

Unexpected population genetic structure of European roe deer in Poland: an invasion of the mtDNA genome from Siberian roe deer

MACIEJ MATOSIUK,* ANETTA BORKOWSKA,* MAGDALENA ŚWISŁOCKA,* PAWEŁ MIRSKI,* ZBIGNIEW BOROWSKI,† KAMIL KRYSIUK,‡ ALEKSEY A. DANILKIN,§ ELENA Y. ZVYCHAYNAYA,§ ALEXANDER P. SAVELJEV¶ and MIROŚŁAW RATKIEWICZ*

*Institute of Biology, University of Białystok, Świerkowa 20B, 15-950 Białystok, Poland, †Department of Forest Ecology, Forest Research Institute, Braci Leśnej 3, Sękocin Stary, 05-090 Raszyn, Poland, ‡Department of Genetics and Animal Breeding, Faculty of Animal Science, Warsaw University of Life Sciences, Ciszewskiego 8, 02-786 Warsaw, Poland, §Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninskii prospekt 33, 119071 Moscow, Russia, ¶Department of Animal Ecology, Russian Research Institute of Game Management and Fur Farming, Russian Academy of Sciences, Preobrazhenskaya 79, 610000 Kirov, Russia

Abstract

Introgressive hybridization is a widespread evolutionary phenomenon which may lead to increased allelic variation at selective neutral loci and to transfer of fitness-related traits to introgressed lineages. We inferred the population genetic structure of the European roe deer (*Capreolus capreolus*) in Poland from mitochondrial (CR and cyt *b*) and sex-linked markers (ZFX, SRY, DBY4 and DBY8). Analyses of CR mtDNA sequences from 452 individuals indicated widespread introgression of Siberian roe deer (*C. pygargus*) mtDNA in the European roe deer genome, 2000 km from the current distribution range of *C. pygargus*. Introgressed individuals constituted 16.6% of the deer studied. Nearly 75% of them possessed haplotypes belonging to the group which arose 23 kyr ago and have not been detected within the natural range of Siberian roe deer, indicating that majority of present introgression has ancient origin. Unlike the mtDNA results, sex-specific markers did not show signs of introgression. Species distribution modelling analyses suggested that *C. pygargus* could have extended its range as far west as Central Europe after last glacial maximum. The main hybridization event was probably associated with range expansion of the most abundant European roe deer lineage from western refugia and took place in Central Europe after the Younger Dryas (10.8–10.0 ka BP). Initially, introgressed mtDNA variants could have spread out on the wave of expansion through the mechanism of gene surfing, reaching high frequencies in European roe deer populations and leading to observed asymmetrical gene flow. Human-mediated introductions of *C. pygargus* had minimal effect on the extent of mtDNA introgression.

Keywords: genetic structure, hybridization, mtDNA introgression, phylogenetic discordance, roe deer, species distribution modelling

Received 25 June 2013; revision received 31 March 2014; accepted 31 March 2014

Introduction

Introgressive hybridization usually occurs between closely related, evolutionarily young taxa with a sym-

patric or parapatric distribution (Irwin *et al.* 2009; Cabria *et al.* 2011). Species that do not hybridize at present may have done so in the past, possibly leaving genetic signal (Melo-Ferreira *et al.* 2005; Osada *et al.* 2010). Even a small amount of interbreeding can lead to significant levels of introgression (Currat *et al.* 2008), so gene introgression can occur between species

Correspondence: Maciej Matosiuk, Fax: (+48) 85 745 73 01; E-mail: maciej.matosiuk@gmail.com

not known to hybridize extensively (Chan & Levin 2005).

Depending on the inheritance mode, some types of genetic markers introgress more readily than others. In animals, maternally inherited mitochondrial DNA (mtDNA) often displays a substantially higher rate of introgression than biparentally and paternally inherited markers, leading to an asymmetric distribution of the different marker types undergoing introgression (Currat *et al.* 2008; Petit & Excoffier 2009). Toews & Brelsford (2012) underlined the role of differential demography, sex-biased dispersal and behavioural mating preference in creating biogeographical disparities. Asymmetric introgression of genetic markers can also arise as a by-product of hybrid zone movement, as a moving hybrid zone can, hypothetically, leave a tail of introgressed alleles (Buggs 2007). The paucity of available data (Boratynski *et al.* 2011; Derr *et al.* 2012) prevents us from making unequivocal statements about the adaptive properties of mtDNA, which usually is seen as evolutionarily neutral (Avice 2004). However, recent studies have shown that the sequence variation of human

mtDNA worldwide was shaped by climate (Balloux *et al.* 2009) and point to the strong effect of mitochondrial genetic variants on male fertility components and reproductive fitness (Smith *et al.* 2010; Innocenti *et al.* 2011).

The European (*Capreolus capreolus*) and Siberian (*C. pygargus*) roe deer are young sibling species from *Cervidae* family inhabiting the Northern Hemisphere from western Europe to the Pacific coast with a narrow contact zone around the banks of the rivers Volga and lower Don in Russia (Danilkin & Hewison 1996; Randi *et al.* 1998; Fig. 1). Researchers have separated them as independent species due to the lack of evidence for interspecific mixing, their barely overlapping distributions together with a discontinuity of morphometric traits and differences in karyotype and in behavioural ecology (Hewison & Danilkin 2001). So far, there has been no evidence of their interbreeding and hybridization in nature, although experimental crossings have been performed (Danilkin & Hewison 1996). Phylogeographical studies of *C. capreolus*, very limited in the number of samples from Central and Eastern Europe,

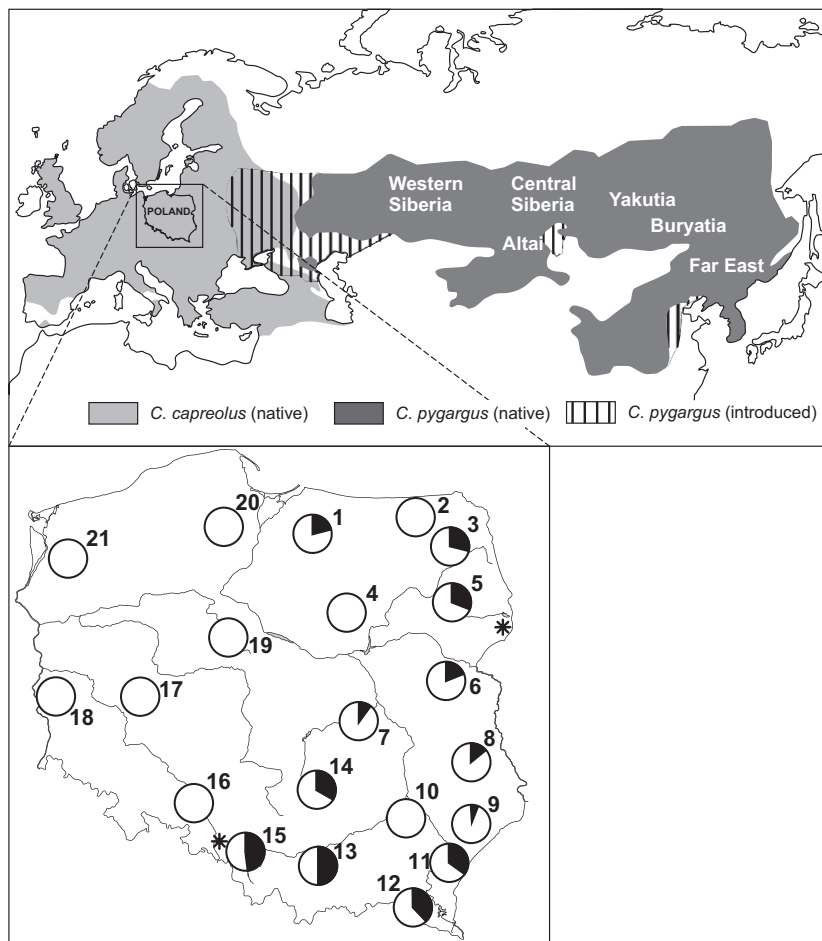


Fig. 1 Distribution of European *Capreolus capreolus* and Siberian *C. pygargus* roe deer in Eurasia (Gonzalez & Tsytulina 2008; Lovari *et al.* 2008) and sampling locations in Poland. Diagrams indicate the proportion of mtDNA CR haplotypes belonging to European (open circle) and Siberian (black circle) roe deer lineages in the populations. Names of sampling localities and sample size in Table S1, Supporting information. *, Place of Siberian roe deer introduction in Poland.

also did not show any signs of mtDNA introgression (Vernesi *et al.* 2002; Randi *et al.* 2004). However, more recent analysis of the European roe deer population from Moscow area has revealed extensive introgression of Siberian roe deer mtDNA (Zvychnayaya *et al.* 2011b), implying that the earlier considerations about possibility of these two species interbreeding were too cautious. Unexpectedly, our population genetic study of European roe deer indicated that introgression of *C. pygargus* mitochondrial genome is widespread in Poland, nearly 2000 km westward from the contact zone of the species.

Besides natural processes, human activities have an accumulative impact on the genetic integrity of populations and species through translocation and introduction of foreign species, often bringing them into contact with previously isolated groups (Crispo *et al.* 2011). These activities, often involving species of commercial value and game species, multiply hybridization events and increase preferential gene flow, blurring the evolutionarily shaped large-scale phylogenetic structure of native species (Excoffier *et al.* 2009). The roe deer also experienced large-scale human-caused alterations. During the early twentieth century, both species underwent major declines in numbers and shifts in their distribution, mainly due to deforestation and overhunting. Changes in management practices enabled these two species to restore their continuous ranges naturally and re-establish a previously existing contact zone (Danilkin & Hewison 1996). These important changes were accompanied by introductions and translocations of nearly 3000 animals of both species, mostly in western Russia and Eastern Europe (Pavlov 1999), but only one introduction of the Siberian roe deer, near Stavropol in Russia, is known to have been successful (Danilkin & Hewison 1996). There are only two available records of introduction of the Siberian roe deer in Poland. Eleven animals were released in the Bialowieza Primeval Forest in 1891 (Pavlov 1999) and several in Silesia in 1909 (Gleiss 1967).

In this study, we address the consequent questions about the possible character of hybridization event(s), focusing on genetic signatures in mtDNA and sex-linked markers that may quantify the impact of natural or human-mediated hybridization on present genetic structure of European roe deer populations in Poland. We also investigate possible mechanisms underlying widespread mtDNA introgression and the occurrence of Siberian roe deer mtDNA so far beyond the present contact zone of the species. We have performed species distribution modelling to infer past range dynamics of both species and the possibility of their co-occurrence (simultaneous or consecutive) in Central Europe. Ecological niche modelling procedures gave insights into potential colonization (*C. pygargus*) and recolonization

(*C. capreolus*) routes in Central Europe, thus allowing us to revisit phylogeographical patterns proposed for European roe deer (Randi *et al.* 2004; Sommer *et al.* 2009).

Material and methods

Sample collection

The study covered the whole area of Poland (Fig. 1). Muscle or skin samples from 452 legally shot European roe deer (*C. capreolus*) were received from hunters during the 2008–2010 hunting seasons. Samples consisting of 10–37 individuals per population were collected from 21 populations (Table S1, Supporting information). The reference samples for this study came from 16 Siberian roe deer (*C. pygargus*) from western and central Siberia. Species of the specimens were identified by morphological traits (Danilkin & Hewison 1996).

Laboratory procedures

Total DNA was extracted from frozen tissues using the DNeasy Blood & Tissue Kit (Qiagen). The mtDNA control region (CR; 525 bp) was amplified with primers PRO and PHE (Wood & Phua 1996) for all individuals studied. The complete cytochrome *b* gene (cyt *b*; 1140 bp) was amplified using primers ML103 and MH104 (Chikuni *et al.* 1995) for 184 European roe deer and 16 Siberian roe deer. Roe deer were also sequenced at four sex-linked nuclear markers. Sex of the individuals was assessed using ZFX/ZFY gene system. The ninth intron of the ZFX gene was amplified using primers ZFX_F (5'-ATAGCATGGGCAGCAGCTTA-3') and ZFX_R (5'-TCCATCAGAATTATTACCTAACAATGTGCA-3') for 157 European roe deer (125 females and 32 males) and eight Siberian roe deer females. Y chromosome-specific fragments were amplified for 58 males: 51 European and seven Siberian roe deer. We used primers SRY_F (5'-TGTTTCAGAGTATTGAACGATGATGT-3') and SRY_R (5'-TATTGAAAATAAGCGCAAGAAATCCAGGCT-3') to amplify a 621-bp fragment of the SRY gene and the primers developed by Hellborg & Ellegren (2003) to amplify the DBY4 and DBY8 fragments of Y chromosome introns. ZFX and SRY gene primers were designed using Primer3 (Rozen & Skaletsky 1999) and FastPCR (Kalendar *et al.* 2009) software. The details of the PCRs and specific amplification procedures are given in Appendix S1 (Supporting information). DNA sequences were aligned with BIOEDIT v7.0.4 (Hall 1999) and revised manually. ZFX alleles were assessed using the algorithms provided in PHASE as implemented in DNASP v5.10 (Librado & Rozas 2009).

Population genetics and identification of introgressed individuals

All the roe deer occurring in Poland are classified by morphology as European roe deer, *Capreolus capreolus* (Pucek 1981). Individuals possessing Siberian roe deer mtDNA were designated as introgressed individuals, and their frequency was estimated in each population. The number of polymorphic sites and haplotypes (N_h), haplotype diversity (h) and nucleotide diversity (π) were calculated for species-specific mtDNA lineages (separately for CR and *cyt b*) and for concatenated mtDNA sequences, using ARLEQUIN v3.5.1.2 (Excoffier & Lischer 2010).

Analysis of molecular variance (AMOVA) was used to assess the partitioning of genetic differences among populations and species-specific CR mtDNA lineages (Excoffier *et al.* 1992). Spatial analysis of molecular variance (SAMOVA; Dupanloup *et al.* 2002) was carried out to identify genetically distinct populations based on CR mtDNA haplotype data for comparison with their geographical location. AMOVA and SAMOVA analyses were conducted using only the samples from Poland (Appendix S1, Supporting information).

Phylogenetic and phylogeographical analyses

Phylogenetic analyses were conducted to assess whether any of the European roe deer from Poland had introgressed mtDNA or sex-linked genes from Siberian roe deer. Phylogenetic relationships within the genus *Capreolus* were reconstructed using the Bayesian approach implemented in BEAST v1.7.2 (Drummond *et al.* 2012) for two data sets: control region alone and concatenated sequences of CR and *cyt b*. We used a single accession of each haplotype found in this study, additional mtDNA sequences from *C. capreolus* and *C. pygargus* downloaded from GenBank and the sequence of the water deer (*Hydropotes inermis*, EU315254.1) as out-group. In these analyses, we used nucleotide substitution models selected under the Akaike information criterion (AIC) implemented in jMODELTEST 0.1.1 (Posada 2008): HKY+I+G for CR and HKY+I for *cyt b*. For our analyses, we used the Yule process tree prior (Gernhard 2008). Each Bayesian analysis run consisted of an MCMC chain with 30 000 000 iterations sampled every 3000th generation; the first 10% of the iterations were the burn-in. The sampled trees were annotated in TreeAnnotator v1.7.2 (BEAST software) and visualized in Figtree v1.3.1 (<http://tree.bio.ed.ac.uk/software/figtree>).

To explore the haplotype genealogies for mtDNA markers, ZFX and Y-specific chromosome sequences, median joining networks were constructed in NETWORK

v4.6.1.0 (<http://www.fluxus-engineering.com>). To check whether CR mtDNA sequences were noted previously in the literature, we compared our data with sequences of *C. capreolus* and *C. pygargus* CR mtDNA obtained from GenBank (Table S1, Supporting information).

To test the hypothesis of recent population growth, we used Fu's (1997) F_S -statistics, the raggedness index and the pairwise mismatch distribution method. The statistics were calculated in ARLEQUIN. Departure from a model of sudden expansion was tested for CR mtDNA clades of *C. capreolus* and pooled *C. pygargus* mtDNA haplotypes found in the European roe deer genome in Poland. This inferred the expansion of European roe deer possessing *C. capreolus* or *C. pygargus* haplotypes. The squared differences (SSD) between expected and observed mismatch distributions were summed, and parametric bootstrapping with 1000 replicates was used (Excoffier 2004). For a haplogroup for which an expansion model could not be rejected and the distribution was unimodal, we estimated the expansion time from the mismatch distribution (τ) as $t = \tau/2u$, where t is the expansion time in generations and u is the substitution rate per generation for the whole sequence (Rogers 1995). We assumed a mutation rate of 0.04–0.08 substitutions $\times 10^{-6}$ (Randi *et al.* 1998) $\times n$ nucleotides per generation time $g = 4$ years (Nilssen *et al.* 2009). To test whether species-specific *cyt b* amino acid sequences show signs of adaptive protein evolution, we performed the McDonald–Kreitman test (McDonald & Kreitman 1991) in DNASP v5.10.

We used the Bayesian approach implemented in BEAST to calculate the time of divergence from the most recent common ancestor (TMRCA), for mtDNA CR and *cyt b* separately. We applied a strict molecular clock with previously reported mutation rates for the control region (4% and 8% per Ma) and cytochrome *b* (2% per Ma) in ungulates (Randi *et al.* 1998). The computations proceeded as for phylogenetic relationship reconstruction described above. The results of independent runs were combined using LOGCOMBINER v1.7.2 (BEAST software).

Species distribution modelling

Both species distribution modelling was conducted to estimate the probability of their co-occurrence in Central Europe in the past (simultaneously or consecutively). Maximum entropy method using MAXENT 3.3.3k software (Phillips *et al.* 2006; Elith *et al.* 2011) was used to build the models of distribution at the present conditions, last glacial maximum (LGM, c. 21 kyr BP) and last interglacial (LIG, c. 120–140 kyr BP; details in Appendix S1, Supporting information).

Distribution of species-specific mtDNA vs. environmental variables

To investigate the possible impact of environmental factors on the distribution of roe deer possessing mtDNA of *C. pygargus* in Poland, we performed multivariate multiple regression using DISTLM v5 (Anderson 2004). The frequencies of roe deer possessing *C. pygargus* mtDNA haplotypes for all 21 sampling locations were used to construct the response matrix and tested against the following predictor matrices: geographical distance, average temperature in January, number of days with temperature below 0 °C, number of days with snow cover, average snow cover depth, average annual rainfall and sample position with respect to the Vistula River. Selected climatic factors are considered to be the most important, limiting the distribution of both roe deer species (Danilkin & Hewison 1996), and the Vistula River was chosen as potential latitudinal environmental barrier in Poland. Marginal tests of each predictor were made, followed by conditional tests where latitude and longitude were included as covariables to the predictor variables. Then, sequential tests were run, employing a forward selection procedure to produce a combined model of *C. pygargus* mtDNA distribution among the studied roe deer populations, using DISTLM forward v1.3 (Anderson 2003). The *P* values were obtained from 9999 permutations. Environmental variables were taken from the *Climate Atlas of Poland* (Lorenc 2005).

Results

Level of mtDNA introgression and phylogenetic analyses

The CR sequences were generated for 452 roe deer hunted in Poland. The alignment showed 60 different haplotypes (GenBank accession nos KJ558223–KJ558287). The majority of sequences (377) gave 51 haplotypes and formed a European roe deer lineage. Interestingly, the remaining nine haplotypes belonged to the Siberian roe deer lineage (Table S1, Supporting information). Alignment of *cyt b* sequences (GenBank

nos KJ558288–KJ558338) gave fewer haplotypes (46) than CR mtDNA in both the *C. capreolus* and the *C. pygargus* mtDNA lineages. Genetic diversity indices for combined and separate mtDNA markers found in European roe deer in Poland are given in Table 1.

Bayesian phylogenetic reconstruction based on CR mtDNA roe deer sequences from Poland confirmed a clear division into two different clades corresponding to European and Siberian roe deer at observed 4.2% nucleotide divergence and an average 26.1 nucleotide substitutions (Fig. 2). High intraspecific similarity and large interspecific differences at mitochondrial markers ruled out incomplete lineage sorting as a possible scenario explaining the genetic variability found in European roe deer. The CR mtDNA of *C. capreolus* fell into three haplogroups coinciding with those named by Randi *et al.* (2004) as Clade Central, Clade West and Clade East, with the majority (65% of the individuals) grouped in Clade Central. The results for concatenated phylogenetic analysis of CR and *cyt b* data (Fig. S1, Supporting information) were broadly consistent with the CR tree (Fig. 2) and revealed higher divergence of three *C. capreolus* clades. The less abundant Clade West, with only four haplotypes (5.3% of individuals), was significantly specified (Fig. S1, Supporting information). Among the 37 identified CR mtDNA haplotypes from Clade Central, the two most frequent haplotypes (H1 and H10) were present in almost every population (19.9% and 7.7% of the individuals studied, respectively). The haplotypes from Clade West were distributed in the north and southeast of Poland, with frequencies ranging from 3.2% to 22.6%. CR mtDNA haplotypes belonging to Clade East were identified in 13 populations from northeastern, southeastern and central Poland, with frequencies up to 30–40% (Fig. S2, Supporting information).

Foreign CR mtDNA haplotypes were present in 12 of the 21 populations. Introgressed individuals constituted 16.6% of the deer studied; they were found in the east and southeast of the study area at frequencies ranging from 7.1% to 50% (Fig. 1). Using Siberian roe deer CR mtDNA accessions from GenBank, we were able to determine the approximate geographical origin of those

Table 1 Summary of genetic diversity for different markers found in European roe deer in Poland

Marker	<i>N</i>	mtDNA <i>C. capreolus</i>				mtDNA <i>C. pygargus</i>			
		<i>N</i> _{ind}	<i>N</i> _h	<i>h</i> (SD)	π (SD)	<i>N</i> _{ind}	<i>N</i> _h	<i>h</i> (SD)	π (SD)
Control region	452	377	51	0.92 (0.01)	0.009 (0.005)	75	9	0.71 (0.03)	0.006 (0.004)
Cytochrome <i>b</i>	184	114	42	0.93 (0.02)	0.006 (0.003)	70	4	0.42 (0.06)	0.002 (0.001)
CR + <i>cyt b</i>	184	114	67	0.98 (0.05)	0.008 (0.004)	70	10	0.72 (0.04)	0.004 (0.002)

N, sample size; *N*_{ind}, number of individuals with mtDNA of *C. capreolus* or *C. pygargus*; *N*_h, number of haplotypes; *h*, haplotype diversity; π , nucleotide diversity; SD, standard deviations.

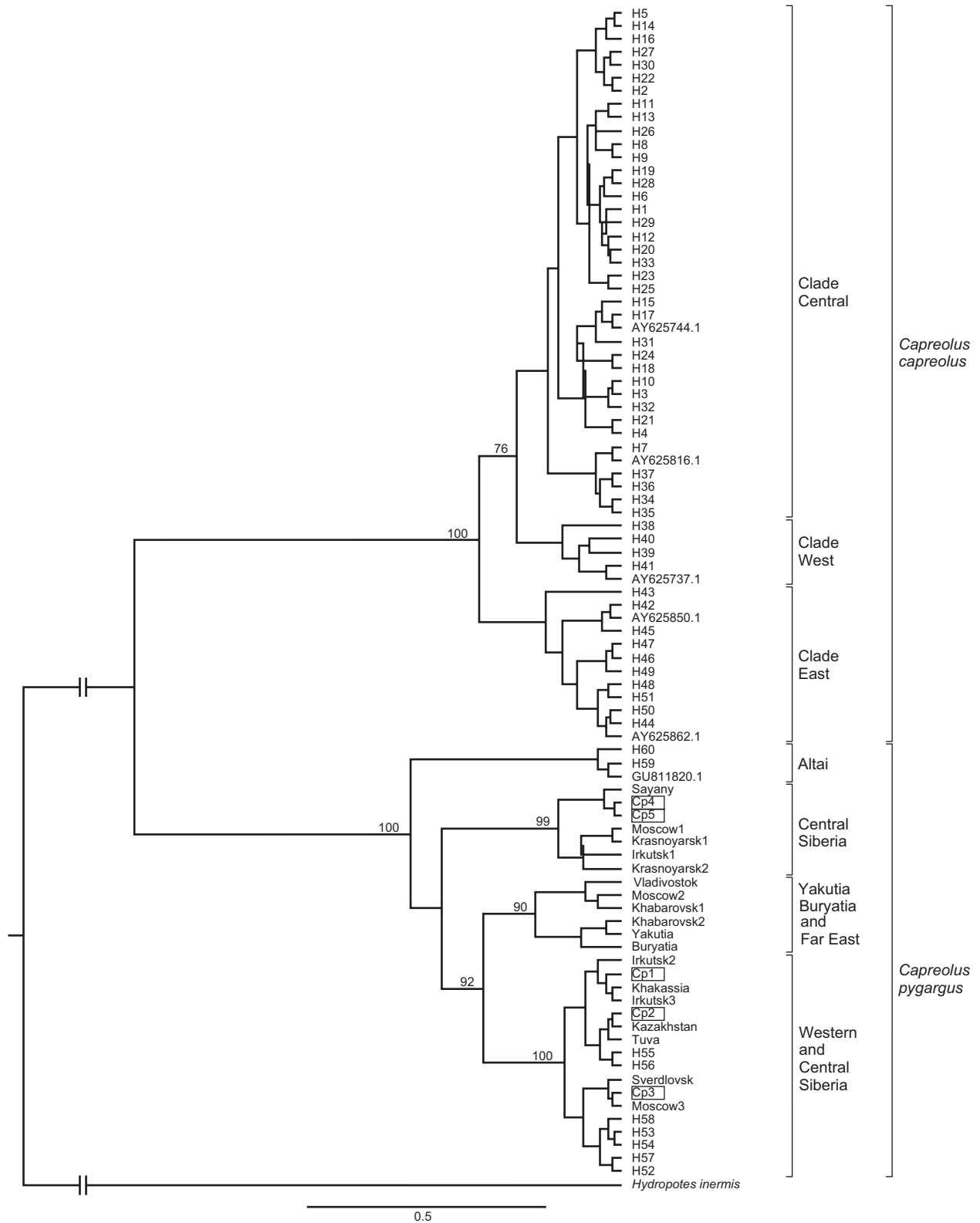


Fig. 2 Bayesian tree based on mtDNA control region sequences found in Polish roe deer populations (haplotypes H1–H60). Division of the main *C. capreolus* mtDNA clades (Randi *et al.* 2004) was ascertained using five CR mtDNA sequences from GenBank. CP1–CP5, haplotypes of reference *C. pygargus* samples from western and central Siberia. Remaining *C. pygargus* haplotypes from Zvychnayaya *et al.* (2011a). Bayesian posterior probability values shown at nodes.

haplotypes found in the European roe deer genome. Seven of them (haplotypes H52–H58) were grouped with haplotypes from western and central Siberia. The other two (H59 and H60) had their origin in the Altai Mountains. Despite strong statistical support for subclades within the Siberian roe deer lineage, both CR mtDNA and concatenated CR and *cyt b* trees showed no clear geographical structure, as the haplotypes from central Siberia clustered with those from western Siberia or formed a separate group (Fig. 2 and Fig. S1, Supporting information).

The time to TMRCA for the clades containing the two roe deer species was estimated at 0.81 Ma (95% highest posterior density interval, HPD 0.35–1.21 Ma) using concatenated CR and *cyt b* sequences. The age estimate for the divergence of Clade West and Clade East from Clade Central of *C. capreolus* was 0.29–0.35 Ma, similar to TMRCA estimated for the *C. pygargus* lineages (0.37 Ma, 95% HPD 0.17–0.53 Ma; Fig. S2, Supporting information).

The two species lineages identified by Bayesian phylogeny reconstruction were recovered in the networks constructed using CR mtDNA (not shown) and *cyt b* mtDNA. The species-specific lineages were completely separated by at least 34 mutation steps. The *cyt b* network also delineated three separate haplogroups within the European roe deer, suggesting the presence of a strong phylogeographical structure. The most geographically widespread Clade Central is characterized by starlike shape and 30 haplotypes that differ by 1–3 mutation steps from the most frequent haplotype. The Siberian roe deer *cyt b* lineage is reticulated and contains four subclades with many distinct haplotypes. Two of those subclades are composed of haplotypes found in the European roe deer mtDNA genome in Poland (Fig. S3, Supporting information).

Sex-linked markers

Unlike the mtDNA results, sex-specific markers did not show signs of introgression. The sequences of the two roe deer species were distinguished by two (ZFX gene) or four (concatenated Y chromosome markers) specific nucleotide differences (Fig. 3). Our alignment of the ZFX gene (465 bp) from 157 European roe deer and eight Siberian roe deer includes five alleles (GenBank accession nos KJ558339–KJ558343). Two ZFX alleles (A1 and A2) found in European roe deer were not shared with Siberian roe deer. Additionally, allele A1 was noted in *C. capreolus* and introgressed individuals (Fig. 3A). The alignment of concatenated Y chromosome markers (880 bp) from 51 European roe deer and seven Siberian roe deer males revealed four haplotypes (GenBank accession nos KJ558344–KJ558347). Likewise, for Y

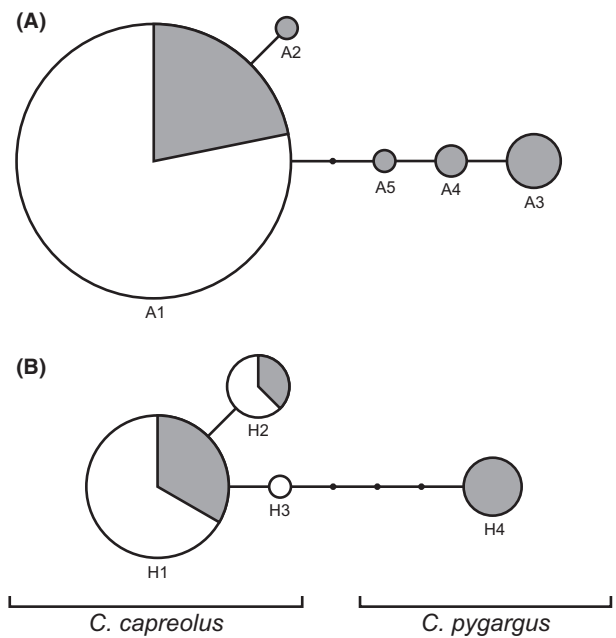


Fig. 3 Median joining networks based on ZFX gene (A) and three concatenated Y chromosome-specific fragments: SRY gene, DBY4 and DBY8 (B) in European and Siberian roe deer. Each circle represents a single allele/haplotype and is scaled by its frequency in the whole sample. Proportion of individuals possessing mtDNA of *C. pygargus* is marked in grey.

chromosome markers, haplotypes H1, H2 and H3 were only found in the European roe deer. Two of them (H1 and H2) were found in European roe deer males despite their mtDNA origin (Fig. 3B).

MtDNA diversity and population structure

The mtDNA haplotype phylogeny of roe deer in Poland reflects a high level of variation (Fig. 2). In the light of this, we carried out AMOVA of two species-specific mtDNA lineages. Genetic variation was significantly subdivided into three components. Due to the level of introgression, most of the genetic diversity reflects differences between species-specific mtDNA lineages ($\Phi_{CT} = 0.84$) and within populations ($\Phi_{ST} = 0.87$). The smallest portion of total molecular variance was distributed between populations within species-specific lineages (Table 2). Interestingly, SAMOVA of both species-specific lineages gave the same picture of genetic differentiation among population groups. Analyses on *C. capreolus* mtDNA indicated that the major increase in Φ_{CT} was for $K = 3$ ($\Phi_{CT} = 0.157$, $P < 0.05$), where three populations formed two separate groups (9–11 and 12) opposite the group consisting of the other 18 populations. When SAMOVA was carried out on *C. pygargus* mtDNA and 12 introgressed populations, genetic differentiation among groups also was maximal when $K = 3$ ($\Phi_{CT} = 0.593$, $P < 0.05$), with

Table 2 Results of hierarchical analysis of molecular variance (AMOVA)

Variance component	Variance	% total	<i>P</i>	Φ -statistics	
Between species-specific mtDNA lineages	12.02	84.12	<0.001	Φ_{CT}	0.84
Between populations within species-specific mtDNA lineages	0.39	2.75	<0.001	Φ_{SC}	0.17
Within populations	1.88	13.14	<0.001	Φ_{ST}	0.87

P, probability of obtaining a greater-than-observed variance value by chance alone. Significance for Φ -statistics ($P < 0.05$, 1000 permutations) is in bold type.

southeastern populations 11 and 12 forming two separate groups opposite the rest (1-3-5-6-7-8-9-13-14-15). The grouping of populations obtained with $K = 3$ was in agreement with founding CR mtDNA haplotypes of *C. pygargus* originating from the Altai Mountains and rare haplotypes of *C. capreolus* from Clade East in populations 11 and 12 (Fig. 2, Table S1, Supporting information).

Demographic expansion and spatial expansion were detected only in Clade Central of *C. capreolus*. Fu's value computed from CR mtDNA was negative for this clade ($F_s = -24.84$, $P < 0.0001$). Tajima's neutrality test was also negative but not significant ($D = -1.29$, $P > 0.05$). For *cyt b*, however, the D value was significant ($D = -2.34$, $P < 0.001$). Still, the McDonald-Kreitman test did not show significant signs of adaptive protein evolution and indicated only two fixed nonsynonymous mutations between European and Siberian roe deer, as compared to 21 fixed synonymous substitutions between species. Additionally, the mismatch distribution of Clade Central fits the unimodal curves expected under a sudden expansion model (SDD = 0.001, $P > 0.05$, $R = 0.025$, $P < 0.05$; Fig. S4, Supporting information). The main demographic expansion event occurred at $\tau = 2.55$ and involved a population change from an initial $\theta_0 = 0.03$ (95% C.I. 0–0.78) to a final $\theta_1 = 20.58$ (95% C.I. 6.92–9999.99). The spatial expansion occurred at $\tau = 2.49$ ($\theta = 0.01$, 95% C.I. 0–1.23). Assuming a CR mtDNA mutation rate of 0.04–0.08 substitutions per site per million years and a generation time of 4 years, both the demographic expansion and spatial expansion of Clade Central took place between 7.4 ka BP and 15.1 ka BP.

Species distribution modelling

Models of present roe deer distribution fitted the actual species distribution. The mean area under the receiver

operating curve (AUC) reached 0.941, which indicated high predictive power of the *C. capreolus* model. Potential suitable habitats for Siberian roe deer were located along the northern edge of the actual range and in greater part of China, and the AUC value was also high (0.881). Additionally, present distribution model for Siberian roe deer showed potential habitats of suitability 0.2–0.4 spanning westward to Central Europe, whereas no suitable habitats were found eastward from the present distribution for European roe deer (Fig. S5A, Supporting information).

Jackknife test suggested the annual mean temperature as the most important factor in the *C. pygargus* distribution model and both the minimum temperature of the coldest month and the annual mean temperature as the most important factors in the *C. capreolus* model. Other temperature variables were significant for both species, while precipitation appeared to be less relevant.

Distribution of species-specific mtDNA vs. environmental variables

A test for the influence of geographical coordinates on genetic population structure showed that frequency of introgressed individuals in a population depended on longitude ($P = 0.004$) but not on latitude ($P = 0.072$; Table 3). Temperature in January, snow cover depth, days with snow cover, number of days below 0 °C and location with respect to the Vistula River also had significant effects on maintenance of Siberian roe deer mtDNA in the European roe deer genome. When geographical coordinates were incorporated as covariables in a conditional test, none of the analysed factors remained significant, but a sequential test fitted temperature in January prior to longitude and all other climatic factors (42.04% of variation explained, $P = 0.003$; Table 3).

Discussion

Levels of introgression

Our finding of unexpectedly high genetic variability of mitochondrial markers in the European roe deer is a compelling example of introgressive hybridization of the mtDNA genome between closely related mammal species, the European and Siberian roe deer. The foreign mtDNA observed in *C. capreolus* from Poland breaches the distribution range of the European roe deer, crossing the current westward distribution range of Siberian roe deer by approximately 2000 km (Fig. 1). It covers <50% of the longitudinal range of *C. capreolus*, assuming consistency of introgression cline in the eastward direction. The level of introgression reached 23% when only the eastern and southeastern Polish

Table 3 Effects of abiotic and climatic factors on *C. pygargus* mtDNA distribution in 21 studied roe deer populations in Poland

Factor	Marginal tests		Conditional tests		Sequential tests	
	% var	<i>P</i>	% var	<i>P</i>	% var	<i>P</i>
Temperature in January	42.04	0.003	5.33	0.180	42.04	0.003
Latitude	14.09	0.072	—	—	9.53	0.089
Snow cover depth	41.23	0.003	10.81	0.053	5.85	0.157
Annual rainfall	13.72	0.106	3.82	0.288	2.70	0.290
No. of days below 0 °C	34.53	0.008	3.64	0.285	0.51	0.656
Vistula River	28.42	0.013	0.65	0.678	0.45	0.684
Days with snow cover	41.08	0.002	2.66	0.338	0.27	0.797
Longitude	40.22	0.004	—	—	0.27	0.742
Coordinates	46.46	0.009	—	—	—	—

% var, percentage of genetic variation explained by the particular variable; *P*, probability values.

Marginal and sequential tests of the forward selection procedure are reported. The top-down sequence of variables corresponds to the sequence indicated by the forward selection procedure (sequential tests).

populations of European roe deer were included. These two results suggest that this unique example should not be considered as massive introgression by the means of scoring system applied by Toews & Brelsford (2012).

The described widespread introgression of *C. pygargus* mtDNA into the *C. capreolus* genome was not accompanied by introgression of sex-linked markers. Reduced introgression of sex-linked markers is consistent with greater abundance of species incompatibilities on sex chromosomes (Baack & Rieseberg 2007). Mechanism of primary ancient introgression during postglacial range expansion of *C. capreolus* with subsequent *C. pygargus* range replacement, as proposed below, should drive geographical mito-nuclear discordance by preferential introgression of genes experiencing low intraspecific gene flow (mtDNA). Preliminary results of our genome-wide analysis using paired-end RAD-tag sequencing (unpublished data) also indicated very limited introgression of nuclear genes and confirmed that recent European and Siberian roe deer hybrids are not present in the populations studied.

Pathway of hybridization and mechanism of mtDNA introgression

Moderate-to-high levels of introgression and regularity of Siberian roe deer mtDNA cline allow us to hypothesize

that the current genetic structure of *C. capreolus* in eastern and southeastern Poland represents mostly a tail of natural postglacial introgressive hybridization, the remnant of an ancient range extension of *C. pygargus*. It would require co-occurrence of these two species in Central Europe after LGM, although there is no fossil evidence for Siberian roe deer appearing further west than Kiev, Ukraine (Danilkin & Hewison 1996). Our species distribution modelling analyses could not yield direct proof of such co-occurrence, because little past climatic predictions restrained to LGM and LIG conditions are available. Nevertheless, distribution of European roe deer tends to be more affected by cold glacial cycles and shows no present eastward colonization potential, whereas distribution of Siberian roe deer tends to be less restrained by harsh climate, displaying low potential for colonization of Eastern Europe territories in the present.

The analyses also indicated that during LIG (ca. 120–140 kyr BP), *C. pygargus* could have extended its range as far west as Central Europe (Fig. 4B). Thus, we suggest that after the LGM, the Siberian roe deer most likely established populations in Eastern Europe, probably reaching the territory of present-day Poland. Shortly thereafter, however, postglacial climate amelioration enabled more temperate species such as the European roe deer to expand eastward. Fu's *F_s* test and the mismatch distribution of *C. capreolus* Clade Central correspond to spatial expansion of the population in this area at 7.4–14.8 ka BP, probably after the Younger Dryas (10.8–10.0 ka BP). This range expansion of *C. capreolus* likely brought the two roe deer species into broader contact and led to interbreeding.

The observed asymmetrical gene flow resembles many cases of interspecific introgression, when a species expanding its range is massively introgressed by local genes, mostly by those experiencing reduced gene flow (Currat *et al.* 2008). It has been suggested that such asymmetry can arise due to a demographic imbalance between two species at the wave front, where the population of the expanding species initially is at lower density. As a consequence, genes introgressed on the wave front of range expansion can reach very high frequencies, even near-complete fixation, when new habitats are colonized (Currat *et al.* 2008; Excoffier *et al.* 2009). Most probably, a similar scenario prevailed during the range expansion of the European roe deer; although characterized by low dispersal potential, its average density and whole population number (indicating rapid growth potential) exceed those of Siberian roe deer by a factor of 10 (Danilkin & Hewison 1996). It is supported by recent studies of European roe deer populations inhabiting the Moscow area: the frequency of foreign mtDNA in *C. capreolus*

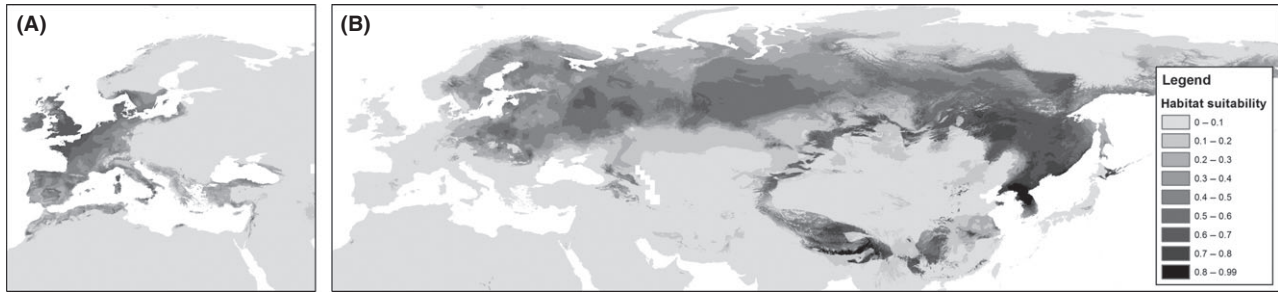


Fig. 4 Habitat suitability prediction for the Last Interglacial (ca. 120–140 kyr BP) of European (A) and Siberian roe deer (B).

populations was even higher (80%) than that in Poland (Zvychnaynaya *et al.* 2011b).

In our opinion, the mechanism of the described widespread mtDNA introgression was mostly driven by demographic disparities between retracting Siberian and expanding European roe deer. In the particular case of the introgressive hybridization at the wave front, hybrid matings should occur predominantly between females of the rare species and males of the common species (Chan & Levin 2005). Present species demography suggests that Siberian roe deer females have been in minority, therefore pointing to the direction of introgression asymmetry. Between the two species, body size differences should not be decisive in hybridization limitation, as about 20% of artificial interspecific crosses resulted in birth of live offspring without any form of assistance (Hewison & Danilkin 2001).

Phylogeny reconstruction in the genus *Capreolus* has also yielded some insights into possible scenario of hybridization. Among the CR mtDNA haplotypes of *C. pygargus* found in the European roe deer populations in Poland, we identified a relatively young, phylogenetically distinct group of haplotypes (H52–H54, H57 and H58; Fig. 2). It arose 23 kyr ago at the latest, and the haplotypes were possessed by most of the introgressed individuals (nearly 75%) and were found in every admixed population. The haplotypes are closely related to those found in Siberian roe deer from western and central Siberia but have not been detected within its natural distribution range in previous species-wide studies (Vorobieva *et al.* 2011; Zvychnaynaya *et al.* 2011a), suggesting ancient origin of present introgression. The same unique *C. pygargus* haplotypes found in Poland were also predominant in the Moscow area (Zvychnaynaya *et al.* 2011b), which supports the proposed gene surfing on the eastward expansion wave of European roe deer after hybridization event(s) with Siberian roe deer in Central Europe. A similar pattern was described in *Myodes gapperi* populations where introgressed individuals possessed distinct genetic signature from *M. rutilus*, not shared with pure populations of *M. rutilus* (Runck *et al.* 2009).

Human-mediated introductions of *C. pygargus* could have only complemented the extent of mainly natural introgression. Genetic legacy of nearly 1500 Siberian roe deer introduced in Eastern Europe during the twentieth century can be easily spotted. However, in most cases, their success was largely hindered by small group size and unfavourable sex ratio (Danilkin & Hewison 1996) and interbreeding with introgressed *C. capreolus* individuals prevailing in Eastern Europe. The presence of two haplotypes (H59 and H60) belonging to a group currently found only in the Altai Mountains (Vorobieva *et al.* 2011) is most likely the result of introductions of Siberian roe deer in nineteenth and twentieth centuries in Poland, Slovakia or Ukraine (Danilkin & Hewison 1996), as the haplotypes' occurrence was limited to the southeastern Poland.

European roe deer main recolonization routes revisited

The European roe deer was present in Europe at least 600 000 years ago. Since then, its distribution range has been pushed and pulled by the numerous interglacial and glacial phases of the Pleistocene, causing major contractions and expansions during the Late Quaternary (Sommer *et al.* 2009). The first Europe-wide assessment of *C. capreolus* phylogeography suggested the presence of multiple glacial refugia, with Iberia and the Balkans as the most important, that could have contributed to the postglacial colonization of Central Europe (Randi *et al.* 2004). In contrast, more recent review by Sommer *et al.* (2009) combining patterns of genetic data and fossil records suggests that recolonization process of Central Europe was completed by the second wave of advance of European roe deer, coinciding with climate amelioration after the Younger Dryas. In the light of the available genetic data, findings of roe deer fossils in the Carpathians ring region led to the hypothesis that this common ungulate colonized most parts of European lowlands from Carpathian and/or eastern European refugia.

With our new results from Poland, a previously underexamined but phylogeographically important area

where many intra- and interspecific lineages met after LGM (Hewitt 2004), we were able to revisit the recolonization patterns proposed for European roe deer. Our species distribution modelling to LGM under MI-ROC model was highly consistent with fossil findings of roe deer in the Carpathian region (Sommer *et al.* 2009), indicating suitable habitats exactly in Carpathian Mountains and present-day Bulgaria which could have served as refugia (Fig. 5). Most importantly, this model showed possible western refugia in France and even north from the Alps, in southern Germany. We suggest that measurable effect of colonization from southern refugia was most likely shifted and suppressed by range expansion from western refuge as mtDNA haplotypes of Clade East originating from southern refugia were limited in distribution mostly to southeastern Poland (Fig. S2, Supporting information). Only, the eastward recolonization of presently most abundant European roe deer mtDNA lineage (Clade Central) could have enabled the spread of initial introgression from Siberian roe deer straight up to Moscow region (as shown in Fig. 5).

Distribution of species-specific mtDNA vs. environmental variables

The proposed mechanism of widespread mtDNA introgression during the range expansion of European roe

deer does not comprehensively explain the persistent maintenance of Siberian roe deer mtDNA variants over time. The nature of selection process could not be easily inferred, as both neutral and advantageous mutations can spread in the invading species, creating very similar spatial distribution of introgressed alleles (Excoffier *et al.* 2009). We did not find an unusual pattern of Siberian roe deer mtDNA in Poland and did not detect any signs of positive selection on the cytochrome *b* gene in the McDonald–Kreitman test. However, multivariate multiple regression analysis not incorporating geographical coordinates suggested that a number of factors including temperature in January, snow cover depth and days with snow cover could influence the maintenance of Siberian roe deer mtDNA in the European roe deer genome. Snow cover depth and the duration of the snow period along with the predominant vegetation type were ranked as the most important factors limiting the distribution of both roe deer species. The first two factors negatively affect roe deer survival by exposing them to carnivores and reducing food resource availability (Danilkin & Hewison 1996). Hybridization with the Siberian roe deer, the bigger of the two species and better adapted to severe weather conditions, could have led to introgression of functional parts of the genome including quantitative traits enabling survival in eastern and southeastern Poland where the harsh continental climate creates unfavourable

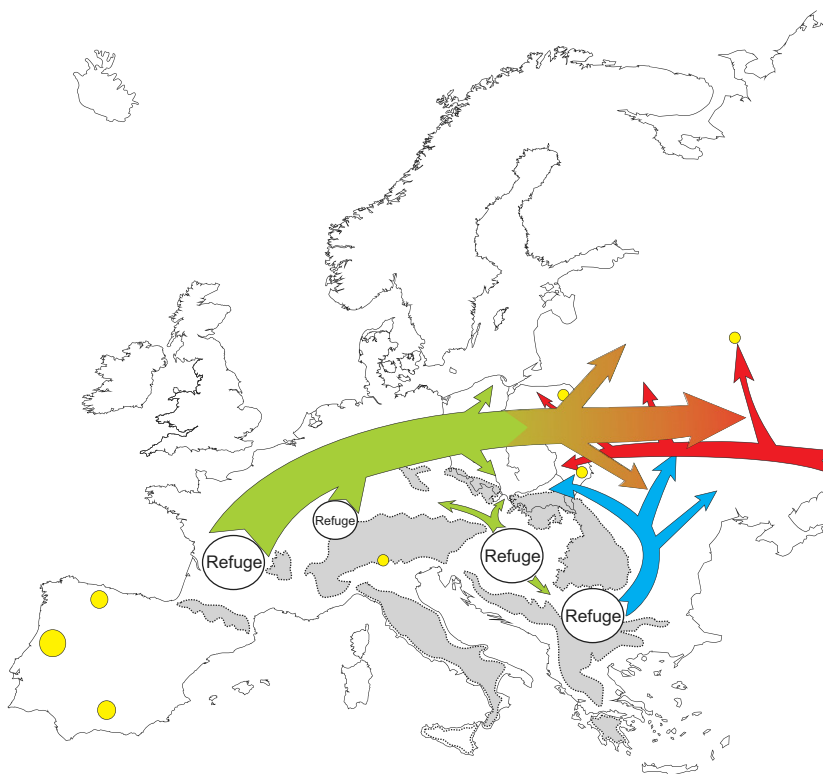


Fig. 5 Hypothesized postglacial recolonization routes of Central Europe by roe deer. Siberian roe deer lineage in red. European roe deer lineages: Clade Central in green, Clade East in blue and Clade West in yellow. Records of Clade West in Italian Alps and Iberian Peninsula after Randi *et al.* (2004).

conditions for temperate species. However, none of the environmental or climatic variables had a significant effect on the distribution of species-specific mtDNA lineages when geographical coordinates were incorporated as covariables. It is most likely an outcome of geographical distribution of species-specific mtDNA lineages which follow east–west cline. More analyses are needed to confirm whether maintenance of *C. pygargus* mtDNA is a result of adaptive introgression dependent on environmental variables or a by-product of stochastic demographic changes occurring in natural populations.

Acknowledgements

We thank Prof. Ettore Randi (Istituto Nazionale per la Fauna Selvatica, Ozzano, Italy) and Dr Sergey Anchugov (Kurgan, Russia) for providing samples of *C. capreolus* and *C. pygargus*. We are also grateful to Magdalena Czajkowska, Norbert Duda and Piotr Rode for their help in collecting tissue of *C. capreolus* from Poland. Michael Jacobs helped edit the article. This work was financially supported by the National Science Centre in Poland (Grant No. N N304 390838).

This article is a part of the dissertation research of M.M., which is focused on hybridization between roe deer species. A.B., M.M. and M.R. designed the study, and A.B. managed the project; M.M., M.S., Z.B. and K.K. collected samples of European roe deer, and A.P.S. provided samples of Siberian roe deer and an information about roe deer translocations in Europe; M.M. and M.S. conducted the laboratory work; A.A.D. and E.Y.Z. provided mtDNA sequences of *C. pygargus* from Russia and articles published in Russian. M.M., A.B., P.M. and M.R. implemented the methods and analysed the results. M.M. and A.B. wrote the article. All authors read, discussed and approved this version of the manuscript.

Data accessibility

Mitochondrial and nuclear DNA sequences are deposited in GenBank under the following accessions: KJ558223–KJ558287 (CR), KJ558288–KJ558338 (cyt *b*), KJ558339–KJ558343 (ZFX) and KJ558344–KJ558347 (SRY, DBY4 and DBY8). Alignment files, tree files, DISTLM v5 and DISTLM forward v1.3 input files are deposited under DRYAD entry doi: 10.5061/dryad.hn11 m.

References

Anderson MJ (2003) *DISTLM Forward: A FORTRAN Computer Program to Calculate a Distance-Based Multivariate Analysis for*

a Linear Model Using Forward Selection. Department of Statistics, University of Auckland, Auckland, New Zealand.

Anderson MJ (2004) *DISTLM v.5: A FORTRAN Computer Program to Calculate a Distance-Based Multivariate Analysis for a Linear Model*. Department of Statistics, University of Auckland, Auckland, New Zealand.

Avice JC (2004) *Molecular Markers, Natural History, and Evolution*, 2nd edn. Sinauer Associates, Sunderland, Massachusetts.

Baack EJ, Rieseberg LH (2007) A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development*, **17**, 513–518.

Balloux F, Handley LJJ, Jombart T, Liu H, Manica A (2009) Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **276**, 3447–3455.

Boratynski Z, Alves PC, Berto S, Koskela E, Mappes T, Melo-Ferreira J (2011) Introgression of mitochondrial DNA among *Myodes* voles: consequences for energetics? *BMC Evolutionary Biology*, **11**, 355.

Buggs RJA (2007) Empirical study of hybrid zone movement. *Heredity*, **99**, 301–312.

Cabria MT, Michaux JR, Gómez-Moliner BJ *et al.* (2011) Bayesian analysis of hybridization and introgression between the endangered European mink (*Mustela lutreola*) and the polecat (*Mustela putorius*). *Molecular Ecology*, **20**, 1176–1190.

Chan KMA, Levin SA (2005) Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution*, **59**, 720–729.

Chikuni K, Mori Y, Tabata T, Saito M, Monma M, Kosugiyama M (1995) Molecular phylogeny based on the kappa-casein and cytochrome *b* sequences in the mammalian suborder Ruminantia. *Journal of Molecular Evolution*, **41**, 859–866.

Crispo E, Moore JS, Lee-Yaw JA, Gray SM, Haller BC (2011) Broken barriers: human-induced changes to gene flow and introgression in animals. *BioEssays*, **33**, 508–518.

Curat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908–1920.

Danilkin AA, Hewison AJM (1996) *Behavioural Ecology of Siberian and European Roe Deer*. Chapman and Hall, London.

Derr JN, Hedrick PW, Halbert ND *et al.* (2012) Phenotypic effects of cattle mitochondrial DNA in American bison. *Conservation Biology*, **26**, 1130–1136.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571–2581.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, **13**, 853–864.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 481–501.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
- Gernhard T (2008) The conditioned reconstruction process. *Journal of Theoretical Biology*, **253**, 769–778.
- Gleiss HGW (1967) *Unter Robben, Gnus und Tigerschlangen. Chronik des Zoologischen Gartens Breslau 1865-1965*. Natura et Patria Verlag, Wedel, Germany.
- Gonzalez T, Tsytsulina K (2008) *Capreolus pygargus*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. www.iucnredlist.org.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.
- Hellborg L, Ellegren H (2003) Y chromosome conserved anchored tagged sequences (YCATS) for the analysis of mammalian male-specific DNA. *Molecular Ecology*, **12**, 283–291.
- Hewison AJM, Danilkin AA (2001) Evidence for separate specific status of European (*Capreolus capreolus*) and Siberian (*C. pygargus*) roe deer. *Mammalian Biology*, **66**, 13–21.
- Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B*, **359**, 183–195.
- Innocenti P, Morrow EH, Dowling DK (2011) Experimental evidence supports a sex-specific selective sieve in mitochondrial genome evolution. *Science*, **332**, 845–848.
- Irwin DE, Rubtsov AS, Panov EN (2009) Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*Emberiza leucocephalos*) (Aves: Passeriformes). *Biological Journal of the Linnean Society*, **98**, 422–438.
- Kalendar R, Lee D, Schulman AH (2009) FastPCR Software for PCR primer and probe design and repeat search. *Genes, Genomes and Genomics*, **3**, 1–14.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- Lorenc H (2005) *Climate Atlas of Poland*. Institute of Meteorology and Water Management, Warsaw. [in Polish].
- Lovari S, Herrero J, Conroy J *et al.* (2008) *Capreolus capreolus*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. www.iucnredlist.org.
- McDonald JH, Kreitman M (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature*, **351**, 652–654.
- Melo-Ferreira J, Boursot P, Suchentrunk F, Ferrand N, Alves PC (2005) Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology*, **14**, 2459–2464.
- Nilsen EB, Gaillard JM, Andersen R *et al.* (2009) A slow life in hell or a fast life in heaven: demographic analyses of contrasting roe deer populations. *Journal of Animal Ecology*, **78**, 585–594.
- Osada N, Uno Y, Mineta K, Kameoka Y, Takahashi Terao K (2010) Ancient genome-wide admixture beyond the current hybrid zone between *Macaca fascicularis* and *M. mulatta*. *Molecular Ecology*, **19**, 2884–2895.
- Pavlov MP (1999) Akklimatizatsia okhotnich'e-promylovyykh zverei i ptits v SSSR. Chast' 3. Kopytnye (Acclimation of game mammals and birds in the Soviet Union: Part 3. Ungulates). Russian Research Institute of Game Management and Fur Farming, Kirov. [in Russian].
- Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology and Evolution*, **24**, 386–393.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.
- Pucek Z (1981) *Keys to Vertebrates of Poland. Mammals*. Polish Scientific Publishers, Warsaw.
- Randi E, Pierpaoli M, Danilkin A (1998) Mitochondrial DNA polymorphism in populations of Siberian and European roe deer (*Capreolus pygargus* and *C. capreolus*). *Heredity*, **80**, 429–437.
- Randi E, Alves PC, Carranza J, Milosevic-Zlatanovic S, Sfougaris A, Mucci N (2004) Phylogeography of roe deer (*Capreolus capreolus*) populations: the effects of historical genetic subdivisions and recent nonequilibrium dynamics. *Molecular Ecology*, **13**, 3071–3083.
- Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608–615.
- Rozen S, Skaletsky H (1999) Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365–386.
- Runck AM, Matocq MD, Cook JA (2009) Historic hybridization and persistence of a novel mito-nuclear combination in red-backed voles (genus *Myodes*). *BMC Evolutionary Biology*, **9**, 114.
- Smith S, Turbill C, Suchentrunk F (2010) Introducing mother's curse: low male fertility associated with an imported mtDNA haplotype in a captive colony of brown hares. *Molecular Ecology*, **19**, 36–43.
- Sommer RS, Fahlke JM, Schmolcke U, Benecke N, Zachos FE (2009) Quaternary history of the European roe deer *Capreolus capreolus*. *Mammal Review*, **39**, 1–16.
- Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.
- Vernesi C, Pecchioli E, Caramelli D, Tiedemann R, Randi E, Bertorelle G (2002) The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. *Molecular Ecology*, **11**, 1285–1297.
- Vorobieva NV, Sherbakov DY, Druzhkova AS *et al.* (2011) Genotyping of *Capreolus pygargus* fossil DNA from Denisova Cave reveals phylogenetic relationships between ancient and modern populations. *PLoS One*, **6**, e24045.
- Wood NJ, Phua SH (1996) Variation in the control region sequence of the sheep mitochondrial genome. *Animal Genetics*, **27**, 25–33.
- Zvychnayaya EY, Danilkin AA, Kholodova MV, Sipko TP, Berber AP (2011a) Analysis of the variability of the control

region and cytochrome *b* gene of mtDNA of *Capreolus pygargus* Pall. *Biology Bulletin*, **38**, 434–439.

Zvychnaynaya EY, Kiryakulov VM, Kholodova MV, Danilkin AA (2011b) Roe deer (*Capreolus*) from Moscow area: analysis of mitochondrial control region polymorphism. *Vestnik okhotovedeniya*, **8**, 168–172. [in Russian with English summary].

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Extended methods.

Fig. S1 Bayesian tree based on concatenated mtDNA sequences (CR and *cyt b* gene) found in Polish roe deer populations.

Fig. S2 Proportion of mtDNA control region haplotypes belonging to the main clades of European and Siberian roe deer in Polish populations.

Fig. S3 Median joining network of cytochrome *b* mtDNA haplotypes of *Capreolus capreolus* and *C. pygargus* found in sample set of European roe deer from Poland.

Fig. S4 Mismatch distribution of pairwise nucleotide differences for each mtDNA lineage found in European roe deer genome in Poland.

Fig. S5 Results of the species distribution modelling for European and Siberian roe deer.

Table S1 Numbers of individuals possessing different CR mtDNA haplotypes and respective *cyt b* variants found in European roe deer populations in Poland.