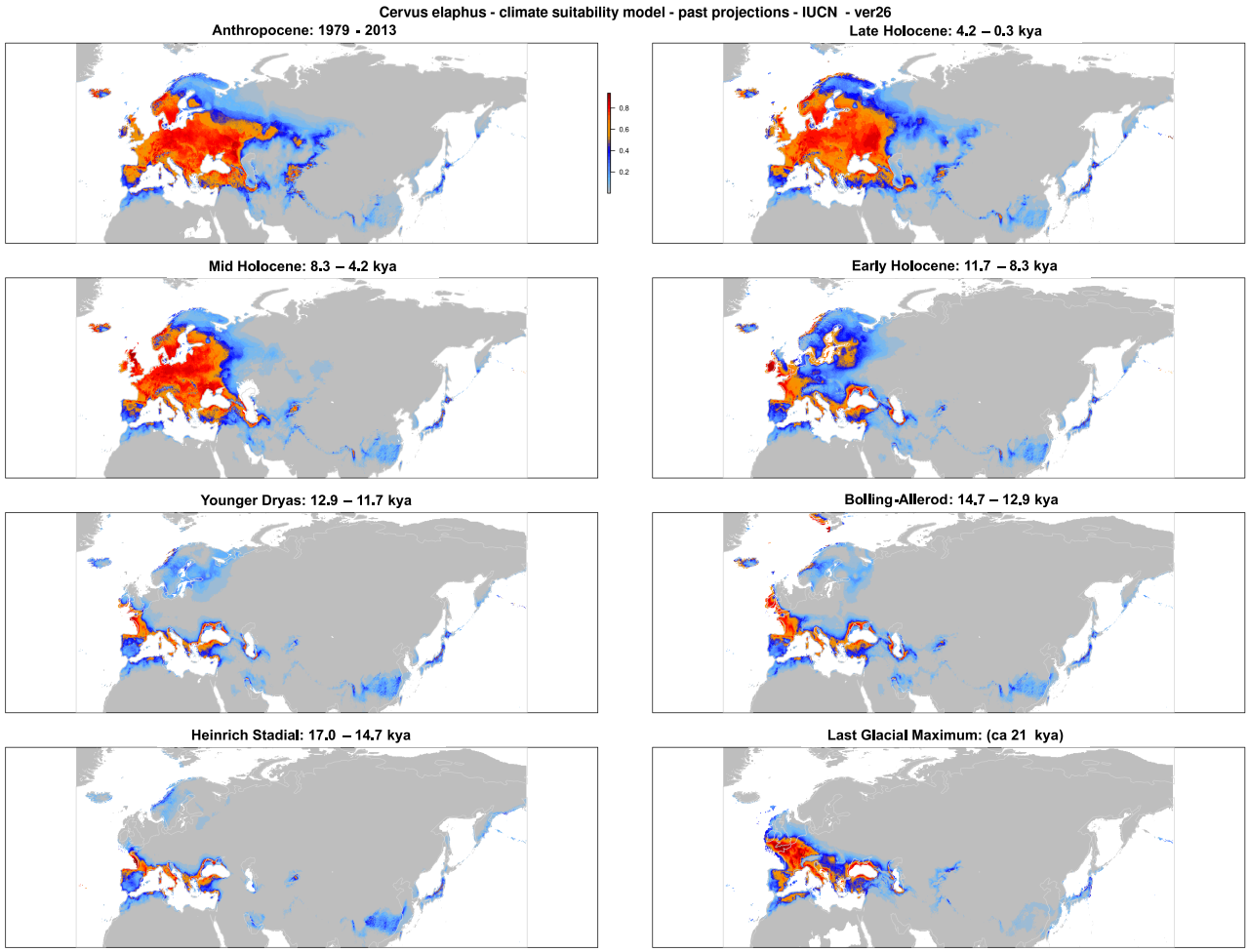

Fig. 7. Climate suitability modeling suggests postglacial recolonization only truly commenced after the Younger Dryas and proceeded proceeded over the North European plain. Projections of red deer climatic suitability ranges during various past time periods. Potential additional limitations caused by either biotic factors, such as vegetation, or migration barriers, such as water barriers and mountain ranges, have not been taken into consideration.

(Fig. 1d). But the X-chromosomal data set suggests differently and separates peripheral individuals from more interior individuals, the former clustering with Caspian red deer ("Caucasus" and "Turkey") and the latter clustering with northwestern Europe (Fig. 1g).

Although conclusions may be partly confounded by recent translocations (Niedziałkowska et al. 2011; Valnisty et al. 2024), it appears that much of present-day Belarus and the Baltic States was originally recolonized by the western lineage (mtDNA lineage A), whereas present-day European Russia was recolonized by Caspian deer (mtDNA lineage E), which migrated out of the Caucasus (Doan et al. 2022; Valnisty et al. 2024). This dual heritage of red deer in northeastern Europe has been preserved by X-chromosomal and mtDNA loci (Doan et al. 2022), but is less clear from autosomal loci due to introgression of allochthonous DNA exchanged through male-mediated gene flow.

During the advance of mtDNA lineage A along the northern European coastline, a subgroup of individuals split off in the direction of Scandinavia, entering southern Sweden before the collapse of the Öresund land bridge (Björck 1995; Höglund et al. 2013; Bailey et al. 2020; Nilsson et al. 2020). The X-chromosomal dendrogram shows this ancestry of Scandinavian deer and groups them together with northwestern European deer (Fig. 1g). In contrast, in the autosomal dendrogram, the Scandinavian deer cluster separately from northwestern European individuals (Fig. 1d), toward Spanish, British, and Irish deer, and therefore away from eastern lineages. Again, this discrepancy can be explained by assuming male-mediated gene flow following secondary contact between the western and eastern lineages. This genetic exchange transferred relatively more autosomal loci than X-chromosomal loci from the eastern lineage into the western lineage. The Scandinavian, British, and Irish deer were shielded from this influx from eastern red deer lineages by water barriers and could retain their original genetic identity, unlike their mainland relatives.

Mesola deer, from the Po Valley in northeastern Italy, are another interesting case. These individuals, the only native deer left on the peninsula, carry mtDNA haplogroup D, a haplogroup that used to be widespread in southeastern Europe as recently as the mid-Holocene (Doan et al. 2022). While autosomal and Y-chromosomal data indicate that these individuals are related to Sardinian and Barbary deer (Figs. 1, 4e, and 5) (Hajji et al. 2007), X-chromosomal data instead reveal closer similarity to individuals from southeastern Europe (Fig. 1g). The latter finding supports the hypothesis that the ancestors of Mesola deer arrived in the Po Valley during the Holocene (Doan et al. 2017, 2022), likely from southwestern Europe. The present-day near confinement of mtDNA haplogroup D

to Mesola Forest conceals the shared ancestry of Mesola deer and southeastern European red deer (Borowski et al. 2016). Upon arrival in the Po Valley, the ancestors of Mesola deer likely interbred with Italian deer of mtDNA lineage B (Doan et al. 2017). Male-mediated gene flow can explain why the mitogenome and X chromosomes of Mesola deer remained similar to southeastern European deer, while their autosomes and Y chromosomes became to resemble the donor population.

We emphasize that the observed inconsistencies between the autosomal and X-chromosomal data sets are unlikely to reflect artifacts. Unlike mtDNA data (Carranza et al. 2016), the X-chromosomal data set is a multilocus data set and hence inferences are not confounded by incomplete lineage sorting. The X-chromosomal distances were inferred from almost 2.5 million variable sites and were ultrametric by approximation. This resulted in a X-chromosomal dendrogram with nearly equal root-to-tip distances, suggesting that single-locus stochastics were indeed canceled out (de Jong et al. 2024).

Recent Admixture Events

The observed isolation-by-distance trend (Fig. 1d and g; supplementary fig. S4, Supplementary Material online), despite the presence of (X-chromosomal) phylogenetic breaks (Figs. 4 and 6), fits a scenario of reticulate evolution with repeated fission–fusion of geographically adjacent populations. Least affected by the admixture events are the peripheral populations in Iberia, the British Isles, and the Caucasus (supplementary fig. S5b, Supplementary Material online). This scenario is supported by the projections of species distribution modeling, which indicate range oscillations associated with cycles of stadials and interstadials separating the end of the LGM from the start of the Holocene (Fig. 7). Given the complexity of the continuous range contractions and expansions, and with each consequent admixture event partly obscuring the signals of previous admixture, we can only hope to reconstruct the most recent of these admixture events (Carranza et al. 2024).

One such a relatively recent admixture event appears to have occurred in western Europe. Whereas Y-chromosomal and mtDNA data cluster northwestern European deer with Spanish deer (Figs. 4 and 5), the autosomal and X-chromosomal data sets leave no doubt that all northwestern European deer, including Scandinavian deer, are genetically more similar to Barbary and Sardinian deer (Figs. 1 and 6A). Because ordination analyses are not robust to population bottlenecks, this genetic similarity between northwestern European deer and Barbary/Sardinian deer is not evident from the PCoA plots (Fig. 1c and f). However, it can be readily inferred from the dendrograms by comparing tip-to-tip distances (Fig. 1d and g), which correspond to genetic distances between sample pairs in the underlying distance matrices (Fig. 1b and e; supplementary fig. S2, Supplementary Material online).

While being closer to the latter, allele frequencies of northwestern European populations are roughly intermediate between Spanish deer and Sardinian/Barbary deer (Fig. 6b and c). Genetic distances between these populations are furthermore not ultrametric and hence do not fit a binary splitting model (supplementary fig. S5, Supplementary Material online). This suggests an admixture event prior or during the recolonization of northwestern Europe, with some contribution coming from relatives of Spanish deer (mtDNA lineage A), but with most contribution coming from relatives of Barbary/Sardinian deer (mtDNA lineage B). The ancestors of British and Irish deer appear to have received little gene flow from Barbary/Sardinian deer and these populations therefore share relatively more ancestry with Spanish deer (Figs. 1 and 6a).

The recolonization of northern latitudes culminated in the next admixture event, as it brought the western hybrid lineage into secondary contact with eastern lineages. At present, millennia later, the hybridization zone is still evident from the parapatric distributions of mtDNA and Y-chromosomal haplotypes and from the geographical ranges of X-chromosomal clusters (Fig. 4). While the phylogeographic break is less prominent from autosomal data (owing to male-mediated Holocene gene flow and/or slower lineage sorting in glacial refugia due to higher N_e), this data set does reveal that the allele frequencies of populations at the contact zone are intermediate between those at either side of it (Fig. 6b and c).

Assuming that secondary contact was established the latest in the mid-Holocene (~5 kya) (Doan et al. 2022), and assuming an average generation time of five years (Pérez-Espona et al. 2009; Borowik et al. 2016), genetic exchange has been ongoing for a minimum period of approximately 1,000 generations. Theory predicts that the hybridization cline widens over time, with the initial abrupt discontinuity gradually converting into an isolation-by-distance trend (Barton 1979; Barton and Hewitt 1985; Kierepka et al. 2023; Preckler-Quisquater et al. 2023).

In the absence of selective pressures, the width (w) of a hybridization cline is thought to be given by the following formula: $w = 2.51 \times d \times t^{0.5}$, in which d denotes dispersal distance (in km/generation) and in which t represents the time since first contact in number of generations (Arntzen et al. 2017; Stankowski et al. 2023). For red deer, evidence exists that long-distance dispersal out of home ranges is predominantly carried out by young males, which in later years might or might not become dominant stags (Kropil et al. 2015). Assuming a mean dispersal distance of 25 km per generation (Jarnemo 2008; Loe et al. 2009; Niedziałkowska et al. 2012), after a time period of 1,000 generations, the expected cline width approaches 1,500 km. This would imply that almost all red deer populations—except for the most remote ones (e.g. Spanish deer) or those separated by water barriers (e.g. Scandinavian, British, and Irish deer)—have been affected by introgression to some extent. If, instead, we assume a mean dispersal distance of only five km per generation (Nussey et al. 2005; Kamler et al. 2007; Pérez-Espona et al. 2008; Jarnemo 2011), the present-day expected cline width would not exceed 400 km.

The Y-chromosomal data set appears to be more in agreement with the second scenario. The Y chromosomes of a few stags (e.g. "Netherlands3," "Iran7," "Greece1," "Greece2," and "Greece3") cluster aberrantly (Figs. 4e and 5), but their presence likely dates back to documented human-made translocation events (Karaiskou et al. 2014; de Jong et al. 2020). Ignoring these individuals, the geographic boundaries between Y-chromosomal clades are less abrupt than the mtDNA boundaries (Fig. 4d and e), but only marginally so. For instance, the most westerly located male individual with a southeastern Y-chromosomal haplotype was found close to the German–Czech border (Fig. 4e), and even this occurrence might be due to a recent translocation (Krojerová-Prokešová et al. 2015).

Subspecies Designation

The renewed understanding of the demographic processes underlying the present-day red deer population structure, coupled with refined estimates of genetic distances, calls for a reevaluation of subspecies designations (Meiri et al. 2017).

Many currently accepted subspecies are closely related to the Central European red deer (*Cervus elaphus hippelaphus*: "FranceCentral," "NW-Europe," "NE-Europe1") and only distinct in terms of relative genetic distance (i.e. high F_{ST} but low D_{xy}). This includes Scottish deer (*Cervus elaphus scoticus*), Scandinavian deer (*Cervus elaphus atlanticus*: "Norway", "Sweden"), Sardinian deer (*C. e. corsicanus*), and Barbary deer (*C. e. barbarus*) (Zachos et al. 2016; Meiri et al. 2017). These relatively young lineages, which separated in the Late Pleistocene or Early Holocene, barely had time to accumulate private mutations. Instead of amassing novel genetic variation, and rather on the contrary, these lineages lost genetic variation owing to population size reductions, which made them genetically more homogenous (i.e. decrease of π) and more divergent in relative terms (increase of F_{ST}) but not more divergent in absolute terms (i.e. no increase of D_{xy}) (Figs. 1d and 3). On the other hand, insularity allowed these populations to preserve otherwise lost ancestral genetic identities, spanning a gradient of varying levels of introgression between Spanish deer and Sardinian/Barbary deer. In similar vein, Mesola deer (*C. e. italicus*) are unique not because of a long independent evolutionary history, but because of their hybrid origin.

Certain peripheral lineages, in contrast, do stand out because of a long independent history and are minimally affected by admixture, as shown by high absolute genetic distances relative to other populations (i.e. high D_{xy} values). This is true for the Carpathian deer (*Cervus elaphus pannoniensis*: "SE-Europe") in terms of X-chromosomal distances (Meiri et al. 2017), but only Caspian red deer (*Cervus elaphus maral*: "Caucasus" and "Turkey,") and Spanish red deer (*Cervus elaphus hispanicus*) meet this condition for X-chromosomal and autosomal data alike.

The coalescence time estimates of Y-chromosomal and mtDNA clades (Figs. 4 and 5) suggest that Spanish red deer descend from a lineage that migrated into western Europe before the LGM (Queirós et al. 2019; Doan et al. 2022; Fernández-Garcia et al. 2024; Carranza et al. 2024). Unlike their relatives north of the Pyrenees, they seem to have been minimally affected by subsequent gene flow from post-LGM waves and also not have contributed to the subsequent recolonization of northern latitudes (Queirós et al. 2019; Carranza et al. 2024). Testimony to their long independent evolutionary history is a suite of endemic phenotypic traits, including the absence of neck manes, a unique vocalization mechanism, and pronounced tongue protrusion (Geist 1999; Frey et al. 2012; Passilongo et al. 2013; Volodin et al. 2013; Della Libera et al. 2015). It thus appears that Spanish deer are among the many endemic varieties that characterize the flora and fauna of the Iberian Peninsula, therewith underlining its conservation value (Ramos et al. 2001; Nieto Feliner 2014; McDevitt et al. 2022; Çoraman et al. 2019).

Conclusion

Red deer autosomal population structure differs in several aspects from X-chromosomal population structure. The geographical ranges of X-chromosomal clusters are largely congruent with the parapatric distributions of Y-chromosomal and mtDNA haplotypes. We postulate that the inheritance properties of the X chromosome allowed it to better capture and subsequently preserve the genetic signals of population stratification in Late Glacial refugia.

Supplementary Material

Supplementary material is available at *Molecular Biology and Evolution* online.

Acknowledgments

For help with the collection of samples, we are grateful to Hugh Rose, Gerrit-Jan Spek, Vadim E. Sidorovich, and Rimvydas Juškaitis (eastern Europe), and Anders Jarnemo and Carl-Gustaf Thulin (Sweden). We thank Daniela Kalthoff (NHM Stockholm), Christiane Funk (MfN Berlin), and Irina Ruf and Kathrin Krohmann (Senckenberg Museum Frankfurt) for providing museum samples, even if not included in the final data set. We thank Elif Ertas and Katharina Geiss for help with DNA extractions.

Funding

This work was mainly supported by Hesse's funding program LOEWE and the Leibniz Association. Additional financial support included grants PRG1209 and TK215 from the Estonian Ministry of Education and Research (awarded to U.S.), grant MZE-RO0723 from the Ministry of the Czech Republic (awarded to L.B.), a grant from the Polish National Science Centre (2016/23/N/NZ8/03995), and the Danish National Research Foundation (grant DNRF173); the latter two awarded to M.S.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability

The short-read data have been deposited to the NCBI SRA repository (BioProject: PRJNA1187868). Genotype files used for population genetic analyses are available from the Dryad repository (under same title as manuscript), together with the bioinformatic workflows.

References

Arntzen JW, de Vries W, Canestrelli D, Martínez-Solano I. Hybrid zone formation and contrasting outcomes of secondary contact over transects in common toads. *Mol Ecol*. 2017:26(20):5663–5675. https://doi.org/10.1111/mec.14273.

Bailey G, Andersen SH, Maarleveld TJ. Denmark: mesolithic coastal landscapes submerged. In: Bailey G, Galanidou N, Peeters H, Jöns H, Mennenga M, editors. *The archaeology of Europe's drowned landscapes*. Cham: Springer International Publishing; 2020. p. 39–76.

Barton NH. The dynamics of hybrid zones. *Heredity (Edinb)*. 1979:43(3):341–359. https://doi.org/10.1038/hdy.1979.87.

Barton NH, Hewitt GM. Analysis of hybrid zones. *Annu Rev Ecol Syst*. 1985:16(1):113–148. https://doi.org/10.1146/annurev.es.16.110185.000553.

Bertl J, Ringbauer H, Blum MGB. Can secondary contact following range expansion be distinguished from barriers to gene flow? *PeerJ*. 2018:6:e5325. https://doi.org/10.7717/peerj.5325.

Bhagwat SA, Willis KJ. Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? *J Biogeogr*. 2008:35(3):464–482. https://doi.org/10.1111/j.1365-2699.2007.01861.x.

Bivand R, Keitt T, Rowlingson B, Pebesma E, Sumner M, Hijmans R, Baston D, Rouault E, Warmerdam F, Ooms J, *et al*. 2021. rgdal: Bindings for the "Geospatial" Data Abstraction Library. [accessed 2024 April]. https://CRAN.R-project.org/package=rgd.

Björck S. A review of the history of the Baltic Sea, 13.0-8.0 ka BP. *Quat Int*. 1995:27:19–40. https://doi.org/10.1016/1040-6182(94)00057-C.

Borowik T, Wawrzyniak P, Jędrzejewska B. Red deer (*Cervus elaphus*) fertility and survival of young in a low-density population subject to

predation and hunting. *J Mammal*. 2016:97(6):1671–1681. https://doi.org/10.1093/jmammal/gyw133.

Borowski Z, Świsłocka M, Matosiuk M, Mirski P, Krysiuk K, Czajkowska M, Borkowska A, Ratkiewicz M. Purifying selection, density blocking and unnoticed mitochondrial DNA diversity in the red deer, *Cervus elaphus*. *PLoS One*. 2016:11(9):e0163191. https://doi.org/10.1371/journal.pone.0163191.

Britton T, Oxelman B, Vinnersten A, Bremer K. Phylogenetic dating with confidence intervals using mean path lengths. *Mol Phylogenet Evol*. 2002:24(1):58–65. https://doi.org/10.1016/S1055-7903(02)00268-3.

Carden RF, McDevitt AD, Zachos FE, Woodman PC, O'Toole P, Rose H, Monaghan NT, Campana MG, Bradley DG, Edwards CJ. Phylogeographic, ancient DNA, fossil and morphometric analyses reveal ancient and modern introductions of a large mammal: the complex case of red deer (*Cervus elaphus*) in Ireland. *Quat Sci Rev*. 2012:42:74–84. https://doi.org/10.1016/j.quascirev.2012.02.012.

Carranza J, Pérez-González J, Anaya G, de Jong M, Broggini C, Zachos FE, McDevitt AD, Niedziałkowska M, Sykut M, Csányi S, *et al*. Genome-wide SNP assessment of contemporary European red deer genetic structure highlights the distinction of peripheral populations and the main admixture zones in Europe. *Mol Ecol*. 2024:33(18):e17508. https://doi.org/10.1111/mec.17508.

Carranza J, Salinas M, de Andrés D, Pérez-González J. Iberian red deer: paraphyletic nature at mtDNA but nuclear markers support its genetic identity. *Ecol Evol*. 2016:6(4):905–922. https://doi.org/10.1002/ece3.1836.

Çoraman E, Dietz C, Hempel E, Ghazaryan A, Levin E, Presetnik P, Zagmajster M, Mayer F. Reticulate evolutionary history of a Western Palaearctic Bat Complex explained by multiple mtDNA introgressions in secondary contacts. *J Biogeogr*. 2019:46(2):343–354. https://doi.org/10.1111/jbi.13509.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al*. The variant call format and VCFtools. *Bioinformatics*. 2011:27(15):2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

de Jong JF, van Hooft P, Megens H-J, Crooijmans RPMA, de Groot GA, Pemberton JM, Huisman J, Bartoš L, Iacolina L, van Wieren SE, *et al*. Fragmentation and translocation distort the genetic landscape of ungulates: red deer in the Netherlands. *Front Ecol Evol*. 2020:8:535715. https://doi.org/10.3389/fevo.2020.535715.

de Jong MJ, de Jong JF, Hoelzel AR, Janke A. Sambar: an R package for fast, easy and reproducible population-genetic analyses of biallelic SNP data sets. *Mol Ecol Resour*. 2021:21(4):1369–1379. https://doi.org/10.1111/1755-0998.13339.

de Jong MJ, Niamir A, Wolf M, Kitchener AC, Lecomte N, Seryodkin IV, Fain SR, Hagen SB, Saarma U, Janke A. Range-wide whole-genome resequencing of the brown bear reveals drivers of intraspecies divergence. *Commun Biol*. 2023:6(1):153. https://doi.org/10.1038/s42003-023-04514-w.

de Jong MJ, van Oosterhout C, Hoelzel R, Janke A. Calculating and interpreting FST in the genomics era. bioRxiv 614506. https://doi.org/10.1101/2024.09.24.614506, 24 September 2024, preprint: not peer reviewed.

Della Libera M, Passilongo D, Reby D. Acoustics of male rutting roars in the endangered population of Mesola red deer *Cervus elaphus italicus*. *Mamm Biol*. 2015:80(5):395–400. https://doi.org/10.1016/j.mambio.2015.05.001.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011:43(5):491–498. https://doi.org/10.1038/ng.806.

Doan K, Niedziałkowska M, Stefaniak K, Sykut M, Jędrzejewska B, Ratajczak-Skrzatek U, Piotrowska N, Ridush B, Zachos FE, Popović D, *et al*. Phylogenetics and phylogeography of red deer mtDNA lineages during the last 50 000 years in Eurasia. *Zool J Linn Soc*. 2022:194(2):431–456. https://doi.org/10.1093/zoolinnean/zlab025.

Doan K, Zachos FE, Wilkens B, Vigne J-D, Piotrowska N, Stanković A, Jędrzejewska B, Stefaniak K, Niedziałkowska M. Phylogeography of the Tyrrhenian red deer (*Cervus elaphus corsicanus*) resolved using ancient DNA of radiocarbon-dated subfossils. *Sci Rep*. 2017:7(1):2331. https://doi.org/10.1038/s41598-017-02359-y.

Eckert AJ, Carstens BC. Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Mol Phylogenet Evol*. 2008:49(3):832–842. https://doi.org/10.1016/j.ympev.2008.09.008.

Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016:32(19):3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

Fernández-Garcia M, Pederzani S, Britton K, Agudo-Perez L, Cicero A, Geiling J, Daura J, Sanz-Borras M, Marin-Arroyo AB. Palaeoecology of ungulates in northern Iberia during the Late Pleistocene through isotopic analysis of teeth. *Biogeosci Discuss*. 2024;21(19):4413–4437. https://doi.org/10.5194/bg-21-4413-2024.

Fickel J, Bubliy OA, Stache A, Noventa T, Jirsa A, Heurich M. Crossing the border? Structure of the red deer (*Cervus elaphus*) population from the Bavarian–Bohemian forest ecosystem. *Mamm Biol*. 2012:77(3):211–220. https://doi.org/10.1016/j.mambio.2011.11.005.

Fordham DA, Saltré F, Haythorne S, Wigley TML, Otto-Bliesner BL, Chan KC, Brook BW. PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography*. 2017:40(11):1348–1358. https://doi.org/10.1111/ecog.03031.

Frantz AC, Hamann J-L, Klein F. Fine-scale genetic structure of red deer (*Cervus elaphus*) in a French temperate forest. *Eur J Wildl Res*. 2008:54(1):44–52. https://doi.org/10.1007/s10344-007-0107-1.

Frey R, Volodin I, Volodina E, Carranza J, Torres-Porras J. Vocal anatomy, tongue protrusion behaviour and the acoustics of rutting roars in free-ranging Iberian red deer stags (*Cervus elaphus hispanicus*). *J Anat*. 2012:220(3):271–292. https://doi.org/10.1111/j.1469-7580.2011.01467.x.

Frichot E, François O. LEA: an R package for landscape and ecological association studies. *Methods Ecol Evol*. 2015:6(8):925–929. https://doi.org/10.1111/2041-210X.12382.

Geist V. *Deer of the world: their evolution, behaviour and ecology*. 1st ed. Shrewsbury: Swan Hill Press; 1999.

Geptner VG; Nasimovich. 1988. Mammals of the Soviet Union v. 2, pt. 1a (1998). [accessed 2024 April]. https://library.si.edu/digital-library/book/mammalsofsov211998gept.

Haanes H, Røed KH, Perez-Espona S, Rosef O. Low genetic variation support bottlenecks in Scandinavian red deer. *Eur J Wildl Res*. 2011:57(6):1137–1150. https://doi.org/10.1007/s10344-011-0527-9.

Hajji GM, Zachos FE, Charfi-Cheikrouha F, Hartl GB. Conservation genetics of the imperilled Barbary red deer in Tunisia. *Anim Conserv*. 2007:10(2):229–235. https://doi.org/10.1111/j.1469-1795.2007.00098.x.

Hewitt GM. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc Lond*. 1996:58(3):247–276. https://doi.org/10.1006/bijl.1996.0035.

Hewitt GM. Post-glacial re-colonization of European biota. *Biol J Linn Soc Lond*. 1999:68(1-2):87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x.

Hijmans RJ. 2019. raster: Geographic Data Analysis and Modeling. [accessed 2024 April]. https://CRAN.R-project.org/package=raster.

Hijmans RJ, Phillips S, Leathwick J, Elith J. 2020. dismo: Species Distribution Modeling. [accessed 2024 April]. https://CRAN.R-project.org/package=dismo.

Höglund J, Cortazar-Chinarro M, Jarnemo A, Thulin C-G. Genetic variation and structure in Scandinavian red deer (*Cervus elaphus*): influence of ancestry, past hunting, and restoration management. *Biol J Linn Soc Lond*. 2013:109(1):43–53. https://doi.org/10.1111/bij.12049.

Holsinger KE, Weir BS. Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nat Rev Genet*. 2009:10(9):639–650. https://doi.org/10.1038/nrg2611.

Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. *Genetics*. 1992:132(2):583–589. https://doi.org/10.1093/genetics/132.2.583.

Jarnemo A. Seasonal migration of male red deer (*Cervus elaphus*) in southern Sweden and consequences for management. *Eur J Wildl Res*. 2008:54(2):327–333. https://doi.org/10.1007/s10344-007-0154-7.

Jarnemo A. Male red deer (*Cervus elaphus*) dispersal during the breeding season. *J Ethol*. 2011:29(2):329–336. https://doi.org/10.1007/s10164-010-0262-9.

Jombart T. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008:24(11):1403–1405. https://doi.org/10.1093/bioinformatics/btn129.

Kamler JF, Jędrzejewska B, JĘdrzejewski W. Factors affecting daily ranges of red deer *Cervus elaphus* in Białowieża Primeval Forest, Poland. *Acta Theriol*. 2007:52(2):113–118. https://doi.org/10.1007/BF03194206.

Karaiskou N, Tsakogiannis A, Gkagkavouzis K, Operator of Parnitha National Park; Papika S, Latsoudis P, Kavakiotis I, Pantis J, Abatzopoulos TJ, Triantaphyllidis C, *et al*. Greece: a Balkan subrefuge for a remnant red deer (*Cervus elaphus*) population. *J Hered*. 2014:105(3):334–344. https://doi.org/10.1093/jhered/esu007.

Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. Climatologies at high resolution for the earth's land surface areas. *Sci Data*. 2017:4(1):170122. https://doi.org/10.1038/sdata.2017.122.

Kierepka EM, Preckler-Quisquater S, Reding DM, Piaggio AJ, Riley SPD, Sacks BN. Genomic analyses of gray fox lineages suggest ancient divergence and secondary contact in the southern Great Plains. *J Hered*. 2023:114(2):110–119. https://doi.org/10.1093/jhered/esac060.

Krojerová-Prokešová J, Barančeková M, Koubek P. Admixture of Eastern and Western European red deer lineages as a result of postglacial recolonization of the Czech Republic (central Europe). *J Hered*. 2015:106(4):375–385. https://doi.org/10.1093/jhered/esv018.

Kropil R, Smolko P, Garaj P. Home range and migration patterns of male red deer *Cervus elaphus* in Western Carpathians. *Eur J Wildl Res*. 2015:61(1):63–72. https://doi.org/10.1007/s10344-014-0874-4.

Laan M, Wiebe V, Khusnutdinova E, Remm M, Pääbo S. X-chromosome as a marker for population history: linkage disequilibrium and haplotype study in Eurasian populations. *Eur J Hum Genet*. 2005:13(4):452–462. https://doi.org/10.1038/sj.ejhg.5201340.

Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 2011a:27(21):2987–2993. https://doi.org/10.1093/bioinformatics/btr509.

Li H. Tabix: fast retrieval of sequence features from generic TAB-delimited files. *Bioinformatics*. 2011b:27(5):718–719. https://doi.org/10.1093/bioinformatics/btq671.

Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009:25(14):1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009:25(16):2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Liu J, Shen Q, Bao H. Comparison of seven SNP calling pipelines for the next-generation sequencing data of chickens. *PLoS One*. 2022:17(1):e0262574. https://doi.org/10.1371/journal.pone.0262574.

Loe LE, Mysterud A, Veiberg V, Langvatn R. Negative density-dependent emigration of males in an increasing red deer population. *Proc R Soc Lond B Biol Sci*. 2009:276(1667):2581–2587. https://doi.org/10.1098/rspb.2009.0224.

Ludt CJ, Schroeder W, Rottmann O, Kuehn R. Mitochondrial DNA phylogeography of red deer (*Cervus elaphus*). *Mol Phylogenet Evol*. 2004:31(3):1064–1083. https://doi.org/10.1016/j.ympev.2003.10.003.

Mackiewicz P, Matosiuk M, Świsłocka M, Zachos FE, Hajji GM, Saveljev AP, Seryodkin IV, Farahvash T, Rezaei HR, Torshizi RV, *et al*. Phylogeny and evolution of the genus Cervus (Cervidae, Mammalia) as revealed by complete mitochondrial genomes. *Sci Rep*. 2022:12(1):16381. https://doi.org/10.1038/s41598-022-20763-x.

Masonbrink RE, Alt D, Bayles DO, Boggiatto P, Edwards W, Tatum F, Williams J, Wilson-Welder J, Zimin A, Severin A, *et al*. A pseudomolecule assembly of the Rocky Mountain elk genome. *PLoS One*. 2021:16(4):e0249899. https://doi.org/10.1371/journal.pone.0249899.

McDevitt AD, Coscia I, Browett SS, Ruiz-González A, Statham MJ, Ruczyńska I, Roberts L, Stojak J, Frantz AC, Norén K, *et al*. Next-generation phylogeography resolves post-glacial colonization patterns in a widespread carnivore, the red fox (*Vulpes vulpes*), in Europe. *Mol Ecol*. 2022:31(3):993–1006. https://doi.org/10.1111/mec.16276.

McFarlane SE, Hunter DC, Senn HV, Smith SL, Holland R, Huisman J, Pemberton JM. Increased genetic marker density reveals high levels of admixture between red deer and introduced Japanese sika in Kintyre, Scotland. *Evol Appl*. 2020:13(2):432–441. https://doi.org/10.1111/eva.12880.

Meiri M, Kosintsev P, Conroy K, Meiri S, Barnes I, Lister A. Subspecies dynamics in space and time: a study of the red deer complex using ancient and modern DNA and morphology. *J Biogeogr*. 2017:1–14. https://doi.org/10.1111/jbi.13124.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020:37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.

Nabholz B, Glémin S, Galtier N. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. *Mol Biol Evol*. 2008:25(1):120–130. https://doi.org/10.1093/molbev/msm248.

Niedziałkowska M, Doan K, Górny M, Sykut M, Stefaniak K, Piotrowska N, Jędrzejewska B, Ridush B, Pawełczyk S, Mackiewicz P, *et al*. Winter temperature and forest cover have shaped red deer distribution in Europe and the Ural Mountains since the Late Pleistocene. *J Biogeogr*. 2021:48(1):147–159. https://doi.org/10.1111/jbi.13989.

Niedziałkowska M, Fontaine MC, Jędrzejewska B. Factors shaping gene flow in red deer (*Cervus elaphus*) in seminatural landscapes of central Europe. *Can J Zool*. 2012:90(2):150–162. https://doi.org/10.1139/z11-122.

Niedziałkowska M, Jędrzejewska B, Honnen A-C, Otto T, Sidorovich VE, Perzanowski K, Skog A, Hartl GB, Borowik T, Bunevich AN, *et al*.Molecular biogeography of red deer *Cervus elaphus* from eastern Europe: insights from mitochondrial DNA sequences. *Acta Theriol (Warsz)*. 2011:56(1):1–12. https://doi.org/10.1007/s13364-010-0002-0.

Nieto Feliner G. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspect Plant Ecol Evol Syst*. 2014:16(5):265–278. https://doi.org/10.1016/j.ppees.2014.07.002.

Nilsson B, Hansson A, Sjöström A. Sweden: submerged landscapes of the early mesolithic. In: Bailey G, Galanidou N, Peeters H, Jöns H, Mennenga M, editors. *The archaeology of Europe's drowned landscapes*. Cham: Springer International Publishing; 2020. p. 77–93.

Nussey DH, Coltman DW, Coulson T, Kruuk LEB, Donald A, Morris SJ, Clutton-Brock TH, Pemberton J. Rapidly declining fine-scale spatial genetic structure in female red deer. *Mol Ecol*. 2005:14(11):3395–3405. https://doi.org/10.1111/j.1365-294X.2005.02692.x.

Nussey DH, Pemberton J, Donald A, Kruuk LEB. Genetic consequences of human management in an introduced island population of red deer (*Cervus elaphus*). *Heredity (Edinb)*. 2006:97(1):56–65. https://doi.org/10.1038/sj.hdy.6800838.

Ortiz EM. 2019. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. [accessed 2024 April]. https://zenodo.org/records/2540861.

Paradis E, Schliep K. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 2018:35(3):526–528. https://doi.org/10.1093/bioinformatics/bty633.

Passilongo D, Reby D, Carranza J, Apollonio M. Roaring high and low: composition and possible functions of the Iberian stag's vocal repertoire. *PLoS One*. 2013:8(5):e63841. https://doi.org/10.1371/journal.pone.0063841.

Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. Ancient admixture in human history. *Genetics*. 2012:192(3):1065–1093. https://doi.org/10.1534/genetics.112.145037.

Pérez-Espona S, Pérez-Barbería FJ, Goodall-Copestake WP, Jiggins CD, Gordon IJ, Pemberton JM. Genetic diversity and population structure of Scottish Highland red deer (*Cervus elaphus*) populations: a mitochondrial survey. *Heredity (Edinb)*. 2009:102(2):199–210. https://doi.org/10.1038/hdy.2008.111.

Pérez-Espona S, Pérez-Barbería FJ, Mcleod JE, Jiggins CD, Gordon IJ, Pemberton JM. Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). *Mol Ecol*. 2008:17(4):981–996. https://doi.org/10.1111/j.1365-294X.2007.03629.x.

Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. *Ecol Modell*. 2006:190(3-4):231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026.

Preckler-Quisquater S, Kierepka EM, Reding DM, Piaggio AJ, Sacks BN. Can demographic histories explain long-term isolation and recent pulses of asymmetric gene flow between highly divergent grey fox lineages? *Mol Ecol*. 2023:32(19):5323–5337. https://doi.org/10.1111/mec.17105.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007:81(3):559–575. https://doi.org/10.1086/519795.

Queirós J, Acevedo P, Santos JPV, Barasona J, Beltran-Beck B, González-Barrio D, Armenteros JA, Diez-Delgado I, Boadella M, de Mera IF, *et al*. Red deer in Iberia: molecular ecological studies in a southern refugium and inferences on European postglacial colonization history. *PLoS One*. 2019:14(1):e0210282. https://doi.org/10.1371/journal.pone.0210282.

Ramos MA, Lobo JM, Esteban M. Ten years inventorying the Iberian fauna: results and perspectives. *Biodivers Conserv*. 2001:10(1):19–28. https://doi.org/10.1023/A:1016658804566.

R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. *Nature*. 2009:461(7263):489–494. https://doi.org/10.1038/nature08365.

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

Shi M-M, Chen X-Y. Leading-edge populations do not show low genetic diversity or high differentiation in a wind-pollinated tree. *Popul Ecol*. 2012:54(4):591–600. https://doi.org/10.1007/s10144-012-0332-7.

Skog A, Zachos FE, Rueness EK, Feulner PGD, Mysterud A, Langvatn R, Lorenzini R, Hmwe SS, Lehoczky I, Hartl GB, *et al*. Phylogeography of red deer (*Cervus elaphus*) in Europe. *J Biogeogr*. 2009:36(1):66–77. https://doi.org/10.1111/j.1365-2699.2008.01986.x.

Sommer RS, Nadachowski A. Glacial refugia of mammals in Europe: evidence from fossil records. *Mamm Rev*. 2006:36(4):251–265. https://doi.org/10.1111/j.1365-2907.2006.00093.x.

Sommer RS, Zachos FE. Fossil evidence and phylogeography of temperate species: 'glacial refugia' and post-glacial recolonization. *J Biogeogr*. 2009:36(11):2013–2020. https://doi.org/10.1111/j.1365-2699.2009.02187.x.

Sommer RS, Zachos FE, Street M, Jöris O, Skog A, Benecke N. Late Quaternary distribution dynamics and phylogeography of the red deer (*Cervus elaphus*) in Europe. *Quat Sci Rev*. 2008:27(7-8):714–733. https://doi.org/10.1016/j.quascirev.2007.11.016.

Stankowski S, Chase MA, McIntosh H, Streisfeld MA. Integrating top-down and bottom-up approaches to understand the genetic architecture of speciation across a monkeyflower hybrid zone. *Mol Ecol*. 2023:32(8):2041–2054. https://doi.org/10.1111/mec.16849.

Valnisty AA, Homel KV, Kheidorova EE, Molchan VO, Nikiforov MY. Between the lines: mitochondrial lineages in the heavily managed red deer population of Belarus. *Mamm Biol*. 2024:104(2):205–214. https://doi.org/10.1007/s42991-023-00397-w.

Vigne J-D. Zooarchaeology and the biogeographical history of the mammals of Corsica and Sardinia since the last ice age. *Mamm Rev*. 1992:22(2):87–96. https://doi.org/10.1111/j.1365-2907.1992.tb00124.x.

Volodin I, Volodina E, Frey R, Carranza J, Torres-Porras J. Spectrographic analysis points to source–filter coupling in rutting roars of Iberian red deer. *acta ethol*. 2013:16(1):57–63. https://doi.org/10.1007/s10211-012-0133-1.

Volodin IA, Nahlik A, Tari T, Frey R, Volodina EV. Rutting roars in native Pannonian red deer of Southern Hungary and the evidence of acoustic divergence of male sexual vocalization between Eastern and Western European red deer (*Cervus elaphus*). *Mamm Biol*. 2019:94:54–65. https://doi.org/10.1016/j.mambio.2018.10.009.

Webster TH, Wilson Sayres MA. Genomic signatures of sex-biased demography: progress and prospects. *Curr Opin Genet Dev*. 2016:41:62–71. https://doi.org/10.1016/j.gde.2016.08.002.

Xia Q, Guo Y, Zhang Z, Li D, Xuan Z, Li Z, Dai F, Li Y, Cheng D, Li R, *et al*. Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (Bombyx). *Science*. 2009:326(5951):433–436. https://doi.org/10.1126/science.1176620.

Zachos FE, Frantz AC, Kuehn R, Bertouille S, Colyn M, Niedziałkowska M, Pérez-González J, Skog A, Sprěm N, Flamand M-C. Genetic structure and effective population sizes in European red deer (*Cervus elaphus*) at a continental scale: insights from microsatellite DNA. *J Hered*. 2016:107(4):318–326. https://doi.org/10.1093/jhered/esw011.

Zachos FE, Hajji GM, Hmwe SS, Hartl GB, Lorenzini R, Mattioli S. Population viability analysis and genetic diversity of the endangered red deer *Cervus elaphus* population from Mesola, Italy. *Wildlife Biol*. 2009:15(2):175–186. https://doi.org/10.2981/07-075.

Zink RM. Drawbacks with the use of microsatellites in phylogeography: the song sparrow *Melospiza melodia* as a case study. *J Avian Biol*. 2010:41(1):1–7. https://doi.org/10.1111/j.1600-048X.2009.04903.x.