

Genetic diversity and structure of the hedgehogs *Erinaceus europaeus* and *Erinaceus roumanicus*: evidence for ongoing hybridization in Eastern Europe

KATERINA I. ZOLOTAREVA¹, MARIANA M. BELOKON², YURI S. BELOKON², MARINA V. RUTOVSKAYA³, LUDMILA A. HLYAP³, VLADIMIR P. STARYKOV⁴, DMITRY V. POLITOV², VLADIMIR S. LEBEDEV⁵, and ANNA A. BANNIKOVA^{1*}

¹Lomonosov Moscow State University, Leninskie Gory 1, 119991 Moscow, Russia

²Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin str. 3, 119991 Moscow, Russia

³Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninskii pr., 33, 119071 Moscow, Russia

⁴Surgut State University, Lenin str. 1, 628400 Surgut, Russia

⁵Zoological Museum, Moscow State University, B.Nikitskaya 6, 125009 Moscow, Russia

Received 1 June 2020; revised 2 August 2020; accepted for publication 7 August 2020

Secondary contact zones between related species are key to understanding speciation mechanisms. The Central European sympatry zone of West European (*Erinaceus europaeus*) and northern white-breasted (*Erinaceus roumanicus*) hedgehogs is well studied, whereas data on the Eastern European sympatry zone are scarce. We examined the genetic variation in Russian populations using the mitochondrial *Cytb* gene, *TTR* intron 1 and 11 microsatellites to assess genetic variability and distribution patterns. In contrast to the Central European sympatry zone, we found evidence of ongoing hybridization between the two species in the sympatry zone of European Russia, where the proportion of individuals with mixed ancestry was *c.* 20%. Our data indicate bi-directional mtDNA introgression, but with a higher frequency of *E. europaeus* haplotypes in hybrids. The proportion of pure specimens with introgressed mitotypes is higher in *E. roumanicus* than in *E. europaeus*. Nuclear data showed the prevalence of the genetic contribution from *E. roumanicus* in admixed individuals. Demographic analyses indicated recent population growth in *E. europaeus* and little change in *E. roumanicus*, suggesting that *E. europaeus* colonized East Europe later than *E. roumanicus*.

ADDITIONAL KEYWORDS: demographic history – interspecies hybridization – phylogeography – postglacial colonisation – secondary contact zone.

INTRODUCTION

Speciation under conditions of allopatry and sympatry with accompanying hybridization and introgression events is a focus of a diverse range of modern population genetic and phylogeographical studies (e.g. Brown *et al.*, 2010; Cabria *et al.*, 2011; Coyner *et al.*, 2016; Mastrantonio *et al.*, 2016; Poplavskaya *et al.*, 2017; Scordato *et al.*, 2017; Quilodr n *et al.*, 2019). Understanding the evolutionary and ecological processes that contribute to reproductive

isolation in contact zones with different hybridization patterns is of particular importance. Such patterns of population history are often associated with the glacial history of species and can be used to generate speciation hypotheses.

Hedgehogs of the genus *Erinaceus* have played an important role in helping us to reveal the influence of Quaternary geography on extant species in Europe. They provide classic examples of postglacial recolonization patterns and the subsequent formation of secondary sympatry zones in response to climate-dependent changes during the Pleistocene (Santucci *et al.*, 1998; Hewitt, 1999;

*Corresponding author. E-mail: hylomys@mail.ru

Seddon *et al.*, 2001; Berggren *et al.*, 2005). However, to our knowledge, all phylogeographic and population studies to date have concerned only the Western and Central European parts of their range. The genetic structure of the eastern populations must also have been influenced by the dramatic Pleistocene range dynamics.

Among the four species of *Erinaceus*, two are distributed in the European part of Eurasia, and one is distributed in Western Siberia. The West European hedgehog *E. europaeus* Linnaeus, 1758 is common in Western, Central and Eastern Europe, and is known from the islands of Ireland, Great Britain, Corsica, Sardinia, Sicily and the Azores (Holz & Niethammer, 1990; Mathias *et al.*, 1998; Mitchell-Jones *et al.*, 1999), and it has been introduced to New Zealand (King, 1990). The southern white-breasted hedgehog *E. concolor* Martin, 1838 inhabits Asia Minor and the South Caucasus (Hutterer, 2005). The range of the northern white-breasted hedgehog *E. roumanicus* Barret-Hamilton, 1900 covers Central Europe from Poland to Austria and Slovenia, the Balkan states, the Greek and Adriatic isles, Turkish Thrace, and eastwards into parts of Russia, Ukraine to the North Caucasus, Western Siberia and the River Ob (Hutterer, 2005; Zaitsev *et al.*, 2014).

Two zones of sympatry are known between *E. europaeus* and *E. roumanicus*: in Central Europe (Poland, the Czech Republic, Austria and Italy) and in north-eastern Europe (Latvia, Estonia and European Russia eastwards to the Ural Mountains). Despite there being few hybrid individuals (based on morphological characters) from the Central European sympatry zone (Ruprecht, 1973; Kratochvil, 1975; Bauer, 1976), or obtained in laboratory crossing experiments (Poduschka & Poduschka, 1983), the question of natural hybridization remains unanswered. In the Central European sympatry zone, no evidence for interspecies hybridization has been detected (Bolfíková & Hulva, 2012), with the exception of Trenčianske Teplice, Slovakia, where a backcrossed hybrid of *E. europaeus* × *E. roumanicus* with mitochondrial DNA originating from *E. roumanicus* was recorded (Bolfíková *et al.*, 2017). Two more potential hybrids were found recently in the vicinity of Linz, Austria (Curto *et al.*, 2019). Both records correspond to the westernmost edge of the range of *E. roumanicus* in Central Europe. Preliminary studies of the Moscow region population have demonstrated that, in the Eastern European sympatry zone, both species, together with individuals of a mixed genotype, occur (Bogdanov *et al.*, 2009; Bannikova *et al.*, 2010). Owing to the ambiguity of their morphological identification, especially if such identifications are based solely on external morphology (Kratochvil, 1975; Zaitsev,

1984), genetic analysis should play a crucial role in determining the parental forms and their hybrids in sympatric populations and help define the distribution boundaries of the species.

In contrast to the data on morphological variability (Ruprecht, 1973; Kratochvil, 1975; Bauer, 1976; Holz, 1978; Krystufek, 1983, 2002) and phylogeography (Santucci *et al.*, 1998; Seddon *et al.*, 2001) that are well documented for Western and Central European populations, there remains a dearth of knowledge on the genetic structure of *E. europaeus* and *E. roumanicus* in Eastern Europe and Western Siberia. The aim of our study was to evaluate the pattern of co-distribution of *E. europaeus* and *E. roumanicus*, to examine genetic variation and analyse the demographic history of the two species in European Russia and Western Siberia, and to estimate the intensity and direction of gene flow between species in the Eastern European sympatry zone. In particular, we aimed to test the hypothesis that most of the gene flow is from *E. europaeus* to *E. roumanicus*, as predicted from experimental hybridization data (Poduschka & Poduschka, 1983).

MATERIAL AND METHODS

TISSUE COLLECTION AND SAMPLING

Our material consisted of 221 specimens of hedgehogs from 57 localities across the ranges of *E. europaeus* and *E. roumanicus* in Eastern Europe and Western Siberia (Fig. 1; Table 1). The preliminary taxonomic identification of hedgehogs was determined on exterior characteristics (fur coloration and spine striping, presence or absence of a mask on the muzzle, and presence and shape of a white spot on the ventral side), either directly (if the animals were available) or from photographs taken by collectors (Supporting Information, Table S1). For some road-killed specimens or those killed by dogs, taxonomic diagnosis was by cranial characters (shape of nasal bones, structure of maxillary-premaxillary suture) as described by Zaitsev (1984). The zone of sympatry was considered the territory where the two species coexisted, i.e. the Moscow region in our study.

For 181 specimens (20 *E. europaeus* and 59 *E. roumanicus* from the allopatric populations and 102 hedgehogs from the sympatry zone), 11 microsatellite loci, a 755-bp fragment of *TTR* intron 1 and the complete mitochondrial *Cytb* gene were analysed. For an additional 40 specimens from different geographical localities, only *Cytb* and *TTR* ($n = 37$) or only *Cytb* ($n = 3$) were sequenced. In the phylogeographic analysis, 67 sequences of *Cytb* were retrieved from GenBank (Supporting Information 1). Since the data on *Cytb* in GenBank are represented by haplotypes rather than

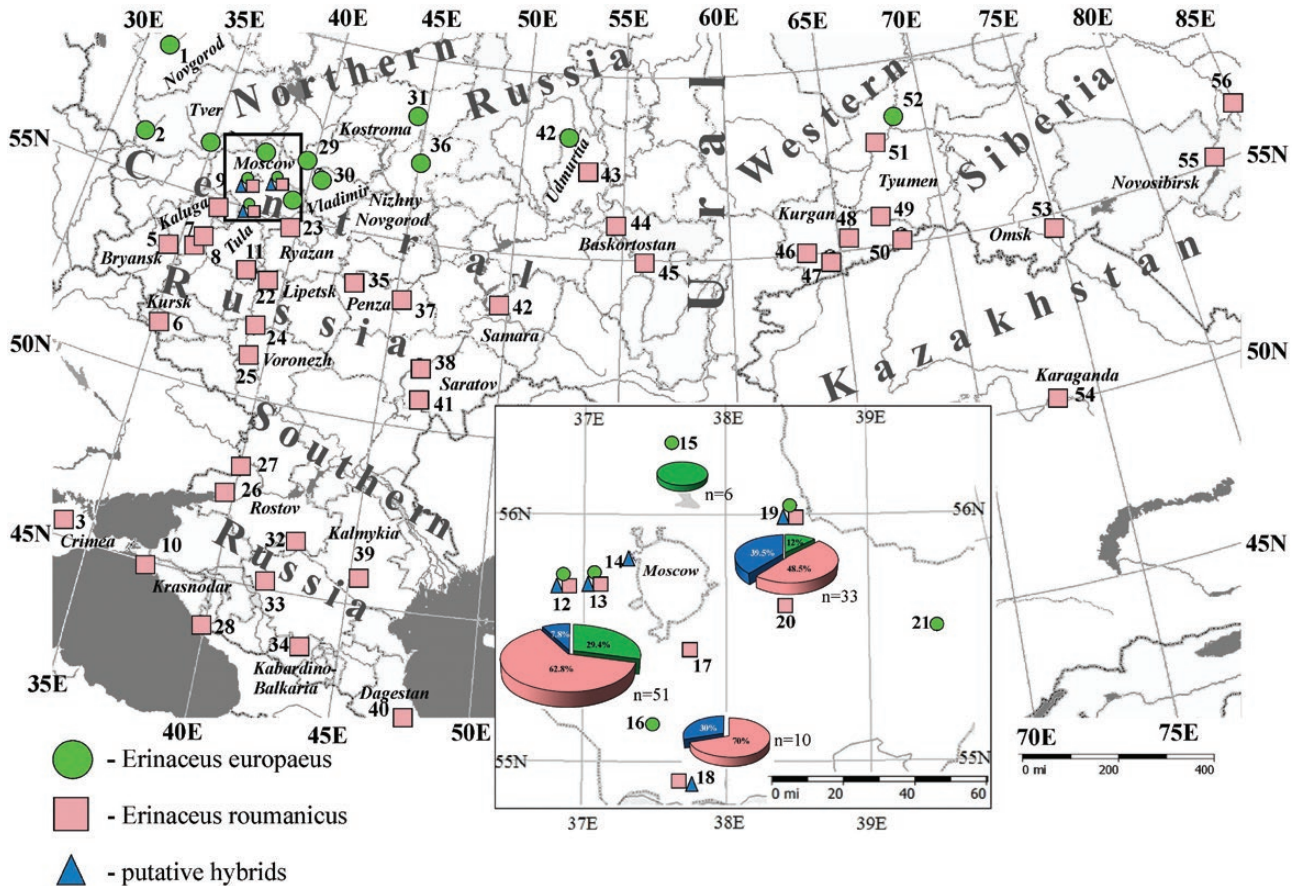


Figure 1. The sampling localities of our material based on the results of genotyping. Locality names and detailed geographic information are provided in Table 1.

individuals, in the phylogeographical analysis the number of haplotypes was converted to the number of individuals with a specific haplotype according to the data from published sources (Santucci, 1998; Seddon *et al.*, 2001, 2002).

EXPERIMENTAL PART

DNA was extracted mainly from small ear biopsies of live-trapped animals or from muscle tissue collected from road-killed animals. Part of the samples comprised dry muscle samples from the Zoological Museum of Moscow University collection. Genomic DNA was extracted using a standard protocol of proteinase K digestion, phenol-chloroform deproteinization and isopropanol precipitation (Sambrook *et al.*, 1989).

Amplification was performed on the My Cycler (Bio-Rad, Singapore, USA) and Mastercycler Gradient (Eppendorf, Hamburg, Germany) thermocyclers. General methods for the amplification and combination of primers for the complete sequence of cytochrome *b* (*Cytb*, 1140 bp) and fragment of intron 1 of transthyretin

(*TTR*, ~755 bp) were amplified using the protocols and primer combination described by Bannikova *et al.* (2014). The sequencing of *Cytb* was performed with the internal primers specially designed for the present study, which are provided in the Supporting Information (Table S2). PCR products were sequenced on the autosequencing system ABI 3100-Avant using ABI PRISM® BigDye™ Terminator v.3.1 (Applied Biosystems, Foster City, CA, USA). All sequences were aligned using DNASTar Lasergene SeqMan Pro v.7.1.0 (Burland, 2000) and BioEdit software v.7.0.9.0 (Hall, 1999) followed by eye finishing (alignment has been manually adjusted based on visual analysis). The sequences obtained in the present study can be accessed via GenBank (Accession numbers: *Cytb* haplotypes: MT856912–MT856943).

MICROSATELLITE AND TTR GENOTYPING

In the microsatellite analysis, we used primers developed for *E. europaeus* (Becher & Griffiths, 1997; Henderson *et al.*, 2000). After preliminary

Table 1. Geographic information, designation of localities and number of specimens used in the study. Collecting locality codes are the same as those in Fig. 1.

Collecting region	No. of localities, and code as in Figure 1	No. of specimens
<i>E. roumanicus</i> from allopatric populations ($N = 88$)		
Krasnodar region: Primernoie and Maliy Utrish	2 (10, 28)	10
South Ural, Bashkortostan: Noviy Burtyuk and Revet	2 (44, 45)	3
Bryansk region: Bryansk	1 (5)	4
Voronezh region: Divnogorye, Voronezh	2 (24, 25)	3
Republic of Crimea, Primorskiy	1 (3)	4
Republic of Dagestan: Madjalis	1 (40)	1
Republic of Kalmykia, Zulturgan	1 (39)	1
Kaluga region: Steklozavod, Yagodnoye and Kaluzhskye Zaseki reserve	3 (7–9)	3
Kurgan region: Al'menevo, Berezo, Oktyabr'skoye, Odino, Ketovo	5 (46–50)	9
Kursk region: Zaoleshenka	1 (6)	1
Lipetsk region: Lebedyan'	1 (22)	1
Republic of Kabardino-Balkaria: Nalchik	1 (34)	1
Novosibirsk region: Novosibirsk	1 (55)	3
Samara region: Zhigulevsk reserve	1 (42)	1
Saratov region: Lugovskoye and D'yakovka Road	2 (38, 41)	4
Omsk region: Zolotukhino	1 (53)	1
Penza Region: Krasnaya Polyana and Chemodanovka	2 (35, 37)	2
Rostov region: Rostov-on-Don, Krasny Sulin, Rostov reserve (Manych, Volochayevskiy)	3 (26, 27, 32)	22
Ryazan region: Ryazan	1 (23)	1
Stavropol region: Stavropol	1 (33)	3
Tomsk region: Vershinino	1 (56)	1
Tula region: Bol'shaya Sukhotinka	1 (11)	1
North Kazakhstan, Karaganda	1 (54)	2
Udmurtia: Yakshur-Bod'ya	1 (43)	1
Tyumen region: Shatanova	1 (51)	4
Bulgaria: Sinemorets	1	1
<i>Erinaceus europaeus</i> from allopatric populations ($N = 21$)		
Kostroma region: Manturovo area, Shilovo vill.	1 (31)	1
Tver region: Zhelnino and Krutitsy vill.	2 (2, 4)	13
Vladimir region: Vladimir and Yur'ev-Pol'sky	2 (29, 30)	3
Novgorod region: Valdai Upland	1 (1)	1
Tyumen region: Tobolsk	1 (52)	1
Udmurtia: Vortsa	1 (42)	1
Nizhny Novgorod region: Yakushevo	1 (36)	1
<i>E. roumanicus</i> and <i>E. europaeus</i> from zone of sympatry ($N = 112$)		
East of Moscow region: Chernogolovka, Gzhel, Shatura	3 (19–21)	37
West of Moscow region: Zvenigorod, Krasnogorsk, Nikolina Gora	3 (12–14)	55
South of Moscow region: Domodedovo, Chekhov, Prioksko-Terrasny Reserve (PTZ)	3 (16–18)	14
North of Moscow region: Dmitrov	1 (15)	6
Total		221

analyses, we selected 11 polymorphic loci containing dinucleotide repeats (*EEU1*, *EEU2*, *EEU3*, *EEU4*, *EEU5*, *EEU6*, *EEU12*, *EEU36*, *EEU37*, *EEU43* and *EEU54*). PCR was performed using the GenPak®

PCR-Core Kit (Galart-Diagnosticum production), according to the manufacturer's instructions. Products of amplification of microsatellite loci were separated in 6% polyacrylamide gel. The *E. coli* DNA plasmid

pBR322 treated with *HpaII* restriction endonuclease (SibEnzyme) was used as a standard length marker for fragment size control. The resulting mix contained fragments of the following sizes: 622, 527, 404, 307, 242, 238, 217, 201, 190, 180, 160, 146, 147, 123, 110, 90, 76, 67 and 34 bp. Electrophoresis was performed in a standard Tris-borate buffer for 3–4 h under an electric field voltage of 300 V and electric current of 55 mA per gel. Results of electrophoresis were dyed with ethidium bromide and visualized and photographed using a gel documentation system (Doc-print II, Vilber Lourmat, Marne-la-Vallée, France) in a transilluminator under 312 nm UV light. Images were analysed using the Photo-Capt program and also manually by comparing fragments obtained for each sample with the lengths of fragments in the *pBR322/HpaII* mix. In addition, samples were constantly compared to each other (several samples from the previous batch were obligatorily loaded on every new gel). *TTR* genotypes were phased into haplotypes based on the presence of characteristic deletions at positions 536–540 and 697–700 in the alleles of *E. europaeus*.

ANALYSES OF NUCLEAR VARIATION

The analysis of genetic variability was performed separately for allopatric populations of the two species and for the population from the sympatry zone. Allele frequencies, the number of alleles per locus (N_A), observed (H_O) and expected (H_E) heterozygosity, and fixation index (F_{st}) were evaluated in MS Excel for GenAlEx v.6.501 (Peakall & Smouse, 2006, 2012). Deviation from the Hardy–Weinberg equilibrium was calculated in Arlequin v.3.5.1.3 (Excoffier & Lischer, 2010). The presence of null alleles (false homozygotes) was tested with Micro-Checker (Van Oosterhout *et al.*, 2004) using the method of Brookfield (1996). Allele frequency heterogeneity for combined samples of two species of hedgehogs was tested using the program POPGENE v.1.32 (Yeh *et al.*, 1998).

The genetic structure of *E. roumanicus* and *E. europaeus* was evaluated in Structure v.2.3.4 (Pritchard *et al.*, 2000; Falush *et al.*, 2003) and BAPS v.5.2 (Corander *et al.*, 2006, 2008a, b) on the basis of combined data on the distribution of microsatellites (simple sequence repeats) and *TTR* alleles (SSR+*TTR*) within populations. Structure and BAPS analyses included individuals for which all alleles of the selected loci were characterized ($N = 181$).

To detect genotype admixture in the dataset using Structure v.2.3.4 we applied an admixture model of independent frequencies, using population information. The parameters of the final run were as follows: 1 000 000 MCMC steps, with burn-in-period 100 000 and three iterations for each K. The number of tested populations was within the $K = 2–15$ interval.

The optimal number of groups was selected in the program Structure Harvester (Earl & von Holdt, 2012), using the Evanno method (Evanno *et al.*, 2005), and resulted in $K = 3$. In further analyses, we assumed a threshold for pure species of $Q = 0.9$.

The BAPS analysis pipeline was as follows: at the first stage, the program divided the entire studied set of populations/specimens into two clusters (mixture stage); on the basis of the obtained division into two forms, the mixed origin was tested (admixture analysis). The portion of alleles originating from each of the two groups was evaluated for every specimen, and on the basis of the distribution of genotypes of pure forms obtained with Monte Carlo simulations, a more stringent test of mixed origin that makes it possible to conservatively estimate the number of hybrid specimens was conducted. The parameters of analysis were as follows: minimum size of population = 1, number of iterations = 200, number of reference specimens = 1000, and number of repeated iterations = 20. Additionally, analysis with unfixed number of clusters was performed in BAPS, allowing the program to select it (Corander & Marttinen, 2006).

NewHybrids v.1.1 beta (Anderson & Thompson, 2002) was used to identify admixture and classify hybrid status to classes beyond the F1 generation. This method identifies hybrid individuals on the basis of the posterior probability of belonging to different pure parental or hybrid categories. Simulations were run under default options with 124 470 sweeps, which, as following from preliminary runs, far exceeded the value necessary for convergence.

Microsatellite analysis of genetic diversity was conducted using the dataset divided into five groups: (1) *E. roumanicus* from allopatric populations; (2) *E. europaeus* from allopatric populations; (3) *E. roumanicus* from the sympatry zone; (4) *E. europaeus* from the sympatry zone; and (5) specimens from the sympatry zone with mixed (hybrid) genotype. Groups 3–5 were identified based on the results of Structure and BAPS analyses.

Principal component analysis (PCA) was used to examine the pattern of variation in the combined sample of the two species based on the genotypes of 11 microsatellite loci and *TTR*. The analysis was performed in STATISTICA for Windows v.8 (StatSoft, 1998).

ANALYSIS OF MITOCHONDRIAL VARIATION

Complete or nearly complete *Cytb* sequences (1101–1140 bp) were obtained for 221 specimens. The final alignment employed for evaluation of genetic diversity and estimation of demographic parameters was truncated to 1101 bp. For 16 specimens, only short fragments (< 600 bp) could be sequenced; these

data were used for identification of the mitotype group at the species level only. The number of haplotypes (H), haplotype diversity (h), nucleotide diversity (π), Tajima's D and Fu's F_s neutrality tests and mismatch distribution parameters (τ values, SSD for the sudden expansion model, raggedness index and corresponding P -values) were calculated using Arlequin v.3.5.1.3 (Excoffier & Lischer, 2010). Empirical distribution of statistics produced by the mismatch distribution analysis was obtained based on 1000 bootstrap pseudoreplicates.

To estimate the change in effective population size over time, a general pattern of demographic history for the main populations was described using the Bayesian Skyline Plot method (Drummond *et al.*, 2005) as implemented in BEAST v.1.8.4 (Drummond & Rambaut, 2007). Stepwise-constant skyline model with 8–10 groups was employed. The alignment was partitioned into 1st+2nd and 3rd codon positions. The optimum models were selected under the BIC criterion using the standard routine implemented in Treefinder (Jobb, 2011). The MCMC procedure was run for each species or population group with 100 000 000 iterations, and the genealogy and parameters of the model were stored every 20 000 steps. Tracer v.1.6 (Rambaut *et al.*, 2014) was used to assess the convergence of chains and to generate the Bayesian skyline plots. The burn-in was set to 10 000 000 iterations in each run. The *Cytb* substitution rate applicable for recent events in hedgehog evolutionary history is unknown; therefore, the skyline analyses produced relative time estimates measured in substitutions per site and a relative population size measured in units of $\theta = 2N_e\mu$ (where $2N_e$ is the effective size of mitochondrial population and μ is the substitution rate).

Relationships among haplotypes were illustrated using median-joining networks constructed in NETWORK v.4.5.0.0 (Bandelt *et al.*, 1999) and POPART v.1.7 (Leigh & Bryant, 2015) under default options. Phylogenetic neighbour-joining analysis was performed using PAUP* v.4.0b10 (Swofford, 2003). To assess clade stability, 1000 bootstrap pseudoreplicates were analysed.

RESULTS

CHARACTERISTICS OF MICROSATELLITE LOCI

In the microsatellite analysis, 181 specimens were genotyped. Characteristics of 10 microsatellite loci and a fragment of *TTR* intron 1 are presented in Supporting Information (Table S3). No evidence of genotyping error owing to stuttering or large allele dropout was observed; however, the analysis indicated the presence of null alleles at some loci. In allopatric populations, Micro-Checker detected null

alleles at two microsatellite loci (*EEU6* and *EEU43* with frequencies of 0.245 and 0.179, respectively) for *E. europaeus* and eight loci (*EEU1*, *EEU2*, *EEU4*, *EEU5*, *EEU6*, *EEU37*, *EEU43*, and *EEU54* with frequencies ranging from 0.0669 to 0.160) for *E. roumanicus*. In the sympatry zone for *E. europaeus* at the loci *EEU2*, *EEU6*, *EEU12* and *EEU37*, the presence of null alleles with probable frequencies ranging from 0.124 to 0.333 was assumed. For *E. roumanicus* from the sympatry zone, the presence of null alleles in the loci *EEU2*, *EEU4*, *EEU5*, *EEU37*, *EEU43* and *EEU54* was assumed with probable frequencies of 0.088–0.264.

In the allopatric populations, no homozygotes for a null allele were detected. In the sympatry zone, specimens homozygous for a null allele were found at the locus *EEU12*. Owing to the fact that in the sample of individuals with a mixed genotype at the *EEU12* locus, a highly probable presence of the null allele and a high percentage of homozygous for the null allele (~10.5%) were observed, this locus was excluded from further analysis.

RESULTS OF STRUCTURE AND BAPS ANALYSES

Allelic frequency analysis of the entire sample ($N = 181$) based on 10 microsatellite loci and a fragment of *TTR* intron 1 was performed in the Structure and BAPS programs with a different number of groups (from 2 to 15). With the number of groups $K = 2$, results indicated the presence of two species, *E. roumanicus* and *E. europaeus*, as well as individuals with a mixed genotype (Fig. 2).

With the optimal number of groups ($K = 3$) defined by the Structure Harvest program (Earl & von Holdt, 2012), the three “pure” populations correspond to: (1) *E. roumanicus* from South Russia (Krasnodar, Rostov and the Stavropol regions; Dagestan, Kalmykia, Crimea and Kabardino-Balkaria) and south-central Russia (Penza and Saratov), south-west Ural and West Siberia; (2) *E. roumanicus* from Central Russia (sympatry zone of the Moscow region; Bryansk, Kaluga, Ryazan and Tula) and south-central Russia (Lipetsk); (3) *E. europaeus* from all the studied regions. Thus, *E. roumanicus* was divided into two groups in this analysis, whereas *E. europaeus* appeared to be homogeneous. Additionally, one can recognize hybrid genotypic categories comprising individuals with admixed ancestry from groups A and B on the one hand and from groups C and B on the other (Supporting Information, Fig. S1B).

Based on the results of Structure with $K = 2$, we attributed 21 specimens (12%) from the sympatry zone of *E. roumanicus* and *E. europaeus* to the hybrid category. Among them, in ten specimens, the 95% confidence intervals for Q (the proportion of ancestry)

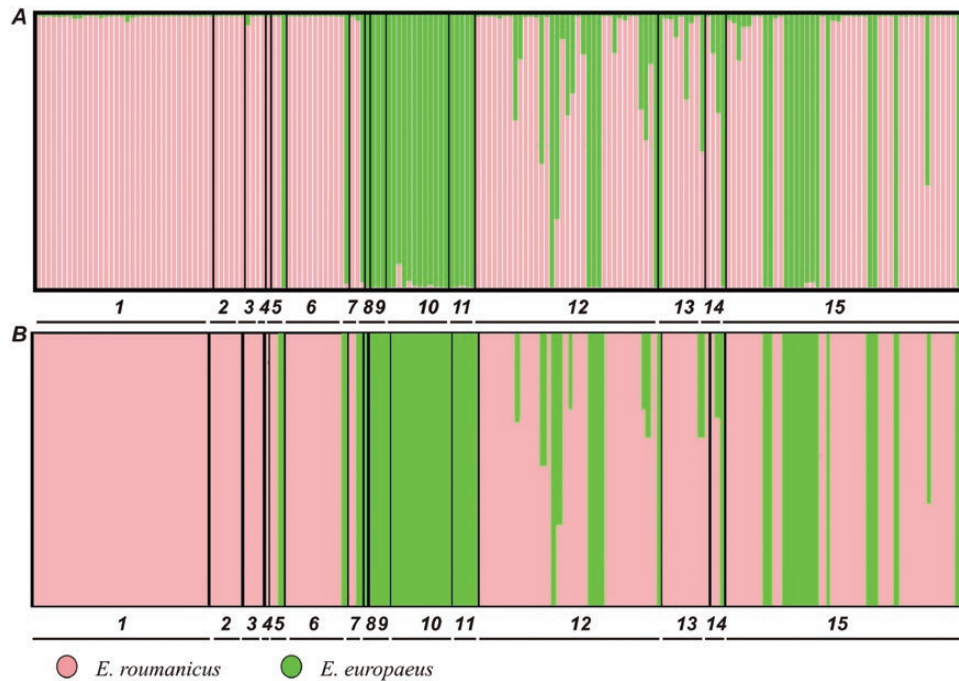


Figure 2. The probability that each individual from different parts of the range belonged to a particular cluster identified by the analysis of frequencies of alleles of microsatellite and *TTR* loci in (A) Structure v.2.3.4 and (B) BAPS v.5.2 under $K = 2$. The geographic information about specimens is as follows (Fig. 1): 1 – South Russia: Krasnodar (loc. 10, 28), Rostov (loc. 26, 27, 32), and Stavropol (loc. 33) regions, Dagestan (loc. 40), Kalmykia (loc. 39), Crimea (loc. 3) and Kabardino-Balkaria (loc. 34); 2–5 – Central Russia: Bryansk (loc. 5), Kaluga (loc. 7–9), Ryazan (loc. 23), Tula (loc. 11) and Nizhny Novgorod (loc. 36), Lipetsk (loc. 22), Saratov (loc. 38, 41) and Penza (loc. 35, 37) regions; 6 – Siberia: Kurgan (loc. 46–50), Omsk (loc. 53), Novosibirsk (loc. 55) and Tyumen (loc. 51) regions; 7 – West Ural: Bashkortostan (loc. 44, 45) and Udmurtia (loc. 42, 43); 8–10 – north and north-west Russia: Valdai (loc. 1), Kostroma (loc. 31) and Tver (loc. 2–4) regions; 11 – Central Russia, north Moscow region (loc. 15); 12 – Central Russia, east Moscow region (loc. 19–21); 13 – Central Russia, south Moscow region (loc. 16–18); 14–15 – Central Russia, west of Moscow region (loc. 12–14).

produced by Structure did not include 0 or 1; therefore these animals were classified as definite hybrids. Eleven other specimens with $Q < 0.92$ were classified as suspected hybrids. This threshold was based on the observation that, in pure populations, the value of Q is close to 1 (mean values are 0.992 and 0.989 for *E. roumanicus* and *E. europaeus*, respectively) with the lowest value being 0.922. In subsequent analyses, the two categories of hybrids were combined in one.

Genotyping of each of the 181 specimens using nuclear loci and *Cytb* finally resulted in a total sample containing 45 *E. europaeus* (including 25 individuals from the sympatry zone), 115 *E. roumanicus* (including 56 individuals from the sympatry zone) and 21 hybrids (Supporting Information, Table S1).

Using NewHybrids v.1.1 beta (Anderson & Thompson, 2002), six distinct genotype frequency classes were simulated, including pure species I (*E. europaeus*), pure species II (*E. roumanicus*), F1 *E. europaeus* × *E. roumanicus* hybrids, F2 hybrids (F1 hybrid × F1 hybrid) and backcrosses to pure species

(F1 hybrid × *europaeus* and F1 hybrid × *roumanicus*). As a result, among the 102 animals from the contact zone, a total of 19 specimens were classified as probable hybrids under the criterion that the posterior probability of belonging to pure species is below 0.1. Within these, 16 hybrids were attributed to some particular hybrid category with posterior probabilities of 0.72 to 0.97 (Table 2, marked in bold). Four specimens had a relatively high probability (> 0.72) of being classified as F1 *E. europaeus* × *E. roumanicus* hybrids, one appeared as a F2 hybrid ($P = 0.88$), and 11 were classified as backcrosses to *E. roumanicus* ($P > 0.76$). No backcrosses to *E. europaeus* were detected unambiguously; however, few specimens could potentially belong to this genotype category, albeit with a rather low probability ($0.2 < P < 0.5$). Genotypes of three putative hybrids could not be unambiguously assigned to any of the four hybrid categories (Table 2, probabilities are shown in bold italics). Finally, for some specimens considered to be “pure” based on Structure results (shown in italics in Table 2), there

Table 2. Summary of microsatellite+*TTR1* and mtDNA genetic data for hedgehogs from the Moscow region contact zone. The **Q** values from Structure and posterior probability assignments to different genotype frequency classes obtained with NewHybrids (F1, F2, backcross to *E. europaeus* (BxEe) and backcross to *E. roumanicus* (BxEr)) are provided. The 95% credibility intervals generated by Structure are shown in brackets. Bold text indicates values that definitely confirm hybrid origin of specimen. The values of mixed but unclassified genotypes (maximum $P < 0.7$) are marked in bold italics. The values for admittedly pure specimens but with a probability of admixture of > 0.1 are shown in italics. Hybrids identified by Structure are marked with asterisks. The data on specimens that are unambiguously ($P > 0.9$) classified as pure *E. roumanicus* or pure *E. europaeus* are omitted

ID	NewHybrids posterior probabilities of class membership									
	Structure: Q (95%CI)									
	<i>E. europaeus</i>	<i>E. roumanicus</i>	<i>E. europaeus</i>	<i>E. roumanicus</i>	F1	F2	BxEe	BxEr	<i>Cytb</i>	
Ch4	0.013 (0.000–0.141)	0.987 (0.859–1.000)	0.000	0.670	0.000	0.003	0.000	0.327	Ee	
Ch13	0.009 (0.000–0.103)	0.991 (0.897–1.000)	0.000	0.898	0.000	0.0006	0.000	0.101	Er	
PTZ11-1	0.033 (0.000–0.266)	0.967 (0.734–1.000)	0.000	0.837	0.000	0.002	0.000	0.161	Ee	
ZBS07-2	0.032 (0.000–0.265)	0.968 (0.735–1.000)	0.000	0.883	0.000	0.002	0.000	0.115	Er	
ZBS07-4	0.046 (0.000–0.318)	0.954 (0.682–1.000)	0.000	0.709	0.000	0.007	0.000	0.284	Er	
ZBS09-1	0.047 (0.000–0.290)	0.953 (0.710–1.000)	0.000	0.143	0.000	0.009	0.000	0.848	Er	
ZBS10-4	0.974 (0.780–1.000)	0.026 (0.000–0.220)	0.230	0.000	0.000	0.342	0.428	0.000	Ee	
Ch6m*	0.739 (0.387–1.000)	0.261 (0.000–0.613)	0.0003	0.000	0.382	0.318	0.282	0.018	Er	
Ch8f*	0.394 (0.085–0.726)	0.606 (0.274–0.915)	0.000	0.001	0.007	0.230	0.0003	0.762	Ee	
Ch8m*	0.173 (0.000–0.477)	0.827 (0.523–1.000)	0.000	0.002	0.000	0.025	0.000	0.973	Ee	
Ch12f*	0.556 (0.271–0.840)	0.444 (0.160–0.729)	0.000	0.000	0.840	0.061	0.022	0.077	Er	
Ch15m*	0.766 (0.438–1.000)	0.234 (0.000–0.562)	0.0001	0.000	0.916	0.028	0.045	0.011	Er	
Ch16*	0.77 (0.338–1.000)	0.23 (0.000–0.662)	0.001	0.000	0.750	0.083	0.044	0.124	Ee	
Ch17*	0.088 (0.000–0.431)	0.912 (0.569–1.000)	0.000	0.729	0.000	0.010	0.000	0.261	Ee	
Ch18f*	0.398 (0.127–0.718)	0.602 (0.282–0.873)	0.000	0.000	0.048	0.080	0.0004	0.872	Ee	
Ch18m*	0.289 (0.035–0.590)	0.711 (0.410–0.965)	0.000	0.001	0.004	0.033	0.000	0.963	Ee	
Ch24*	0.159 (0.000–0.485)	0.841 (0.515–1.000)	0.000	0.026	0.001	0.054	0.000	0.920	Er	
Ch51*	0.144 (0.000–0.499)	0.856 (0.501–1.000)	0.000	0.033	0.001	0.053	0.0001	0.914	Ee	
Ch69*	0.35 (0.096–0.663)	0.65 (0.337–0.904)	0.000	0.000	0.023	0.063	0.001	0.914	Ee	
Ch70*	0.464 (0.172–0.802)	0.536 (0.198–0.828)	0.0002	0.000	0.005	0.888	0.055	0.052	Ee	
Ch71*	0.186 (0.000–0.486)	0.814 (0.514–1.000)	0.000	0.012	0.0001	0.043	0.000	0.945	Er	
Krasng*	0.145 (0.000–0.485)	0.855 (0.515–1.000)	0.000	0.689	0.000	0.114	0.000	0.197	Ee	
NK2013*	0.36 (0.096–0.677)	0.64 (0.323–0.904)	0.000	0.001	0.001	0.507	0.001	0.491	Ee	
PTZ09-2*	0.317 (0.041–0.638)	0.683 (0.362–0.959)	0.000	0.001	0.010	0.054	0.0004	0.936	Ee	
PTZ13-2*	0.502 (0.176–0.841)	0.498 (0.159–0.824)	0.000	0.000	0.720	0.045	0.009	0.226	Er	
PTZ08-2*	0.084 (0.000–0.418)	0.916 (0.582–1.000)	0.000	0.048	0.0002	0.033	0.000	0.919	Ee	
ZBS07-3*	0.17 (0.000–0.489)	0.83 (0.511–1.000)	0.000	0.099	0.000	0.030	0.000	0.871	Ee	
ZBS12-4*	0.623 (0.325–0.954)	0.377 (0.046–0.675)	0.000	0.000	0.105	0.640	0.250	0.006	Er	

was a non-negligible probability (> 0.1) that they still contained an admixture from an alien species (e.g. Ch4 and ZBS07-4). These individuals were not treated as hybrids in subsequent analyses because NewHybrids did not reject the hypotheses that they belonged to the pure class while the values of Q inferred for them by Structure were low.

In summary, in the sample from the Moscow sympatry zone ($N = 102$), the proportion of admixed individuals was $\sim 20\%$. Among them, the most numerous category comprised backcrosses to *E. roumanicus* ($\sim 12\%$); however, F1 *E. europaeus* \times *E. roumanicus* hybrids were also found, and their proportion was tentatively estimated as 4%.

MICROSATELLITE ANALYSIS OF GENETIC DIVERSITY

We analysed genetic diversity separately for the samples of allopatric populations of the two species and for the populations from the sympatry zone as identified from the results of Structure. Each of the ten microsatellites was polymorphic and had 2–12 alleles for *E. europaeus*. In *E. roumanicus*, one locus (*EEU36*) was monomorphic, and the rest were polymorphic with 3–16 alleles. On average, the microsatellite diversity in *E. roumanicus* (9.27 alleles per locus) was higher than that in *E. europaeus* (5.91 alleles per locus).

In the sympatry zone, both species were polymorphic at all ten loci with 2–16 alleles. Allelic richness of

E. europaeus was higher in the sympatry zone than in allopatric populations for six out of 11 loci. Allelic composition of *E. roumanicus* was higher in the sympatry zone for four loci, whereas it was higher in allopatric populations for five loci. However, on average, in both species, the number of alleles per locus in the sympatry zone and in allopatric populations did not significantly differ (Supporting Information, Tables S4–S7). For more detailed analytical results, see Supporting Information 2.

The pattern of distribution of genotypes in the plane of the first two principle components showed two well-defined clusters corresponding to the two species. The species were separated along the first PC, which accounts for 23.2% of the total variation, while the other PCs were responsible for less than 5% each. Individuals from the sympatry zone with admixed genotypes were placed between these two groups (Fig. 3), and their distribution was apparently biased towards *E. roumanicus*.

MITOCHONDRIAL VARIABILITY AND INTROGRESSION

The distribution of *E. europaeus* and *E. roumanicus* mitotypes among genotypes found in the SSR+*TTR* analysis using Structure is shown in Figures 3B and 4. Mitochondrial neighbour-joining reconstruction revealed two groups, mainly corresponding to *E. europaeus* and *E. roumanicus* (Fig. 4). However, two

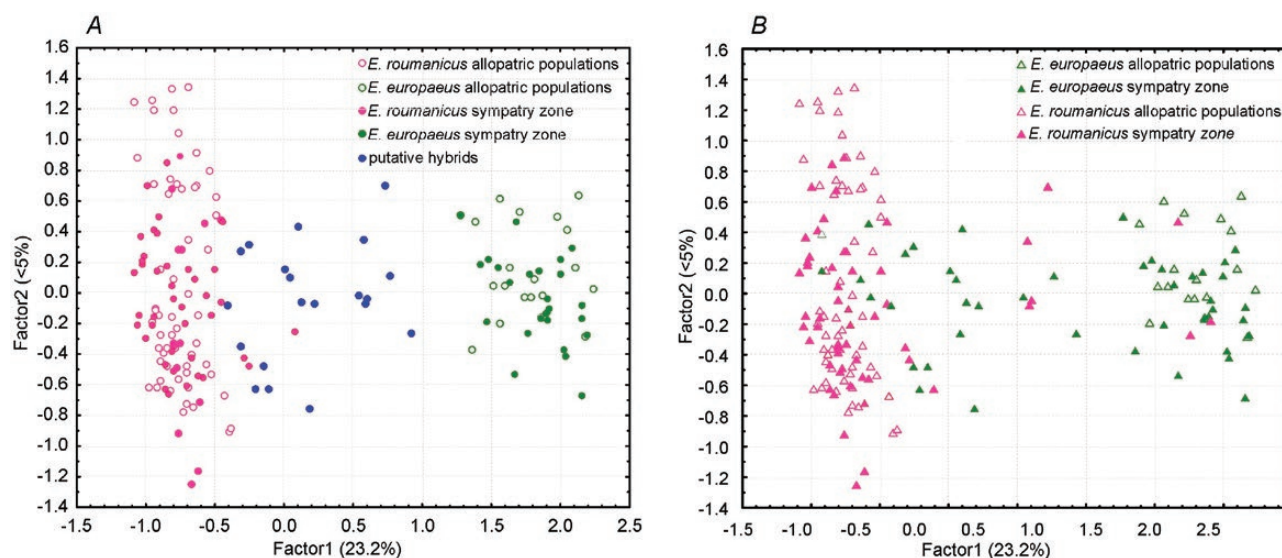


Figure 3. Ordination of genotypes based on the analysis of 11 microsatellites and the *TTR* loci as produced by the PCA. A, strong differentiation among *E. europaeus* (green) and *E. roumanicus* (pink) at nuclear loci was observed. Putative hybrids (blue) were localized at an intermediate position. B, correspondence of *Cytb* mitotypes to the identified groups of nuclear genotypes. With a clear distribution of specimens among mitogroups of *E. europaeus* and *E. roumanicus*, hybrid specimens from the sympatry zone mainly possess the mitotype of *E. europaeus*; introgressed mitotypes are found mostly in *E. roumanicus*.

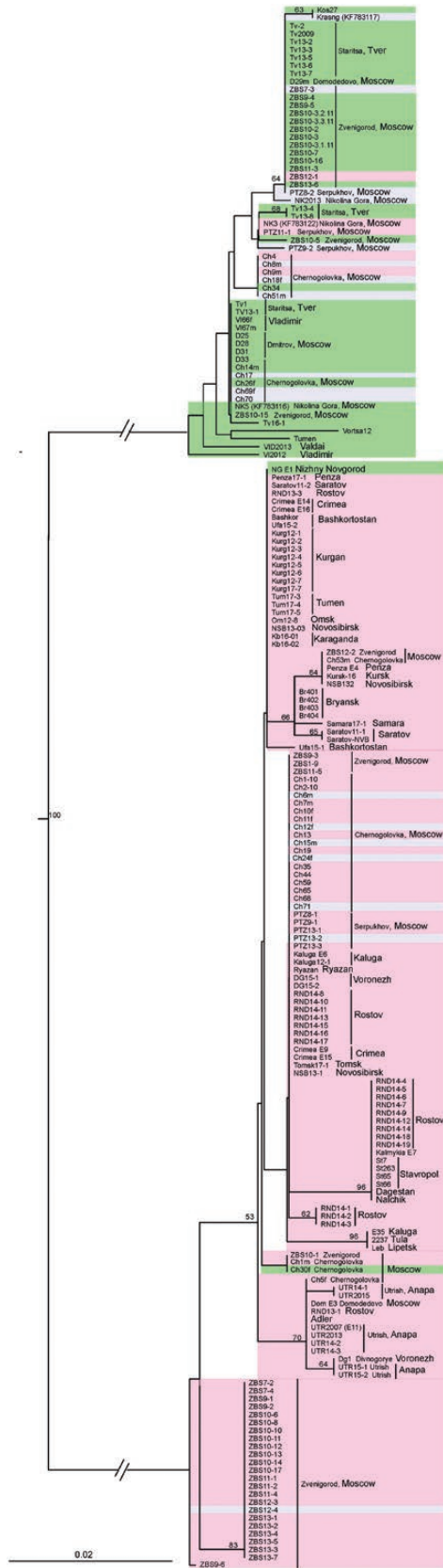


Figure 4. A neighbour-joining tree illustrating the differentiation of 196 specimens of *E. europaeus* and

specimens identified as *E. europaeus* in the SSR+*TTR1* analysis clustered with *E. roumanicus* on the *Cytb* tree, and five specimens identified as *E. roumanicus* clustered in the *E. europaeus* *Cytb* group. Thus, about 13% of our sample of *E. roumanicus* and *E. europaeus* possessed the alien mitotype. Among the 21 hybrids found in the SSR+*TTR* analysis, 14 individuals (66.7%) had a *Cytb* of *E. europaeus*, and only seven hybrids (33.3%) had a *Cytb* of *E. roumanicus*. The majority of specimens with the alien mitotype were found in the sympatry zone of *E. roumanicus* and *E. europaeus*, mainly in the east (loc. 19) and south (loc. 18) of the Moscow region. Only four hybrids and two *Cytb* introgressed specimens were found in the west of the Moscow region (locs. 12–14). It is worth mentioning that a specimen with an alien mitotype was also found at the site located ~ 350 km east of Moscow: *E. europaeus* from the Nizhny Novgorod region (loc. 36) possessed the mitotype of *E. roumanicus*. Detailed information on the geographical distribution of the *E. europaeus* and *E. roumanicus* *Cytb* haplotypes in Russia is presented in Figure 4.

CORRESPONDENCE OF MORPHOLOGICAL AND GENETIC DIAGNOSTICS

The genotyping of each of the 181 specimens using nuclear loci and *Cytb* resulted in a total sample of $N = 181$ containing 45 *E. europaeus* (including two specimens with the alien mitotype), 115 *E. roumanicus* (including five with the alien mitotype) and 21 hybrids (Supporting Information, Table S1). For 79 specimens from the allopatric populations and 71 hedgehogs from the zone of sympatry (with 11 hybrids among them), species affiliation was identified by morphological characteristics and compared with the results of genotyping. In the allopatric populations, neither *E. roumanicus* ($N = 59$) nor *E. europaeus* ($N = 20$) deviated from what we expected from the phenotype. In the zone of sympatry, morphological identification deviated from the genetic one in 19% of cases (14 of 73). Most of these individuals were genetically identified as putative hybrids; however, two individuals were recognized as pure species (D29 and NK3) by the Structure and NewHybrids programs. Among the putative hybrids, 73% had a morphotype of *E. roumanicus*, and only one hedgehog (Ch18f) was characterized by intermediate external and skull morphology. On

E. roumanicus based on the mitochondrial *Cytb* gene. Numbers above branches denote bootstrap support (1000 pseudoreplicates). Colours designate species identification based on nuclear data: green – pure *E. europaeus*, pink – pure *E. roumanicus*, blue – putative hybrids.

Table 3. Morphotype and genotype matching variants and their proportion in the zone of sympatry ($N = 73$)

Variants of morphotype and genotype matching	n
Pure <i>E. europaeus</i> with morphotype of <i>E. europaeus</i>	22
Pure <i>E. roumanicus</i> with a morphotype of <i>E. roumanicus</i>	37
Pure <i>E. roumanicus</i> with a morphotype of <i>E. europaeus</i>	1
Pure <i>E. roumanicus</i> with an intermediate morphotype	1
Pure <i>E. europaeus</i> with an intermediate morphotype	1
Hybrids with a morphotype of <i>E. roumanicus</i>	8
Hybrids with a morphotype of <i>E. europaeus</i>	2
Hybrids with an intermediate morphotype	1

the whole, an intermediate morphotype was found in three specimens (*E. europaeus* D29, *E. roumanicus* NK3 and a backcross to *E. roumanicus* Ch18f) from the total sample of 73 (Table 3).

PHYLOGEOGRAPHICAL STRUCTURE

The relationships among the *Cytb* haplotypes were examined using median-joining networks (Fig. 5; Supporting Information, Fig. S2). To evaluate the genetic diversity of species according to the most complete geographical selection, using all sequences available from the GenBank, and to compare our data with those for hedgehogs from Western Europe, our original sequences were truncated to 382 bp, which corresponds to the length of the alignments from Santucci *et al.* (1998) and Seddon *et al.* (2001, 2002).

The phylogeographic structure of *E. europaeus* (Fig. 5A) was clear and demonstrated the presence of monophyletic haplogroups. The first haplogroup (I) contained only Western European haplotypes from Spain northwards through France and the Netherlands and into the UK and Ireland and corresponded to the E2 clade of mitotypes in Seddon *et al.* (2001). It was subdivided into two subgroups: Ia produced a star-like pattern around the haplotype that is common in Western, Central and north-central Europe; and Ib included the majority of the Iberian haplotypes and two haplotypes shared with individuals from other parts of Western Europe. One more haplotype from Sicily appeared to be far removed from both subgroups (12 substitutions) and was related to the E3 clade in Seddon *et al.* (2001). Finally, the largest and most distant (14 substitutions) haplogroup II combined hedgehogs from Eastern Europe and Western Siberia with populations of

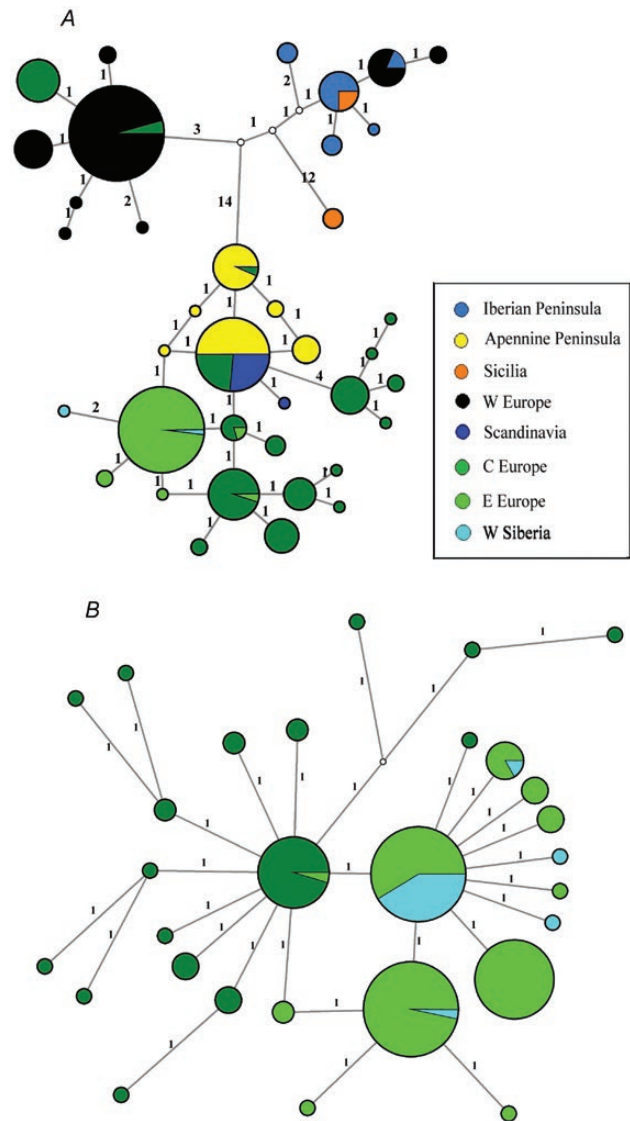


Figure 5. Median-joining network of the *Cytb* haplotypes for (A) *E. europaeus* ($N = 304$) and (B) *E. roumanicus* ($N = 162$) of Europe and Western Siberia based on the 382 bp. Haplotypes are represented by circles, which are proportional to the haplotype frequency. The number of substitutions between haplotypes is shown on each branch of the net. Geographical locations of haplotypes are grouped as follows: (1) Iberian Peninsula (Spain and Portugal); (2) Apennine Peninsula; (3) Sicily; (4) Western Europe (Great Britain, France, Ireland and Jersey); (5) Scandinavia (Sweden, Norway and Finland); (6) Central Europe (Germany, North Italy, Czech Republic, the Netherlands, Belgium, Poland, Switzerland, Hungary, Austria and Estonia); (7) Eastern Europe (European Russia); and (8) Western Siberia.

Central and north-central Europe, and the Apennine and Scandinavian Peninsulas, and corresponded to the E1 clade in Seddon *et al.* (2001).

The median-joining network for *E. roumanicus* showed little structure (Fig. 5B), but some trends were apparent. The haplotypes from Western Europe from one side and Eastern Europe and Siberia from the other side tended to group together but with a low level of differentiation. None of the haplotypes from European Russia and Siberia were present in the Western European populations.

To evaluate the genetic diversity of species from Russia and the adjacent eastern part of their ranges, the relationships among the *Cytb* haplotypes were examined based on 1101 bp (Supporting Information, Fig. S2). Among 15 haplotypes of *E. europaeus*, several haplotypes from North Russia (e2, e14 and e15), and haplotypes from Siberia (e1) and South Ural (e3) are the most divergent (Supporting Information, Fig. S2A). In populations of *E. roumanicus* from Russia and Kazakhstan, 16 haplotypes are found (Supporting Information, Fig. S2B). The most divergent East European haplotype (r4) belongs to the populations from the west of Moscow region (Fig. 1: loc. 12). Several haplotypes are unique for South Russia (Rostov region: loc. 26, 27 and 32); however, no haplotypes endemic to Siberia were revealed.

GENETIC VARIABILITY AND DEMOGRAPHICAL ANALYSIS

In Eastern Europe and Western Siberia, we identify four subpopulations of *E. roumanicus* and two of *E. europaeus* (Table 4). This subdivision scheme is partly based on the results of the analysis in Structure (see above) but contains additional geographical subsamples to present a more detailed pattern of geographic variation.

For *E. roumanicus*, the four geographical subpopulations correspond to: (1) the sympatry zone

(Moscow region); (2) South Russia (Stavropol region, Rostov region, Crimea, Krasnodar region, Kalmykia, Dagestan and Kabardino-Balkaria); (3) South Russia and Central Russia without the sympatry zone (Tula, Saratov, Ryazan, Penza, Kaluga and Bryansk regions); and (4) Western Siberia + Northern Kazakhstan (Novosibirsk, Omsk and Kurgan regions; Karaganda province).

For *E. europaeus*, two geographical subpopulations correspond to: (1) North Russia (Tver, Kostroma and Vladimir regions); and (2) the Moscow region sympatry zone.

The genetic variability indices are presented in Table 4. Demographical parameters are provided for each species in their sympatric and allopatric areas (Table 5; Fig. 6). All analyses were also performed for the total samples of both species in Russia (Table 5; Figs 6, 7).

As a result, comparison of summary statistics of *Cytb* diversity in the populations of both species from Western and Central Europe from one side (Bolfíková & Hulva, 2012; Bolfíková et al., 2017) and Eastern Europe (European Russia) from the other side (Table 4) revealed significantly higher haplotype diversity in Western and Central Europe for *E. europaeus* but not for *E. roumanicus*. Tests of neutrality yielded negative values indicating recent population growth for all geographical populations of both species.

Bayesian skyline plots indicated a recent increase both in sympatric (Fig. 6A) and allopatric (Fig. 6C) populations of *E. europaeus*. In contrast to that, no expansion is evident in *E. roumanicus* (Fig. 6B, D). Both Fu's *F_s* and Tajima's *D* neutrality statistics are significantly negative for the East European *E. europaeus* (Table 4), thus indicating population growth in this species. In contrast, neither of the

Table 4. The summary statistics of *Cytb* diversity in the populations of *E. roumanicus* and *E. europaeus* and the results of neutrality tests: number of individuals (*n*), haplotype diversity (*h*), nucleotide diversity (π), Tajima's *D*, Fu's *F_s* and *p*-values. Neutrality statistics significant at $P < 0.05$ are marked with bold text

Sample	<i>n</i>	<i>h</i>	π	Tajima's <i>D</i>	Tajima's <i>D</i> p-value	Fu's <i>F_s</i>	Fu's <i>F_s</i> p-value
<i>E. europaeus</i>							
N Russia	17	0.7941 ± 0.0783	0.0020 ± 0.0014	-1.1950	0.1090	-1.2707	0.2010
Moscow region	42	0.7526 ± 0.0443	0.0014 ± 0.0008	-0.1813	0.4690	-2.5764	0.0890
Russia total (1101 bp)	61	0.7787 ± 0.0395	0.0018 ± 0.0010	-1.6852	0.0250	-6.4893	0.0070
<i>E. roumanicus</i>							
Moscow region	62	0.6378 ± 0.0366	0.0026 ± 0.0015	1.2539	0.9120	2.9120	0.8750
South Russia	41	0.7817 ± 0.0438	0.0027 ± 0.0016	1.2292	0.9100	1.4704	0.7860
S+C Russia without Moscow	65	0.8692 ± 0.0225	0.0028 ± 0.0016	-0.2221	0.4510	-1.5487	0.2960
Siberia+N Kazakhstan	18	0.3072 ± 0.1316	0.0004 ± 0.0004	-1.4014	0.0680	0.2390	0.0002
Russia+Kazakhstan (1101 bp)	145	0.8389 ± 0.0176	0.0029 ± 0.0016	-0.3468	0.4470	-1.5877	0.3360

Table 5. The results of mismatch analysis based on *Cytb* data in the populations of *E. roumanicus* and *E. europaeus*: number of individuals (*n*), time in substitutions (*tau*), relative *tau* (substitutions per site), SSD, Raggedness index (Rg) and p-values

Sample	<i>n</i>	<i>tau</i> and 90% CI	Relative <i>tau</i> (substitutions per site)	SSD	SSD p-value	Rg	Rg p-value
<i>E. europaeus</i>							
N Russia	17	2.7 (0.652–4.340)	0.0024	0.0537	0.1150	0.1837	0.0760
Moscow region	42	2.1 (0.746–3.631)	0.0009	0.0122	0.2820	0.0500	0.6260
Russia total (1101 bp)	61	2.4 (0.811–4.170)	0.0010	0.0152	0.2280	0.0589	0.3890
<i>E. roumanicus</i>							
Moscow region	62	6.0000 (0.014–14.234)	0.0027	0.1500	0.0660	0.3909	0.0020
South Russia	41	4.8000 (1.129–7.834)	0.0021	0.0441	0.1150	0.1179	0.0950
S+C Russia without Moscow	65	4.1000 (1.518–6.240)	0.0018	0.0143	0.2330	0.0490	0.2410
Siberia+N Kazakhstan	18	3.0000 (0.473–3.500)	0.0013	0.0014	0.6270	0.2637	0.6040
Russia+Kazakhstan (1101 bp)	145	4.7000 (1.264–7.728)	0.0021	0.0074	0.6350	0.0255	0.6770

tests are significant for *E. roumanicus*, which is consistent with population stability as demonstrated by the skylines. Considering sympatric populations of the two species, the mismatch distribution is pronouncedly multimodal in *E. roumanicus* and close to unimodal in *E. europaeus* (Fig. 7), which supports population expansion in the latter but not in the former species. The values of *tau* are lower for *E. europaeus* compared to *E. roumanicus* (Table 5), thus suggesting that the expansion in the eastern populations of the West European hedgehog could be a relatively recent event. It is necessary to note that the results of demographic analyses should be treated with caution because of their sensitivity to population structuring, departure from neutrality and other factors (Grant, 2015).

DISCUSSION

GEOGRAPHIC DISTRIBUTION OF HEDGEHOGS BASED ON GENETIC IDENTIFICATION

Morphological variability within species of *Erinaceus* often confuses species identification, especially when based solely on external characteristics, which can vary considerably by individual and age (Zaitsev, 1984, 2002). Our data demonstrated a high probability of a match between morphological and genetic diagnosis in the allopatric populations. However, the situation was more complex in the zone of sympatry. In the Moscow contact zone, intermediate morphotypes were found, and the morphologically ‘pure’ species were genotyped as hybrids. Thus, our analyses showed that the identification of species in the zone of sympatry

can be challenging, and genotyping becomes necessary. Furthermore, genotyping should not be limited to one or two loci. Multilocus analysis is essential to facilitate species distinction, as well as detection of hybrids and their classification.

Genotyping allowed us to determine accurately the distribution boundaries of species in the sympatry zone and at the eastern end of the range. The first report on the coexistence of *E. europaeus* and *E. roumanicus* in the territory of European Russia referred to the Surazhsky district of the Chernigov province (Ognev, 1928). A collection of both species sampled in Odintsovo in the vicinity of Zvenigorod (Lutsino village by V.V. Kucheruk in May 1941 and near the village of Nikolina Gora by N.A. Formozov in 2006) as well as from the vicinity of Noginsk (Chernogolovka village by M. Rutovskaya in 2010–2013) are stored in the Zoological Museum of Moscow State University. The presence of both *E. europaeus* and *E. roumanicus* among specimens sampled near Chernogolovka (Noginsky district) was confirmed from cytogenetic methods (Sokolov *et al.*, 1991). Our data also confirm the sympatry of these species in the Moscow region, both in the current study and our previous study (Bogdanov *et al.*, 2009).

The presented data on the Eastern European zone of sympatry refer to a territory in the Moscow region of about 50 000 km². In reality, the zone of sympatric distribution of hedgehogs must be wider since specimens genotyped as *E. europaeus* were found in three localities far to the east of the Moscow region (the Nizhny Novgorod region, Central Udmurtia and Tumen). The specimen from the Nizhny Novgorod region has a mitotype of *E. roumanicus*.

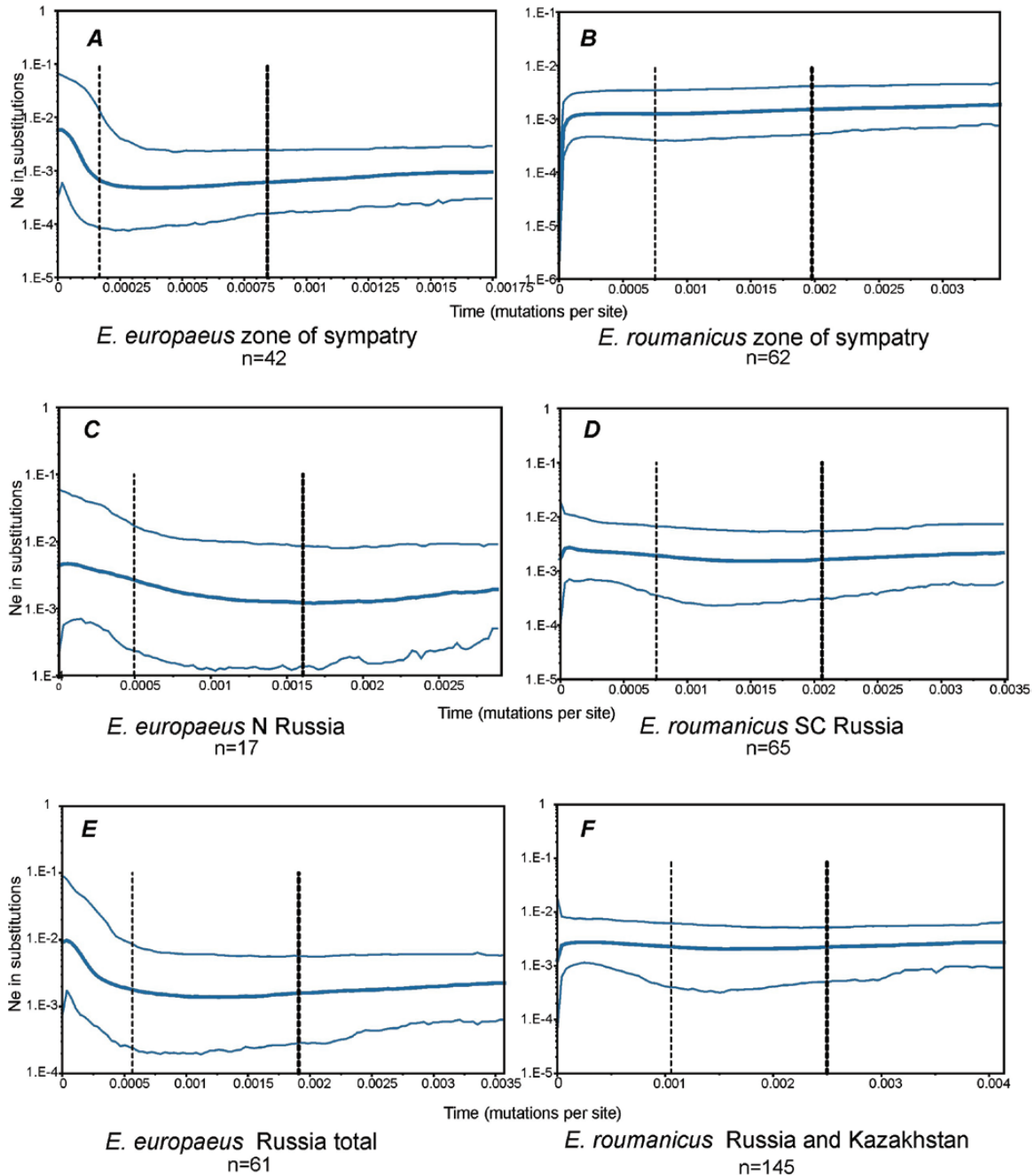


Figure 6. Bayesian skyline plots for large geographical samples of *Erinaceus* spp. based on *Cytb* sequences (1101 bp). The central thick line shows the median of the N_e estimation, and the 95% confidence intervals are indicated by the pale lines. Populations of the Moscow region sympatry zone of *E. europaeus* (A) and *E. roumanicus* (B), allopatric populations of *E. europaeus* of north-central Russia (C), allopatric population of *E. roumanicus* of south and central Russia (D), *E. europaeus* of Russia in total (E), *E. roumanicus* of Russia and Kazakhstan in total (F).

To date, published information on the boundaries of the ranges of hedgehogs has been based only on morphological data (Ognev, 1928; Zaitsev *et al.*, 2014), with the least studied distribution of both species

in the Urals and Western Siberia. In our study, the hedgehogs from South Ural and Western Siberia were genotyped for the first time, and both species were found, but not in the same localities. Among

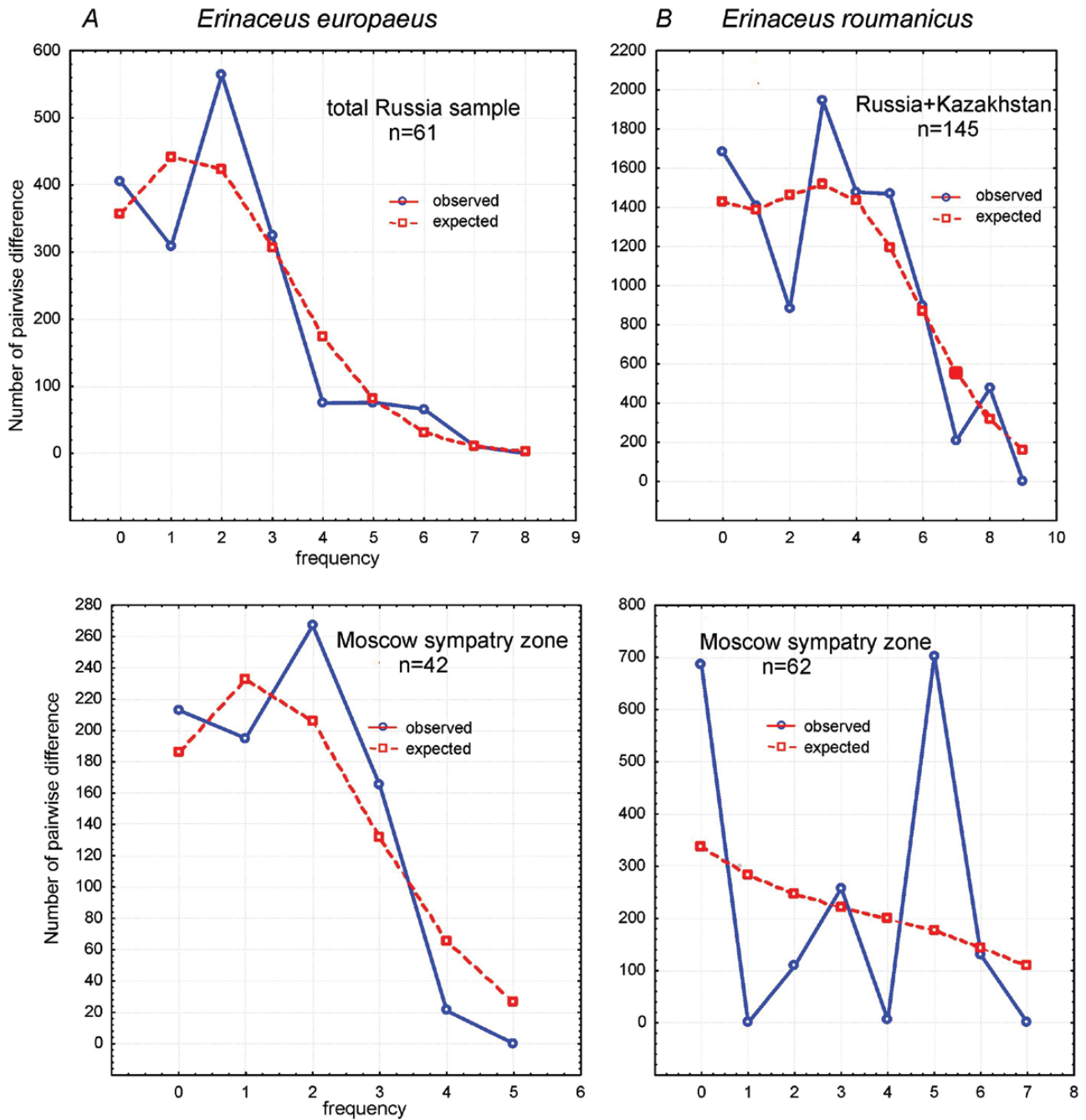


Figure 7. Mismatch distribution graphs for samples from the Moscow sympatry zone and the total samples of *E. europaeus* (A) and *E. roumanicus* (B).

them, *E. roumanicus* predominates. *E. europaeus* was present in the samples from Udmurtia and in the Western Siberia Tumen region.

HYBRIDIZATION IN THE ZONE OF SYMPATRY

Earlier, we have shown that *E. europaeus* and *E. roumanicus* diverged approximately 0.6–1.7

million years ago and are not the closest sister species of *Erinaceus*, with the *Cytb* K2P genetic distance between them being ~14% (Bannikova *et al.*, 2014). Nonetheless, the most important finding of our study is the evidence of hybridization of *E. europaeus* and *E. roumanicus* in the East European sympatry zone, in the territory of the Moscow region. In the sample of 102 specimens, 25 individuals were pure *E. europaeus*

(24.5%), 56 individuals were pure *E. roumanicus* (54.9%) and 21 individuals (20.6%) possessed mixed genotypes. Among pure specimens of *E. europaeus* and *E. roumanicus*, individuals with introgressed mitotypes were found, suggesting the occurrence of past introgression events. Thus, it is evident that the relatively long time since divergence of *E. europaeus* and *E. roumanicus* (c. 1 Mya) was, however, insufficient for the formation of effective reproductive barriers in allopatry. We may conclude that, among the three defined types of interspecific hybridization (Quilodrán *et al.*, 2014), the case in hedgehogs is of the type characterized by F1 hybrids undergoing recombination between homologous chromosomes during meiosis, resulting in genetic introgression from one species into the other and the formation of a hybrid zone. The high frequency of pure parental forms characterizes the hybrid zone as bimodal (Jiggins & Mallet, 2000), a pattern associated with the existence of pre-mating or post-mating barriers (Redenbach & Taylor, 2003; Bailey *et al.*, 2004; Minder *et al.*, 2007).

The majority of the genetically detected admixed individuals were confirmed with reliable accuracy as backcrosses to *E. roumanicus*. However, four of the 102 specimens corresponded to the first-generation (F1) hybrid class. This result means that, in the East European sympatry zone, there are footprints of not only ancient but also recent hybridization events.

Of the four individuals diagnosed as definite F1 hybrids based on NewHybrids analysis of nuclear data,

three presented *E. roumanicus* mitotypes. By contrast, backcrosses to *E. roumanicus* ($N = 12$) predominantly showed the mitotype of *E. europaeus* ($N = 10$; 83%). On the whole, 66.7% of all hybrids found in the Structure analysis ($N = 21$) had a *Cytb* of *E. europaeus*. In the Moscow zone of sympatry, the proportion of *E. roumanicus* ($N = 56$) with the *Cytb* of *E. europaeus* was two times higher ($N = 5$; 9%) than *E. europaeus* ($N = 25$) with the *Cytb* of *E. roumanicus* ($N = 1$; 4%). However, the sample size meant that the difference was not statistically significant ($P = 0.66$; Fisher's exact test).

Thus our data indicate predominantly reciprocal transfer of mtDNA in the two common hedgehogs and, at the same time, suggest partial asymmetry of mtDNA introgression [the prevalence of the *E. europaeus* mitotype in hybrids (Fig. 8C) and a greater number of *E. roumanicus* with the *E. europaeus* mitotype than vice versa (Fig. 8A, B, D)], which, however, should be verified with a larger sample. We note that pooling the samples of pure individuals and backcrosses (15/68 and 1/24 in *E. roumanicus* and *E. europaeus*, respectively) gives $P = 0.06$.

The proportion of different genotype classes in the analysis using NewHybrids and the Q value in Structure indicate that the majority of mixed genotypes were closer to *E. roumanicus* than to *E. europaeus* (expressed in the asymmetric distribution of Q values and proportion of backcross classes). This finding partly agrees with the results of

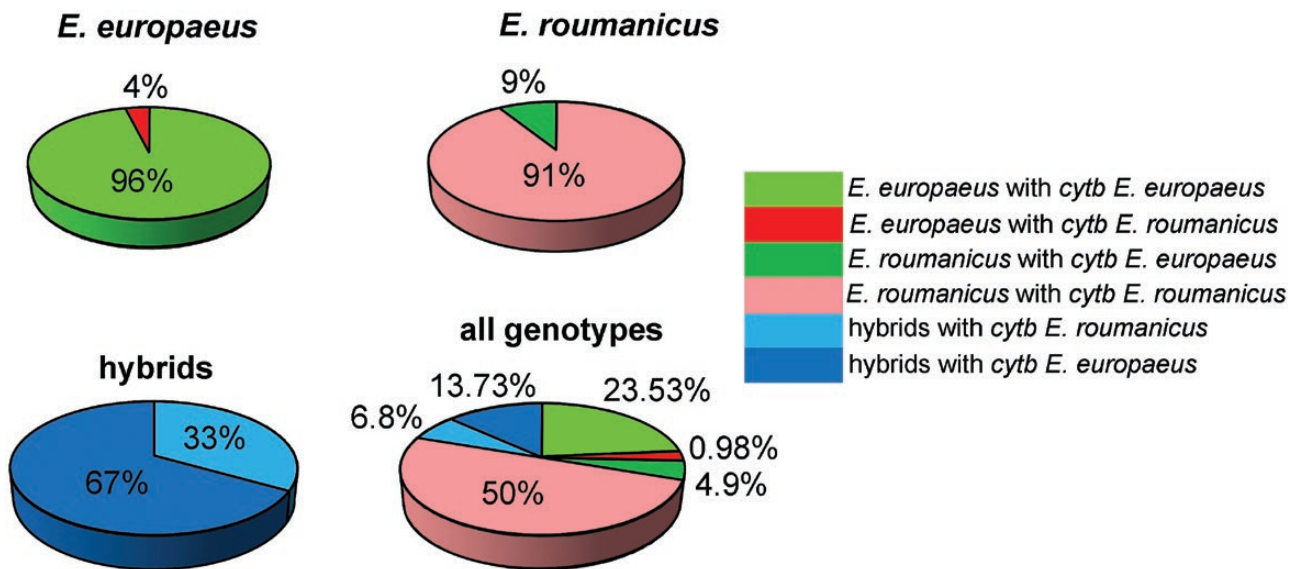


Figure 8. The ratio of different genotypes in the Moscow region sympatry zone: (A) *E. europaeus*, (B) *E. roumanicus*, (C) hybrids, (D) all genotypes.

hybridization experiments between *E. europaeus* and *E. roumanicus*, in which backcrosses were obtained by backcrossing ♀F1 hybrids and ♂*E. roumanicus* [♀F(♀*E. europaeus* × ♂*E. roumanicus*) × ♂*E. roumanicus*], whereas crossing between F1 hybrids and ♂*E. europaeus* was abortive (Poduschka & Poduschka, 1983). These authors also obtained a single F2 hybrid litter, which confirms that both sexes in the F1 are potentially fertile. The molecular data demonstrate that F1 females are fertile regardless of the direction of crossing; however, it remains unclear whether the same is true for males.

The predominance of hybrid nuclear genotypes with a higher proportion of *E. roumanicus* suggests that hybrids are more likely to join the *E. roumanicus* than the *E. europaeus* population. The observed pattern apparently deviates from the simplest model whereby an F1 hybrid chooses a mate randomly from the pool of pure specimens of the two species. In this case, one would expect that ~30% (25/81) of non-F1 admixed individuals will be closer to *E. europaeus* than to *E. roumanicus*, which is significantly different from the observed ratio of 0/12 ($P = 0.032$; Fisher's exact test). This phenomenon can be explained by: (1) asymmetric assortative mating in pure individuals or F1 hybrids as described for different vertebrates (Coyner *et al.*, 2015; Peters *et al.*, 2017; for review see Wirtz, 1999; Shurtliff, 2011); (2) dependence of survival and/or fertility of F1 hybrids (or backcrosses) on the direction of crossing as was suspected for the house mouse hybrid zone (Good *et al.*, 2008); and (3) the effect of drift following the wavefront model (Currat *et al.*, 2008) in which *E. roumanicus* is the expanding species.

In summary, our results indicate that interspecies hybridization in hedgehogs results in bi-directional introgression of both nuclear and mitochondrial genes. This pattern contrasts sharply to the numerous cases of unidirectional introgression of mtDNA, for example, from *Clethrionomys rutilus* to *Clethrionomys glareolus* (Tegelström, 1987; Deffontaine *et al.*, 2005; Boratyński *et al.*, 2011, 2014), from *Eptesicus nilssoni* to *Eptesicus serotinus* (Artyushin *et al.*, 2009), and from *Lepus timidus* to *Lepus europaeus*, *Lepus granatensis* and *Lepus castroviejoi* (Melo-Ferreira *et al.*, 2005), as well as several other hare species in Asia and North America (Alves *et al.*, 2008), and from different species of short-tailed ground squirrel *Spermophilus spp.* to *S. major* (Ermakov *et al.*, 2002; Spiridonova *et al.*, 2006). The latter phenomena are often explained by selective advantage of the introgressed mitotypes (Melo-Ferreira *et al.*, 2005; Ropiquet *et al.*, 2006; Boratyński *et al.*, 2014) being better adapted to a local environment or less affected by mutational load compared to the native mitotypes (Hill, 2019).

Additional experimental evidence is necessary to evaluate the fitness cost of mitochondrial introgression in hedgehogs (Boratynski *et al.*, 2014 or Bize *et al.*, 2018). Meanwhile, one should note that the available data on hedgehogs do not conform to the mitonuclear compatibility species concept advocated by Hill (2017), which suggests a principal role of mitonuclear coevolution in speciation.

THE TREND IN THE DISTRIBUTION OF PURE AND HYBRID FORMS

Our data did not allow us to perform a spatial analysis of the hybrid zone (there was a lack of representative samples on transects). However, we attempted to identify the main trends in the distribution of pure and hybrid forms. We found that the proportion of species differed in different parts of the sympatry zone. The proportion of *E. europaeus* increased in samples northwards compared with the samples from the southern parts of the sympatry zone. In the vicinity of Vladimir (loc. 29, 30), Dmitrov (loc. 15) and further north, only *E. europaeus* were found. The most southern locality for *E. europaeus* was in the western surroundings of Moscow (Fig. 1: loc. 12, 13), where *E. europaeus* occurred less often (~30%) than *E. roumanicus* (~63%). In this location, five *E. roumanicus* individuals possessed the mitotype of *E. europaeus* and two *E. europaeus* individuals had the mitotype of *E. roumanicus*.

On the other hand, specimens from the south of the contact zone predominantly belonged to *E. roumanicus*. Starting from the south of the Moscow region (loc. 18) and further southwards, our samples contained only *E. roumanicus*. Interestingly, in the sample from loc. 18 ($N = 9$), three hybrids and six *E. roumanicus* individuals were found. Pure *E. europaeus* was absent, but one *E. roumanicus* individual and three hybrids possessed the *Cytb* of *E. europaeus*. Most of the hybrids (67%) were found in the eastern part of the Moscow region (Fig. 1: loc. 19).

DIFFERENCE BETWEEN CENTRAL AND EASTERN EUROPEAN SYMPATRIC ZONES

Currently, only a small number of hybrids are known in the Central European sympatry zone: among 260 genotyped samples only one backcrossed hybrid with a mitotype of *E. roumanicus* was found in Slovakia (Bolfiková *et al.*, 2017) and two individuals out of 82 were detected as potential hybrids in Austria (Curto *et al.*, 2019). In contrast, according to our data, 21 specimens out of 102 in the Russian contact zone turned out to be plausible hybrids. The reason for the low number of hybrids of the common hedgehog in the Western European sympatry zone remains unclear.

Hybridization between *E. europaeus* and *E. roumanicus* might be frequency-dependent, with males from a high-density population preferentially hybridizing with females from a low-density population (Wirtz, 1999). Potentially, the *E. europaeus* populations in Russia have a lower density than sympatric *E. roumanicus*, and, in addition, their numbers may vary considerably over time. It is worth mentioning that reliable data on the species ratio in the East European contact zone are currently not available. Nevertheless, the numerical predominance of *E. roumanicus* over *E. europaeus* in our dataset can logically be attributed to its greater abundance in East Europe. In our sample from the sympatry zone, the proportion of *E. europaeus* and *E. roumanicus* was 24.8% and 63.5%, respectively, whereas in the Central European zone of sympatry, the species ratio shifted towards *E. europaeus*, e.g. 75.25% of *E. europaeus* and 24.75% of *E. roumanicus* in the Czech Republic according to Bolfiková & Hulva (2012). Thus, one may suppose that interspecific hybridization between *E. europaeus* females and *E. roumanicus* males is perhaps more common when *E. europaeus* is in the minority, for example at the margin of its range.

A similar situation has been described in hares (Thulin *et al.*, 2006a). Both examples provide empirical evidence for an increase in hybridization due to the skewed abundances of two species – a pattern referred to as Hubbs' Principle (Hubbs, 1995; Grant & Grant, 1997). Given the findings of Poduschka & Poduschka, (1983) concerning the viability of different forms of hybrids, we hypothesize that these two factors, the higher density of *E. roumanicus* in European Russia and the predominant success of backcrossing between female F1 hybrids and male *E. roumanicus*, complement each other and explain the difference between Central and Eastern European zones of sympatry.

The examination of several contact zones with different hybridization patterns is of particular interest. Thulin *et al.* (2006a, b) discussed the differences in the frequency of hybridization between the hares *Lepus timidus* and *Lepus europaeus* in Scandinavia and European Russia. Another relevant case is that of the karyomorphs of the striped hamster *Cricetulus barabensis tuvunicus* and *Cricetulus b. pseudogriseus*, which had a signature of hybridization in one contact zone but were completely isolated in the other (Poplavskaya *et al.*, 2017). In the case of hedgehogs, the two divergent species are sympatric in two contact zones without visible environmental barriers between them. However, the signals of ongoing hybridization are only found in Eastern Europe. Our study clearly demonstrates that the structure and intensity of hybridization in different zones of sympatry may be different, being modulated by species density

and/or other possible factors. If so, species delimitation may fail when relying on data for a single sympatric population.

PATTERNS OF EXPANSION

The differences in the genetic structure of closely related species provides insights into speciation and colonization history. The western part of the range of *E. europaeus* and *E. roumanicus* has been thoroughly studied (Santucci *et al.*, 1998; Seddon *et al.*, 2001; Bolfiková & Hulva, 2012), and the existence of three Pleistocene Mediterranean refugia (the Iberian Peninsula, the Apennine Peninsula and the Balkans), with northwards expansion from each of them, has been postulated. The refugial area of the West European hedgehog is presumed to have been in the Iberian and Apennine peninsulas, whereas the Balkan Peninsula (including adjacent islands), has been suggested as the main refugium for *E. roumanicus* (Santucci *et al.*, 1998; Hewitt, 2001; Seddon *et al.*, 2001; Bolfiková *et al.*, 2017). Skyline plots for these species in Central Europe indicated a nearly constant population size in the recent past for *E. europaeus*, whereas the population size of *E. roumanicus* has slightly increased (Bolfiková & Hulva, 2012). This suggests that *E. roumanicus* occupied Central Europe later than *E. europaeus*. It also seems probable that *E. roumanicus* colonized its Eastern European range at a later time and replaced *E. europaeus* through competitive exclusion in the Baltic area, which may explain the current gap in the distribution of *E. europaeus* in the South Baltic region and north-western Russia.

In contrast, our comparison of the genetic structure of both species in European Russia indicates a recent population expansion for *E. europaeus* and, at the same time, suggests stability for *E. roumanicus*. Thus, one may conclude that *E. europaeus* colonized this region later than *E. roumanicus*. Furthermore we hypothesize that *E. europaeus* may have colonized the Eastern Baltic area and European Russia from Scandinavia and would have appeared in the Eastern European contact zone relatively recently. Therefore, the Baltic gap in the *E. europaeus* range would reflect a primary rather than a secondary event. A comparison of the genetic structure of both species suggests that *E. roumanicus* in fact colonized Central Europe later than *E. europaeus* (Bolfiková & Hulva, 2012). However, in European Russia, *E. europaeus* arrived later than *E. roumanicus*, as it required more time to disperse southwards from Scandinavia. Our hypothesis regarding the recent invasion of *E. europaeus* in Russia is consistent with the recent formation of the Eastern European sympatry zone and explains the ongoing hybridization here in contrast to the Central European sympatry zone as was proposed

by Bolfíková & Hulva (2012). One should note that if the suggested scenario of a recent expansion of *E. europaeus* into the range of *E. roumanicus* in East Europe is correct, then the observed pattern of introgression contradicts the predictions of the hypothesis of Currat *et al.* (2008), namely that the directions of gene flow is usually from the resident species into the genome of the colonizing species.

CONCLUSION

The most important finding of our study is the existence of intensive hybridization between the Western European hedgehog and the northern white-breasted hedgehog in the East European zone of sympatry, thus contrasting with the pattern observed in the Central European zone of sympatry. Gene flow between the two species is bidirectional. However, the introgression is asymmetric, with a higher frequency of backcrossing to *E. roumanicus* than *E. europaeus*. The nature of this asymmetry requires further study.

ACKNOWLEDGEMENTS

We thank Prof. V.V. Rozhnov for active collaboration and organization of sample collection. Many thanks to V.M. Malygin, N.A. Formozov and B.I. Sheftel for their assistance in collecting specimens used in this study. Three reviewers provided helpful comments on a previous version of the manuscript. The work was supported by the Russian Foundation for Basic Research, project 20-04-00081a.

REFERENCES

- Alves PC, Melo-Ferreira J, Freitas H, Boursot P. 2008. The ubiquitous mountain hare mitochondria: multiple introgressive hybridization in hares, genus *Lepus*. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**: 2831–2839.
- Anderson EC, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**: 1217–1229.
- Artyushin IV, Bannikova AA, Lebedev VS, Kruskop SV. 2009. Mitochondrial DNA relationships among North Palaearctic *Eptesicus* (Vespertilionidae, Chiroptera) and past hybridization between common serotine and northern bat. *Zootaxa* **2262**: 40–52.
- Bailey RI, Thompson CD, Butlin RK. 2004. Premating barriers to gene exchange and their implications for the structure of a mosaic hybrid zone between *Chorthippus brunneus* and *C. jacobsi* (Orthoptera: Acrididae). *Journal of Evolutionary Biology* **17**: 108–119.
- Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37–48.
- Bannikova AA, Lebedev VS, Rutovskaya MV, Hlyap LA, Rozhnov VV. 2010. Genetic identification and hybridization of common hedgehogs in Central Russia. In: *Integrity of the species in mammals. Isolation barriers and hybridization*. Abstracts of international meeting, Peterhof: , Russia, May 12–17, 2010 KMK Press, 9.
- Bannikova AA, Lebedev VS, Abramov AV, Rozhnov VV. 2014. Contrasting evolutionary history of hedgehogs and gymnures (Mammalia: Erinaceomorpha) as inferred from a multigene study. *Biological Journal of the Linnean Society* **112**: 499–519.
- Bauer VK. 1976. Der Braunbrustigel *Erinaceus europaeus* L. *Annalen des Naturhistorischen Museums in Wien. Bd. 80*: 273–280.
- Becher SA, Griffiths R. 1997. Isolation and characterization of six polymorphic microsatellite loci in European hedgehog *Erinaceus europaeus*. *Molecular Ecology* **6**: 89–90.
- Berggren KT, Ellegren H, Hewitt GM, Seddon JM. 2005. Understanding the phylogeographic patterns of European hedgehogs, *Erinaceus concolor* and *E. europaeus* using the MHC. *Heredity* **95**: 84–90.
- Bize P, Lowe I, Hürlimann LM, Heckel G. 2018. Effects of the mitochondrial and nuclear genomes on nonshivering thermogenesis in a wild derived rodent. *Integrative and Comparative Biology* **58**: 532–543
- Bogdanov AS, Bannikova AA, Pirusskii YuM, Formozov NA. 2009. Genetic evidence of hybridization between West European and northern white-breasted hedgehogs (*Erinaceus europaeus* and *E. roumanicus*) in Moscow Region. *Biology Bulletin* **36**: 647–651.
- Bolfíková B, Hulva P. 2012. Microevolution of sympatry: landscape genetics of hedgehogs *Erinaceus europaeus* and *E. roumanicus* in Central Europe. *Heredity* **108**: 248–255.
- Bolfíková BČ, Eliášová K, Loudová M, Kryštufek B, Lymberakis P, Sándor AD, Hulva P. 2017. Glacial allopatry vs. postglacial parapatry and peripatry: the case of hedgehogs. *PeerJ* **5**: e3163.
- Boratyński 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.
- Boratyński Z, Melo-Ferreira J, Alves PC, Berto S, Koskela E, Pentikäinen OT, Tarroso P, Ylilauri M, Mappes T. 2014. Molecular and ecological signs of mitochondrial adaptation: consequences for introgression? *Heredity* **113**: 277–286.
- Brookfield JFY. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* **5**: 453–455.
- Brown RM, Nichols RA, Faulkes CG, Jones CG, Bugoni L, Tatayah V, Gottelli D, Jordan WC. 2010. Range expansion and hybridization in Round Island petrels (*Pterodroma* spp.): evidence from microsatellite genotypes. *Molecular Ecology* **19**: 3157–3170.

- Burland TG. 2000.** DNASTAR's Lasergene sequence analysis software. *Methods in Molecular Biology* **132**: 71–91.
- Cabria MT, Michaux JR, Gómez-Moliner BJ, Skumatov D, Maran T, Fournier P, López de Luzuriaga J, Zardoya R. 2011.** Bayesian analysis of hybridization and introgression between the endangered European mink (*Mustela lutreola*) and the polecat (*Mustela putorius*). *Molecular Ecology* **20**: 1176–1190.
- Corander J, Marttinen P. 2006.** Bayesian identification of admixture events using multi-locus molecular markers. *Molecular Ecology* **15**: 2833–2843.
- Corander J, Sirén J, Arjas E. 2008b.** Bayesian spatial modelling of genetic population structure. *Computational Statistics* **23**: 111–129.
- Corander J, Marttinen P, Sirén J, Tang J. 2008a.** Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9**: 539.
- Coyner BS, Murphy PJ, Matocq MD. 2016.** Hybridization and asymmetric introgression across a narrow zone of contact between *Neotoma fuscipes* and *N. macrotis* (Rodentia: Cricetidae). *Biological Journal of the Linnean Society* **115**: 162–172.
- Currat M, Ruedi M, Petit R, Excoffier L. 2008.** The hidden side of invasions: massive introgression by local genes. *Evolution* **62**: 1908–1920.
- Curto M, Winter S, Seiter A, Schmid L, Scheicher K, Barthel LMF, Plass J, Meimberg H. 2019.** Application of a SSR-GBS marker system on investigation of European hedgehog species and their hybrid zone dynamics. *Ecology and Evolution* **9**: 2814–2832.
- Deffontaine V, Libois R, Kotlík P, Sommer R, Nieberding C, Paradis E, Searle JB, Michaux JR. 2005.** Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). *Molecular Ecology* **14**: 1727–1739.
- Drummond AJ, Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**: 214.
- Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005.** Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* **5**: 1185–1192.
- Earl DA, von Holdt BM. 2012.** Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.
- Ermakov OA, Surin VL, Titov SV, Tagiev AF, Luk'yanenko AV, Formozov NA. 2002.** A molecular genetic study of hybridization in four species of ground squirrels (*Spermophilus*: Rodentia, Sciuridae). *Russian Journal of Genetics* **38**: 796–809.
- Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.
- Excoffier L, Lischer H. 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.
- Falush D, Stephens M, Pritchard JK. 2003.** Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.
- Good JM, Dean MD, Nachma MW. 2008.** A complex genetic basis to X-linked hybrid male sterility between two species of house mice. *Genetics* **179**: 2213–2228.
- Grant WS. 2015.** Problems and cautions with sequence mismatch analysis and Bayesian Skyline Plots to infer historical demography. *Journal of Heredity* **106**: 333–346.
- Grant P, Grant B. 1997.** Hybridization, sexual imprinting, and mate choice. *American Naturalist* **149**: 1–28.
- Hall TF. 1999.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95–98.
- Henderson M, Becher SA, Doncaster CP, Maclean N. 2000.** Five new polymorphic microsatellite loci in the European hedgehog, *Erinaceus europaeus*. *Molecular Ecology* **9**: 1949–1950.
- Hewitt GM. 1999.** Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**: 87–112.
- Hewitt GM. 2001.** Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology* **10**: 537–49.
- Hill GE. 2017.** The mitonuclear compatibility species concept. *Auk* **134**: 393–409.
- Hill GE. 2019.** Reconciling the mitonuclear compatibility species concept with rampant mitochondrial introgression. *Integrative and Comparative Biology* **59**: 912–924.
- Holz H. 1978.** Studien an Europäischen Igel. *Journal of Zoological Systematics and Evolutionary Research* **16**: 148–165.
- Holz H, Niethammer J. 1990.** *Erinaceus europaeus* Linnaeus, 1758–Braunbrustigel, Westigel. In: Niethammer J, Krapp F, eds. *Handbuch der Säugetiere Europas, 3/1: Insektenfresser, Herrentiere*. Wiesbaden: Aula-Publisher, 26–49.
- Hubbs CL. 1955.** Hybridization between fish species in nature. *Systematic Zoology* **4**: 1–20.
- Huttenner R. 2005.** Order Erinaceomorpha. In: Wilson D, Reeder DM, eds. *Mammal species of the world, 3rd edn, Vol. 1*. Baltimore: Johns Hopkins University Press, 212–219.
- Jiggins CD, Mallet J. 2000.** Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution* **15**: 250–255.
- Jobb G. 2011.** *TREEFINDER version of March 2011*. Germany: Munich. Available at: www.treefinder.de
- King CM, ed. 1990.** *The handbook of New Zealand mammals*. Melbourne: Oxford University Press, 1–600.
- Kratochvil J. 1975.** Zur Kenntnis der Igel der Gattung *Erinaceus* in der SSR (Insectivora, Mamm.). *Zoologické Listy* **24**: 297–312.
- Krystufek B. 1983.** The distribution of hedgehogs (*Erinaceus* L., 1758, Insectivora, Mammalia) in Western Yugoslavi. *Biosistematika* **9**: 71–78.
- Krystufek B. 2002.** Cranial variability in the eastern hedgehog *Erinaceus concolor* (Mammalia: Insectivora). *Journal of Zoology* **258**: 365–373.
- Leigh JW, Bryant D. 2015.** PopART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**: 1110–1116.

- Mastrantonio V, Porretta D, Urbanelli S, Crasta G, Nascetti G. 2016.** Dynamics of mtDNA introgression during species range expansion: insights from an experimental longitudinal study. *Scientific Reports* **6**: 30355.
- Mathias ML, Ramalhinho MG, Santos-Reis M, Petrucci-Fonseca F, Libois R, Fons R, de Carvalho GF, Oom MM, Collares-Pereira M. 1998.** Mammals from the Azores islands (Portugal): an updated overview. *Mammalia* **62**: 397–407.
- 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.
- Minder AM, Rothenbuehler C, Widmer A. 2007.** Genetic structure of hybrid zones between *Silene latifolia* and *Silene dioica* (Caryophyllaceae): evidence for introgressive hybridization. *Molecular Ecology* **16**: 2504–2516.
- Mitchell-Jones AJ, Amori G, Bogdanowicz W, Kryštufek B, Reijnders PJH, Spitzenberger F, Stumme M, Thissen JBM, Vohralík V, Zima J. 1999.** *The atlas of European mammals*. London: T & AD Poyser Natural History, Academic Press.
- Ognev SI. 1928.** *Zveri Vostochnoi Evropy i Severnoi Azii (Animals of Eastern Europe and Northern Asia), Vol. 1*. Moscow: Glavnauka, 631.
- Peakall R, Smouse PE. 2006.** GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288–295.
- Peakall R, Smouse PE. 2012.** GenALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**: 2537–2539.
- Peters KJ, Myers SA, Dudaniec RY, O'Connor JA, Kleindorfer S. 2017.** Females drive asymmetrical introgression from rare to common species in Darwin's tree finches. *Journal of Evolutionary Biology* **30**: 1940–1952.
- Poduschka W, Poduschka C. 1983.** Kreuzungsversuche an mitteleuroäischen. *Igeln Säugetierkd Mitt* **3**: 1–12.
- Poplavskaya NS, Lebedev VS, Bannikova AA, Belokon MM, Belokon YuS, Pavlenko MV, Korablev VP, Kartavtseva IV, Bazhenov YuA, Surov AV. 2017.** Microsatellite loci variation and investigation of gene flow between two karyoforms of *Cricetulus barabensis* sensu lato (Rodentia, Cricetidae). *Russian Journal of Genetics* **53**: 76–90.
- Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Quilodrán CS, Currat M, Montoya-Burgos JI. 2014.** A general model of distant hybridization reveals the conditions for extinction in Atlantic salmon and Brown trout. *PLoS ONE* **9**: e101736.
- Quilodrán CS, Nussberger B, Montoya-Burgos JI, Currat M. 2019.** Hybridization and introgression during density-dependent range expansion: European wildcats as a case study. *Evolution* **73**: 750–761.
- Rambaut A, Suchard M, Drummond A. 2014.** *MCMC trace analysis tool: Tracer v.1.6.0*. Available at: <http://tree.bio.ed.ac.uk/software/tracer/>
- Redenbach Z, Taylor EB. 2003.** Evidence for bimodal hybrid zones between two species of char (Pisces: *Salvelinus*) in northwestern North America. *Journal of Evolutionary Biology* **16**: 1135–1148.
- Ropiquet A, Hassanin A. 2006.** Hybrid origin of the Pliocene ancestor of wild goats. *Molecular Phylogenetics and Evolution* **41**: 95–104.
- Ruprecht AL. 1973.** O rozmieszeniu przedstawicieli rodzaju *Erinaceus* Linnaeus, 1758 w Polsce. *Przegląd Zoology* **27**: 81–86.
- Sambrook J, Fritsch EF, Maniatis T. 1989.** *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press.
- Santucci F, Emerson BC, Hewitt GM. 1998.** Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology* **7**: 1163–1172.
- Scordato ESC, Wilkins MR, Semenov G, Rubtsov AS, Kane NC, Safran RJ. 2017.** Genomic variation across two barn swallow hybrid zones reveals traits associated with divergence in sympatry and allopatry. *Molecular Ecology* **26**: 5676–5691.
- Seddon JM, Santucci F, Reeve NJ, Hewitt GM. 2001.** DNA footprints of European hedgehogs, *Erinaceus europaeus* and *E. concolor*: Pleistocene refugia, postglacial expansion and colonization routes. *Molecular Ecology* **10**: 2187–2198.
- Seddon JM, Santucci F, Reeve N, Hewitt GM. 2002.** Caucasus mountains divide postulated postglacial colonization routes in the white-breasted hedgehog, *Erinaceus concolor*. *Journal of Evolutionary Biology* **15**: 463–467.
- Shurtliff QR. 2011.** Mammalian hybrid zones: a review. *Mammal Review* **43**: 1–21.
- Sokolov VY, Aniskin VM, Lukyanova IV. 1991.** Karyological differentiation of two hedgehog species in genus *Erinaceus* (Insectivora, Erinaceidae) in the USSR. *Zoologicheskii Zhurnal* **70**: 111–120.
- Spiridonova LN, Chelomina GN, Tsuda K, Yonekawa H, Starikov VP. 2006.** Genetic evidence of extensive introgression of short-tailed ground squirrel genes in a hybridization zone of *Spermophilus major* and *S. erythrogenys*, inferred from sequencing of the mtDNA cytochrome b gene. *Russian Journal of Genetics* **42**: 802–809.
- StatSoft I. 1998.** *STATISTICA for Windows [computer program manual]*. Tulsa: StatSoft, Inc.
- Swofford DL. 2003.** *PAUP* – phylogenetic analysis using parsimony (*and other methods), version 4*. Sunderland: Sinauer Associates.
- Tegelström H. 1987.** Transfer of mitochondrial DNA from the northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*). *Journal of Molecular Evolution* **24**: 218–227.
- Thulin CG, Fang M, Averianov A. 2006a.** Introgression from *Lepus europaeus* to *L. timidus* in Russia revealed by

- mitochondrial single nucleotide polymorphisms and nuclear microsatellites. *Hereditas* **143**: 68–76.
- Thulin CG, Stone J, Tegelström H, Walker CW. 2006b.** Species assignment and hybrid identification among Scandinavian hares *Lepus europaeus* and *L. timidus*. *Wildlife Biology* **12**: 29–38.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004.** Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535–538.
- Wirtz P. 1999.** Mother species-father species: unidirectional hybridization in animals with female choice. *Animal Behaviour* **58**: 1–12.
- Yeh FC, Rong-Cai Y, Boyle T. 1998.** *POPGENE version 1.31*. Edmonton, Alberta: University of Alberta, Center for International Forestry Research.
- Zaitsev MV. 1984.** On the taxonomy and diagnostic of hedgehogs of the subgenus *Erinaceus* (Mammalia, Erinaceinae) of the Fauna of the USSR. *Zoologicheskij Zhurnal* **63**: 720–730.
- Zaitsev MV. 2002.** Taxonomic patterning – a new approach to the study of geographic variation in mammals. *Proceedings of Zoological Institute of Russian Academy of Science* **296**: 163–170.
- Zaitsev MV, Voyta LL, Sheftel BI. 2014.** *The Mammals of Russia and adjacent territories: Lipotyphlans*. St. Petersburg: Nauka.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. The comparison of the results of analysis of frequencies of alleles of microsatellite and *TTR* loci in Structure v.2.3.4 under $K = 2$ (A) and $K = 3$ (B). The geographic information about specimens is as in [Figure 2](#).

Figure S2. Median-joining network (POPART v.1.7) of the *Cytb* haplotypes for (A) – *E. europaeus* of Russia and Western Europe ($N = 64$) and (B) – *E. roumanicus* of Russia and Kazakhstan ($N = 145$) based on the 1101 bp. Haplotypes are represented by circles, which are proportional to the haplotype frequency. Geographical locations of haplotypes are grouped as follows: (1) Northern Russia (for *E. europaeus*), (2) Central Russia, (3) Central Russia without Moscow contact zone (for *E. roumanicus*), (4) Moscow region separately (for *E. roumanicus*), (5) Southern Russia (for *E. roumanicus*), (6) Western Siberia, (7) Central Europe, (8) United Kingdom (for *E. europaeus*).

Supporting Information 1. Sequences retrieved from GenBank and used in our study.

Supporting Information 2. Details of the microsatellite analysis of genetic diversity.

Table S1. Summary results of specimen identification by different types of data.

Table S2. The original primers used for sequencing of *Cytb* in *Erinaceus* spp.

Table S3. Individual multilocus genotypes (SSR+*TTR* alleles). Three-digit numbers indicate the sizes of microsatellite alleles; *TTR* alleles are designated as 1, 3 – *E. europaeus* and 2 – *E. roumanicus*.

Table S4. The range of fragment lengths and the genetic diversity of species in allopatric populations.

Table S5. Results of the χ^2 -test for allele frequency heterogeneity for pooled pure samples of two species of hedgehogs.

Table S6. The values of genetic diversity of *E. europaeus* and *E. roumanicus* from the sympatry zone of the Moscow region.

Table S7. Results of the χ^2 -test for allele frequency differences in the samples of each species in pairs allopatric and sympatric populations: *E. europaeus* – left side of the table and *E. roumanicus* – right side.