Molecular Ecology (2004) **13**, 1469–1480 doi: 10.1111/j.1365-294X.2004.02157.x

Mitochondrial phylogeography of the moor frog, *Rana arvalis*

W. BABIK,* W. BRANICKI,† M. SANDERA,‡ S. LITVINCHUK,§ L. J. BORKIN,¶ J. T. IRWIN** and J. RAFIŃSKI*

**Institute of Zoology, Jagiellonian University, Ingardena 6, 30–060 Kraków, Poland, †Institute of Forensic Research, Kraków, Poland, ‡Museum of Nature 'Bohemian Paradise', Jičín, Czech Republic, §Institute of Cytology and ¶Zoological Institute, Russian Academy of Sciences, St Petersburg, Russia, **Bucknell University, Lewisburg, PA USA*

Abstract

The moor frog *Rana arvalis* is a lowland species with a broad Eurasiatic distribution, from arctic tundra through forest to the steppe zone. Its present-day range suggests that glacial refugia of this frog were located outside southern European peninsulas. We studied the species-wide phylogeographical pattern using sequence variation in a 682 base pairs fragment of mtDNA cytochrome *b* gene; 223 individuals from 73 localities were analysed. Two main clades, A and B, differing by *c*. 3.6% sequence divergence were detected. The A clade is further subdivided into two subclades, AI and AII differing by 1.0%. All three lineages are present in the Carpathian Basin (CB), whereas the rest of the species range, including huge expanses of Eurasian lowlands, are inhabited solely by the AI lineage. We infer that AII and B lineages survived several glacial cycles in the CB but did not expand, at least in the present interglacial, to the north. The geographical distribution and genealogical relationships between haplotypes from the AI lineage indicate that this group had two glacial refugia, one located in the eastern part of the CB and the other probably in southern Russia. Populations from both refugia contributed to the colonization of the western part of the range, whereas the eastern part was colonized from the eastern refugium only. The effective population size as evidenced by θ_{ML} is an order of magnitude higher in the AI lineage than in the AII and B lineages. Demographic expansion was detected in all three lineages.

Keywords: Carpathian Basin, eastern refugia, mitochondrial DNA, nested clade analysis, phylogeography, *Rana arvalis*

Received 5 December 2003; revision received 23 January 2004; accepted 23 January 2004

Introduction

Pleistocene glaciations profoundly influenced the temperate biota of the northern hemisphere (Andersen & Borns 1997; Hewitt 2000). During long periods of unfavourable climate the ranges of most animal and plant species were restricted to warmer, more southern areas harbouring suitable habitat. A broad consensus based on a combination of genetic and palaeontological data has emerged during the last decade that three southern regions: the Iberian, Italian and Balkan Peninsulas had constituted the main glacial refugia for European fauna and flora and served subsequently as sources of populations that recolonized deglaciated areas (Comes & Kadereit 1998; Taberlet *et al*. 1998; Hewitt 1999, 2000; Petit *et al*. 2003). However, palaeoecological, palaeobotanical and genetic findings from an increasing number of organisms suggest that many species could have survived glacial maxima outside the Mediterranean in other ice-free areas, e.g. in Central and Northern Europe, the southern Urals or even southern Siberia (Taberlet & Bouvet 1994; Bilton *et al*. 1998; Schmitt & Seitz 2001; Stewart & Lister 2001; Jaarola & Searle 2002; Brunhoff *et al*. 2003; Haase *et al*. 2003; Kropf *et al*. 2003). If correct, the routes of colonization of Europe were more complex and the origin of colonizing populations more diverse than thought previously. Phylogeographical studies on species whose present-day distributions suggest non-Mediterranean glacial refugia are especially important in understanding the origin of European fauna and flora (Bilton *et al*. 1998; Jaarola & Searle 2002; Brunhoff *et al*. 2003).

Correspondence: W. Babik. Fax: + 48 12 634 37 16; E-mail: babik@zuk.iz.uj.edu.pl

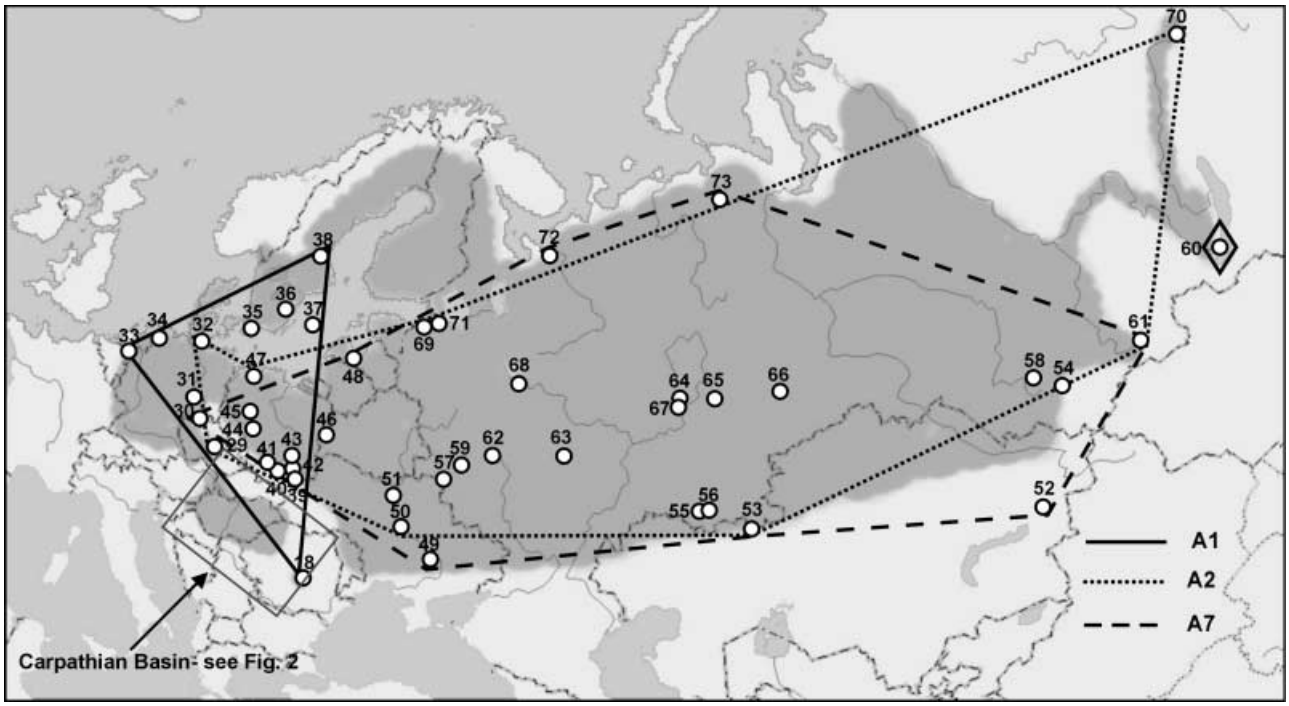


Fig. 1 Populations of *R. arvalis* outside the Carpathian Basin (CB) sampled for variation in a 682 bp mt cytochrome *b* fragment; the locality codes correspond with those in Table 1. The distribution of *R. arvalis* in Eurasia (Borkin *et al*. 1984; Ischenko 1997; Kuzmin 1999) is shaded. The polygons show the distribution of the three major haplotypes from the AI group and their putative descendants (also see Fig. 4). For detailed map of the CB see Fig. 2.

The moor frog, *Rana arvalis* Nilsson, distributed widely in Eurasia from the southern tundra across various kinds of forests to steppe, is a species well suited for such studies. It ranges from the Netherlands in the west to the Altai Mountains, Baikal Lake and Yakutia (approx. 124° E) in the east, where it also inhabits extensive areas of permafrost (Fig. 1; Borkin *et al*. 1984; Ishchenko 1997; Kuzmin 1999). In Europe the moor frog's southern distribution limit lies at about 45° N. The species has a continuous distribution in Europe across the lowlands north of the Carpathians, and most of Scandinavia. A second, relatively small part of its range encompasses lowlands along the middle Danube, centred in the Hungarian Plain. This southern area, further referred to as the Carpathian Basin (CB), is separated from the northern one by the Carpathians in the north and east and by the Alps in the west (Figs 1 and 2). The species is absent from the Balkan and Italian peninsulas. *R. arvalis* is predominantly a lowland species restricted to light soils where it hibernates, but may penetrate into mountainous areas along larger river valleys (Yakovlev 1980; Ishchenko 1989).

A previous study of allozyme variation in Polish, Hungarian and Romanian samples revealed that populations separated by the Carpathians are genetically distinct. Moreover, populations north of the Carpathians exhibit substantially higher levels of genetic variation than those from the CB (Rafiński & Babik 2000). Therefore, a hypothesis was put forward that extensive areas of Europe north of the Carpathians were colonized not from the south, but rather from a large eastern refugium, located in Eastern Europe or Western Siberia/Kazakhstan. A Siberian origin of the European *R. arvalis* populations was postulated previously by Stugren (1966) on purely biogeographical grounds. Also, range expansion was inferred not to have occurred by leptocurtic dispersal but rather through wide-front advance, preventing the loss of genetic variation and causing lack of isolation by distance in northern populations (Rafiński & Babik 2000).

In order to trace the past history of *R. arvalis* and verify previous hypotheses derived from the allozyme data, the pattern of variation in a 682 base pairs (bp) fragment of mitochondrial cytochrome *b* was surveyed. Specifically, we aimed at: (i) assessing species-wide phylogeographical structure; (ii) determining to what extent mitochondrial DNA (mtDNA) variation parallels the differentiation at nuclear loci revealed in Central Europe; and (iii) identifying the historical and demographical factors that have shaped the current pattern of mtDNA variation.

Materials and methods

Samples and laboratory methods

A total of 223 individuals taken from 73 localities covering the entire geographical range of *R. arvalis* was analysed (Table 1, Figs 1 and 2). *R. temporaria*, a sister species to *R.*

Table 1 The sampling localities with their geographical coordinates, sample sizes (*n*) and the haplotypes detected in each locality. Localities were first divided into two groups: within and outside the Carpathian Basin and then according to country of origin from south to north

No	Locality	Coordinates	*n*	Haplotypes
Localities within the Carpathian Basin (CB)				
Austria				
1	Goesseldorfer See	46°34′ N 14°38′ E	1	B2
2	Rabenhof	46°40′ N 15°40′ E	5	B1, B2, B5(2), B7
3	Leibnitz	46°47′ N 15°32′ E	1	B4
4	Weiden	47°53′ N 16°52′ E	2	B1, B5
5	Vienna	48°12′ N 16°21′ E	2	B1, B3
6	Marchegg	48°16′ N 16°54′ E	6	A12 (6)
7	Korneuburg	48°20′ N 16°21′ E	1	B3
8	Stockerau	48°23′ N 16°13′ E	2	A1, A36
9	Schönfeld	48°30′ N 14°56′ E	1	A12
10	Schrems	48°47′ N 15°04′ E	1	A1
11	Thaures	48°53′ N 15°04′ E	3	A1(3)
Hungary				
12	Kaszopuszta	46°19′ N 17°11′ E	4	B1(2), B5, B8
13	Tiszaalpár	46°49′ N 19°59′ E	4	A1(3), A6
14	Ócsa	47°20′ N 19°20′ E	3	B3(3)
15	Debrecen	47°32′ N 21°38′ E	6	A1, A2, A14(4)
16	Fehértó	46°37′ N 20°02′ E	7	B1(2), B2(2), B6(3)
17	Tokaj	48°05′ N 21°24′ E	4	A1, A6, A16, A29
Romania				
18	Reci	45°51′ N 25°55′ E	4	A1(2), A5(2)
Slovakia				
19	Svätý Jur	48°10′ N 17°10′ E	6	A12(3), A25, A35, B3
20	Tajba	48°24′ N 21°47′ E	3	A1, A2, A10
Czech Republic				
21	Lanžhot	48°40′ N 16°58′ E	9	A1, A12(6), A33(2),
22	Hevlín	48°46′ N 16°2′ E	1	B2
23	Lednice	48°48′ N 16°51′ E	8	A1(2), A12(5), B1
24	Božice	48°50′ N 16°19′ E	6	A1(2), A12(4)
25	Nosislav	49°01′ N 16°39′ E	7	A1(3), A12(3), A28
26	Záblatské Louky	49°07′ N 14°39′ E	2	A1, A13
27	Babín	49°33′ N 15°54′ E	2	A1(2)
28	Malé Dářko	49°40′ N 15°53′ E	5	A12(5)
Localities outside the Carpathian Basin				
Czech Republic				
29	Bohdaneĺc	50°05′ N 15°40′ E	4	A1(4)
30	Rybniště	50°53′ N 14°32′ E	5	A1, A22, A23(3)
Germany				
31	Leipzig	51°20′ N 12°21′ E	1	A1
32	Probsteierhagen	54°21′ N 10°17′ E	1	A2
Netherlands				
33	Hilvarenbeek	51°29′ N 5°07′ E	2	A1(2)
34	Zwinderen	52°43′ N 6°40′ E	2	A1(2)
Sweden				
35	Sankt Olof	55°38′ N 14°08′ E	4	A1(4)
36	Tveta	57°19′ N 15°48′ E	3	A1(2) A4
37	Visby	57°30′ N 18°20′ E	1	A1
38	Gäddsjö	60°10′ N 16°51′ E	5	A1(5)
Poland				
39	Walawa	49°52′ N 22°53′ E	5	A1, A2(2), A3(2)
40	Biadoliny	49°59′ N 20°44′ E	2	A1, A34
41	Jeziorzany	49°59′ N 19°46′ E	5	A1(3), A15, A32
42	Dwikozy	50°44′ E 21°47′ E	3	A1(3)
43	Zwoleń	51°21′ N 21°35′ E	3	A1(2), A2
44	Sośnie	51°28′ N 17°37′ E	1	A1
45	Rogaczewo Wielkie	52°01′ N 16°49′ E	2	A1, A2
46	Białowieża	52°42′ N 23°52′ E	5	A2, A7(3), A11
47	Kołobrzeg	54°10′ N 15°34′ E	4	A2(2), A30, A31

Table 1 *Continued*

No	Locality	Coordinates	*n*	Haplotypes
Latvia				
48	Tukuny	56°58′ N 23°10′ E	2	A7, A24
Ukraine				
49	Dnepropetrovsk	48°27′ N 35°03′ E	2	A7, A21
50	Smela	49°13′ N 31°52′ E	1	A2
51	Chernobyl′	51°16′ N 30°14′ E	2	A7, A26
Kazakhstan				
52	Karakil′tas	46°55′ N, 80°45′ E	1	A7
53	Peremyotnoe	51°12′ N 50°49′ E	3	A2, A7(2)
Russia				
54	Altaysky Reserve	50°57′ N 88°09′ E	3	A2
55	Orenburg	51°52′ N 56°11′ E	1	A2
56	Mayorskaya	51°59′ N 55°17′ E	2	A2 (2)
57	Bryansky Les Reserve	52°13′ N 34°04′ E	1	A7
58	Krasnogorskoye	52°17′ N 86°10′ E	1	A2
59	Zhudryo	53°07′ N 35°24′ E	5	A2(5)
60	Goloustnaya	52°02′ N 105°24′ E	3	A1(3)
61	Ust′Uyuk	52°05′ N 94°21′ E	2	A2, A24
62	Skuratovsky	54°13′ N 37°37′ E	3	A2(3)
63	Pushta	54°43′ N 43°14′ E	5	A2(4), A17
64	Izhevsk	56°50′ N 53°12′ E	5	A2(2), A17, A18, A19
65	Krasnoufimsk	56°53′ N 57°32′ E	1	A7
66	Talitsa	57°01′ N 63°43′ E	2	A7, A8
67	Loza	57°22′ N 53°00′ E	1	A9
68	Borok	58°00′ N 38°15′ E	2	A2, A20
69	Narva Reservoir	59°18′ N 28°30′ E	5	A7(4), A27
70	Del′gey	59°54′ N 119°05′ E	3	A2(3)
71	Sankt Petersburg	59°55′ N 30°20′ E	1	A2
72	Archangel	64°35′ N 40°33′ E	1	A7
73	Severnaya Kosheda	67°03′ N 59°25′ E	1	A7

arvalis according to a recent molecular analysis (Veith *et al*. 2003), and *R. asiatica* were used as outgroups.

DNA was extracted by standard proteinase K-phenol–chloroform protocol from frozen or alcohol-preserved tissues. A 791 bp fragment of the mt cytochrome *b* gene was amplified with primers *ralu1* (5′AACCTTATGACCCC-AACAATACG3′) (Bos & Sites 2001) and modified *H15502* (5′GGGTTAGCTGGTGTAAAATTGTCTGGG3′) (Tanaka-Ueno *et al*. 1998). Thirty μL polymerase chain reactions (PCR) contained 3 μL of 10 × PCR buffer with $(NH_4)_2SO_4$ (Fermentas), 2.5 mM $MgCl_2$, 1 mM of the forward and reverse primers, 0.2 mM of each dNTP and 0.5 U of *Taq* polymerase (Fermentas). The cycling scheme was as follows: 94 °C for 2min, 56 °C for 45 s, 72 °C for 2 min, followed by 35 cycles at 94 °C for 30 s, 56 °C for 45 s, 72 °C for 1 min and the final extension step at 72 °C for 10 min. Quality of the PCR product was checked by electrophoresis in 1.5% TAE agarose gels for 15 min. The PCR product was purified with Wizard PCR Preps DNA Purification System (Promega), sequenced using the BigDye Terminator Kit and run on an ABI 3100 automated sequencer (Applied Biosystems). Most of the samples were sequenced using the *ralu1* primer, but in the case of any ambiguity or discovery of a new haplotype the reverse strand was also sequenced with the *H15502* primer. To ensure that the amplified cytochrome *b* fragment represented the mitochondrial gene rather than a nuclear pseudogene, we compared sequences obtained from total genomic DNA and from purified mtDNA of one individual (locality Sośnie, see Table 1). Sequences of a 682 bp fragment of the mt cytochrome *b* gene corresponding to positions 16717–17398 of the complete mitochondrial genome of *R. nigromaculata* (Sumida *et al*. 2001) were obtained for 223 *R. arvalis* individuals and for two outgroup species.

Phylogenetic analyses

Nucleotide sequences were translated into amino acid sequences using DNASP version 3.53 (Rozas & Rozas 1999). Nucleotide diversities (π), mean between-group sequence divergences (D_{xy}) as well as between-group net sequence divergences, D_a (Nei 1987) were computed with MEGA2 (Kumar *et al*. 2001); standard errors of the estimates were obtained using 1000 bootstrap replicates. The appropriate model of sequence evolution was chosen on the basis of hierarchical likelihood-ratio tests as implemented in

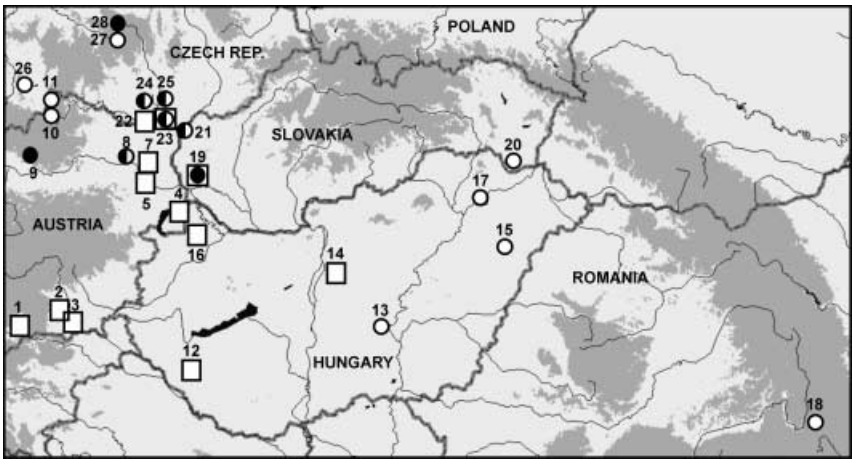


Fig. 2 Populations from the Carpathian Basin sampled for variation in a 682 bp mt cytochrome *b* fragment; the locality codes correspond with those in Table 1. Different symbols are used for three major mtDNA lineages: AI — empty circles; AII — filled circles; B — squares. Superimposed symbols and half-filled circles indicate the presence of haplotypes from different lineages in populations. Shaded areas: above 500 m a.s.l.

MODELTEST 3.06 (Posada & Crandall 1998). The parameters of the models were computed using PAUP* 4.0b10 (Swofford 2002). The model chosen was HKY + Γ with the base frequencies: 0.2355, 0.2884, 0.1597 and 0.3164, for A, C, G and T, respectively, transition/transversion ratio ti/tv = 24.41, and gamma shape parameter $\alpha = 0.3381$. These parameters were used to compute a matrix of pairwise maximum likelihood (ML) distances among haplotypes and then used to construct a neighbour joining (NJ) tree with PAUP*. Robustness of the topology was tested with 1000 bootstrap replicates.

Maximum parsimony (MP) searches were conducted in PAUP* using 1000 random sequence additions and tree-bisection–reconnection (TBR) branch swapping. Characters were treated as unordered and equally weighted. Robustness of the MP tree was tested with 1000 bootstrap replicates.

Rate heterogeneity among lineages was tested by comparing the log likelihoods of ML trees, based on the HKY + Γ model, obtained with and without enforcing molecular clock (Felsenstein 1988).

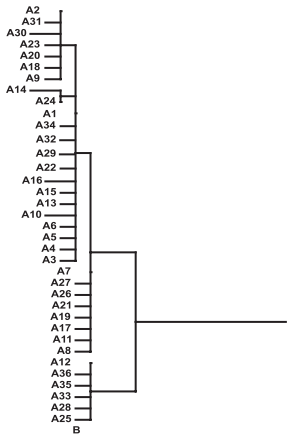
Demographic analyses and nested clade analysis

Two complementary approaches were applied to infer the demographic history of the populations. First, ML estimators of theta (θ_{ML}, $\theta = 2N_e\mu$ for mitochondrial genes, where N_e is female effective population size and μ is mutation rate) and exponential growth parameter (g) in lineages identified by phylogenetic analyses were computed jointly using FLUCTUATE version 1.4 (Kuhner *et al*. 1998). This coalescent-based method takes into account genealogical relationships among sequences and allows explicitly for population-size changes (Kuhner *et al*. 1998).

Second, to gain further insight into the demographic and population history of *R. arvalis*, we applied nested clade analysis (NCA) (Templeton *et al*. 1992; Templeton 1998). A statistical parsimony network was constructed according to the algorithm of Templeton *et al*. (1992) as implemented in TCS version 1.13 (Clement *et al*. 2000). The procedure of nesting haplotypes into higher-level clades followed the rules given in Templeton *et al*. (1992), adapted specifically to DNA sequences by Crandall (1996). The null hypothesis of no association between haplotypes and their geographical locations was tested using GEODIS version 2.0 (Posada *et al*. 2000). On the basis of the nested cladogram, frequencies of the haplotypes and geographical location of each clade, the program calculates the distance statistics — clade distance (D_c), nested clade distance (D_n), the interior-tip distances (I-T_c and I-T_n) and their statistical significance. Random permutations (10 000) of clades vs. sampling localities were used to assess the significance of the associations at the 0.05 level. Clades with significant geographical associations were subjected to a detailed inspection following the rules formulated in the inference key (Templeton 1998; the most recent version available at http://inbio.byu.edu/Faculty/kac/crandall_lab/dposada/documents/NCAkey(24Oct01).pdf). These rules should, in most cases, enable discrimination between contemporary and historical processes which have influenced the genetic structure of the species.

Results

Among 223 *R. arvalis* sequences, 44 haplotypes were identified (GenBank Accession nos AY522383–AY522426), with 65 variable sites, of which 40 were parsimony-informative. All the polymorphic sites exhibited only two base types, so the data complied with the infinite-site model. Eleven variable sites were in the first, one in the second and 53 in the third codon position. We encountered nine nonsynonymous changes in the data set. No insertions, deletions or premature stop codons were observed. This, together with a high ti/tv ratio (24.4 from HKY + Γ model, 24.6 direct count) and identical sequences derived from both purified mitochondria and total genomic DNA of one individual,



ensured that the fragment studied indeed represented a mitochondrial gene rather than its nuclear pseudogene. The complete data set including two outgroup sequences, *R. temporaria* and *R. asiatica* (Accession nos AY522427–AY522428), contained 149 (21.8%) variable positions, of which 75 (11.0%) were parsimony informative.

Phylogenetic analyses

The NJ tree based on the HKY + Γ distance revealed a clear grouping of *R. arvalis* haplotypes into two major clades, A and B (Fig. 3a). Clade A, with bootstrap support 97%, included all the haplotypes found north and east of the Carpathians together with most haplotypes from the Czech Republic, Slovakia, eastern Hungary and Romania. This clade was subdivided further into two subclades: AI with a broad geographical distribution, and AII confined to a small area in the CB — southeastern Czech Republic, northeastern Austria and southwestern Slovakia. Clade B (100% bootstrap support) was detected in the CB — in eastern Austria and western Hungary. Structuring of haplotypes within each of the three lineages was weak and nonsignificant. All three lineages, AI, AII and B, were found in the northwestern part of the CB (Fig. 2). At three localities in Austria, the Czech Republic and Slovakia we found haplotypes from two lineages, AI and AII, one site where AII and B lineages were present and even one locality (Lednice) in which haplotypes from all three lineages were discovered (Table 1, Fig. 2).

MP analysis gave four equally parsimonious trees (193 steps; CI = 0.8342, RI = 0.8991). All were similar, differing only slightly in the branching pattern within the major lineages. One of the most parsimonious trees is shown in Fig. 3b, and confirms the results of the NJ analysis with 100% bootstrap support for B clade, 95% support for A clade and 98% support for the AII subclade.

The hypothesis of clock-like evolution of the sequenced fragment of cytochrome *b* was not rejected by the likelihood ratio test ($\chi^2 = 30.21$, d.f. = 44, $P = 0.94$).

Net sequence divergence (D_a) between lineages AI and B was 3.66%, between AII and B – 3.57% and between AI and AII — 0.99% (Table 2).

Geographic distribution of haplotypes from AI lineage

Unlike groups AII and B, the AI lineage has a very wide geographical distribution and is present both in the CB and over the vast Eurasiatic lowlands. Phylogenetic structuring in this lineage is weak (Fig. 3). However, a statistical parsimony network (Fig. 4) identified three closely related internal haplotypes (A1, A2, A7), each with numerous tip haplotypes separated from a corresponding interior by mainly one mutational step. Such interior, presumably ancestral haplotypes, are the most frequent ones (Table 1),

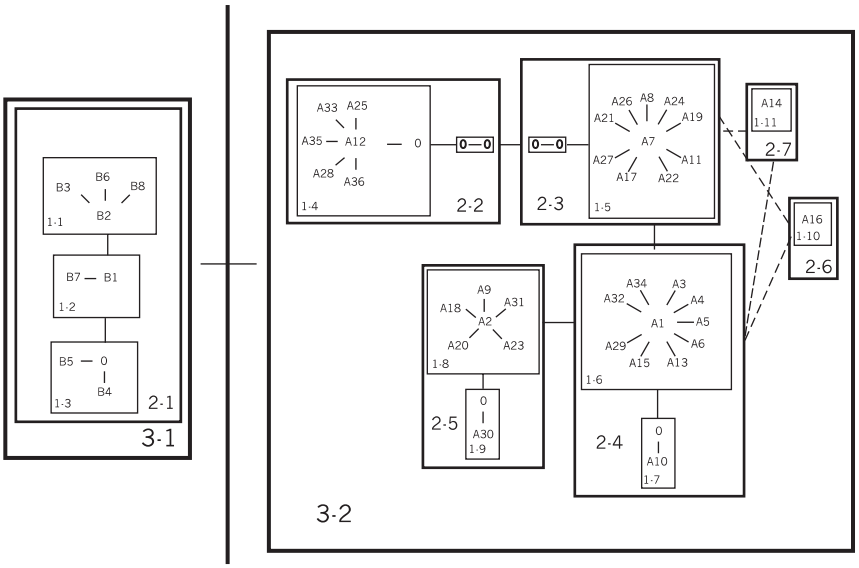


Fig. 4 Nested cladogram for 44 *R. arvalis* haplotypes. The two three-step clades are separated by 22 steps, above the 95% limit of parsimony (11 steps), indicated by a thick vertical line.

and along with their putative descendants show an interesting geographical structure (Figs 1 and 2). The A1 and tip haplotypes related to it are present in the eastern and northwestern part of the CB. They also occur north of the Carpathians, in Poland and further westwards. In the eastern part of the CB, tip, presumably A1-descendant haplotypes separated from A1 by two mutational steps are present, whereas north of the Carpathians no such haplotypes were found. Moreover, none of the tip haplotypes related to A1 are present both north and south of the Carpathians. Surprisingly, the A1 haplotype was also found in one population located at the eastern limit of the species range (Fig. 1, locality 60). The A7 haplotype along with its putative descendants is widespread in the eastern part of the distribution, reaching westwards to easternmost Poland, with one tip haplotype found also in one population on the Czech–German border (Fig. 1, locality 30). Finally, the A2 haplotype and its descendants have the widest distribution spanning 6000 km distance from Germany to Yakutia, overlapping the ranges of both the A1 and A7 (Fig. 1). The A2 haplotype is also present in the two samples from eastern part of the CB; however, no tip haplotypes closely related to it were found in this area (Table 1, Fig. 4).

Demographic analyses and NCA

The values of θ_{ML} and the exponential growth parameter g for the three lineages are given in Table 3. θ_{ML} was significantly higher in AI than in both AII and B lineages (*t*-tests Ps < 0.001). Exponential growth was indicated by high positive g-values in all three lineages with g significantly higher in lineage AII than in either AI or B lineages (*t*-tests Ps < 0.001).

Parsimony was accepted at the 95% level for connection of the haplotypes differing by up to 11 mutational steps.

Table 3 Maximum likelihood estimates of theta θ (θ_{ML}) and exponential growth parameter (g) in three *R. arvalis* phylogenetic lineages. n: number of sequences in lineages. θ_{ML} expressed as percentages. Standard deviations given in parentheses

Lineage	n	θ_{ML}	g
AI	155	10.722 (0.592)	1347.7 (57.4)
AII	39	1.214 (0.254)	7995.8 (936.2)
B	29	0.568 (0.115)	1145.7 (333.0)

The homoplasious loops determined four different cladograms in the plausible set (Fig. 4). A total of nine unambiguous one-step clades were constructed and an additional two constituted degenerate clades according to Templeton & Sing (1993). Five unambiguous two-step clades were revealed, whereas the two degenerate one-step networks were still classified as degenerate clades. At this level, two additional clades (2–2 and 2–3) obtained the degenerate status. The nesting algorithm was terminated at the three-step level, at which no parsimonious connection between clades 3–1 and 3–2 could be established, as the haplotypes belonging to these two clades were separated by at least 22 mutational steps. Clade 3–1 had degenerate status at this level.

The null hypothesis of no geographical association of clades and sampling locations was rejected through contingency tests for two of nine unambiguous one-step clades, as well as for clade 3–2 and the total cladogram (Table 4). Possible reasons for geographical associations drawn from Templeton's key are as follows. For clade 1–4 restricted gene flow with isolation by distance was inferred, whereas clade 1–1 fits the criteria characteristic for past fragmentation. The pattern revealed for the two-step clades grouped

Table 4 Nested contingency analysis of geographical associations for a 682 bp *R. arvalis* cytochrome *b* fragment. Only clades showing genetic and geographical variation are included

Clade	χ^2	*P*
1–1	**35.400**	**0.0294**
1–2	3.938	0.5536
1–3	5.000	0.5990
1–4	**70.023**	**0.0356**
1–5	169.040	0.1863
1–6	260.834	0.5760
1–8	143.504	0.2054
2–1	22.150	0.5331
2–4	37.993	0.2508
2–5	10.984	0.8100
3–2	**547.597**	**0.0050**
Total cladogram	**206.957**	**< 0.0001**

together in network 3–2 indicates fragmentation. The same pattern was inferred for the total cladogram.

Discussion

Dating time of divergence of mtDNA lineages

Veith *et al*. (2003) used the allozyme molecular clock devised for Aegean water frogs (Beerli *et al*. 1996) to calibrate the DNA clock (16S rRNA and a fragment of the rhodopsin gene) for Western Palearctic brown frogs. Although they also sequenced a 465 bp cytochrome *b* fragment, it was not included in the calibration due to signs of saturation in transitional substitutions and some technical problems. The sequences of cytochrome *b* fragments from *R. a. arvalis* and *R. a. wolterstorffi* reported by Veith *et al*. (2003) correspond to our haplotypes A2 and B3 representing lineages AI and B, respectively. Veith *et al*. (2003) estimated the time of separation of these two mitochondrial lineages as 1.03 Myr (95% CI 0.92–1.13 Myr), i.e. predating the onset of most recent climatic oscillations which began about 600 ky BP. The time of divergence between clades A and B estimated from the DNA molecular clock corresponds well with our estimate of divergence of 0.7–1.3 Myr between populations of *R. arvalis* north and south of the Carpathians based on allozymes (Rafiński & Babik 2000). This estimate also relied on the calibration of Beerli *et al*. (1996). However, it is evident from the present study that the distribution of allozyme variation and mtDNA lineages shows striking discordance (see below), so the importance of this finding is not entirely clear. If we assume that the observed divergence of 3.6% between lineages AI and B corresponds to about 1 Myr of separation, the lineages AI and AII would be separated by about 0.3 Myr. Compared to other amphibians, the rate of cytochrome *b* evolution derived from Veith *et al*.'s (2003) calibration is high, implying caution in exact dating of the separation. Nevertheless, it allows setting a relative time frame and strongly suggests divergence before the last glacial period (beginning *c*. 115 ky BP), most probably during one of the previous two or three *c*. 100 ky glacial cycles.

Phylogeographical pattern and glacial refugia

Two of three major mtDNA lineages, AII and B were found exclusively south of the Carpathians. The B lineage inhabits the western part of the CB, i.e. western Hungary, eastern Austria, the southernmost part of the Czech Republic and Slovakia. It is not clear if this group originated in the Carpathian Basin. However, the data allow us to assume that the B lineage, being *c*. 1 Myr old, survived several glacial cycles *in situ* but did not expand to the north, at least not in the present interglacial. The AII group originated *c*. 300 ky BP; according to the applied calibration of molecular clock, it now inhabits only the southern part of the Czech Republic, Slovakia and northern Austria. Ancestors of the AII group could have colonized the CB during former interglacials. This group would also have survived in a glacial refugium located south of the Carpathians, perhaps within its present-day range. Similar to group B, the lineage AII did not expand after the last glacial maximum (LGM) into the lowlands north of the Carpathians.

The postulated survival of *R. arvalis* in the CB throughout the Pleistocene, probably in several glacial refugia (see below), is in line with findings of fossil remains of this species from the lower Pleistocene of Hungary (Venczel 1997). This view is corroborated further by recent data suggesting that the CB could have been an important Pleistocene refugial area, as indicated by the presence of several tree species in Hungary during the LGM (Willis *et al*. 2000), malacological evidence (Sümegi & Krolopp 2002) and the patterns of genetic variation in various animal and plant species (Lagercrantz & Ryman 1990; King & Ferris 1998; Schmitt & Seitz 2001; Stewart & Lister 2001; Jaarola & Searle 2002; Brunhoff *et al*. 2003).

An allozyme study in the butterfly *Erebia medusa* (Schmitt & Seitz 2001) revealed genetic division similar to that of the *R. arvalis* mtDNA lineages in the CB. Butterfly populations from western and eastern Hungary were genetically distinct, the latter clustering with samples from Slovakia and Poland. In a species-wide geographical survey of mtDNA variation in the lizard *Zootoca vivipara*, a haplotype restricted to northern Hungary and Austria was found (Surget-Groba *et al*. 2001). In the adder, *Vipera berus*, survival of two mitochondrial lineages in the CB throughout the glaciations was postulated, one of them east of the Alps, and the other in the eastern part of the CB, or in the Eastern Carpathians (Carlsson 2003).

The geographical structuring of interior, thus presumably ancestral haplotypes from the AI lineage and their

putative descendants (Figs 1 and 2) leads to the following historical scenario for this group. During the last glacial period, *c*. 115–15 ky BP, the A1 haplotype could have survived in the CB and given rise to its southern tip haplotypes, apparently restricted to this area. After the LGM some populations expanded to the north, colonizing the western part of the species range but not spreading to the east. The origin of northern A1-descendant tip haplotypes could have been associated with this expansion. However, the A2 haplotype along with its descendants are also frequent in Poland and westwards. Considering that these haplotypes are widespread in the eastern part of the range of *R. arvalis*, they most probably expanded from the east. The existence of an eastern refugium is corroborated further by the pattern of distribution of the A7 haplotype and its descendants (Fig. 1). The northwestern part of the CB (southern Czech Republic and northern Austria) was probably colonized by the AI lineage after the LGM, as evidenced by presence of the A1 haplotype along with the haplotypes from both AII and B lineages.

If we accept the reasoning above, we need to assume that the A1 haplotype in one locality at the eastern margin of the *R. arvalis* range represents either an ancestral polymorphism retained with low frequency in an eastern refugium, or independent mutation, as A1 is separated by only one mutational step from both the A2 and A7 haplotypes. The same two hypotheses may explain the very rare occurrence of the A2 haplotype, without its immediate descendants, in the eastern part of the CB; alternatively, this haplotype could have entered the area recently.

The data and sampling scheme preclude drawing firm conclusions about the number and exact locations of eastern refugia. We failed to identify regions with a diverse haplotype array in the eastern part of species range. It is even possible that due to climatic and, consequently, vegetational changes, former refugial areas no longer support populations of *R. arvalis*. Lack of latitudinally orientated mountain ranges in European Russia and Western Siberia allows lowland species to shift their ranges freely in response to climatic changes. Thus, identification of refugia on the basis of the present-day distribution of genetic variation may be impossible. Nevertheless, the postulated refugia must have been located in the unglaciated areas of the southeastern part of European Russia or even further east in Western Siberia and/or northern Kazakhstan, as suggested previously (Rafiński & Babik 2000). In spite of cold and arid climatic conditions during the LGM, the presence of woody vegetation is supported by palynological and macrofossil analyses on the northeastern coasts of the Azov and Black Seas (Tarasov *et al*. 2000). Moreover, the southern Urals, hilly eminences of the Russian Plain and the areas of present-day Moscow have been postulated as Pleistocene refugia (Lagercrantz & Ryman 1990; Soffer 1990; Efimik 1996; Markova *et al*. 2002). Recent phylogeographical studies, particularly those conducted on small mammals, point clearly to the importance of southern Russia as a glacial refugium (Taberlet & Bouvet 1994; Bilton *et al*. 1998; Polyakov *et al*. 2000; Jaarola & Searle 2002; Brunhoff *et al*. 2003).

Discordance between allozymes and mtDNA

All populations from the CB scored for variation at allozyme loci clustered together forming a group distinct from populations north of the Carpathians (Rafiński & Babik 2000). However, mtDNA revealed a different pattern. Populations from the eastern part of the CB clearly belong to the same mitochondrial lineage, AI, as those north of the Carpathians. Thus, two aspects of discordance between allozymes and mtDNA need to be addressed. First, why populations in the CB with three distinct mtDNA lineages are similar at allozyme loci and second, why populations belonging to the same mtDNA lineage, AI, are substantially different at the allozyme level.

Allozyme similarity among populations in the CB could have been caused by genetic admixture with unequal nuclear and mitochondrial gene flow during coexistence of different mtDNA lineages throughout the last glaciation. This would result in relative homogenization of allele frequencies at nuclear loci while retaining the original mtDNA (García-París *et al*. 2003). For this explanation to be valid an uneven dispersal rate between the sexes needs to be assumed, with more vagile males responsible for nuclear gene exchange. Biased genetic admixture could be facilitated by reductions of effective population size during the last glaciation.

Dissimilarity at allozyme loci among populations belonging to the AI lineage, separated by the Carpathian mountains, is easily accounted for if two glacial refugia for northern populations are recognized, as suggested above. Source populations probably differed in allele frequencies, and both contributed to the genetic pool of the northern populations, which in consequence differ at allozymes from those from the eastern part of the CB. This would also explain the higher level of genetic variation at allozyme loci in Polish populations compared to the CB (Rafiński & Babik 2000). To verify this hypothesis data on variation at nuclear genes in the eastern part of the *R. arvalis* range are needed.

Demographic history of R. arvalis

In lineages AII and B, found only in the CB, the estimates of θ_{ML} were of an order of magnitude lower than for the AI lineage (Table 3). As no evidence for different rates of mt cytochrome *b* evolution in these lineages exists, it can be assumed that the θ_{ML} values indeed reflect the relative effective population sizes in three lineages, and thus corroborate a larger effective population size in AI as a whole. ML

estimates of population growth rate, *g* (Table 3), show demographic expansion in all three lineages with the highest rate in lineage AII, which may indicate that after the LGM this lineage quickly recovered from a severe bottleneck.

NCA detected past fragmentation for haplotypes grouped in clade 1–1, belonging to the B lineage, even though no apparent geographical barriers exist in this area. Within the AII group, forming clade 1–4, NCA inferred restricted gene flow with isolation by distance. In clade 3–2, comprising both AI and AII lineages, past fragmentation was ascertained. This may reflect separation between AI and AII as well as the postulated range fragmentation within AI during the last glacial period. On the other hand, NCA failed to detect range expansion in the AI group. Range expansion certainly occurred in the AI lineage after the glaciations, as it is now distributed far to the north, including formerly glaciated areas. It is known that NCA performs poorly in detecting range expansion under certain conditions (Templeton 1998; Printzen *et al*. 2003). This occurs commonly when tip haplotypes are found in an expansion area together with the interior, presumably ancestral, haplotypes; in other words, ancestral haplotypes expanded along with their descendants. The distribution of interior haplotypes (A1, A2 and A7) from the AI group (Fig. 1) indicates that such a situation occurred during the postglacial expansion of *R. arvalis*. The contradictory results obtained from the NCA and other methods used to test for demographic expansion emphasize the need for using complementary approaches in identifying demographic and historical population processes.

Taxonomic and conservation implications

Our sampling covered the ranges of all four described subspecies of the moor frog: the nominal form *R. a. arvalis* Nilsson, 1842, *R. a. wolterstorffi* Fejérváry, 1919 thought to inhabit the CB, *R. a. issaltschikovi* Terentjev 1927 from the area of Archangel and *R. a. altaica* Kastschenko, 1899 from the Altai Mountains and areas further east. We did not find any evidence that the ranges of the two latter forms are characterized by distinct mitochondrial lineages. Although additional data from nuclear genes would be highly desirable, we assume tentatively that their morphological differentiation is largely the result of variable ecological conditions as suggested for Central European populations (Babik & Rafiński 2000). The case of *R. a. wolterstorffi* is more complicated. The unexpected pattern of mtDNA variation in the CB, discordant with the results from allozymes (Rafiński & Babik 2000), together with high morphological variablity (Babik & Rafiński 2000), preclude drawing firm conclusions concerning the taxonomic status of these populations, but suggest strongly that populations from the CB harbouring the high mitochondrial and morphological diversity should be a focus of conservation effort.

Acknowledgements

We thank the following people for providing samples: P. Arens, A. Borissovsky, D. Cogălniceanu, V. Ishchenko, T. Knopp, I. Kotserzhinskaya, G. Lada, W. Mayer, K. Milto, M. Pabijan, A. Ruchin, M. Rybacki, J. Szymura and A. Yanchukov. T. E. Matsumoto helped in primer development. M. Pabijan and J. M. Szymura kindly provided purified *R. arvalis* mtDNA. M. Liana offered invaluable help in the field. M. Carlsson made his PhD thesis available to us. W. Mayer, M. Pabijan, J. M. Szymura and four anonymous Referees provided comments and suggestions on earlier drafts of the manuscript. The research was funded by Polish State Committee for Scientific Research grant PB 3 PO4C 010 23 to WB, the work of MS was financed by Foundation 'Nadání Josefa, Marie a Zdeńky Hlávkových' (Prague), LJB and SL were supported by grants from the Russian Foundation for Basic Research (no. 02-04-49631) and from the Russian Federal Program 'Integration' (E-0121) and JTI was supported by a NSERC PDF held with David M. Green (Redpath Museum, McGill University).

References

Andersen BG, Borns HW Jr (1997) *The Ice Age World*. Scandinavian University Press, Oslo.

Babik W, Rafiński J (2000) Morphometric differentiation of the moor frog (*Rana arvalis* Nilss.) in Central Europe. *Journal of Zoological Systematics and Evolutionary Research*, **38**, 239–247.

Beerli P, Hotz H, Uzzell T (1996) Geologically dated sea barriers calibrate a protein clock for aegean water frogs. *Evolution*, **50**, 1676–1687.

Bilton DT, Mirol PM, Mascheretti S *et al*. (1998) Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society of London, Series B*, **265**, 1219–1226.

Borkin LJ, Belimov GT, Sedalishchev VT (1984) New data on distribution of amphibians and reptiles in Yakutia. *Proceedings of the Zoological Institute, USSR Academy of Sciences, Leningrad*, **124**, 89–101 [in Russian, with English summary].

Bos DH, Sites JW (2001) Phylogeography and conservation genetics of the Columbia spotted frog (*Rana luteiventris*; Amphibia, Ranidae). *Molecular Ecology*, **10**, 1499–1513.

Brunhoff C, Galbreath KE, Fedorov VB, Cook JA, Jaarola M (2003) Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late Quaternary biogeography of high latitudes. *Molecular Ecology*, **12**, 957–968.

Carlsson M (2003) *Phylogeography of the adder*, Vipera berus. PhD Thesis, Uppsala University, Uppsala, Sweden.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Sciences*, **3**, 432–438.

Crandall KA (1996) Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. *Molecular Biology and Evolution*, **13**, 115–131.

Efimik VE (1996) Pliocene and Pleistocene relict species in spider fauna of the South Urals. *Zoologichesky Zhurnal*, **75**, 1138–1148 [in Russian, with English summary].

Felsenstein J (1988) Phylogenies from molecular sequences — inference and reliability. *Annual Review of Genetics*, **22**, 521–565.

García-París M, Alcobendas M, Buckley D, Wake DB (2003)

Dispersal of viviparity across contact zones in Iberian populations of fire salamanders (*Salamandra*) inferred from discordance of genetic and morphological traits. *Evolution*, **57**, 129–143.

Haase M, Misof B, Wirth T, Baminger H, Baur B (2003) Mitochondrial differentiation in a polymorphic land snail: evidence for Pleistocene survival within the boundaries of permafrost. *Journal of Evolutionary Biology*, **16**, 415–428.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Ishchenko VG (1989) Population biology of amphibians. *Soviet Scientific Reviews, Series F, Physiology: General Biology*, **3**, 119–155.

Ishchenko V (1997) *Rana arvalis* Nilsson, 1842. In: *Atlas of Amphibians and Reptiles in Europe* (ed. Gasc J-P). Muséum National d'Histoire Naturelle, Paris.

Jaarola M, Searle JB (2002) Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology*, **11**, 2613–2621.

King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.

Kropf M, Kadereit JW, Comes HP (2003) Differential cycles of range contraction and expansion in European high mountain plants during the Late Quaternary: insights from *Pritzelago alpina* (L.) O. Kuntze (Brassicaceae). *Molecular Ecology*, **12**, 931–949.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149**, 429–434.

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*, **17**, 1244–1245.

Kuzmin SL (1999) *The Amphibians of the Former Soviet Union*. Pensoft, Moscow.

Lagercrantz U, Ryman N (1990) Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variation. *Evolution*, **44**, 38–53.

Markova AK, Simakova AN, Pusachenko AY (2002) Ecosystems of Eastern Europe at the epoch of maximum cooling of the Valday glaciation (24–18 thousand years ago) based on floristic and theriological data. *Doklady, Russian Academy of Sciences, Moscow*, **386**, 681–685 [in Russian].

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Petit RJ, Aguinagalde I, de Beaulieu JL *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Polyakov AV, Zima J, Searle JB, Borodin PM, Ladygina T (2000) Chromosome races of the common shrew *Sorex araneus* in the Ural Mts: a link between Siberia and Scandinavia? *Acta Theriologica*, **45** (Suppl. 1), 19–26.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.

Printzen C, Ekman S, Tonsberg T (2003) Phylogeography of *Cavernularia hultenii*: evidence of slow genetic drift in a widely disjunct lichen. *Molecular Ecology*, **12**, 1473–1486.

Rafiński J, Babik W (2000) Genetic differentiation among northern and southern populations of the moor frog *Rana arvalis* Nilsson in central Europe. *Heredity*, **84**, 610–618.

Rozas J, Rozas R (1999) DNASP, version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioiformatics*, **15**, 174–175.

Schmitt T, Seitz A (2001) Intraspecific allozymatic differentiation reveals the glacial refugia and the postglacial expansions of European *Erebia medusa* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, **74**, 429–458.

Soffer O (1990) The Russian Plains at the Last Glacial Maximum. In: *The World at 18 000 BP 1. High Latitudes* (eds Soffer O, Gamble C). Unwin-Hyman, London.

Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16**, 608–613.

Stugren B (1966) Geographic variation and distribution of the moor frog, *Rana arvalis* Nilss. *Annales Zoologici Fennici*, **3**, 29–39.

Sümegi P, Krolopp E (2002) Quatermalacological analyses for modelling of the Upper Weichselian palaeoenvironmental changes in the Carpathian Basin. *Quaternary International*, **91**, 53–63.

Sumida M, Kanamori Y, Kaneda H *et al.* (2001) Complete nucleotide sequence and gene rearrangement of the mitochondrial genome of the Japanese pond frog *Rana nigromaculata*. *Genes and Genetic Systems*, **76**, 311–325.

Surget-Groba Y, Heulin B, Guilamme C-P *et al.* (2001) Intraspecific phylogeography of *Lacerta vivipara* and the evolution of viviparity. *Molecular Phylogenetics and Evolution*, **18**, 449–459.

Swofford DL (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland.

Taberlet P, Bouvet J (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proceedings of the Royal Society of London, Series B*, **255**, 195–200.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Tanaka-Ueno T, Matsui M, Sato T, Takenaka S, Takenaka O (1998) Local population differentiation and phylogenetic relationships of Russian brown frog, *Rana amurensis* inferred by mitochondrial cytochrome *b* gene sequences (Amphibia, Ranidae). *Japanese Journal of Herpetology*, **17**, 91–97.

Tarasov PE, Volkova VS, Webb T III *et al.* (2000) Last glacial maximum biomes reconstructed from pollen and plant macrofossil data from northern Eurasia. *Journal of Biogeography*, **27**, 609–620.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic association with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic association with haplotypes inferred from restriction endonuclease

mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659–669.

Veith M, Kosuch J, Vences M (2003) Climatic oscillations triggered post-Messinian speciation of Western Palearctic brown frogs (Amphibia, Ranidae). *Molecular Phylogenetics and Evolution*, **26**, 310–327.

Venczel M (1997) Amphibians and reptiles from the lower Pleistocene of Osztramos (Hungary). *Nymphaea*, **23–25**, 77–88.

Willis KJ, Rudner E, Sümegi P (2000) The full-glacial forests of central and southeastern Europe. *Quaternary Research*, **53**, 203–213.

Yakovlev VA (1980) On vertical distribution and breeding of the moor frog in the Altai Nature Reserve. *Ecologiya, Sverdlovsk*, **4**, 89–90 [in Russian].

W. Babik uses molecular tools to study the evolution of amphibians and genetic analyses of natural hybrid zones; W. Branicki is interested in the application of molecular methods to human and animal forensic studies; M. Sandera is interested in Palearctic herpetofauna, especially the families Ranidae and Agamidae, with an emphasis on field studies; S. N. Litvinchuk's research focuses mainly on speciation of amphibians of eastern Europe; L. J. Borkin is working on systematics, zoogeography, speciation, ecology and conservation of amphibians and reptiles of Eurasia; J. T. Irwin is an ecophysiologist interested in applying molecular techniques to study the evolution of complex physiological adaptations; the late J. Rafiński worked on various aspects of amphibian evolution with special emphasis on sexual selection and sperm competition.
