Multilocus phylogeny of arvicoline voles (Arvicolini, Rodentia) shows small tree terrace size

Authors: Martínková, Natália, and Moravec, Jiří

Source: Folia Zoologica, 61(3–4) : 254-267

Published By: Institute of Vertebrate Biology, Czech Academy of Sciences

URL: https://doi.org/10.25225/fozo.v61.i3.a10.2012

Folia Zool. – 61 (3–4): 254–267 (2012)

Multilocus phylogeny of arvicoline voles (Arvicolini, Rodentia) shows small tree terrace size

Natália MARTÍNKOVÁ[1,2]* and Jiří MORAVEC[2]

[1] *Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, v.v.i., Květná 8, 603 65 Brno, Czech Republic; e-mail: martinkova@ivb.cz*
[2] *Institute of Biostatistics and Analyses, Masaryk University, Kamenice 3, 625 00 Brno, Czech Republic; e-mail: jirka678@seznam.cz*

Received 4 January 2012; Accepted 25 June 2012

Abstract. We combined mitochondrial (*cyb*, control region, *coi*, *nd4*) and nuclear (*irbp*, *ghr*, *sry*, *lcat*) DNA sequence data to infer phylogenetic relationships of arvicoline voles. The concatenated supermatrix contained 72.8 % of missing data. From this dataset, Bayesian inference showed close relationships of *Arvicola* and *Chionomys*, *Proedromys* with *Lasiopodomys* and *Microtus gregalis*, *Phaiomys* with *Neodon* and *M. clarkei*. Genus *Microtus* formed a supported group with *Blanfordimys* and *N. juldaschi*. The gene partition taxon sets were explained in the multilocus phylogeny in such a way that the resulting Bayesian inference tree represented a unique solution on a terrace in the tree space. This means that although the supermatrix contained a large proportion of missing data, it was informative in retrieving a phylogeny with a unique optimality score, tree likelihood.

Key words: divergence, evolutionary history, supertree, supermatrix, phylogenetic tree terrace, *Microtus*, Arvicolinae

Introduction

Arvicoline voles (Rodentia, Arvicolini) are a young group of small rodents distributed on the northern hemisphere. They started to diverge probably as recently as two to three million years ago, but in the short time frame, they speciated into one of the most speciose mammalian groups (Wilson & Reeder 2005). Today, the species show rapid temporal changes in genetic composition of populations (Bryja et al. 2007, Oliver et al. 2009, Rudá et al. 2010; but see Spaeth et al. 2009) and fast karyotype reorganisation (Mazurok et al. 2001, Mekada et al. 2002, Sitnikova et al. 2007, Mitsainas et al. 2010) coupled with gene reorganisation between mitochondrial and nucleotide genomes (Triant & DeWoody 2007, 2008). Populations of arvicoline voles diverge quickly when they become fragmented in refugia, leading at times to speciation in these areas, and refugia become speciation traps (Martínková & Dudich 2003, Martínková et al. 2007, Tougard et al. 2008, Kryštufek et al. 2009, Haring et al. 2011). Multiple vole species co-occur in many regions and habitat partitioning was found with sympatric occurrence (Jurdíková et al. 2000, Santos et al. 2011). Yet, they are morphologically similar with small differences between distantly related taxa (Fraguedakis-Tsolis et al. 2009).

Reconstruction of phylogenetic relationships in arvicoline voles is complicated by morphological similarities and rapid karyotype rearrangements, making the group an ideal model for molecular genetic studies. Both mitochondrial (mtDNA) and nuclear (nucDNA) markers have been extensively used to resolve relationships within the group and to study evolutionary history of different taxa. The relative merit of information value of mtDNA and nucDNA markers overlaps in arvicoline rodents. MtDNA sequences provide valuable information in phylogeny reconstruction at specific and intrageneric level, contributing to resolving phylogeographic histories of taxa (e.g. Conroy & Cook 2000a, Jaarola & Searle 2002, Fink et al. 2004, Brunhoff et al. 2006, Fan et al. 2011), identification of putative cryptic species

* *Corresponding Author*

(Kefelioğlu & Kryštufek 1999, Hellborg et al. 2005, Castiglia et al. 2008, Conroy & Neuwald 2008, Weksler et al. 2010) and phylogenetic placement of taxa with unstable position based on other data (Macholán et al. 2001, Jaarola et al. 2004, Martínková et al. 2007, Kryštufek et al. 2009, Bannikova et al. 2010). The phylogenetic signal of mtDNA and its ability to resolve relationships decrease at higher level of phylogenies that exhibits as a rapid burst of diversification (Jaarola et al. 2004). However, also nuclear markers, either DNA sequence data or AFLP markers (Galewski et al. 2006, Abramson et al. 2009, Fink et al. 2010), fail to fully resolve the signal of the rapid diversification of voles at the base of the tree. Incongruence of sampling between studies further complicates interpretation of relationships within the group (Bužan et al. 2008, Haring et al. 2011).

Here, we combine available mitochondrial and nuclear DNA sequence markers to reconstruct phylogenetic relationships of arvicoline rodents and to assess stability of the retrieved model. We use Bayesian inference analysis of a concatenated supermatrix and SuperTriplets supertree reconstruction to estimate parent trees. These we then use to establish the size of the terrace where trees will have the same likelihood with the dataset with large content of missing data.

Material and Methods

Sequences were downloaded from GenBank for 74 species belonging to genera *Arvicola*, *Blanfordimys*, *Chionomys*, *Lasiopodomys*, *Microtus*, *Neodon*, *Phaiomys* and *Proedromys* (Table 1). Herewith, we accept species designation of Wilson & Reeder (2005) with the addition of recently established species *Microtus gromovi* (Bannikova et al. 2010) and *Proedromys liangshanensis* (Liu et al. 2007). Alignments were constructed in Geneious 5.4 (Drummond et al. 2011) with sequences from mitochondrial genes for cytochrome *b* (*cyb*), control region (CR), cytochrome *c* oxidase subunit I (*coi*), NADH dehydrogenase subunit 4 (*nd4*) and nuclear genes for interphotoreceptor retinoid-binding protein, exon 1 (*irbp*), growth hormone receptor, exon 10 (*ghr*), sex-determining region Y (*sry*), lecithin: cholesterol acyl transferase, exons 2 through 5 (*lcat*). The alignments were reduced to contain at least three sequences at every base. In the *sry* gene, the microsatellite $(TC)_n(TG)_n$ (Acosta et al. 2010) could not be aligned unambiguously, and the region was deleted. The data and results are available through TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S12667).

Two additional loci have good taxonomic sampling. However, both loci in the *avpr1a* gene, its upstream region and exon 1, were previously documented to be disparate with the species tree (Fink et al. 2007, 2010), and some sequences of the *avpr1a* exon 1 were shared between distantly related species (Fink et al. 2007). To reduce conflict between gene trees in our analyses, we chose to omit these loci.

Optimal substitution model was estimated in MrModeltest 2.3 (Nylander 2004) with Akaike Information Criterion (AIC) and applied to the individual gene alignments and partitions of the concatenated alignment. Where the parameters of the selected model were extreme, a simpler model was used. Bayesian Inference (BI) was conducted in MrBayes 3.1 (Ronquist & Huelsenbeck 2003) with Markov chain Monte Carlo (MCMC) parameters set to 2-5 million steps sampled every 1000th, five to six chains in two runs with chain temperature 0.08-0.12 and chain swapping attempted once every third generation. The MCMC runs were optimised to mix and ideally to finish with average standard deviation of split frequencies below 0.01, potential scale reduction factor for model parameters approaching 1.000 and proportion of successful chain stage swaps between 0.4 and 0.7. BI is robust in recovering the correct tree topology, but it might fail to establish appropriate branch lengths with default branch length prior (Marshall 2010). The 95 % credibility intervals of the tree lengths from BI were compared to the tree length obtained from maximum likelihood (ML) analysis. The ML tree was calculated in RAxML 7.2 (Stamatakis 2006). The trees were re-rooted to midpoint root to allow for uncertainty in the phylogenetic position of *Arvicola* (Galewski et al. 2006, Bužan et al. 2008, Abramson et al. 2009, Bannikova et al. 2009).

Divergence events between all taxa were investigated in two ways; from a combination of gene trees and directly from the concatenated supermatrix. The gene trees were combined into a SuperTriplets supertree (Ranwez et al. 2010). The method breaks down the gene trees to their smallest components containing three taxa, where any two taxa are more closely related than either is to the third. Supertree then contains medians of relationships from the triplets as they were found in the gene trees. Edge support in the SuperTriplets analysis is the proportion of triplets that support a given edge.

The supermatrix was resolved with partitioned BI ran for 6 million generations with five heated and one cold MCMC, chain temperature set to 0.09 and one chain swap attempted every third generation. Partition rates were allowed to vary.

Table 1. *Accession numbers of DNA sequences used in this study to reconstruct phylogenetic relationships.*

Species	Distrib.	mitochondrial				nuclear			
		cyb	CR	*coi*	*nd4*	*irbp*	*ghr*	*sry*	*lcat*
Arvicola									
A. amphibius (=terrestris)	Europe	AF119269	AY948543	AY332681	AF128938	AY277407	AM392380	FN433499	GQ267511
A. sapidus	Europe	FJ539345	FJ502319					FN433498	
Blanfordimys									
B. afghanus	Asia	EF599109							
B. bucharensis	Asia	AM392369					AM392392		
Chionomys									
C. gud	Asia	EU700087							GQ267512
C. nivalis	Europe	AY513845	AF267284	AY332687		AM919424	AM392378	FN433493	AH005248
C. roberti	Asia	AY513850							
Lasiopodomys									
L. brandtii	Asia						GQ142008		FN401371
L. mandarinus	Asia	FJ986322				AM919413	AM392396		GQ267514
Microtus									
M. abbreviatus	N America	AF163890							
M. agrestis	Europe	FJ619786	AF267270		AF128940	AM919427	AM910792	FN433492	
M. anatolicus	Asia	FJ767740							
M. arvalis	Europe	AY220789	AF267285	AY332683		AM919416	AM392386	FN433490	GQ267517
M. bavaricus	Europe	GQ243218	AF267277						
M. brachycercus	Europe	AY513828							
M. breweri	N America							FN433501	
M. cabrerae	Europe	AY513788							
M. californicus	N America	AF163891						FN433510	
M. canicaudus	N America	AF163892							
M. chrotorrhinus	N America	AF163893				AM919403	AM392383	FN433503	
M. clarkei	Asia	AY641526							
M. daghestanicus	Asia	AY513790					GQ142009		GQ267518
M. dogramacii	Asia	AY513795							
M. duodecimcostatus	Europe	AJ717744					AM392400	FN433496	
M. evoronensis	Asia		HM135862						
M. felteni	Europe	AY513798							
M. fortis	Asia	AF163894	JF261174	JF261174	JF261174				
M. gerbei	Europe	AY513799							
M. gregalis	Asia	AF163895					GQ142007		GQ267513
M. gromovi	Asia	FJ986319	HM135891						
M. guatemalensis	C America	AF410262							
M. guentheri	Eurasia	AY513804				AM919420	AM392397	FN433488	
M. ilaeus	Asia	AY513809							
M. irani	Asia	FJ767748							

M. kikuchii	Asia	AF348082	AF348082	AF348082	AF348082	AM919410	AM392385		
M. levis	Eurasia	DQ015676	DQ015676	DQ015676	DQ015676			FN433489	
M. liechtensteini	Europe	AY513811	AF267281						
M. limnophilus	Asia	FJ986323				AM919426			
M. longicaudus	N America	AF119267			AF128936	AM919414	AM392379		
M. lusitanicus	Europe	AY513812						FN433497	
M. majori	Europe	AY513814				AM919409	AM910796		
M. maximowiczii	Asia	FJ986311	HM135863	HM137737					
M. mexicanus	N America	AF163897	AF251260						
M. middendorffii	Asia	HM119493	HM135899	HM137752		AM919419	AM392390		GQ267516
M. miurus	N America	GU809171							
M. mongolicus	Asia	FJ986305		HM137754					
M. montanus	N America	AF119280	GU394082		AF128946			FN433513	
M. montebelli	Asia	AF163900				AM919421	AM910793		
M. mujanensis	Asia		HM135854						
M. multiplex	Europe	AY513815	AF267276						
M. oaxacensis	C America	AF410260							
M. ochrogaster	N America	AF163901				AM919423	AM392389	FN433506	
M. oeconomus	Holarctic	AB372207	AJ616853			AM919418	AM392388	FN433491	GQ267515
M. oregoni	N America	AF163903						FN433516	
M. pennsylvanicus	N America	AF119279	AY369777	JF443830	AF128945	AM919415	AF540633		
M. pinetorum	N America	AF163904							
M. quasiater	C America	AF410259							
M. richardsoni	N America	AF163905				AM919404	AM392387	FN433517	
M. sachalinensis	Asia	FJ986318	HM135900						
M. savii	Europe	AY513824						FN433495	
M. schelkovnikovi	Asia	AM910619				AM919408	AM910794		
M. socialis	Asia	AY513829				FM162055	FM162073		
M. subterraneus	Europe	AY513832	AF267271	AY332685					
M. tatricus	Europe	AY513837	AF267267						
M. thomasi	Europe	AY513840	AY560558			AM919422	AM910797	FN433494	
M. townsendii	N America	AF163906							
M. transcaspicus	Asia					AM919405	AM910795		
M. umbrosus	C America	AF410261							
M. xanthognathus	N America	AF163907							
Neodon									
N. irene	Asia	AM392370				AM919412	AY294924		
N. juldaschi	Asia	AY513808							
N. sikimensis	Asia					AY163593			
Phaiomys									
Ph. leucurus	Asia	AM392371				AM919400	AM392394		
Proedromys									
Pr. liangshanensis	Asia	FJ463038	FJ463038	FJ463038	FJ463038				

Given the fractional nature of the supermatrix, multiple distinct trees were likely to display the same set of subtrees representing taxa sampled per gene. Such trees will have the same log-likelihood for a partitioned analysis (Sanderson et al. 2011). A set of trees that display the same set of subtrees for sampled loci and have the same log-likelihood is called a terrace. The trees from a terrace are derived from each other by nearest-neighbour interchange (NNI) rearrangement, and, in a dataset with considerable missing data content, the terraces might contain many trees. The size of terraces for our dataset was assessed with perl scripts from the PhyloTerraces package (Sanderson et al. 2011). SuperTriplets supertree and BI tree based on the supermatrix were used as parent trees. Terrace identification requires binary (fully resolved) trees. The BI tree was resolved by accepting all relationships resolved in the posterior sample. The SuperTriplets supertree was resolved using relationships from the *cyb* tree. In terrace identification analysis, the parent tree was broken to subtrees, where each subtree represented relationships of taxa sampled for the respective gene as resolved in the parent tree. All relationships in the subtrees were further characterized by triplets. From these, all alternative parent trees that contain the subtrees were constructed.

Results

The dataset contained 1143 base-pairs (bp) long alignment with 68 taxa for *cyb*, CR alignment was 1025 bp long and contained 25 taxa, *coi* was 1545 bp long with 12 taxa, *nd4* was 1378 bp long with 9 taxa, *irbp* was 1181 bp long with 24 taxa, *ghr* was 911 bp long with 27 taxa, *sry* was 908 bp long with 19 taxa and *lcat* was 590 bp long with 10 taxa. The concatenated supermatrix had 72.8 % missing data composed of missing sequences of individual genes, alignment gaps and unknown nucleotides.

Substitution models selected by AIC for each gene were GTR + Γ + I for *cyb*, HKY + Γ + I for CR, GTR + Γ + I for *coi*, GTR + Γ for *nd4*, HKY + Γ for *irbp*, GTR + Γ + I for *ghr*, HKY + I for *sry* and HKY + I for *lcat*. As the GTR model requires estimation of the rate matrix, a simpler HKY model was tested for *coi*, *nd4* and *ghr* genes. The difference between log-likelihood based on selected and tested model was 10.8 for *coi*, 1.8 for *nd4* and 1.1 for *ghr* and the HKY model was used for *nd4* and *ghr* genes. The resulting trees were similar, and MCMC convergence was faster with the simpler model. The results from the simpler models are reported (Table 2, Figs. 1-2).

The 95 % credibility interval (CI) of the BI tree length based on the partitioned concatenated dataset was 7.89-15.12 with default rate parameter of the exponential branch length prior (λ = 10.0). This CI of the BI tree length did not contain the tree length 3.86 estimated from the ML analysis. Increasing the rate by increments of 10.0 to the final value of 50.0, the tree length decreased to 4.0-4.59, but we did not further test the branch length prior because of decrease in node support for high λ. Node support improved with optimisation of the branch length prior when λ = 20.0, and subsequently decreased (Fig. 3). We further analyse the BI tree with the highest average node support.

The BI supermatrix phylogeny re-rooted with midpoint root showed two initial groups (Fig. 4). The root separated genera *Arvicola* and *Chionomys* from *Microtus*, *Blanfordimys*, *Phaiomys*, *Proedromys* and *Lasiopodomys*. *Proedromys liangshanensis* was a sister species to a well-supported group containing *Lasiopodomys* and *Microtus gregalis*. Similarly, *Phaiomys leucurus* was a sister taxon to a supported group with unresolved internal relationships containing *Neodon irene*, *N. sikimensis* and *M. clarkei*. Remaining taxa formed a monophyletic group with high Bayesian posterior probability (BPP). It contained species currently attributed to genus *Microtus* not mentioned above, genus *Blanfordimys* and *N. juldaschi*. The first group that diverged at this level was the subgenus *Microtus* (*Alexandromys*) predominantly distributed in the eastern Palaearctic. *Microtus fortis* group was well differentiated from the basal species of *Alexandromys*, *M. kikuchii*, *M. montebelli* and *M. oeconomus*. We confirmed position of *M. gromovi* as a distinct taxon rather than a subspecies of *M. maximowiczii*.

Further notable group consisted of *N. juldaschi* with *Blanfordimys afghanus* and *B. bucharensis*. *M. agrestis* was a sister species to this group, but the relationship was unsupported. Nearctic species formed an unsupported group with *M. cabrerae*. Within the Nearctic group, three pairs of sister taxa had significant node support. Subgenera *Microtus* (*Terricola*) and *Microtus* (*Microtus*) were sister groups that were most derived in the BI phylogeny (Fig. 4). In the subgenus *Microtus*, *M. arvalis* group and *M. socialis/guentheri* group were separated, but classification of *M. schelkovnikovi* to the *M. socialis/guentheri* group had low support (BPP = 0.85). Subgenus *Terricola* was separated to the eastern and western clade, where the eastern clade containing *M. majori*, *M. subterraneus* and *M. daghestanicus* had BPP = 0.94.

Table 2. *Substitution models used for separate gene tree analyses and partitions of the supermatrix. Model parameters are means estimated from sampled posterior distribution of the gene trees after burn-in.* κ *– transition/transversion rate ratio,* r *– substitution rate,* f *– base frequency,* α *– shape parameter of the* Γ *distribution, I – proportion of invariable sites, n/a – not available.*

Parameter	*cyb* GTR + Γ + I	CR HKY + Γ + I	*coi* GTR + Γ + I	*nd4* HKY + Γ	*irbp* HKY + Γ	*ghr* HKY + I	*sry* HKY + I	*lcat* HKY + I
κ	n/a	3.0452	n/a	14.1872	4.0786	4.8725	3.3133	6.5095
r(A ↔ C)	0.0176	n/a	0.0236	n/a	n/a	n/a	n/a	n/a
r(A ↔ G)	0.4088	n/a	0.4862	n/a	n/a	n/a	n/a	n/a
r(A ↔ T)	0.0402	n/a	0.0515	n/a	n/a	n/a	n/a	n/a
r(C ↔ G)	0.0065	n/a	0.0070	n/a	n/a	n/a	n/a	n/a
r(C ↔ T)	0.4714	n/a	0.4152	n/a	n/a	n/a	n/a	n/a
r(G ↔ T)	0.0555	n/a	0.0165	n/a	n/a	n/a	n/a	n/a
f(A)	0.3713	0.3262	0.3006	0.3421	0.2205	0.2611	0.2909	0.2015
f(C)	0.3596	0.2586	0.2736	0.3032	0.2809	0.2877	0.2640	0.2728
f(G)	0.0774	0.1141	0.1478	0.0976	0.2909	0.2291	0.2166	0.2766
f(T)	0.1918	0.3012	0.2780	0.2572	0.2077	0.2221	0.2285	0.2491
α	0.6046	1.1420	8.9089	0.2172	0.4326	n/a	n/a	n/a
I	0.4968	0.3503	0.6647	n/a	n/a	0.7059	0.3133	0.6735

The SuperTriplets supertree agreed with the BI tree in distinguishing groups *Arvicola*, *Microtus* (*Terricola*), *Microtus* (*Microtus*), *Microtus* (*Pitymys*) with *M. guatemalensis*, *Blanfordimys* with *N. juldaschi* and a separate group including taxa from the subgenus *Microtus* (*Alexandromys*) (data available at TreeBASE). The other groups were distorted due to different position of *Chionomys nivalis*, *N. sikimensis* and *Lasiopodomys brandtii*. In the supertree, *C. nivalis* was placed as a basal taxon after diversification of *Arvicola*. *Neodon sikimensis* formed a polyphyly with *C. gud* and *C. roberti*. *Lasiopodomys brandtii* was placed within the group containing Nearctic species.
Phylogenetic terraces where the trees belonged to were small. The terrace with the BI tree used as the parent tree consisted of a single tree, and the terrace with the supertree consisted of 15 trees.

Discussion

Tree space of the concatenated dataset

We found that by optimising branch length prior of the Bayesian inference analysis of the multilocus phylogeny of arvicoline rodents with missing data comprising nearly 73 % of the concatenated supermatrix we were able to retrieve a phylogeny that is unique on a terrace. This means that there are no trees with alternative topology that would explain sets of taxa from individual gene partitions that would be present in the BI phylogeny. For the supertree approach, the terrace size was also small, and it contained 15 trees with alternative topology that explained the relationships of gene tree datasets as depicted on the supertree.

The topology of our phylogeny that well explained gene sets in the final trees was not reflected in similar congruence in branch length estimations. Our Bayesian phylogeny with highest node support was longer than the ML tree as is known to occur in partitioned datasets (Marshall et al. 2006, Brown et al. 2010, Marshall 2010). The cause of this discrepancy was recently identified to be a branch length prior that places too much probability density on large tree lengths (Rannala et al. 2012). The default on branch lengths in MrBayes assigns independent and identical exponential priors for individual branch lengths. As the default initial value for branch lengths is 0.1, the MCMC starts from very long trees for large datasets with many taxa, which is often unrealistic. The prior then places too much influence on the posterior. The effect is exacerbated for large datasets where there are partitions with low variability or correlations in rate variation or substitution models in the posterior (Rannala et al. 2012). This seems to be the case in our dataset. We optimised the branch length prior in our partitioned analyses, assuming that by setting the branch length exponential prior mean closer to mean branch length estimated from the ML tree length, the posterior tree length would be more similar to the ML tree length (Zhang et al. 2012). This is not a suitable approach in the Bayesian framework (c.f. Zhang et al. 2012). Our data showed that decreasing the tree length in this way lead to decrease of node support for high values of rate parameter of the exponential distribution of the uncorrelated branch length prior. Using compound Dirichlet priors on branch lengths in

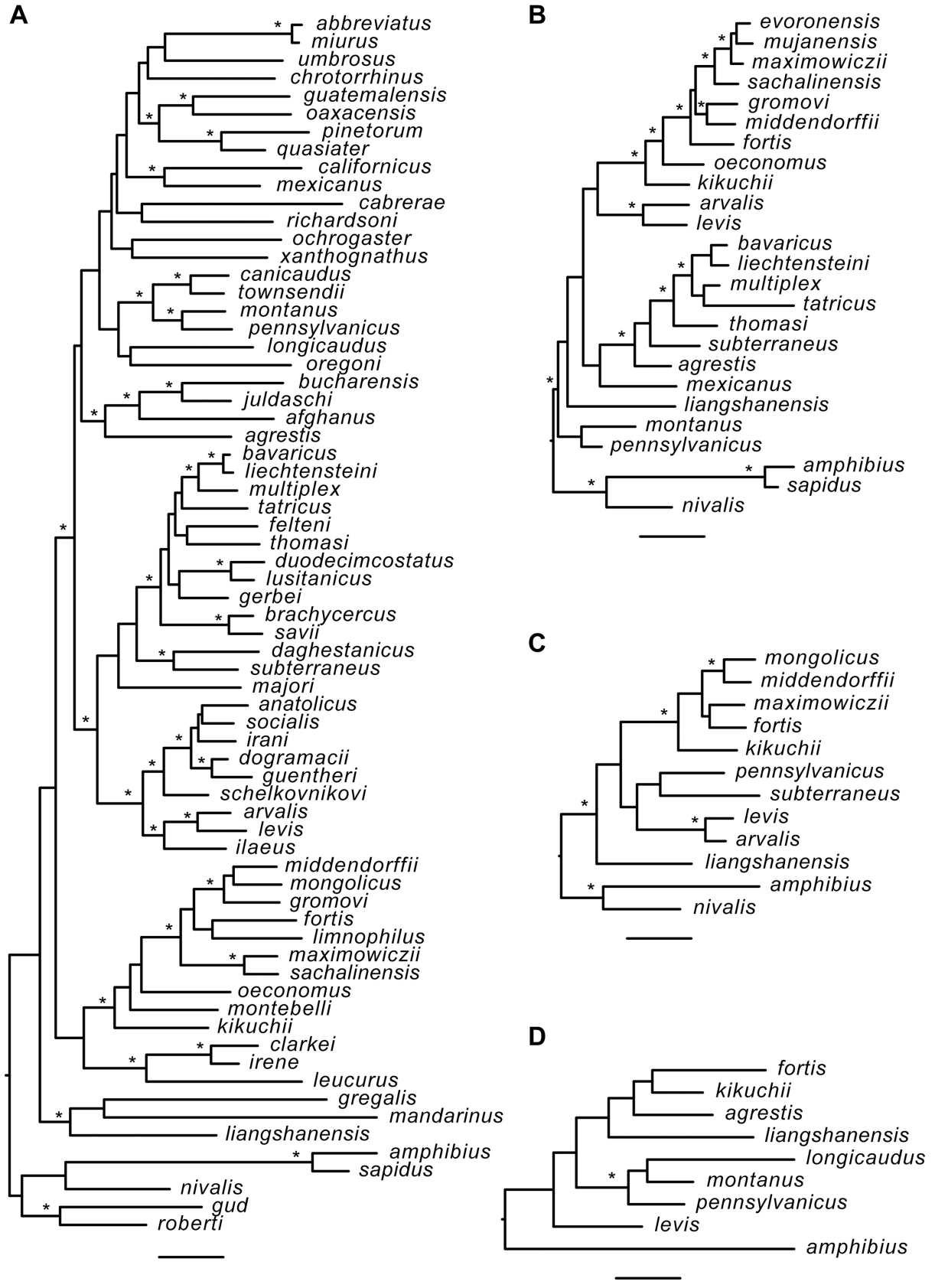

Fig. 1. *Bayesian inference phylogenetic trees based on complete mitochondrial sequences for* cyb *(A), and partial sequences for control region (B),* coi *(C) and* nd4 *(D) genes. All relationships are shown that were compatible with the consensus tree from the posterior sample of trees after burn-in. Scale bars indicate 0.1 substitutions per site, asterisk denotes nodes with Bayesian posterior probability (BPP) ≥ 0.95.*

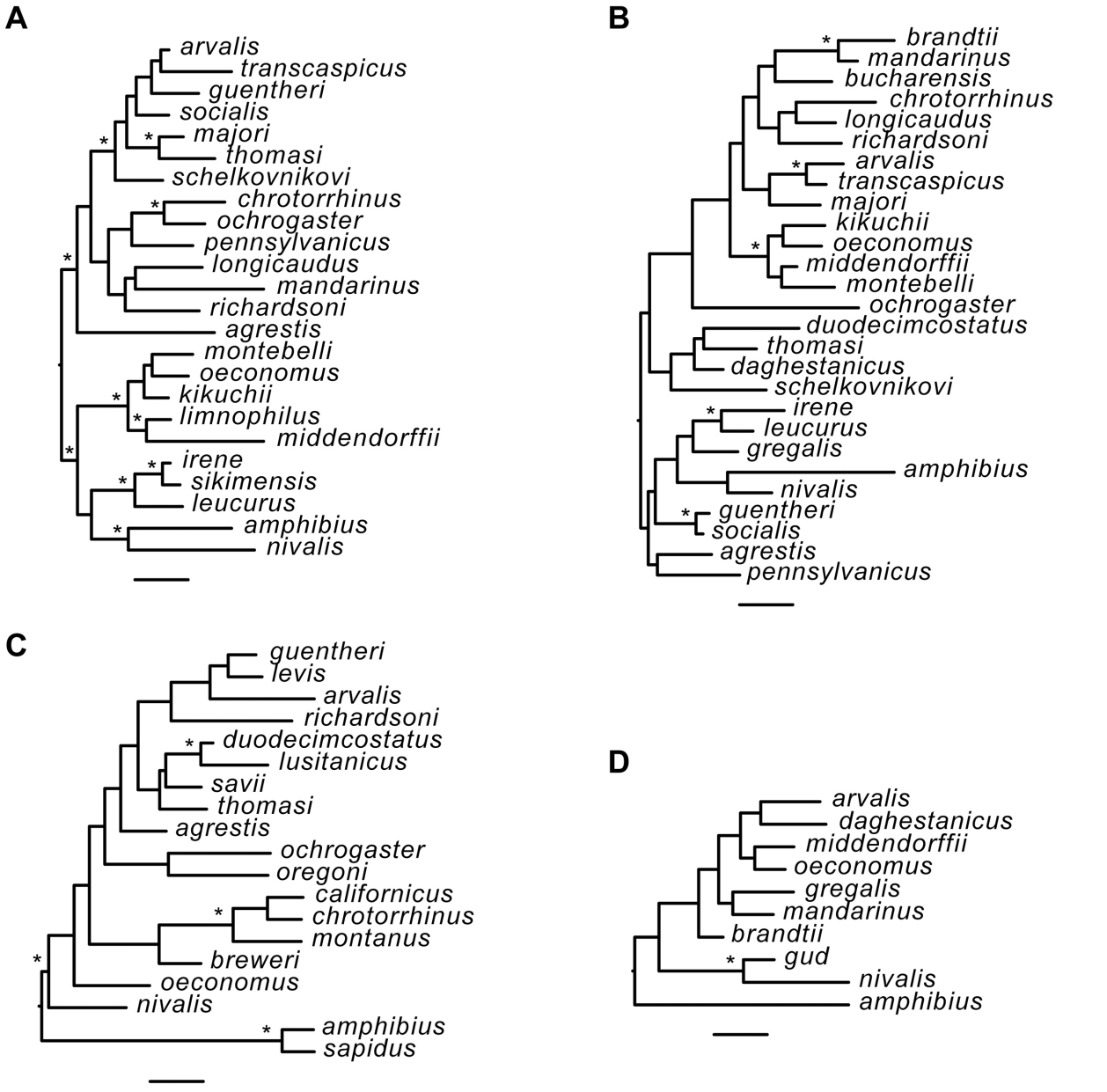

Fig. 2. *Bayesian inference phylogenetic trees based on nuclear sequences for partial* irbp *(A),* ghr *(B),* sry *(C) and* lcat *(D) genes. All relationships are shown that were compatible with the consensus tree from the posterior sample of trees after burn-in. Scale bars indicate 0.01 substitutions per site, asterisk denotes nodes with BPP ≥ 0.95.*

modified MrBayes 3.1 (Zhang et al. 2012), we were not able to obtain a tree with CI of tree length that would include the ML estimate and the average BPP remained lower than in our optimal tree.

Arvicoline phylogeny

Phylogenetic position of *Arvicola* within subfamily Arvicolinae is unstable in studies utilising mtDNA (Bužan et al. 2008, Bannikova et al. 2009) in comparison with studies that use nucDNA (Galewski et al. 2006, Abramson et al. 2009). We also observed this in gene trees. Our supermatrix results show that by rooting the tree of the Arvicolini tribe *sensu* Galewski et al. (2006) with midpoint root, *Arvicola* forms a supported group with *Chionomys* at the base of the tree.

Microtus gregalis represented a phylogenetic enigma. In early molecular phylogenies, it was placed distantly from other supposedly related species at the base of the phylogeny of *Microtus*, but its basal position was unsupported (Conroy & Cook 2000b, Conroy et al. 2001, Jaarola et al. 2004). It was later retrieved as a sister taxon to *Chionomys* based on mtDNA (Bužan &

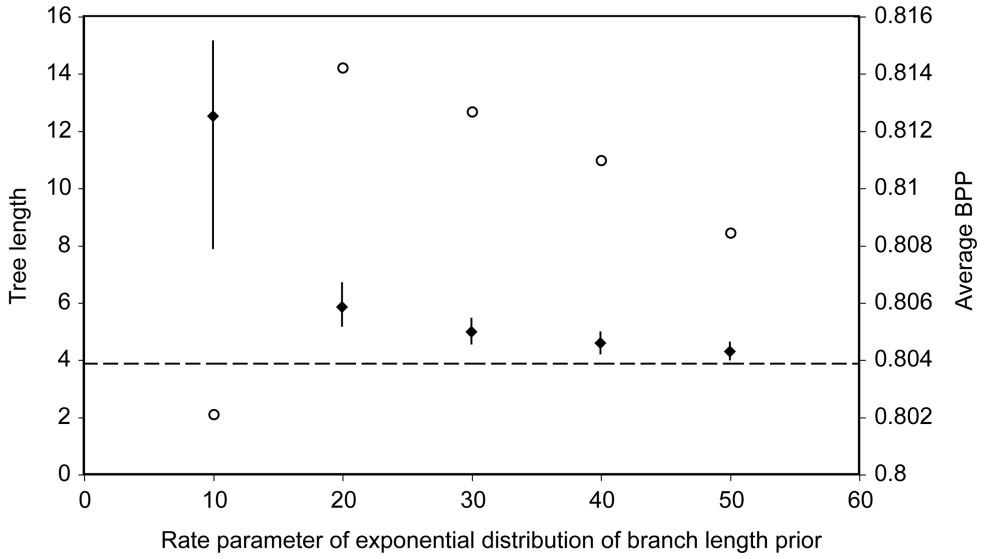

Fig. 3. *Changes of the tree length (black diamonds, primary axis) and average Bayesian posterior probability of node support (empty circles, secondary axis) with increasing shape parameter of the exponential distribution of uncorrelated branch lengths prior. Dashed line indicates tree length estimated from maximum likelihood analysis with GTR + Γ substitution model for each partition. Tree length is given with 95 % credibility interval.*

Kryštufek 2008), but nucDNA grouped *M. gregalis* with *Lasiopodomys* (Abramson et al. 2009). This grouping elucidates towards rapid karyotype rearrangements between species, as *M. gregalis* has 36 chromosomes (Martínková et al. 2004), whereas *L. mandarinus* chromosome number varies between 47 and 52 (Liu et al. 2010). If the ancestral karyotype of the group was $2n = 54$, the rearrangements leading to *M. gregalis* that branches close to the base of the tree were extensive (c.f. Lemskaya et al. 2010). The phylogenetic position of *M. gregalis* in the vicinity of *Lasiopodomys* was confirmed once *Lasiopodomys* was sequenced for the mtDNA (Bannikova et al. 2010). Our combined phylogeny placed *M. gregalis* close to the base of the trees in a well-supported group with *L. mandarinus* and *L. brandtii* in accordance with recent studies (Abramson et al. 2009, Bannikova et al. 2010). The sister taxon of the group is *Proedromys liangshanensis* that was described recently (Liu et al. 2007). Based on phylogeny of complete mtDNA, *Pr. liangshanensis* is a sister species to *Microtus* (Hao et al. 2011). Our analysis included more comprehensive sampling, and we found that the species forms a supported sister relationship with *Lasiopodomys* and *M. gregalis*.

The genus *Neodon* was polyphyletic with *N. juldaschi* grouping with *Blanfordimys* deeper in the phylogeny than other species attributed to *Neodon*, *N. irene* and *N. sikimensis*. The latter two species consistently belonged to the *Phaiomys*/*Neodon* lineage (Galewski et al. 2006, Robovský et al. 2008, Bannikova et al. 2009, 2010) that included also *M. clarkei* in our analyses. Its relationship with *Neodon* was not resolved, forming a strongly supported trichotomy, where BPP for the monophyly of *Neodon* within this lineage was as low as 0.41.

The phylogenetic position of *M. agrestis* was unstable in the gene trees, and it was placed as an unsupported sister taxon to the *N. juldaschi*/*Blanfordimys* group in our multilocus phylogeny. *Microtus agrestis* split from the common ancestors of *Microtus* early in the radiation of the genus, but its closest relatives might not be identifiable today similarly as in the case of *M. cabrerae*. Interestingly, *M. agrestis* from the Iberian peninsula, where *M. cabrerae* is also distributed, forms a phylogenetic lineage distinct from other *M. agrestis* populations. This divergence is present both in mitochondrial and nuclear phylogenies and might represent a cryptic species that was not formally described to date (Jaarola & Searle 2004, Hellborg et al. 2005). The erratic placement of *M. agrestis* in different gene trees and unsupported position of *M. cabrerae* with North American *Microtus* indicates that these species represent relicts of a very early colonization of Arvicolini to Western Europe. Phylogenies of Arvicolidae improve with more comprehensive sampling (Bužan et al. 2008), and based on the fact that we analysed majority of species in the tribe Arvicolini, we are confident to state that the close relatives of *M. agrestis* and *M. cabrerae* are extinct today and their phylogenetic position is influenced more by stochastic processes in DNA sequence evolution such as saturation or

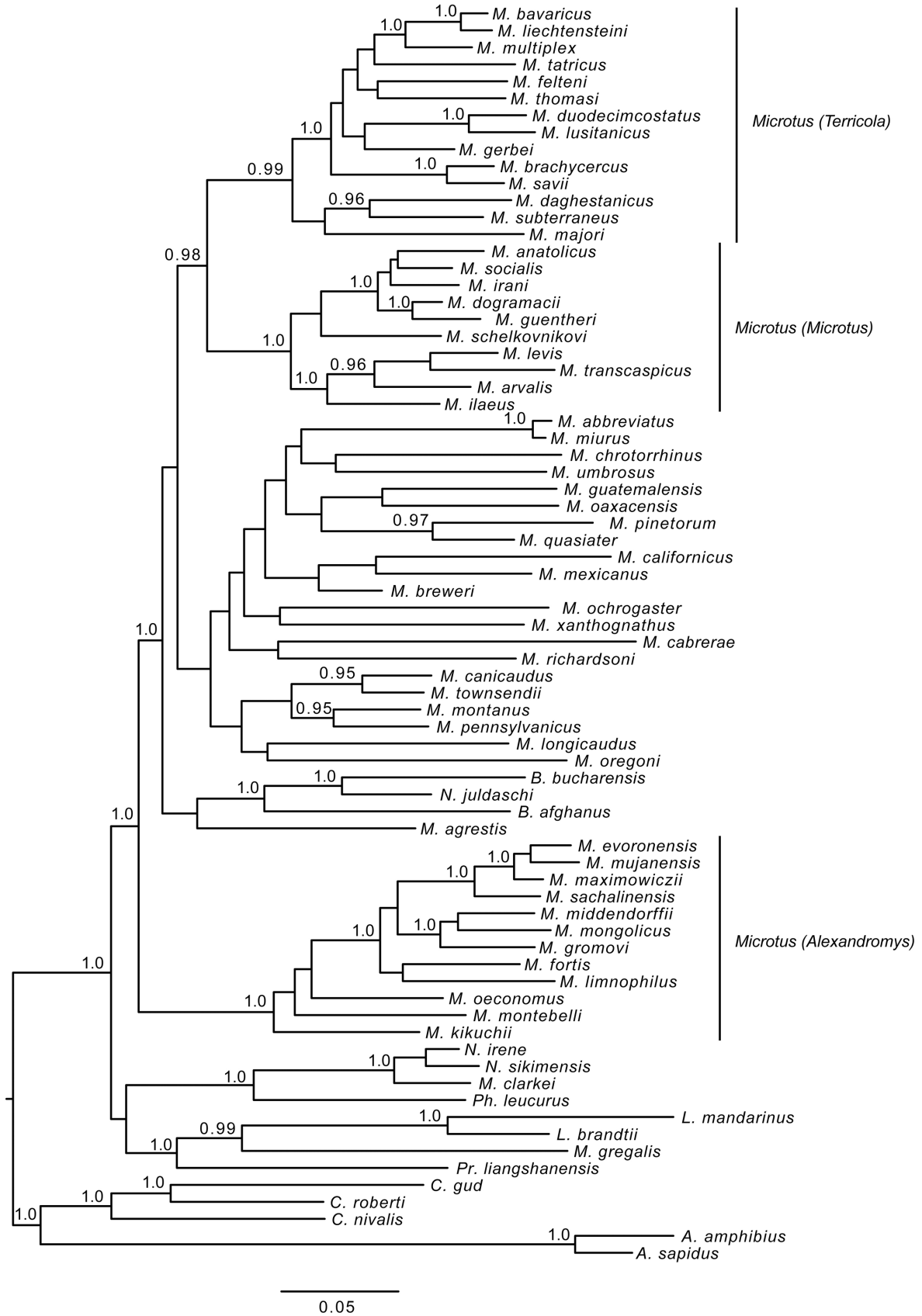

Fig. 4. *Bayesian inference phylogenetic tree based on concatenated sequence of the genes depicted in Figs. 1–2. All relationships are shown that were compatible with the consensus tree from the posterior sample of trees after burn-in. Scale bar is in substitutions per site, BPP ≥ 0.95 is shown.*

convergence and by computational artefacts in phylogeny reconstruction.

Within *Microtus*, we retrieved three groups that represent subgenera recognised by Wilson & Reeder (2005) with minor changes. Subgenus *Alexandromys* was supported without *M. clarkei* as per Wilson & Reeder (2005), *Microtus* without *M. cabrerae* and *Terricola* with *M. tatricus*. The species within subgenera showed relationships established in previous studies (Jaarola et al. 2004, Martínková et al. 2007, Kryštufek et al. 2009, Bannikova et al. 2010, Haring et al. 2000, 2011).

Nearctic *Microtus* consistently suffer from lack of resolution of many relationships (Conroy & Cook 2000b, Conroy et al. 2001, Jaarola et al. 2004, Fink et al. 2010), although recent studies show that these taxa are also strongly influenced by rapid diversification in speciation traps. Weksler et al. (2010) found *M. abbreviatus* from Wrangell Mts. to be divergent and potentially merit species status, and Conroy & Neuwald (2008) distinguished two species within *M. californicus*. In our multilocus study, phylogenetic position of *M. breweri* was particularly unstable. This was probably in lieu of the fact that only the *sry* gene sequence was available for this species.

Species with small ranges were often part of rapidly differentiating groups. This leads to an assumption that geographic isolation in small refugia triggers diversification on both molecular and morphological levels. In *Microtus*, the results of phylogeography couple with species phylogenies where phylogeography nowadays indicates regions and populations that might give rise to new species in the future.

Acknowledgements

We appreciate M. Macholán's invitation to participate in the issue of Folia Zoologica in tribute to Jan Zima's anniversary. N. Martínková is thankful to J. Zima for his guidance and support at the beginning of her career. The bioinformatic analyses were conducted at the computational cluster of the Institute of Vertebrate Biology and at the MetaCentrum. The access to the MetaCentrum computing facilities was provided under the programme "Projects of Large Infrastructure for Research, Development, and Innovations" LM2010005 funded by the Ministry of Education, Youth, and Sports of the Czech Republic. This study was supported by the Grant Agency of the Academy of Sciences of the Czech Republic, grant no. IAA600930609 and with institutional support RVO: 68081766.

Literature

Abramson N.I., Lebedev V.S., Tesakov A.S. & Bannikova A.A. 2009: Supraspecies relationships in the subfamily Arvicolinae (Rodentia, Cricetidae): an unexpected result of nuclear gene analysis. *Mol. Biol. 43: 834–846. (in Russian)*

Acosta M.J., Marchal J.A., Romero-Fernández I., Megías-Nogales B., Modi W.S. & Sánchez Baca A. 2010: Sequence analysis and mapping of the *sry* gene in species of the subfamily Arvicolinae (Rodentia). *Sex. Dev. 4: 336–347.*

Bannikova A.A., Lebedev V.S. & Golenishchev F.N. 2009: Taxonomic position of Afghan vole (subgenus *Blanfordimys*) by the sequence of the mitochondrial *cytb* gene. *Russ. J. Genet. 45: 91–97.*

Bannikova A.A., Lebedev V.S., Lissovsky A.A., Matrosova V., Abramson N.I., Obolenskaya E.V. & Tesakov A.S. 2010: Molecular phylogeny and evolution of the Asian lineage of vole genus *Microtus* (Rodentia: Arvicolinae) inferred from mitochondrial cytochrome *b* sequence. *Biol. J. Linn. Soc. 99: 595–613.*

Brown J.M., Hedtke S.M., Lemmon A.R. & Lemmon E.M. 2010: When trees grow too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. *Syst. Biol. 59: 145–161.*

Brunhoff C., Yoccoz N.G., Ims R.A. & Jaarola M. 2006: Glacial survival or late glacial colonization? Phylogeography of the root vole (*Microtus oeconomus*) in north-west Norway. *J. Biogeogr. 33: 2136–2144.*

Bryja J., Charbonnel N., Berthier K., Galan M. & Cosson J.-F. 2007: Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of voles. *Mol. Ecol. 16: 5084–5097.*

Bužan E.V. & Kryštufek B. 2008: Phylogenetic position of *Chionomys gud* assessed from a complete cytochrome *b* gene. *Folia Zool. 57: 274–282.*

Bužan E.V., Kryštufek B., Hänfling B. & Hutchinson W.F. 2008: Mitochondrial phylogeny of Arvicolinae using comprehensive taxonomic sampling yields new insights. *Biol. J. Linn. Soc. 94: 825–835.*

Castiglia R., Annesi F., Aloise G. & Amori G. 2008: Systematics of the *Microtus savii* complex (Rodentia, Cricetidae) via mitochondrial DNA analyses: paraphyly and pattern of sex chromosome evolution. *Mol. Phylogenet. Evol. 46: 1157–1164.*

Conroy C.J & Cook J.A. 2000a: Phylogeography of a post-glacial colonizer: *Microtus longicaudus* (Rodentia: Muridae). *Mol. Ecol. 9: 165–175.*

Conroy C.J. & Cook J.A. 2000b: Molecular systematics of a holarctic rodent (*Microtus*: Muridae). *J. Mammal. 81: 344–359.*

Conroy C.J., Horteland Y., Cervantes F.A. & Cook J.A. 2001: The phylogenetic position of southern relictual species of *Microtus* (Muridae: Rodentia) in North America. *Mamm. Biol. 66: 332–344.*

Conroy C.J. & Neuwald J.L. 2008: Phylogeographic study of the California vole, *Microtus californicus*. *J. Mammal. 89: 755–767.*

Drummond A.J., Ashton B., Buxton S., Cheung M., Cooper A., Duran C., Field M., Heled J., Kearse M., Markowitz S., Moir R., Stones-Havas S., Sturrock S., Thierer T. & Wilson A. 2011: Geneious v5.4. *Available from http://www.geneious.com/*

Fan Z., Liu S., Liu Y., Zhang X. & Yue B. 2011: How Quaternary geologic and climatic events in the southeastern margin of the Tibetan Plateau influence the genetic structure of small mammals: inferences from phylogeography of two rodents, *Neodon irene* and *Apodemus latronum*. *Genetica 139: 339–351.*

Fink S., Excoffier L. & Heckel G. 2004: Mitochondrial gene diversity in the common vole *Microtus arvalis* shaped by historical divergence and local adaptations. *Mol. Ecol. 13: 3501–3514.*

Fink S., Excoffier L. & Heckel G. 2007: High variability and non-neutral evolution of the mammalian *avpr1a* gene. *BMC Evol. Biol. 7, 176.*

Fink S., Fischer M.C., Excoffier L. & Heckel G. 2010: Genomic scans support repetitive continental colonization events during the rapid radiation of voles (Rodentia: *Microtus*): the utility of AFLPs versus mitochondrial and nuclear sequence markers. *Syst. Biol. 59: 548–572.*

Fraguedakis-Tsolis S.E., Chondropoulos B.P., Stamatopoulos C.V. & Giokas S. 2009: Morphological variation of the five vole species of the genus *Microtus* (Mammalia, Rodentia, Arvicolinae) occurring in Greece. *Acta Zool. 90: 254–264.*

Galewski T., Tilak M., Sanchez S., Chevret P., Paradis E. & Douzery E.J.P. 2006: The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies. *BMC Evol. Biol. 6, 80.*

Hao H., Liu S., Zhang X., Chen W., Song Z., Peng H., Liu Y. & Yue B. 2011: Complete mitochondrial genome of a new vole *Proedromys liangshanensis* (Rodentia: Cricetidae) and phylogenetic analysis with related species: are there implications for the validity of the genus *Proedromys*? *Mitochondr. DNA 22: 28–34.*

Haring E., Herzig-Straschil B. & Spitzenberger F. 2000: Phylogenetic analysis of Alpine voles of the *Microtus multiplex* complex using the mitochondrial control region. *J. Zool. Syst. Evol. Res. 38: 231–238.*

Haring E., Sheremetyeva I.N. & Kryukov A.P. 2011: Phylogeny of Palearctic vole species (genus *Microtus*, Rodentia) based on mitochondrial sequences. *Mamm. Biol. 76: 258–267.*

Hellborg L., Gündüz İ. & Jaarola M. 2005: Analysis of sex-linked sequences supports a new mammal species in Europe. *Mol. Ecol. 14: 2025–2031.*

Jaarola M., Martínková N., Gündüz İ., Brunhoff C., Zima J., Nadachowski A., Amori G., Bulatova N.S., Chondropoulos B., Fraguedakis-Tsolis S., González-Esteban J., López-Fuster M.J., Kandaurov A.S., Kefelioğlu H., da Luz Mathias M., Villate I. & Searle J.B. 2004: Molecular phylogeny of the speciose vole genus *Microtus* (Arvicolinae, Rodentia) inferred from mitochondrial DNA sequences. *Mol. Phylogenet. Evol. 33: 647–663.*

Jaarola M. & Searle J.B. 2002: Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Mol. Ecol. 11: 2613–2621.*

Jaarola M. & Searle J.B. 2004: A highly divergent mitochondrial DNA lineage of *Microtus agrestis* in southern Europe. *Heredity 92: 228–234.*

Jurdíková N., Žiak D. & Kocian Ľ. 2000: Habitat requirements of *Microtus tatricus*: macrohabitat and microhabitat. In: Urban P. (ed.), Výskum a ochrana cicavcov na Slovensku 4. *ŠOP COPK, Banská Bystrica.*

Kefelioğlu H. & Kryštufek B. 1999: The taxonomy of *Microtus socialis* group (Rodentia: Microtinae) in Turkey, with the description of a new species. *J. Nat. Hist. 33: 289–303.*

Kryštufek B., Bužan E.V., Vohralík V., Zareie R. & Özkan B. 2009: Mitochondrial cytochrome *b* sequence yields new insight into the speciation of social voles in south-west Asia. *Biol. J. Linn. Soc. 98: 121–128.*

Lemskaya N., Romanenko S.A., Golenishchev F.N., Rubtsova N.V., Sablina O.V., Serdukova N.A., O'Brian P.C.M., Fu B., Yiğit N., Ferguson-Smith M.A., Yang F. & Graphodatsky A.S. 2010: Chromosomal evolution

of Arvicolinae (Cricetidae, Rodentia). III. Karyotype relationships of ten *Microtus* species. *Chromosome Res. 18: 459–471.*
Liu S., Sun Z., Zeng Z. & Zhao E. 2007: A new vole (Cricetidae: Arvicolinae: *Proedromys*) from the Liangshan Mountains of Sichuan Province, China. *J. Mammal. 88: 1170–1178.*
Liu H., Yan N. & Zhu B. 2010: Two new karyotypes and bandings in *Microtus mandarinus faeceus* (Rodentia). *Hereditas 147: 123–126.*
Macholán M., Filippucci M.G. & Zima J. 2001: Genetic variation and zoogeography of pine voles of the *Microtus subterraneus/majori* group in Europe and Asia Minor. *J. Zool. 255: 31–42.*
Martínková N. & Dudich A. 2003: The fragmented distribution range of *Microtus tatricus* and its evolutionary implications. *Folia Zool. 52: 11–22.*
Martínková N., Nová P., Sablina O.V., Graphodatsky A.S. & Zima J. 2004: Karyotypic relationships of the Tatra vole (*Microtus tatricus*). *Folia Zool. 53: 279–284.*
Martínková N., Zima J., Jaarola M., Macholán M. & Spitzenberger F. 2007: The origin and phylogenetic relationships of *Microtus bavaricus* based on karyotype and mitochondrial DNA sequences. *Folia Zool. 56: 39–49.*
Marshall D.C. 2010: Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the land of long trees. *Syst. Biol. 59: 108–117.*
Marshall D.C., Simon C. & Buckley T.R. 2006: Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. *Syst. Biol. 55: 993–1003.*
Mazurok N.A., Rubtsova N.V., Isaenko A.A., Pavlova M.E., Slobodyanyuk S.Y., Nesterova T.B. & Zakian S.M. 2001: Comparative chromosome and mitochondrial DNA analyses and phylogenetic relationships within common voles (*Microtus*, Arvicolidae). *Chromosome Res. 9: 107–120.*
Mekada K., Koyasu K., Harada M., Narita Y., Shrestha K.C. & Oda S.-I. 2002: Karyotype and X-Y chromosome pairing in the Sikkim vole (*Microtus* (*Neodon*) *sikimensis*). *J. Zool. 257: 417–423.*
Mitsainas G.P., Rovatsos M.T. & Giagia-Athanasopoulou E.B. 2010: Heterochromatin study and geographical distribution of *Microtus* species (Rodentia, Arvicolinae) from Greece. *Mamm. Biol. 75: 261–269.*
Nylander J.A.A. 2004: MrModeltest v2. Program distributed by the author. *Evolutionary Biology Centre, Uppsala University.*
Oliver M.K., Lambin X., Cornulier T. & Piertney S.B. 2009: Spatio-temporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (*Arvicola terrestris*) metapopulations. *Mol. Ecol. 18: 80–92.*
Rannala B., Zhu T. & Yang Z. 2012: Tail paradox, partial identifiability, and influential priors in Bayesian branch length inference. *Mol. Biol. Evol. 29: 325–335.*
Ranwez V., Criscuolo A. & Douzery E.J.P. 2010: SuperTriplets: a triplet-based supertree approach to phylogenomics. *Bioinformatics 26: i115–i123.*
Robovský J., Řičánková V. & Zrzavý J. 2008: Phylogeny of Arvicolinae (Mammalia, Cricetidae): utility of morphological and molecular data sets in a recently radiating clade. *Zool. Scr. 37: 571–590.*
Ronquist F. & Huelsenbeck J.P. 2003: MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics 19: 1572–1574.*
Rudá M., Žiak D., Kocian Ľ. & Martínková N. 2010: Low genetic variability in a mountain rodent, the Tatra vole. *J. Zool. 281: 118–124.*
Sanderson M.J., McMahon M.M. & Steel M. 2011: Terraces in phylogenetic tree space. *Science 333: 448–450.*
Santos S.M., da Luz Mathias M. & Mira A.P. 2011: The influence of local, landscape and spatial factors on the distribution of the Lusitanian and the Mediterranean pine voles in a Mediterranean landscape. *Mamm. Biol. 76: 133–142.*
Sitnikova N.A., Romanenko S.A., O'Brien P.C.M., Perelman P.L., Fu B., Rubtsova N.V., Serdukova N.A., Golenishchev F.N., Trifonov V.A., Ferguson-Smith M.A., Yang F. & Graphodatsky A.S. 2007: Chromosomal evolution of Arvicolinae (Cricetidae, Rodentia). I. The genome homology of tundra vole, field vole, mouse and golden hamster revealed by comparative chromosome painting. *Chromosome Res. 15: 447–456.*
Spaeth P.A., van Tuinen M., Chan Y.L., Terca C. & Hadly E.A. 2009: Phylogeography of *Microtus longicaudus* in the tectonically and glacially dynamic Central Rocky Mountains. *J. Mamm. 90: 571–584.*

Stamatakis A. 2006: RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics 22: 2688–2690.*

Tougard C., Brunet-Lecomte P., Fabre M. & Montuire S. 2008: Evolutionary history of two allopatric *Terricola* species (Arvicolinae, Rodentia) from molecular, morphological, and palaeontological data. *Biol. J. Linn. Soc. 93: 309–323.*

Triant D.A. & Dewoody J.A. 2007: Extensive mitochondrial DNA transfer in a rapidly evolving rodent has been mediated by independent insertion events and by duplications. *Gene 401: 61–70.*

Triant D.A. & Dewoody J.A. 2008: Molecular analyses of mitochondrial pseudogenes within the nuclear genome of arvicoline rodents. *Genetica 132: 21–33.*

Weksler M., Lanier H.C & Olson L.E. 2010: Eastern Beringian biogeography: historical and spatial genetic structure of singing voles in Alaska. *J. Biogeogr. 37: 1414–1431.*

Wilson D.E. & Reeder D.M. 2005: Mammal species of the world. A taxonomic and geographic reference (3rd ed.). *John Hopkins University Press, Baltimore.*

Zhang C., Rannala B. & Yang Z. 2012: Robustness of compound Dirichlet priors for Bayesian inference of branch lengths. *Syst. Biol. 61: 779–784.*