Molecular Ecology (2007) doi: 10.1111/j.1365-294X.2007.03526.x

# Genetic variation and phylogeography of free-living mouse species (genus *Mus*) in the Balkans and the Middle East

M. MACHOLÁN,*† M. VYSKOČILOVÁ,†‡ F. BONHOMME,§ B. KRYŠTUFEK,¶ A. ORTH§ and V. VOHRALÍK**

**Laboratory of Mammalian Evolutionary Genetics, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, Veveří 97, CZ-60200 Brno, Czech Republic, †Institute of Botany and Zoology, Masaryk University, Brno, Czech Republic, ‡Department of Population Biology, Institute of Vertebrate Biology, ASCR, Studenec, Czech Republic, §Biologie Intégrative, Institut des Sciences de l'Evolution CNRS — Université de Montpellier 2, Montpellier Cedex 05, France, ¶Slovene Museum of Natural History, Ljubljana, Slovenia, **Department of Zoology, Charles University, Prague, Czech Republic*

**Abstract**

**This work presents a study of the distribution and pattern of variation throughout the ranges of three free-living mouse species of the genus *Mus* — *M. macedonicus*, *M. spicilegus*, and a *M. cypriacus* — based on sequencing of two segments of the mitochondrial DNA (mtDNA) control region. The study shows a similar level of variability in the three species and suggests their recent population expansion. The highest proportion of variation is found within populations indicating low genetic structuring. Phylogenetic analysis confirms the significant divergence of a mitochondrial lineage of *M. macedonicus* from Israel, recently described as a new subspecies, *M. macedonicus spretoides*. Conversely, no genetic hiatus is revealed between European and Asian populations of *M. macedonicus macedonicus*. Although phylogenetic relationships among *M. spicilegus* populations could not be unravelled precisely, the results suggest a recent westward expansion of the species. The mtDNA divergence between *M. macedonicus* and *M. spicilegus* is 7.3%, suggesting their split between *c.* 700 000 and 1 million years ago. These dates correspond with a coalescent estimate about 720 000 years ago. On the other hand, *M. cypriacus* appeared almost twice as divergent from the former species (4.5%) as from the latter (8.8%) suggesting a divergence of *c.* 430 000–610 000 years ago (coalescent ≈ 490 000 years ago) and 830 000–1.2 million years ago (coalescent ≈ 780 000 years ago), respectively. Approximate times of population expansion have also been estimated for all taxa and groups of populations. Existence of several glacial refuges and various colonization scenarios are discussed; since all estimated divergence times fall within interglacial periods it seems that climatic oscillations did not play a crucial role in the evolution of the three species.**

*Keywords*: glacial refugia, *M. cypriacus*, *M. spicilegus*, Mediterranean region, mtDNA, *Mus macedonicus*

*Received 19 April 2007; revision accepted 1 August 2007*

## Introduction

The last few decades have witnessed the increasing application of molecular markers to infer historical patterns of species' expansion from refugia after glaciation episodes (Avise 2000; Hewitt 2000). These dramatic climatic oscillations appear to have begun about 650 000 years ago with the first cold stage, and was followed by another three glacial maxima interrupted by warmer periods, or interglacials. The last glacial maximum (LGM) ended 11 550 years ago. The relatively rapid oscillations in temperature are known to have caused fluctuations in habitat expansions and contractions and faunal communities have long been considered major agents influencing the origin and/or evolution of recent mammal species. However, the vision of glacial refuges as single homogeneous entities has recently been brought into question by the enhanced resolution of

Correspondence: Miloš Macholán, Fax: +420–541212988; E-mail: macholan@iach.cz

phylogeographical studies (e.g. Branco *et al*. 2002; Petit *et al*. 2002). Moreover, it has been shown that the Mediterranean areas (the Iberian, Apennine and Balkan peninsulas and the Middle East), regarded as glacial refugia for European temperate species, may have not been the only sanctuaries and sources of postglacial expansion (Bilton *et al*. 1998; Kotlík *et al*. 2006).

Given their wide distribution, eastern free-living, 'outdoor' or 'aboriginal' (*sensu* Sage 1981), mouse species, *Mus macedonicus* and *M. spicilegus*, would appear to be excellent models for studies of refugial structure and colonization patterns after the LGM. The current western range of the Balkan short-tailed mouse, *M. macedonicus* Petrov & Ružić (1983), is limited by the Balkan and Caucasian mountains and this species occurs in the Near East, Transcaucasia, Asia Minor, and southern Balkans (Macholán 1999a). The first molecular studies of *M. macedonicus* revealed little phylogenetic structure in this species (Prager *et al*. 1996, 1998; Gündüz *et al*. 2000); however, these previous surveys were limited both in the number of sample sites and individuals analysed. Recently, Orth *et al*. (2002) have shown that *M. macedonicus* populations in Israel are highly distinct from those in remaining parts of the species' range, based on D-loop sequence data for 19 specimens from 12 localities; they also noted slightly lower but significant differentiation of nuclear genes. These results led the authors to describe the Israeli clade as a separate subspecies, *M. macedonicus spretoides*; all remaining populations analysed were included into a nominal subspecies, *M. macedonicus macedonicus*. Moreover, Orth *et al*. (2002) hypothesized the existence of at least two glacial refuges south of the Caucasus for the Balkan short-tailed mouse. Thus, the species appears to have a more complex phylogeographical history than previously envisioned.

The steppe mouse, *M. spicilegus* Petényi, 1882, is presently confined to European lowlands from Austria and Slovakia in the north, to Bulgaria in the south and the Ukraine in the east, including Serbia, Hungary, Romania, and Moldova (Macholán 1999b). In general, this species is allopatric with *M. macedonicus* the only known exception being a restricted area of parapatric contact in eastern Bulgaria (Orth *et al*. 2002). An isolated population of *M. spicilegus* was reported (under the name *M. hortulanus*) near the town of Ulcinj, Montenegro (Petrov & Ružić 1983, 1985; Petrov 1992); as a result of substantial morphometric differences in comparison to other conspecific populations, it was described as a new subspecies, *M. spicilegus adriaticus* (Kryštufek & Macholán 1998). Later, populations from Albania and western Greece were found to be morphologically very similar to the new subspecies (Macholán & Vohralík 1997). These findings raise questions about the origin of the Adriatic populations and their genetic relationships to populations from the main species' distribution area, separated by the wide Balkan mountain range.

A third member of the eastern group of free-living mouse species has recently been discovered on the Mediterranean island of Cyprus (Cucchi *et al*. 2002; Bonhomme *et al*. 2004). This species, described formally as *Mus cypriacus* by Cucchi *et al*. (2006), was shown to be genetically similar to *M. macedonicus* but almost equidistant from *M. spicilegus* (Bonhomme *et al*. 2004; Cucchi *et al*. 2006). Cucchi *et al*. (2002, 2006) and Macholán *et al*. (2007) have described morphometric variation in this species as well as its morphological relationships to other aboriginal species.

This study describes an intensive study of the distribution and pattern of mitochondrial DNA (mtDNA) variation throughout the ranges of *M. macedonicus*, *M. spicilegus*, and *M. cypriacus*. More specifically, and for the first time, we (i) analyse genetic variation and phylogeography of *M. macedonicus* on a pan-European scale, including Macedonia, Greece, Bulgaria, Turkey, Iran, Georgia, Dagestan, Syria, and Israel. The systematic status of populations from Syria are particularly useful for estimating the northern and northeastern border of *M. m. spretoides*; (ii) assess the genetic relationship of the *M. spicilegus adriaticus* population from the type locality with other steppe mouse populations and establish the identity of recently discovered populations of this subspecies from nortwestern and western Greece and from Peloponnesus; and (iii) perform a more complete analysis of the genetic variation and phylogeography of *M. cypriacus* by sampling nine sites scattered across the whole Cyprus, including the Turkish Republic of Northern Cyprus. These analyses allow us to test several null hypotheses:

1 Are both the *M. macedonicus* and *M. spicilegus* subspecies (*macedonicus–spretoides* and *spicilegus–adriaticus*, respectively) reciprocally monophyletic?
2 Is the distribution of *M. m. spretoides* limited to Israel?
3 Did *M. macedonicus* survive the LGM in several glacial refugia?
4 Did *M. spicilegus* colonize Europe from the south, or followed a northern route?
5 Did *M. cypriacus* or its ancestor colonize Cyprus during the Middle Pleistocene?

Finally, by estimating divergence times between all groups, we tested hypotheses relating to the role of glacial periods in the evolution of the three species under study.

## Materials and methods

### *Mice*

Eighty-nine specimens of *Mus macedonicus* (including *M. macedonicus spretoides*) collected from 40 sites, 27 specimens of *M. spicilegus* (including *M. spicilegus adriaticus*) from 13 sites, and 27 specimens of *M. cypriacus* from nine sites

were analysed (see Appendix). The data set was supplemented with one previously published mtDNA sequence of *M. spretus* retrieved from the GenBank database (Accession no. U47539) and four sequences of commensal house mice (recently published) of the *M. musculus* complex (*musculus*: Botosani, Romania; *domesticus*: Kuli Alireza, Iran; *castaneus*: Asalem, Iran; *gentilulus*: San'a, Yemen, GenBank, Accession no. AF074540), giving a total sample size of 148 specimens (details on the commensal mice are to be published elsewhere). GenBank accession numbers for the new mtDNA sequences are EU106188–EU106322.

Since *M. cypriacus* is still poorly known, we make a few ecological remarks here. Mice were collected between 24 March and the 7 May 2005 at altitudes ranging from 147 to 1605 m above sea level. They were snap-trapped in dry areas typically covered with low grass and the shrubs *Sarcopoterium spinosum*, *Berberis craetica* and *Genista* sp., interspersed with a few trees. Adult animals were brown dorsally with an ochre tinge on the flanks and pure white or greyish-white ventrally. Body measurements ($N = 24$) in millimetres were: tail-relative-to-body length 98.90 ± 6.57 (range 88.5–117.1), hind foot length 17.36 ± 0.80 (15.1–18.5), ear length ($N = 21$) 15.10 ± 0.76 (13.6–16.8).

### *PCR amplifications and sequencing*

DNA was extracted from ethanol-preserved tissue or dried skin using DNeasy Tissue Kit (QIAGEN) following the manufacturer's instructions. Two segments encompassing variable domains of the mtDNA control region and flanking tRNAs were amplified using primer pairs L15320-H15782 and L15911-H00072, respectively (Prager *et al*. 1993, 1996). Aliquots of 50 ng of DNA were amplified in 30 μL of the polymerase reaction (PCR) buffer with 1.5 mM $MgCl_2$, 200 μM dNTPs, 0.5 U of *Taq* polymerase (Fermentas), and 0.5 μM of each primer. Amplifications were carried out in a gradient RoboCycler thermal cycler (Stratagene) with 35 cycles of 94 °C for 40 s, 53 °C (L15320- H15782) or 55 °C (L15911-H00072) for 40 s, and 72 °C for 2 min. PCR products were checked on 1.5% agarose gels and then purified using either QIAquick or MinElute PCR Purification Kit (QIAGEN) and sent to Macrogen (South Korea) for sequencing. Both light and heavy strands were sequenced, for a total length of 950 bp (956 bp in *M. spretus*) which were subsequently used for all analyses.

### *Sequence alignment and genetic analyses*

Sequences were aligned with CLUSTAL X, version 1.83 (Thompson *et al*. 1997) using all default values and checked manually. Since substitution saturation of sequences can affect phylogenetic analyses, the aligned sequences were checked for such saturation in two ways. First, for the whole data set, values of Tamura–Nei (TrN) distances

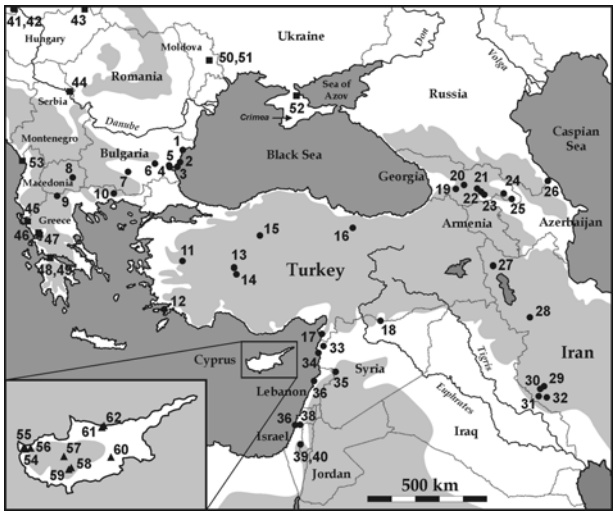


**Fig. 1** Geographical distribution of *Mus macedonicus* (circles), *Mus spicilegus* (squares), and *Mus cypriacus* (triangles) samples analysed. Areas above 500 m above sea level are shaded. Numbers correspond to those in the Appendix.

(with rate heterogeneity and invariant sites, see below) was plotted against *p*-distances, computed with the PAUP* 4.0b10 program package (Swofford 2000), following Philippe & Douzery (1994) and Hassanin *et al*. (1998). A slope of the linear regression ($S$) equal to one indicates no saturation in the data while as saturation increases, the slope approaches zero (Hassanin *et al*. 1998). Secondly, Xia's test implemented in the DAMBE, version 4.2.13 program (Xia & Xie 2001; available at http://web.hku.hk/~xxia/software/software.htm).

Genetic variation was assessed by estimating indices of haplotype diversity, $h$ (Nei 1987), and nucleotide diversity, $\pi$ (Tajima 1983). Population genetic structure was determined by an analysis of molecular variance (AMOVA) using ARLEQUIN 2.000 program (Schneider *et al*. 2000). This analysis was performed among groups of populations, among populations within each group and within each population where the groups were defined according to geographical proximity of individual populations. First, three major groups were considered: *M. m. spretoides*, and the European and Asian group of *M. macedonicus macedonicus* populations, respectively. At the lower level, seven subgroups were defined: aside from Israeli *spretoides* (samples 37–40 in the Appendix; see also Fig. 1), there were two European *macedonicus* subgroups: Bulgarian (samples 1–7) and Greek-Macedonian (samples 8–10), and four Asian *macedonicus* subgroups: Turkish (samples 11–16), Georgian-Dagestanian (samples 19–26), Iranian (samples 27–32), and Turkish-Syrian (samples 17–18, 33–35). In *M. cypriacus*, samples were pooled into four groups: western (samples 54–56 in the Appendix), central (samples 57–59), eastern (sample 60), and northern (samples 61–62). Isolation by distance (Wright 1943) was estimated in *M. macedonicus* and *M. cypriacus*

using the Mantel test in the NTSYS-PC, version 2.01e program (Rohlf 1997): observed and random values of the Mantel *Z* statistic obtained from 5000 permutations of the raw data were compared. Minimum geographical distances between haplotypes were measured using the Microsoft Encarta World Atlas (1998 edition) avoiding large bodies of water as described in Macholán *et al*. (2001). The *M. spicilegus* populations were not analysed with AMOVA or the Mantel test owing to inadequate sampling.

Tajima's *D* statistic was used for estimating potential deviation from selection neutrality and/or recent population expansion or decline (Tajima 1989) within *M. cypriacus*, *M. spicilegus*, and the European and Asian groups of *M. m. macedonicus* populations, respectively (*M. m. spretoides* was excluded from the analysis due to the limited sample size). Subsequently, a mismatch distribution of substitution differences between pairs of haplotypes was calculated. The observed values were compared to the values expected from the population expansion model with parameters estimated using the generalized nonlinear least-squares approach of Schneider & Excoffier (1999; see also Li 1977; Harpending 1994; Rogers 1995) using ARLEQUIN, with parameters estimated from the evolutionary models indicated below.

In order to validate the patterns of population change, we performed a series of maximum likelihood-based analyses to test the fit of the two models: the first assuming stability of population size through time (the null hypothesis), and the second assuming exponential growth or decline of the populations (Kuhner *et al*. 1995, 1998). Since the two models were nested, they were then compared with the likelihood-ratio test (LRT) with one degree of freedom. The analysis was carried out using FLUCTUATE 1.3 (available at http://evolution.genetics.washington.edu/lamarc/index.html). The substitution models given in the next section were used and the parameters θ and *g* (population growth rate) were allowed to vary. We ran the program several times with different numbers of short and long Markov chains for a more efficient examination of the likelihood surface (e.g. three to 10 trials with 20–100 short chains and one to two long chains, followed by three to five runs with none or very few short chains).

### *Phylogenetic analyses and estimation of divergence times*

The whole set of aligned sequences was analysed with the neighbour-joining (NJ) (Saitou & Nei 1987) method under the minimum evolution criterion, the maximum parsimony (MP; Fitch 1971), and Bayesian analysis (BA; Rannala & Yang 1996) methods. A heuristic search with the tree-bisection–reconnection (TBR) branch swapping method was applied in the analyses. In the MP analyses, an ACCTRAN optimization option was used and gaps were treated as a fifth base. The robustness of inferences was assessed by bootstrap resampling (Felsenstein 1985) with 100 or 1000 replicates depending on the method used. Intraspecific trees were inferred using NJ, MP, maximum likelihood (ML; Felsenstein 1981), and BA analyses with *M. spretus* used as an outgroup. All phylogenetic analyses were performed with PAUP* except the Bayesian analysis which was carried out using MRBAYES 3.1.2 (Ronquist & Huelsenbeck 2003). Finally, minimum spanning networks were constructed using ARLEQUIN.

For distance methods, BA and ML analyses, and for all the model-based analyses described above, the most appropriate models were chosen with MODELTEST (Posada & Crandall 1988) using LRT for hypotheses testing. The following substitution models were chosen for individual taxa: *M. macedonicus*: Hasegawa–Kishino–Yano (HKY) model (Hasegawa *et al*. 1985) with unequal rates following the gamma distribution and assuming invariant sites (HKY + Γ + I), the transition/transversion ratio $t_s/t_v = 1.752$, gamma shape parameter alpha = 0.834, proportion of invariant sites $I = 0.720$ (Gu *et al*. 1995; Waddel & Penny 1996); *M. cypriacus*: HKY + Γ + I, $t_s/t_v = 1.883$, alpha = 0.963, $I = 0.875$; *M. spicilegus*: HKY + Γ, $t_s/t_v = 2.069$, alpha = 0.016; all aboriginal species (including the single *M. spretus* specimen): Tamura–Nei (Tamura & Nei 1993) with unequal rates following the gamma distribution and assuming invariant sites (TrN + Γ + I), $t_{s(pur)}:t_{s(pyr)}:t_v = 3.599 : 4.970 : 4.000$, alpha = 0.746, $I = 0.657$. In the BA analyses, values of the parameters were estimated from the data. Bayesian posterior probabilities were estimated from 1 million (*M. macedonicus*) and 2 million (*M. spicilegus*, *M. cypriacus*) generations, respectively, sampled every 1000th generation and excluding a burn-in of 250 000 and 500 000 steps, respectively.

As populations might not be at genetic equilibrium, approximate times of divergence between *M. m. macedonicus* and *M. m. spretoides*, between the European and Asian *M. m. macedonicus* groups of populations, and between the three aboriginal species, were calculated on the basis of the estimated percentage of genetic divergence. The genetic (TrN + Γ + I) distances were corrected for ancestral mtDNA polymorphism according to the formula (Edward 1997): $d = d\text{AB} - 1/2(d\text{A} + d\text{B})$, where $d$AB is the mean genetic distance between lineages A and B, computed from pairwise distances between individuals from different lineages (i.e. A vs. B), and $d$A and $d$B are mean genetic distances within these lineages, respectively. Potential differences in the substitution rates between compared taxa were tested through a series of relative-rate tests with Bonferroni-corrected significance levels using an improved version of the test of Wu & Li (1985) which takes into account taxonomic sampling and phylogenetic relationships (Robinson *et al*. 1998). For this purpose, RRTREE, version 1.1 (Robinson-Rechavi & Huchon 2000) was employed, with the NJ tree chosen as the reference topology and *M. spretus* as the outgroup. Subsequently, divergence times between taxa

**Table 1** Genetic variation within the four groups analysed (E and A denotes, European and Asian group of populations, respectively). The *M. m. spretoides* sample was excluded due to limited sample size

| | No. of populations | No. of haplotypes | Genetic divergence within groups (% TrN dist.) | Haplotype diversity $h \pm SD$ | Nucleotide diversity $\pi \pm SD$ |
|---|---|---|---|---|---|
| *macedonicus* E | 28 | 23 | 0.650 | 0.9788 ± 0.0183 | 0.0061 ± 0.0033 |
| *macedonicus* A | 54 | 41 | 1.067 | 0.9790 ± 0.0104 | 0.0087 ± 0.0046 |
| *cypriacus* | 27 | 26 | 0.810 | 0.9801 ± 0.0169 | 0.0112 ± 0.0059 |
| *spicilegus* | 27 | 19 | 0.893 | 0.9658 ± 0.0205 | 0.0214 ± 0.0109 |

TrN dist., Tamura–Nei distance.

and groups of populations were computed from genetic distances using the *M. musculus musculus–M. m. domesticus* distance as a reference and calibrated with two estimates based on DNA-DNA hybridization data [which assume that the *Mus–Rattus* split occurred 10 million years ago (Ma)]: 350 000 years ago (She *et al*. 1990) and 500 000 years ago (Catzeflis *et al*. 1992; Boursot *et al*. 1993).

Mutation rate and time of divergence from the most recent ancestors ($t_{MRCA}$) were estimated by a Bayesian coalescent analysis under a molecular clock assumption using BEAST version 1.14 (Drummond & Rambaut 2006). The HKY + Γ + I model was used assuming an exponential coalescent model. For estimation of the $t_s/t_v$ ratio, gamma shape parameter and proportion of invariant sites, uniform priors were used, with lower and upper bounds [0,10], [0,10], and [0,1], respectively. For the tree root height parameter, we used uniform priors bounded by 0 and 2 million years (Myr). The parameters were then estimated from a range of plausible alternative trees after two to three independent runs of $10^6$ generations each, sampled every 1000th generation with the first 10% trees discarded as burn-in.

## Results

### *Genetic variation*

Among 143 sequences of the three aboriginal species studied, 70 distinct haplotypes were identified in *Mus macedonicus* (both *macedonicus* and *spretoides*), 23 haplotypes in *M. cypriacus*, and 20 haplotypes in *M. spicilegus* (both *spicilegus* and *adriaticus*; see Appendix). Together with the single sequence of *M. spretus*, the entire 'aboriginal' data set is composed of 114 distinct partial sequences of the mtDNA control region. Out of 972 sites in the input matrix, 230 sites (23.7%) are variable and 123 sites (12.7%) are parsimony informative. The mean transition/transversion ratio is 1.715; however, there are substantial differences between individual substitution types: while CT transitions are the most frequent (43.2% of all substitutions; approximately 208% of the expected value) the lowest proportion of all substitutions are CG transversions (0.6%; about 10% of the expected value).

The saturation analysis revealed weak saturation in the whole set of sequences ($S = 0.640$) but the curve was far from reaching a plateau: when the curve was divided into two halves, even the slope of the second portion was still significantly different from zero ($S = 0.453$). This result was confirmed by Xia's test which shows no significant saturation in the DNA sequences under study.

Genetic variation within the four major groups, i.e. the European and Asian groups of *M. macedonicus. macedonicus*, the *M. cypriacus* group, and the *M. spicilegus* group, is summarized in Table 1. The higher number of haplotypes in Asian *M. m. macedonicus* may be due to the higher number of populations sampled; however, apart from this exception, there are no significant differences in the level of genetic variation between these groups. The AMOVA reveals a lack of geographical structuring of *M. cypriacus* populations with by far the highest proportion of genetic variation (88.4%) occurring within populations. In *M. macedonicus*, 48.2% of the total variation is within populations, whereas 16.3% and 35.5% of variation is found within and among groups of populations, respectively. Among the European *M. m. macedonicus*, Asian *M. m. macedonicus*, and *M. macedonicus spretoides* groups, the proportion of the total variation is 33.5%; however, this value is apparently influenced by a high level of genetic divergence between the two subspecies (*macedonicus* and *spretoides*), since only 16.5% of variation is accounted for by divergence between European and Asian populations of *M. m. macedonicus*, the highest proportion of variation occurring within populations (53.4% in all three group, 66.5% when only the European and Asian groups of *M. m. macedonicus* are considered). The low level of differentiation between European and Asian *M. m. macedonicus* populations is corroborated by the fact that genetic distances between populations from Europe and those from Asia are not significantly higher than distances within the continents. However, the Mantel test reveals significant correlation between genetic and geographical distances in this subspecies

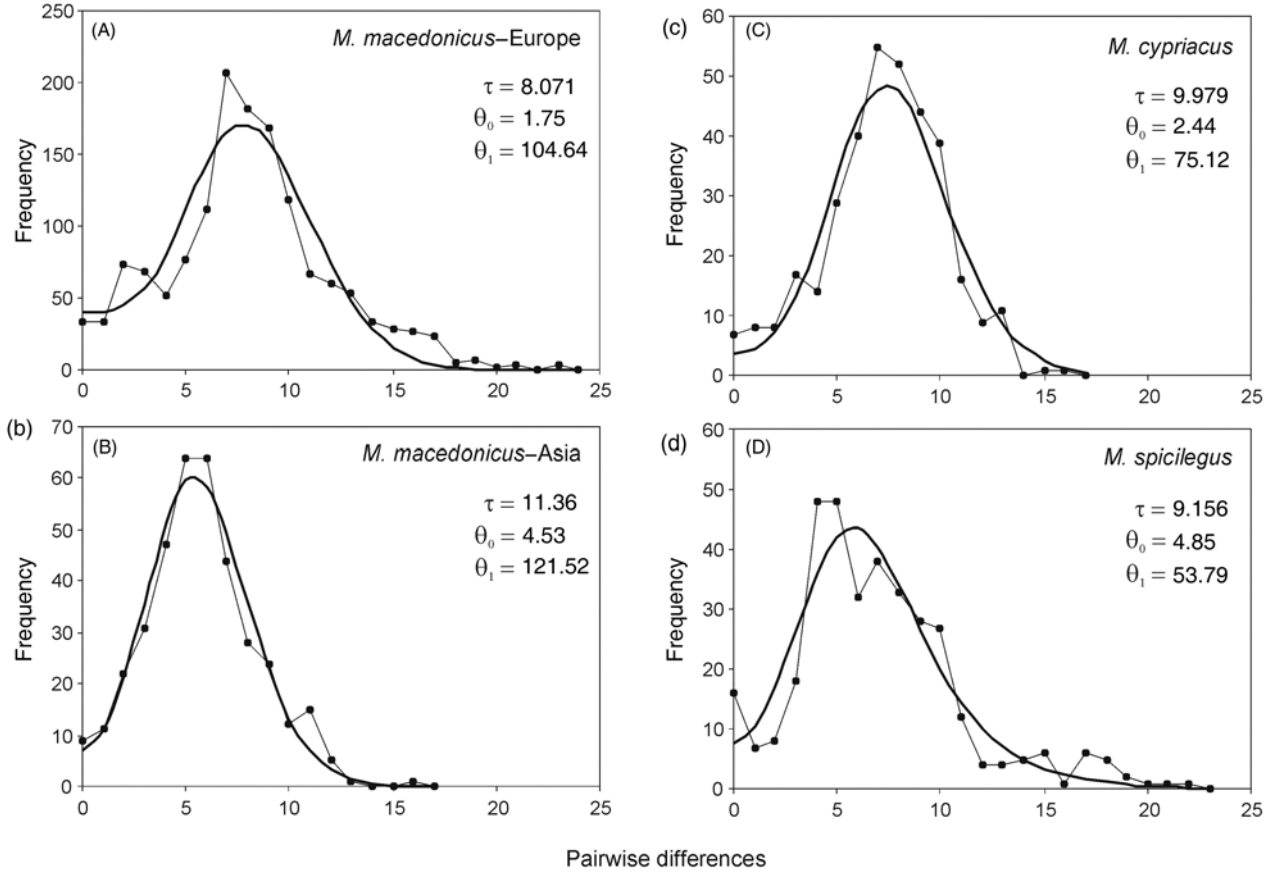


**Fig. 2** Mismatch distribution of mtDNA types in the European and Asian groups of *Mus macedonicus macedonicus* populations (A, B), *Mus cypriacus* (C), and *Mus spicilegus* (D). The observed frequencies (black dots connected with thin lines) are compared to the expected frequencies (thick lines), based on the population expansion function with parameters estimated using a generalized nonlinear least-squares approach. Approximate times of population expansion τ (in $1/2\,u$ units, where $u$ is the mutation rate for the whole sequence) and population sizes before the expansion ($\theta_0$) and at present ($\theta_1$) are given for each group.

($r = 0.2416$, $t = 6.7160$, $p[\text{rand} \geq \text{obs}] = 0.0004$); this correlation is substantially lower, though still significant, in *M. cypriacus* ($r = 0.2071$, $t = 1.9206$, $p[\text{rand} \geq \text{obs}] = 0.0316$).

Tajima's test only shows a significant deviation from neutrality in the Asian group of populations ($P = 0.0386$; 10 000 permutations), indicating either selection acting on the sequences analysed, rate heterogeneity, population growth or a bottleneck effect. Nevertheless, the mismatch distribution for all four groups has a bell-shaped distribution of substitution differences between pairs of haplotypes (Fig. 2) suggesting recent population expansion. The fit of this data with theoretical predictions is significant for all four groups (according both to sum of squared deviations and Harpending's raggedness index; 1000 replicates). Indeed, the coalescence analysis rejects the null hypothesis of a stable population in all groups analysed (Table 2). If we assume that the groups grow exponentially, from the mismatch distribution analysis we can estimate expansion time, τ, expressed as $1/2u$ generations (Li 1977; Rogers & Harpending 1992), and mutation parameters $\theta_0$ and $\theta_1$, given as $\theta_0 = 2\,\mu N_0$ and $\theta_1 = 2\,\mu N_1$, respectively, where $u$ is the mutation rate per *the whole* sequence and $N_0$ and $N_1$ are the population sizes at the time before expansion and at present, respectively (Fig. 2).

**Table 2** Results of the coalescence analysis of the four groups of aboriginal mice studied (E and A denotes European and Asian group of population, respectively). Likelihoods of the two models were compared using the likelihood-ratio tests with 1 degree of freedom

| | Log-likelihood | | | |
|---|---|---|---|---|
| Groups | Stable population model | Exponential change model | $\chi^2$ | $P$ |
| *macedonicus* E | 7.0103 | 15.6726 | 17.3246 | 0.00003 |
| *macedonicus* A | 23.5797 | 82.8268 | 118.4942 | < 0.00001 |
| *cypriacus* | 0.4964 | 6.5925 | 12.1922 | 0.00048 |
| *spicilegus* | 1.7839 | 29.7930 | 56.0182 | < 0.00001 |

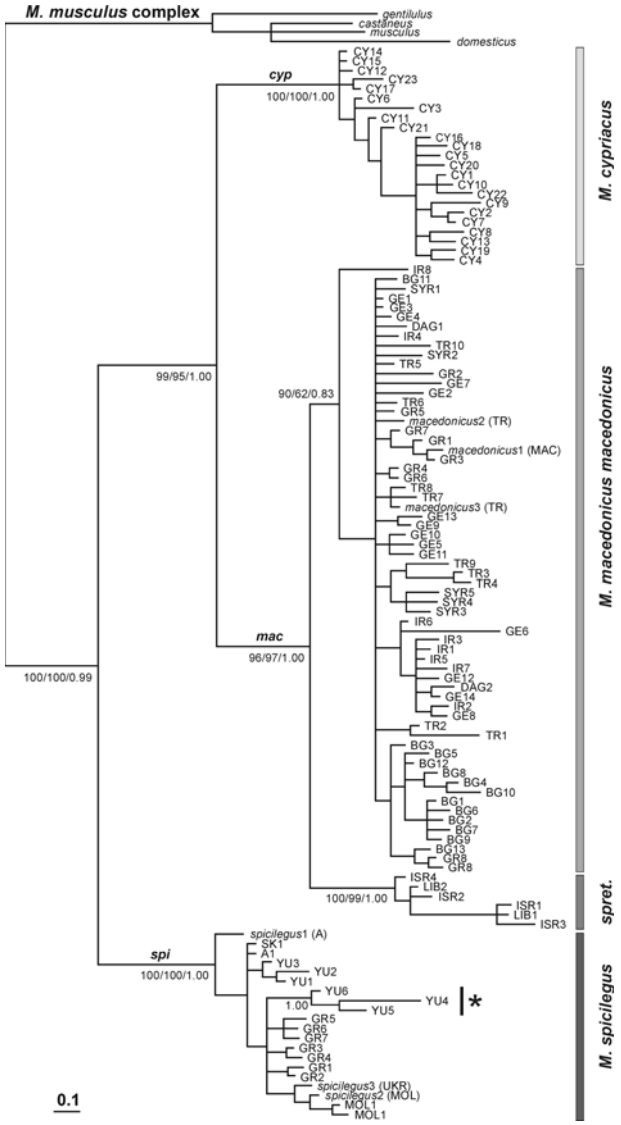


**Fig. 3** Bayesian consensus tree of the three 'aboriginal' mouse species under study and four house mouse haplotypes. The tree was rooted with a single *Mus spretus* sequence. Bayesian posterior probabilities are given for each major clade. Asterisk denotes haplotypes from Ulcinj, Montenegro, the type locality of *Mus spicilegus adriaticus*. Abbreviations: *spret.*, *Mus macedonicus spretoides*; CY, Cyprus; BG, Bulgaria; GR, Greece; TR, Turkey; GE, Georgia; DAG, Dagestan; IR, Iran; SYR, Syria, ISR, Israel; LIB, Lebanon; SK, Slovakia; A, Austria; MOL, Moldova; UKR, Ukraine; YU, Serbia and Montenegro.

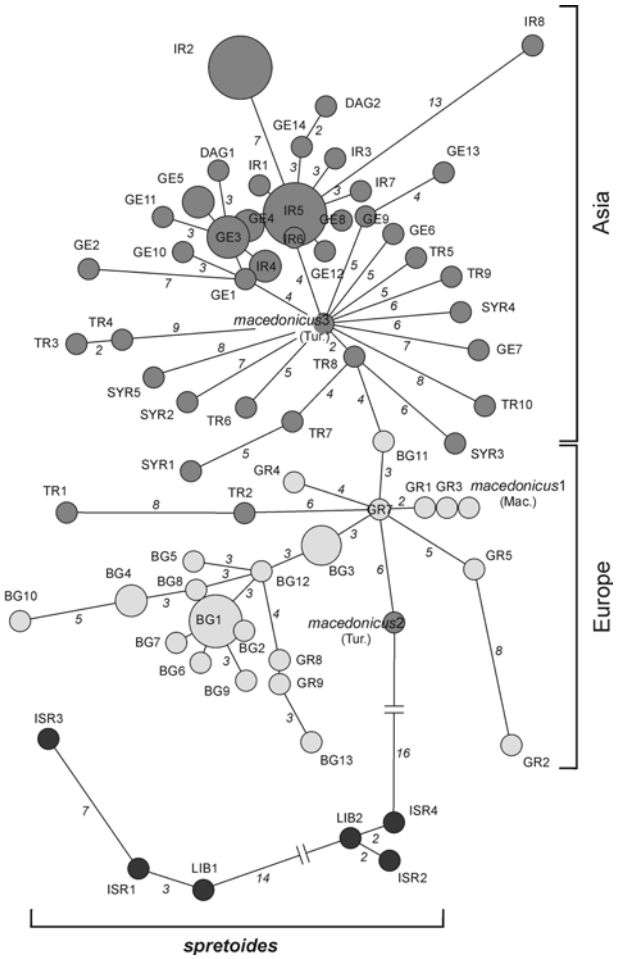


**Fig. 4** A minimum spanning network connecting mitochondrial haplotypes of *Mus macedonicus*. There are 37 alternative connections (not shown). The size of each circle corresponds to the number of haplotypes. See Appendix for the codes of individual haplotypes.

### *Phylogenetic analyses and divergence times*

Phylogenetic relationships among all taxa including all outgroups are depicted in Fig. 3. All specific and/or subspecific clades are clearly distinguished and well supported. *M. cypriacus* and *M. macedonicus* appear to be sister species while *M. spicilegus* is distinct from the other two. The same basic topology is obtained with the ML, NJ and MP analyses. The latter method infers a vast array of trees of length 789 (consistency index: *CI* = 0.497, retention index: *RI* = 0.935, rescaled consistency index: *RC* = 0.464, homoplasy index: *HI* = 0.503). When two randomly chosen sequences per taxon are used, the MP analysis resulted in two most parsimonious trees of length 314 (consistency index: *CI* = 0.736, retention index: *RI* = 0.853, rescaled consistency index: *RC* = 0.628, homoplasy index: *HI* = 0.264). The two MP trees differ in an alternative position of *M. m. castaneus* sequences (grouping either with *M. m. musculus* or *M. m. domesticus*); in both the trees, *M. m. gentilulus* was a sister group of all remaining *M. musculus* subspecies.

Relationships among haplotypes within individual species are shown in minimum spanning networks (MSN) in Figs 4 and 5. In *M. macedonicus* (Fig. 4), haplotypes tend to group into the Asian and European *macedonicus* clades, and the *spretoides* clade but the number of mutation steps

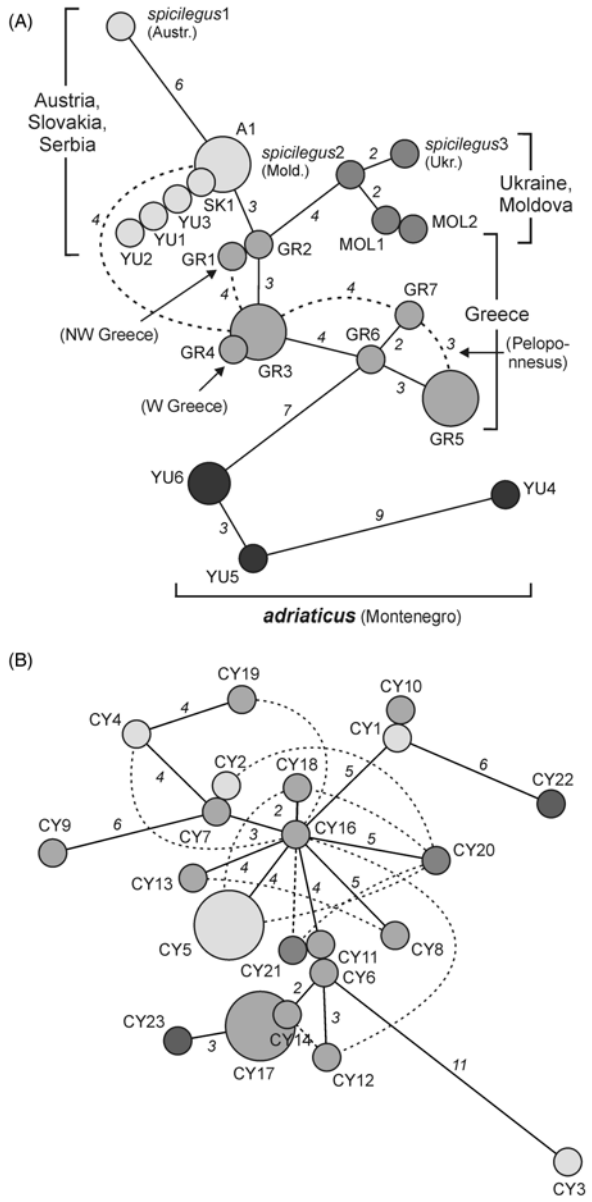


**Fig. 5** Minimum spanning network in *Mus spicilegus* (A) and *Mus cypriacus* (B). Alternative connections are indicated with dashed curves. See Appendix I for the codes of individual haplotypes.

between the former two clades is very low. Moreover, it should be noted that there are as many as 37 alternative connections in the network (not shown). Nevertheless, it seems from the figure that *M. macedonicus* has undergone a recent and rapid radiation in the Middle East as suggested by the star-like pattern among Turkish and Syrian haplotypes followed by two smaller expansions in Georgia and Iran (top of Fig. 4). In Europe, a similar radiation appears in eastern Greece (or in the European part of Turkey which was not sampled in this study). Interestingly, although *spretoides* forms a distinct group separated by as many as 16 steps, this clade does not appear within the basal Middle-Eastern radiation in any alternative topology.

The MSN of *M. spicilegus* haplotypes is shown in Fig. 5(A). Although the haplotypes are grouped according to geographical proximity, the network is rather compact with only weak substructuring suggesting a particular colonization pattern for this species. In *M. cypriacus* (Fig. 5B) the MSN pattern appears completely random since haplotypes from different regions of Cyprus are intermingled and there are 12 alternative connections including rooting with *M. spretus*.

Ambiguous relationships among haplotypes within the aboriginal species are reflected in different results of particular phylogenetic methods and large numbers of trees inferred with MP and ML analyses. The latter approach revealed flat likelihood surfaces in all the three species making any reasonable inference almost impossible. This is also corroborated by the results of Bayesian analysis. In *M. macedonicus* (figure not shown), the tree pattern is highly unresolved, with a few groups of geographically close haplotypes such as (Bulgaria–Greece), (southern Turkey–Syria) (Iran–Georgia–Dagestan), and especially *spretoides* which appear divergent from the *macedonicus* haplotypes. However, it should be noted that divergence *within* the former subspecies is almost as high as between the taxa (cf. Fig. 4). A similar pattern appears in *M. spicilegus*: although geographically close haplotypes form distinct groups, phylogenetic relationships among them are unresolved; the three haplotypes from the type locality of *M. s. adriaticus* in Montengro (YU4-6) form a monophyletic clade, yet its affinity to the haplotypes from nortwestern, western and southwestern Greece is uncertain. All phylogenetic methods used are consistent that they place the Austrian and Slovak haplotypes at the root of the trees. This result is in sharp contrast with the fact that Austria and Slovakia form the northwestern border of the species' range and may be strongly influenced by a single haplotype, *spicilegus* 1 from Halbturn in Austria. Indeed, when this sequence is removed, the topology changes: MP analysis results in a single most parsimonious tree in which the basal group consists of haplotypes from Moldova and Ukraine; in addition, this same basic pattern is revealed by all other methods.

The relative rate test does not detect any significant rate heterogeneity between compared taxa. Considering the mean distance between *M. musculus musculus* and *M. musculus domesticus* ($d = 3.72\%$) and the time since their last common ancestor 350 000 (She *et al*. 1990) to 500 000 years ago (Boursot *et al*. 1993), we can estimate coalescence times for all mtDNA lineages under study. Thus the European and Asian lineages of *M. macedonicus* are estimated to coalesce between 23 000 and 33 000 years ago depending on calibration whereas an order higher estimate is obtained for the coalescence of the two subspecies, *M. macedonicus macedonicus*

| The split | % TrN dist. | Cal. 1 — Cal. 2 | $t_{MRCA}$ |
|---|---|---|---|
| *macedonicus* E–*macedonicus* A | 0.24 | 23–33 | 163 (107–229) |
| *macedonicus–spretoides* | 2.19 | 206–294 | 301 (200–404) |
| *macedonicus–cypriacus* | 4.52 | 426–608 | 491 (344–649) |
| *cypriacus–spicilegus* | 8.77 | 826–1180 | 777 (516–1063) |
| *spicilegus–macedonicus* | 7.28 | 689–979 | 719 (510–927) |

TrN dist., Tamura–Nei distance.

**Table 3** Values of mtDNA divergence expressed as percentage Tamura–Nei distance, times of divergence (in thousand years) between species, subspecies and groups of populations, based on two estimated of the *M. musculus musculus–M. m. domesticus* split (Cal. 1 = 300 000 years; Cal. 2 = 500 000 years), and the time to the most recent common ancestor ($t_{mrca}$) estimated from coalescent trees

and *M. macedonicus spretoides*: 206 000–294 000 years ago (Table 3). Using the same calibration for interspecific divergences we derive 426 000–608 000 years ago for the *M. macedonicus–M. cypriacus* split and 826 000 years ago–1.18 Ma for the *M. spicilegus–M. cypriacus* split (Table 3).

From the estimate of the mutation rate of the two concatenated mtDNA fragments for the whole data set (without outgroups), $\mu = 0.0450$ site$^{-1}$ Myr$^{-1}$ (lower and upper bounds of the 95% highest posterior density interval: 0.0162–0.0783), we estimated $t_{MRCA}$ for all pairs of taxa and/or population groups (Table 3). For the species pairs, $t_{MRCA}$ values fall within the ranges based on TrN distances and delimited by the two calibration points, although they are slightly closer to the lower values (Table 3, right column). This is not surprising since, under the exponential growth model, $t_{MRCA}$ estimates should be lower than those based on the assumption of the constant population size: as we are going back in time, the effective population size decreases and thus the rate of coalescence increases (Nordborg 2003; Felsenstein 2004). Conversely, $t_{MRCA}$ for the two *M. macedonicus* subspecies, as well as for the European and Asian *M. m. macedonicus* populations, are higher than those estimated from the genetic distances. Using the formulas $u = m_T\mu$, where $m_T$ is the number of sites, substituting $\approx$ 1000 bp for $m_T$ and assuming two generations per year (Vohralík *et al*. 1996 and unpublished data) as a conservative estimate, we can approximate times when individual groups started to grow exponentially according to the formula $t_e = \tau/2u$ (Li 1977; Rogers & Harpending 1992), as 111 000 years ago for *M. cypriacus*, 102 000 years ago for *M. spicilegus*, 126 000 years ago for the Asian populations of *M. m. macedonicus*, and 90 000 years ago for the European populations of *M. m. macedonicus*.

## Discussion

The present study substantially increases our knowledge of mtDNA variation throughout the range of *Mus macedonicus.* Our results confirm the hypothesis that the two *M. macedonicus* subspecies, *macedonicus* and *spretoides*, are reciprocally monophyletic. In addition, the 2.2% HKY distance between *M. macedonicus spretoides* and *M. macedonicus macedonicus* is similar to the 2.9% divergence reported by Orth *et al*. (2002) and suggests that the split of the two subspecies occurred between 200 000 and 300 000 years ago depending on the calibration by the *M. musculus musculus–M. m. domesticus* split (350 000 or 500 000 years ago; She *et al*. 1990; Boursot *et al*. 1993). These dates are corroborated by Bayesian coalescent analysis that estimates $t_{MRCA}$ of the subspecies as ~300 000 years ago (Table 3). Increasing the number of sampled populations to cover almost the entire species' range enabled us to address the question of whether this species survived the LGM in more than two glacial refugia as suggested by Orth *et al*. (2002). The MSN in Fig. 4 as well as ML, MP, NJ, and BA phylogenies showed several groupings (e.g. Bulgaria + Greece, Georgia) which could indicate potential refuges within *M. m. macedonicus*. However, all the clusters within the subspecies were very poorly supported and the genetic similarity between populations might reflect recent gene flow among geographically close localities, such as in Bulgaria, Greece and (former Yugoslavian) Macedonia.

The Levant area is known to have been occupied by the Balkan short-tailed mouse for at least the last 100 000 years (Auffray *et al*. 1988, 1990). This means that, given the coalescent time between the two subspecies (Table 3), they must have split about 100 000–200 000 years before *M. m. spretoides* reached the area, and the two *M. macedonicus* subspecies probably diverged well before the LGM. Unfortunately, the latter taxa were not differentiated by Auffray *et al*. (1988, 1990) to evaluate their spatial and/or temporal dynamics across the area. There is an apparent difference between estimates of $t_{MRCA}$ of Asian and European *M. m. macedonicus* populations based on genetic divergence and Bayesian coalescent analysis (Table 3). This discrepancy could be caused, for example, by violation of the molecular clock assumption in late divergences. It should also be noted that if population growth was rapid and recent enough, no scaled time ($\tau$) would pass and no coalescence could occur (Nordborg 2003).

After the LGM, at least one northern refuge could have acted as a source for expansion to the present range. This refuge could have been located in lowlands of western Georgia or eastern Azerbaijan as hypothesized by Orth

*et al*. (2002), yet other areas such as southern Turkey or the lowlands around the Euphrates and Tigris cannot be ruled out. Although not apparent from the phylogenetic trees generated in this study, these refuges should, in theory, be detectable with higher nucleotide and haplotype diversities. Thus, we compared values of $\pi$ and $h$ for the whole Asian *M. m. macedonicus* data set with reduced samples comprising all the sites from Georgia and Daghestan (sites 19–26) and those from southwestern Iran (sites 29–32). Indeed, the Georgian sample reveals increased nucleotide diversity ($\pi = 0.0124 \pm 0.0066$) relative to all-Asian sample ($\pi = 0.0087 \pm 0.0046$; cf. Table 1); however, haplotype diversity does not differ between the two samples (Georgia: $h = 0.9789 \pm 0.0025$; all-Asia: $h = 0.9790 \pm 0.0104$; cf. Table 1). On the other hand, both nucleotide and haplotype diversity of the Iranian sample is lower than those of the whole sample ($\pi = 0.0079 \pm 0.0045$; $h = 0.7556 \pm 0.1295$) but this could be a sampling artefact, since half of this sample consists of individuals trapped at the same place and during the same night and hence, probably represent members of a single family. Interestingly, the approximate time of expansion of European haplotypes (≈ 90 000 years ago) considerably predates the end of the LGM (i.e. the lower limit for the colonization of Europe by the species).

Europe may have been colonized by *M. macedonicus* along the southern edge of the Black Sea which was not connected with the Mediterranean Sea at that time (Hosey 1982; Ryan *et al*. 1997; Aksu *et al*. 1999), and its northerly expansion through the Balkans may have been halted by competitive exclusion with the 'mound-builder' *M. spicilegus* (Ivantcheva & Cassaing 1999; Orth *et al*. 2002). This scenario assumes the 'northern' colonization route of *M. spicilegus*, hypothesized by Kryštufek & Macholán (1998). According to this hypothesis, the steppe mouse arrived in Europe after the LGM along the northern shore of the Black Sea. As a potential eastern refuge, we can assume an area south of the broad belt of woodlands reaching up to the most of the northern shore of the Black Sea (Southwood 2003), e.g. the southernmost part of Crimea (Orth *et al*. 2002), or a region east of the Sea of Azov (Goudie 2000). An alternative scenario assumes two southern colonization waves, the first undergone by *M. spicilegus* and the second wave represented by *M. macedonicus* which subsequently forced the former species farther north to its present range (Kryštufek & Macholán 1998).

An interesting question is the position of the border between the two *M. macedonicus* subspecies. Importantly, both specimens from the vicinity of Byblos in western Lebanon appeared to possess *spretoides* mtDNA haplotypes; thus, the subspecies is not limited to Israel. The site is about 110 km southwest from Qattinah, the nearest *macedonicus* locality. Another two sites harbouring *macedonicus* mice are Jabla and Rabi'ah. It is not clear whether the two taxa are allopatric or parapatric but *M. m. macedonicus* can reach the northeastern part of Lebanon along the Orontes River and the Bekaa Valley, whereas *M. m. spretoides* may be expected to expand northwards from the Jordan River Valley. Hence, if the two forms meet, we can expect a hybrid zone between them somewhere along the northern or northwestern edge of the Lebanon Mountains and/or in the Bekaa Valley.

The evolutionary history of *M. spicilegus* is even more poorly understood. The present data seem to favour the 'northern' hypothesis; however, more conclusive results require analysing more samples from the whole species range. Regardless of the direction of colonization of Europe, the star-like phylogeny pattern of the steppe mouse suggests a rather recent and rapid expansion to its present range. This result is in agreement with recent origin of the Pannonian steppes (Matvejev 1961; Godicl 1980) and indirectly supported by the apparent ability of the steppe mouse to rapidly colonize new areas when new habitat becomes available, as documented by its sudden northerly expansion across the Danube River in Slovakia (cf. Macholán 1996 to Stollmann & Macholán 1999; Krištofík & Danko 2003; Mašán & Stanko 2005). Although the approximate time of exponential population growth of this species (≈ 100 000 years ago) seems to contradict the recent radiation, it should be noted that the pattern of 'expansion' can also be obtained also by recent severe bottleneck (Rogers & Harpending 1992). Thus, we hypothesize a rapid colonization wave originating from a small founder population most probably occupying southeastern Ukraine.

Another striking result is the presence of *M. spicilegus* on Peloponnesus. Contrary to previous studies, the steppe mouse appears to occur along a narrow lowland belt stretching from southeastern Montenegro to the vicinity of the Lake Shköder and Tirana in Albania, and from northwestern (Igoumenitsa) and western (Vlaherna, Komeno) Greece to northwestern Peloponnesus (Patras), whereas presence of *M. macedonicus* has only been evidenced from Macedonia and northeastern Epirus eastwards (Macholán *et al*. 2007). Distribution areas of the two species thus need to be substantially redrawn in this part of the Balkan Peninsula.

The large east Mediterranean island of Cyprus has long been known to harbour two sympatric mouse species, commensal *M. musculus domesticus* and a free-living taxon previously assigned to *M. macedonicus* on the basis of qualitative morphological investigations (but see Kryštufek & Vohralík 2001). However, a thorough morphological analysis of recent and subfossil specimens from Cyprus has recently shown these mice to be distinct from *M. macedonicus* (Cucchi *et al*. 2002). This conclusion was later confirmed by sequencing 310 bp of mitochondrial 16S rDNA, 230 bp of mitochondrial D-loop, and 761 bp of the second intron of the nuclear *Abpa* gene in seven specimens

(Bonhomme *et al*. 2004). The new species appeared almost equidistant from both *M. macedonicus* and *M. spicilegus,* with the time of their divergence estimated at between 500 000 years ago and 1 Ma (Bonhomme *et al*. 2004). In the present study based on a larger sample, we confirmed the separate systematic status of the Cyprus mouse, however, the mitochondrial control region of *M. spicilegus* appeared about twice as divergent (8.8%) from this species as that of *M. macedonicus* (4.5%).

Bonhomme *et al*. (2004) presented two hypotheses for the origin of the wild mouse species on Cyprus. While the first assumes natural colonization of the island long before the LGM (with mice crossing to the island on rafts of vegetation), the second hypothesis proposes an invasion of the island mediated by the first human settlers. In the latter case, the ancestors of the Cyprus mouse should have been present somewhere on the nearby mainland and traces should be present in the subfossil and/or recent material. However, neither a survey of the fossil record (Cucchi *et al*. 2002) nor morphometric studies of recent populations (Macholán *et al.* 2007) indicate the presence of this taxon outside Cyprus after the Middle Pleistocene. Moreover, the level of genetic variation within *M. cypriacus* was found to be comparable to *M. macedonicus* and *M. spicilegus* (Table 1), making a recent founder event improbable. The absence of geographical structuring and lack of significant correlation between geographical and genetic distances could have been caused by sufficiently high gene flow among populations within the relatively small area of the island. Our data presented in this study suggest a colonization of the island by a small group of *M. macedonicus*-like ancestors from the Near East some 500 000 years ago followed by their genetic divergence. Note that due to exponential population growth starting about 100 000 years ago, the time of coalescence of *M. cypriacus* populations (Table 3) is probably underestimated (it should also be kept in mind that the times of expansion are crucially dependent on a precise estimation of the mutation rate, μ).

Divergence times between the three aboriginal species point to the role of Pleistocene temperature oscillations causing fluctuations of woodland areas inhospitable for mice. According to our data, the first split within the eastern group of the three aboriginal mouse species probably appeared either in Asia Minor or in the areas east of this region some 750 000 years ago (700 000 years ago to 1.2 Ma according to the level of genetic divergence), eventually giving rise to *M. spicilegus* and an ancestor of *M. macedonicus* and *M. cypriacus*. Hence, this split predated the onset of the first glacial period. Subsequently, we hypothesize a southward and westward expansion of this ancestor which then colonized Cyprus, accidentally crossing a deep marine strait that separated the island even during the minimum sea levels. This founder event was then followed by the divergence between the island and mainland populations giving rise to *M. cypriacus* and *M. macedonicus* 400 000–600 000 years ago. The last step is accompanied by separation of *M. macedonicus* populations resulting in the origin of the two subspecies, *macedonicus* and *spretoides*. After the LGM, *M. macedonicus* expanded northwards to Europe where it came into contact with *M. spicilegus*. Importantly, all inferred splits are dated to warmer periods (see Vostok ice core, 417 160–2342 BP, http://cdiac.esd.ornl.gov/trends/co2/vostok.htm; and Dome C ice core, 650 000–415 000 BP, ftp://ftp.ncdc.noaa.gov/pub/data/palaeo/icecore/antartica/epica_domec/edc-co2–650 k-390 kxls), i.e. it appears that the glaciations themselves did not play a crucial role in the evolution of this group.

## Acknowledgements

We are grateful to P. Benda, B. Bímová, E. Božíková, J. Červený, S. Fraguedakis-Tsolis, D. Frynta, I. Horáček, S. V. Mezhzherin, L. Mošanský, P. Munclinger, M. Popovici, M. Stanko, and H. M. Steiner for providing tissue samples. Special thanks are to H. C. Hauffe and three anonymous referees for their comments on an earlier version of the manuscript. This work was supported by grant no. A6045307 of the Grant Agency of the Academy of Sciences (to M.M.) and no. 206-05-2334 of the Czech Science Foundation (to V.V.).

## References

Aksu A, Hiscott RN, Yasar D (1999) Oscillating quaternary water levels of the Marmara Sea and vigorous outflow into the Aegean sea from the Marmara Sea-Black Sea drainage corridor. *Marine Geology*, **153**, 275–302.

Auffray J-C, Tchernov E, Bonhomme F, Heth G, Simson S, Nevo E (1990) Presence and ecological distribution of '*Mus spretoides*' and *Mus musculus domesticus* in Israel. Circum-Mediterranean vicariance in the genus *Mus. Zeitschrift für Saügetierkunde*, **55**, 1–10.

Auffray J-C, Tchernov E, Nevo E (1988) Origine du commensalisme de la souris domestique (*Mus musculus domesticus*) vis-à-vis de l'homme. *Comptes Rendus Académie Des Sciences, Paris, Sciences Vie*, **307**, 517–522.

Avise JC (2000) *Phylogeography. The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB (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, Biological Sciences*, **265**, 1219–1226.

Bonhomme F, Orth A, Cucchi T, Hadjisterkotis E, Vigne J-D, Auffray J-C (2004) Découverte d'une nouvelle espèce de souris sur l'île de Chypre. *Comptes Rendus Biologies*, **327**, 501–507.

Boursot P, Auffray J-C, Britton-Davidian J, Bonhomme F (1993) The evolution of the house mouse. *Annual Review of Ecology and Systematics*, **24**, 119–152.

Branco M, Monnerot M, Ferrand N, Templeton A (2002) Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested lineage and mismatch analyses of mitochondrial DNA variation. *Evolution*, **56**, 792–803.

Catzeflis FM, Aguilar JP, Jaeger JJ (1992) Muroid rodents — phylogeny and evolution. *Trends in Ecology & Evolution*, **7**, 122–126.

Cucchi T, Orth A, Auffray J-C *et al*. (2006) A new endemic species of the subgenus *Mus* (Rodentia, Mammalia) on the Island of Cyprus. *Zootaxa*, **1241**, 1–36.

Cucchi T, Vigne J-D, Auffray J-C, Croft P, Peltenburg E (2002) Introduction involontaire de la souris domestique (*Mus musculus domesticus*) à Chypre dès le Néolithique précéramique ancien (fin IXe et VIIIe millénaire av. J.-C.). *Comptes Rendus Palevology*, **1**, 235–241.

Drummond AJ, Rambaut A (2006) Beast V1.4, Bayesian evolutionary analysis sampling trees. Available from: http://beast.bio.ed.ac.uk.

Edward SV (1997) Relevance of microevolutionary processes to higher level molecular systematics. In: *Avian Molecular Evolution and Systematics* (ed. Mindell DP), pp. 251–278. Academic Press, New York.

Felsenstein J (1981) Evolutionary trees from DNA sequences. A maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.

Felsenstein J (1985) Confidence limits on phylogenies with a molecular clock. *Systematic Zoology*, **34**, 152–161.

Felsenstein J (2004) *Inferring Phylogenies*. Sinauer Associates, Sunderland, Massachusetts.

Fitch WM (1971) Towards defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology*, **20**, 406–416.

Godicl L (1980) Steppenflora in Nordwest-Jugoslawien. *Razprave IV. Razreda SAZU, Ljubljana*, **12**, 281–367.

Goudie A (2000) *The Human Impact on the Natural Environment*. Blackwell, Oxford, UK.

Gu X, Fu Y-X, Li W-H (1995) Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution*, **12**, 546–557.

Gündüz İ, Tez C, Malikov V, Vaziri A, Polyakov AV, Searle JB (2000) Mitochondrial DNA and chromosomal studies of wild mice (*Mus*) from Turkey and Iran. *Heredity*, **84**, 458–467.

Harpending RC (1994) Signatures of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

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

Hassanin A, Lecointre G, Tillier S (1998) The 'evolutionary signal' of homoplasy in protein-coding gene sequences and its consequences for a priori weighting in phylogeny. *Comptes Rendus Académie Des Sciences, Paris, Sciences Vie*, **321**, 611–620.

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

Hosey GR (1982) The Bosporus land-bridge and mammalian distributions in Asia Minor and the Balkans. *Säugetierkundliche Mitteilungen*, **30**, 53–62.

Ivantcheva A, Cassaing J (1999) Male–female interactions and socio-sexual isolation of *Mus spicilegus* toward two other species of east-European mice. *M. macedonicus* and *M. m. Musculus. Comptes Rendus Académie Des Sciences, Paris, Sciences Vie*, **322**, 597–605.

Kotlík P, Deffontaine V, Mascheretti S, Zima J, Michaux JR, Searle JB (2006) A northern glacial refugium for bank voles (*Clethrionomys glareolus*). *Proceedings of the National Academy of Sciences, USA*, **103**, 14860–14864.

Krištofík J, Danko Š (2003) Distribution of *Mus spicilegus* (Mammalia: Rodentia) in Slovakia. *Lynx*, **34**, 55–60.

Kryštufek B, Macholán M (1998) Morphological differentiation in *Mus spicilegus* and the taxonomic status of mound-building mice from the Adriatic coast of Yugoslavia. *Journal of Zoology*, **245**, 185–196.

Kryštufek B, Vohralík V (2001) *Mammals of Turkey and Cyprus Introduction, Checklist, Insectivora*. Zgodovinsko društvo za južno Primorsko, Koper.

Kuhner MK, Yamato J, Felsenstein J (1995) Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. *Genetics*, **140**, 1421–1430.

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

Li W-H (1977) Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. *Genetics*, **85**, 331–337.

Macholán M (1996) Morphometric analysis of European house mice. *Acta Theriologica*, **41**, 255–275.

Macholán M (1999a) *Mus macedonicus* Petrov & Ružić, 1983. In: *The Atlas of European Mammals* (eds Mitchell-Jones AJ, Amori J, Bogdanowicz W *et al*.), pp. 284–285. Academic Press, London.

Macholán M (1999b) *Mus spicilegus* Petényi, 1882. In: *The Atlas of European Mammals* (eds Mitchell-Jones AJ, Amori J, Bogdanowicz W *et al*.), pp. 288–289. Academic Press, London.

Macholán M, Filippucci MG, Zima J (2001) Genetic variation and zoogeography of pine voles of the *Microtus subterraneus/majori* group in Europe and Asia Minor. *Journal of Zoology*, **255**, 31–42.

Macholán M, Mikula O, Vohralík V (2007) Morphometric analysis of two eastern-Mediterranean non-commensal mouse species, *Mus macedonicus* and *M. cypriacus* (Rodentia: Muridae), using traditional and geometric approaches. *Zoologischer Anzeiger*, **246**, doi: 10.1016/j.jcz.2007.07.003.

Macholán M, Vohralík V (1997) Note on the distribution of *Mus spicilegus* (Mammalia: Rodentia) in the south-western Balkans. *Acta Societatia Zoologicae Bohemicae*, **61**, 219–226.

Mašán P, Stanko M (2005) Mesostigmatic mites (Acari) and fleas (Siphonaptera) associated with nests of mound-building mouse, *Mus spicilegus* Petényi, 1882 (Mammalia, Rodentia). *Acta Parasitologica*, **50**, 228–234.

Matvejev S (1961) *Biogeography of Yugoslavia*. Biološki institut N.R. Srbije, 9.

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

Nordborg M (2003) Coalescent theory. In: *Handbook of Statistical Genetics*, 2nd edn (eds Balding DJ, Bishop M, Cannings), pp. 602–635. John Wiley and Sons, Inc., Chichester.

Orth A, Auffray J-C, Bonhomme F (2002) Two deeply divergent mitochondrial clades in the wild mouse *Mus macedonicus* reveal multiple glacial refuges south of Caucasus. *Heredity*, **89**, 353–357.

Petit RJ, Brewer S, Bordacs S *et al*. (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology Management*, **156**, 49–74.

Petrov B (1992) *Mammals of Yugoslavia — Insectivores and Rodents*. Supplementum 37. Natural History Museum, Belgrade,

Petrov B, Ružić A (1983) Preliminary report on the taxonomic status of the members of the genus *Mus* in Yugoslavia with description of a new subspecies (*Mus hortulanus macedonicus* ssp. n., Rodentia, Mamm.). *Proceedings of the 2nd Symposium on the Serbian Fauna, Beograd*, 175–178.

Petrov B, Ružić A (1985) Taxonomy and distribution of members of the genus *Mus* (Rodentia, Mammalia) in Yugoslavia. *Zbornik Radova O Fauni SR Srbije, SANU Beograd*, **3**, 209–243.

Philippe H, Douzery E (1994) The pitfalls of molecular phylogeny based on four species as illustrated by the Cetacea/Artiodactyla and postglacial patterns of subdivision in the meadow grasshopper *Chorthippus parallelus*. *Heredity*, **80**, 633–641.

Posada D, Crandall KA (1988) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Prager EM, Orrego C, Sage RD (1998) Genetic variation and phylogeography of Central Asian and other house mice, including a major new mitochondrial lineage in Yemen. *Genetics*, **150**, 835–861.

Prager EM, Sage RD, Gyllensten U *et al*. (1993) Mitochondrial DNA sequence diversity and the colonization of Scandinavia by house mice from East Holstein. *Biological Journal of the Linnean Society*, **50**, 85–122.

Prager EM, Tichy H, Sage RD (1996) Mitochondrial DNA sequence variation in the eastern house mouse, *Mus musculus*: Comparison with other house mouse mice and report of 75-bp tandem repeat. *Genetics*, **143**, 427–446.

Rannala B, Yang Z (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution*, **43**, 304–311.

Robinson M, Gouy M, Gautier C, Mouchiroud D (1998) Sensitivity of the relative-rate test to taxonomic sampling. *Molecular Biology and Evolution*, **15**, 1091–1098.

Robinson-Rechavi M, Huchon D (2000) RRTREE: relative-rate tests between groups of sequences on a phylogenetic tree. *Bioinformatics*, **16**, 296–297.

Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608–615.

Rogers AR, Harpending H (1992) Population growth marks waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rohlf FJ (1997) *NTSYS-PC: Numerical Taxonomy and Multivariate Analysis System*, Exeter Software Version 2.00. Setauket, New York.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Ryan W, Pitman II W, Major C *et al*. (1997) An abrupt drowning of the Black Sea shelf. *Marine Geology*, **138**, 119–126.

Sage RD (1981) Wild mice. In: *The Mouse in Biomedical Research* (eds Foster HL, Small JD, Fox JG), pp. 39–90. Academic Press, New York.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **102**, 1079–1089.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN, Version 2.000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

She JX, Bonhomme F, Boursot P, Thaler L, Catzeflis F (1990) Molecular phylogenies in the genus *Mus*: comparative analysis of electrophoretic scnDNA hybridisation and mtDNA RFLP data. *Biological Journal of the Linnean Society*, **41**, 83–103.

Southwood TRE (2003) *The Story of Life*. Oxford University Press, New York.

Stollmann A, Macholán M (1999) The mound-building mouse is part of the Slovakian fauna — a reply. *Folia Zoologica*, **48**, 237–239.

Swofford DL (2000) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4 0b10. Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**, 437–460.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

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

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.

Vohralík V, Sofianidou TS, Frynta D (1996) Population structure in *Mus macedonicus* (Mammalia: Rodentia) in the Balkans. *Folia Zoologica*, **45**, 97–103.

Waddel PJ, Penny D (1996) Evolutionary trees of apes and humans from DNA sequences. In: *Handbook of Symbolic Evolution* (eds Lock AJ, Peters CR), pp. 53–73. Clarendon Press, Oxford, UK.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Wu C-I, Li WH (1985) Evidence for higher rates of nucleotide substitutions in rodents than in man. *Proceedings of the National Academy of Sciences, USA*, **82**, 1741–1745.

Xia X, Xie Z (2001) DAMBE: Software package for data analysis in molecular biology and evolution. *Journal of Heredity*, **92**, 371–373.

---

Miloš Macholán is interested in systematics and evolutionary genetics of small mammals, evolution of the genus *Mus* and hybridization in the house mouse complex. Martina Vyskočilová is engaged in molecular genetics of varied organisms including mice and fish. Francois Bonhomme leads an evolutionary genetics group at the University of Montpellier II. Annie Orth works in this group where she is specialized on wild mice genetics. Boris Kryštufek and Vladimir Vohralík are interested in taxonomy, zoogeography and biology of small mammals, focusing mainly on the Balkans, Asia Minor and the Near East.

---

## Appendix

Collection sites, sample sizes (*N*), site locations and haplotype information of all samples analysed

| Species/no. | Locality | *N* | GenBank | Haplotype | Latitude/longitude |
|---|---|---|---|---|---|
| *M. macedonicus macedonicus* | | | | | |
| | Bulgaria | | | | |
| 1 | Prilep | 1 | This report | *mac*BG1 | 43°24′N, 27°55′E |
| 2 | Banya | 6 | This report | *mac*BG1-6 | 42°48′N, 27°53′E |
| 3 | Slantchev Briag | 3 | This report | *mac*BG1,7,8 | 42°42′N, 27°43′E |
| 4 | Jitarovo | 2 | This report | *mac*BG9 | 42°38′N, 27°18′E |
| 5 | Iabeltchevo | 2 | This report | *mac*BG1,10 | 42°43′N, 27°16′E |
| 6 | Veltsin | 3 | This report | *mac*BG3,4,11 | 42°45′N, 26°50′E |
| 7 | Partizanin | 2 | This report | *mac*BG12,13 | 42°15′N, 25°15′E |
| | Macedonia | | | | |
| 8 | Gradsko | 1 | U47535 | *macedonicus*1 | |
| | Greece | | | | |
| 9 | Florina | 3 | This report | *mac*GR1-3 | 40°47′N, 21°25′E |
| 10 | Stavroupoli | 6 | This report | *mac*GR4-9 | 41°12′N, 24°43′E |
| | Turkey | | | | |
| 11 | Bardakci | 2 | This report | *mac*TR1,2 | 39°10′N, 28°34′E |
| 12 | Datça | 2 | This report | *mac*TR3,4 | 36°45′N, 27°39′E |
| 13 | Karabulut | 1 | This report | *mac*TR5 | 38°28′N, 31°29′E |
| 14 | Akburun | 1 | this report | *mac*TR6 | 37°47′N, 31°36′E |
| 15 | Lake Emir, Ankara | 2 | AF074546–7 | *macedonicus*2,3 | |
| 16 | Koyulhisar | 2 | This report | *mac*TR7,8 | 40°19′N, 37°38′E |
| 17 | Çevlık (Mağaracık) | 1 | This report | *mac*TR9 | 36°09′N, 35°56′E |
| 18 | Akcakale | 1 | This report | *mac*TR10 | 36°43′N, 38°58′E |
| | Georgia | | | | |
| 19 | Gori | 1 | This report | *mac*GE1 | 41°52′N, 43°30′E |
| 20 | Kareli | 1 | This report | *mac*GE2 | 42°01′N, 43°53′E |
| 21 | Chardakhi | 5 | This report | *mac*GE3-5 | 41°52′N, 44°35′E |
| 22 | Lissi | 4 | This report | *mac*GE4,6,7; *mac*IR5 | 41°44′N, 44°47′E |
| 23 | Krtzanissi | 2 | This report | *mac*GE8,9 | 41°36′N, 44°57′E |
| 24 | Alazani | 4 | This report | *mac*GE10-13 | 41°37′N, 45°58′E |
| 25 | Chirakskaya | 2 | This report | *mac*GE14,15 | 41°22′N, 46°21′E |
| | Dagestan, Russia | | | | |
| 26 | Derbent | 2 | This report | *mac*DAG1,2 | 42°04′N, 48°17′E |
| | Iran | | | | |
| 27 | Basţam | 1 | This report | *mac*IR1 | 38°54′N, 44°59′E |
| 28 | Choplu | 6 | This report | *mac*IR2,3 | 36°28′N, 47°01′E |
| 29 | Hinage | 2 | This report | *mac*IR4 | 34°44′N, 47°57′E |
| 30 | Bisotun | 2 | This report | *mac*IR5,6 | 34°23′N, 47°26′E |
| 31 | Bavineh | 1 | This report | *mac*IR7 | 33°36′N, 47°11′E |
| 32 | Pahlat | 5 | This report | *mac*IR5,8 | 33°29′N, 48°04′E |
| | Syria | | | | |
| 33 | Rabi′ah | 1 | This report | *mac*SYR1 | 35°49′N, 36°02′E |
| 34 | Jabla | 2 | This report | *mac*SYR2,3 | 35°28′N, 35°54′E |
| 35 | Qattinah | 2 | This report | *mac*SYR4,5 | 34°40′N, 36°37′E |
| *M. m. spretoides* | | | | | |
| | Lebanon | | | | |
| 36 | El Fidar, Byblos | 2 | This report | *mac*LEB1,2 | 34°06′N, 35°38′E |
| | Israel | | | | |
| 37 | Dor | 1 | This report | *mac*ISR1 | 32°38′N, 34°56′E |
| 38 | Poriyya | 1 | This report | *mac*ISR2 | 32°42′N, 35°37′E |
| 39 | Jerusalem1 | 1 | This report | *mac*ISR3 | 31°46′N, 35°13′E |
| 40 | Jerusalem2 | 1 | This report | *mac*ISR4 | 32°43′N, 35°03′E |

**Appendix** *Continued*

| Species/no. | Locality | *N* | GenBank | Haplotype | Latitude/longitude |
|---|---|---|---|---|---|
| *M. spicilegus spicilegus* | | | | | |
| | Austria | | | | |
| 41 | Mönchhof | 2 | This report | *spi*AUT1 | 47°53′N, 16°57′E |
| 42 | Halbturn | 1 | U47536 | *spicilegus*1 | 47°52′N, 16°59′E |
| | Slovakia | | | | |
| 43 | Vrbová nad Váhom | 2 | This report | *spi*AUT1, *spi*SK1 | 47°52′N, 18°03′E |
| | Serbia | | | | |
| 44 | Pančevo | 3 | This report | *spi*YU1-3 | 44°52′N, 20°40′E |
| | Greece | | | | |
| 45 | Igoumenitsa | 2 | This report | *spi*GR1,2 | 39°31′N, 20°16′E |
| 46 | Vlaherna | 2 | This report | *spi*GR3,4 | 39°09′N, 20°59′E |
| 47 | Komeno | 2 | This report | *spi*GR3 | 39°03′N, 21°02′E |
| 48 | Patras1 | 3 | This report | *spi*GR5 | 38°15′N, 21°44′E |
| 49 | Patras2 | 2 | This report | *spi*GR6,7 | 38°16′N, 21°45′E |
| | Moldova | | | | |
| 50 | Kishinev1 | 2 | This report | *spi*MOL1,2 | 46°58′N, 28°55′E |
| 51 | Kishinev2 | 1 | U47537 | *spicilegus*2 | |
| | Ukraine | | | | |
| 52 | Dshankoi | 1 | U47538 | *spicilegus*3 | |
| *M.* s. *adriaticus* | | | | | |
| | Montenegro | | | | |
| 53 | Ulcinj | 4 | This report | *spi*YU4-6 | 41°55′N, 19°13′E |
| *M. cypriacus* | | | | | |
| | Cyprus | | | | |
| 54 | Petratis Gorge, Polis | 1 | This report | *cyp*CY1 | 35°00′N, 32°22′E |
| 55 | Neo Chorio | 2 | This report | *cyp*CY2,3 | 35°01′N, 32°20′E |
| 56 | Pelathousa | 2 | This report | *cyp*CY4,5 | 35°02′N, 32°28′E |
| 57 | Mt. Olympus,Troodos | 1 | This report | *cyp*CY6 | 34°56′N, 32°53′E |
| 58 | Apsiou | 7 | This report | *cyp*CY7-13 | 34°48′N, 33°01′E |
| 59 | Paramytha | 10 | This report | *cyp*CY5,14–19 | 34°46′N, 33°00′E |
| 60 | Kornos | 2 | This report | *cyp*CY20,21 | 34°56′N, 33°24′E |
| 61 | St. Hilarion | 1 | This report | *cyp*CY22 | 35°19′N, 33°17′E |
| 62 | Zeytinlik, Girne (Kyreneia) | 1 | This report | *cyp*CY23 | 35°19′N, 33°18′E |