

## 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

SABINE FINK<sup>1</sup>, MARTIN C. FISCHER<sup>1,2</sup>, LAURENT EXCOFFIER<sup>1,2</sup>, AND GERALD HECKEL<sup>1,2,\*</sup>

<sup>1</sup>Department of Biology, Computational and Molecular Population Genetics (CMPG), Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland; and

<sup>2</sup>Swiss Institute of Bioinformatics, Genopode, 1015 Lausanne, Switzerland;

\*Correspondence to be sent to: Department of Biology, Computational and Molecular Population Genetics (CMPG), Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland; E-mail: gerald.heckel@iee.unibe.ch.

Sabine Fink and Martin C. Fischer have contributed equally to this work.

Received 13 May 2008; reviews returned 9 October 2008; accepted 15 March 2010

Associate Editor: Brett Riddle

**Abstract.**—Single locus studies might not resolve phylogenetic relationships and the evolutionary history of taxa. The analysis of multiple markers promises higher resolution, and congruence among loci may indicate that the phylogenies represent the underlying species history. Here, we examine the utility of a genome-wide approach based on amplified fragment length polymorphisms (AFLP) and several DNA sequence markers in resolving phylogenetic signals in the rapidly radiating rodent genus *Microtus* which produced about 70 vole species within the last 1.2–2 myr. The current Holarctic distribution of *Microtus* is assumed to have resulted from three independent colonization events out of Asia to North America, Europe, and northern Asia without subsequent colonization, which would have led to deep splits between species from different continents. We investigated this hypothesis of three single colonization events by reconstructing the phylogenetic relationships among species from all three continents based on data from the first exon of the nuclear arginine vasopressin receptor 1a gene (EXON1), an adjacent noncoding region and the mitochondrial cytochrome *b* gene. The phylogenetic patterns obtained from these sequence markers are contrasted to genome-wide data on more than 1800 amplified fragment length polymorphisms (AFLP) analyzed for the same samples. Our results show that the single sequence markers partially resolve the phylogenetic relationships within *Microtus*, but with some incongruence mostly between EXON1 and the other loci. However, deeper nodes of the radiation are only weakly supported and neither the combination of the markers nor additional nuclear sequences improved the resolution significantly. AFLPs provided much stronger support for major continent-specific clades, and show also that reciprocal monophyly of American and European voles is incomplete. Our results demonstrate that *Microtus* voles colonized the American and European continents each repeatedly in several independent events on similar colonization routes during their radiation. More generally, this study supports the suitability of AFLPs as an alternative to sequence markers to resolve the evolutionary history of rapidly radiating taxa. [AFLP; arginine vasopressin receptor 1a; Arvicolinae; colonization; cytochrome *b*; phylogeny; rodents; voles.]

Phylogenetic estimations based on single-gene analyses might fail to reflect the evolutionary history and the systematic relationship between species and higher taxa. Gene trees can show incongruent topologies compared with species trees mainly because of locus-specific effects such as the influence of selection, recombination, and incomplete lineage sorting (DeSalle and Val Giddings 1986; Seehausen 2004). However, when results among gene phylogenies are congruent, they are likely to represent the underlying species tree (Nichols 2001). Therefore, comparative analyses of several loci and marker types are becoming more abundant for deep-level phylogenies (Reyes et al. 2004; Springer et al. 2004), and also for inferring evolutionary history within a family or genus (e.g. Giannini and Simmons 2003; Jordan et al. 2003; Suzuki et al. 2004; Jansa et al. 2006). Moreover, multiple-gene phylogenies might also help to resolve shallow phylogenetic relationships and associated systematic uncertainties about closely related species (Giannasi et al. 2001).

When different loci reflect different modes of evolution, phylogenetic signals might not lead to a congruent pattern (DeSalle and Val Giddings 1986; Seehausen 2004). Incongruent patterns may be caused for instance by differences between mitochondrial loci with smaller

effective population size leading to the rapid sorting of lineages versus nuclear loci with larger effective population sizes potentially leading to incomplete lineage sorting. Furthermore, combinations of a smaller number of nuclear loci might still not reflect the true evolutionary history of the genome, given the potential influence of selection and recombination on some loci.

Whole-genome sequencing might circumvent some of these problems and may soon become feasible for non-model organisms, but high costs and effort currently prevent its wider usage. AFLP markers are one of the few methods available that assess variation across a very large number of loci in nonmodel organisms (Bensch and Akesson 2005). Although AFLPs are seldom used in systematic studies (Reyes et al. 2004; Springer et al. 2004), they could indeed prove to be the marker of choice for many purposes. Because AFLPs are normally highly variable, they may provide sufficient resolution to reconstruct evolutionarily young splits among closely related species (Althoff et al. 2007). Although AFLP experiments have to be designed cautiously to avoid the screening of nonhomologous fragments (Althoff et al. 2007; McKinnon et al. 2008), their genome-wide distribution minimizes the locus-specific effects that could be problematic for analyses of single or just a few loci

(Buntjer et al. 2002; Ogden and Thorpe 2002; Creer et al. 2004; Kingston and Rosel 2004; Sullivan et al. 2004).

Here we use a combination of AFLP markers and sequence information from mitochondrial DNA (mtDNA) and nuclear DNA (nucDNA) to investigate the phylogenetic relationships and biogeographic history in the rodent genus *Microtus* (Cricetidae: Arvicolinae: voles) that has undergone explosive waves of speciation and colonization, representing the fastest known radiation in extant mammals (Griffiths et al. 2000). The *Microtus* genus forms an excellent model system to study the utility of AFLP genome scans for resolving phylogenetic signal in a rapidly evolving genus with many extant species for which a biogeographic hypothesis has been formulated but could not be fully evaluated with classical phylogenetic markers (Jaarola et al. 2004). The direct comparison of the genomic approach with various single locus analyses helps to elucidate caveats of different methods as well as their potential to reflect the species tree.

#### Taxonomic and Systematic Background

Many systematic uncertainties remain for the *Microtus* genus, and the number of extant species recognized within the genus varies extensively among taxonomists (Nowak 1991; Musser and Carleton 1993). A fossil record is missing for most extant species, or it appears relatively late (Tamarin 1985), but microtine species are thought to have originated 1.2–2 Ma from the fossil genus *Allophaiomys* (Musser and Carleton 1993; Chaline et al. 1999). Starting from its putative origin in southern Asia, its evolutionary history was postulated to have included three major independent colonization events leading to its current distribution (Fig. 1): one colonization wave from southern Asia to northern Asia, another to Europe, and a third over

the Beringian land bridge to North America (Brunet-Lecomte and Chaline 1991; Chaline et al. 1999). This colonization scenario is expected to lead to deep divergence among Asian, European, and American species of the genus, but past molecular systematic studies using the mitochondrial cytochrome *b* (*cytb*) gene have provided no substantial support for these major splits (Jaarola et al. 2004). Previous studies have grouped European species according to the previously described subgenera *Terricola* and *Microtus*, with *Agricola* as the basal clade, which in turn is polyphyletic (*Microtus cabrerae* clusters with American species, whereas *M. agrestis* is more closely related to European species). Given that inter- and intraspecies distances overlap for *cytb* (Jaarola et al. 2004), this marker might not be ideal to resolve the species tree.

The approximately 70 *Microtus* species worldwide are morphologically very similar (Niethammer and Krapp 1982; Nowak 1991; Musser and Carleton 1993; Mitchell-Jones et al. 1999), unlike other prominent examples for rapid radiations at smaller geographic scales (e.g., cichlids, Albertson et al. 1999; Danley and Kocher 2001; Galapagos finches, Sato et al. 2001). Nevertheless, morphological characteristics have been used to describe species groups and other subgeneric relationships in *Microtus* (Nowak 1991; Musser and Carleton 1993), but it is still debated whether some subgenera should be recognized as separate genera (Niethammer and Krapp 1982; Mitchell-Jones et al. 1999). Geographically separated karyotypic forms are frequently found within species (Niethammer and Krapp 1982; Modi 1987; Mitchell-Jones et al. 1999), as are distinct evolutionary lineages of mtDNA (e.g., *M. longicaudus*, Conroy and Cook 2000b; *M. oeconomus*, Brunhoff et al. 2003; Galbreath and Cook 2004; *M. agrestis*, Jaarola and Searle 2004; *M. arvalis*, Fink et al. 2004; Schmidt-Chanasit et al. 2010). These genetic subdivisions could indicate either

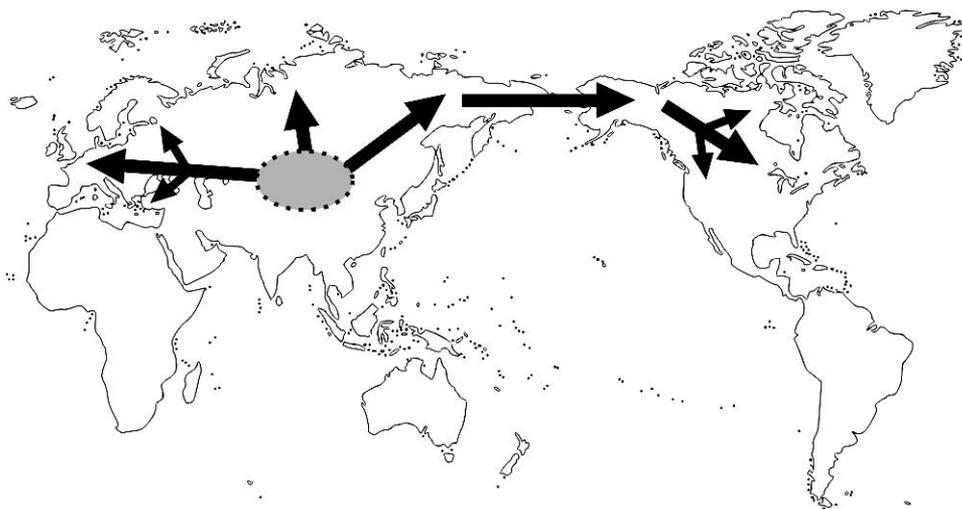


FIGURE 1. Classical colonization hypothesis for the *Microtus* genus based on fossil data (Brunet-Lecomte and Chaline 1991; Chaline et al. 1999). The putative origin in South Asia is represented as a grey oval. Colonization to Europe, North Asia, and North America would have occurred in three independent events indicated with arrows.

ongoing speciation processes (e.g. Hellborg et al. 2005) or currently unrecognized "cryptic species" (see Jaarola and Searle 2004).

Previous analysis of a few *Microtus* species among the subfamily Arvicolinae (Galewski et al. 2006) indicated suitability of nuclear sequence information for phylogenetic resolution within the genus, but noted the necessity of more nucDNA to obtain a supported species tree. In this study, we analyze the first exon (EXON1) of the arginine vasopressin 1a receptor (*avpr1a*) gene, which has been shown to be highly variable at the species level in *Microtus* (Fink et al. 2007). Additionally, we sequenced the flanking noncoding upstream region (UPSTREAM) of EXON1, and we compared the nuclear sequence phylogenies with analyses of the *cytb* gene. The results from DNA sequences were further examined against genome-wide data from AFLP markers in order to evaluate the efficacy of each approach for reconstructing phylogenetic and biogeographic histories.

## MATERIAL AND METHODS

### Samples

We selected species of *Microtus* representing the entire Holarctic distribution of the genus and analyzed multiple individuals per species whenever possible in order to assess intraspecific variation. We analyzed a total of 62 individuals from 24 *Microtus* species across three continents (Europe, North America, Asia; Table 1). One individual per species was analyzed for all three sequence markers and the AFLPs, whereas up to three additional individuals per marker and species were added whenever available to account for intraspecific variability, especially in species with known highly divergent intraspecific lineages (e.g., *M. arvalis*, Fink et al. 2004; *M. agrestis*, Jaarola and Searle 2004). We also obtained molecular data from an individual of *Arvicola terrestris* to use as outgroup—this species belongs to a sister genus of *Microtus*, with 18% sequence divergence for *cytb* (Fink et al. 2006; Galewski et al. 2006). Tissue samples, stored in EtOH absolute or deep frozen, were either obtained by live-trapping voles with Longworth small mammal traps (Penlon Ltd), from ecologists studying a particular species, or from museum collections. Total genomic DNA was extracted using Magnetic beads (MagneSil™ BLUE; Promega) or a standard phenol-chloroform protocol (Sambrook et al. 1989). Additional sequence information was obtained from GenBank for the *avpr1a* gene from two species (*M. ochrogaster* and *M. montanus*, accession numbers AF069304 and AF070010; Young et al. 1999) and for the *cytb* gene from 18 species (accession numbers AF119280, AF163893, AF163890, AF163896, AF163900, AF163901, and AF163903-AF163906, Conroy and Cook 2000a; AF187230, Conroy and Cook 2000b; AY167210, AY513788, AY513798, AY513816, AY513819, AY513829, and AY513837, Jaarola et al. 2004). These sequences are included in multiple individual analyses and are labeled with the suffix GB in Figure 3 a–c.

### DNA Sequencing

We sequenced 783 base pairs (bp) of exon 1 of the arginine vasopressin 1a receptor gene (referred to as EXON1) and 703 bp of the adjacent region upstream of the first exon (referred to as UPSTREAM), as well as 930 bp of the mitochondrial *cytb* gene. A reaction volume of 25  $\mu$ L was used for amplification on a GeneAmp PCR System 9700 (Applied Biosystems) using Qiagen *Taq* polymerase. The reaction conditions for *cytb* and EXON1 are described elsewhere (Fink et al. 2004, 2006, 2007). For the UPSTREAM region, 4 primer pairs were used: V1aR 3'rep-ProtF 5'-GCTGGACATGTGGAATAACCAA-3' and V1aR 3'rep-ProtR 5'-TAGGAACCTCTCGGGAAACTCA-3', and 3-Prot-Mus-Mic-re 5'-CGGGAAACTCATGCTGTGCG-3' together with either 3-Prot-Moc-for 5'-TGACGAAGTC-ACCAGAAAAGA-3', 3-Prot-Mus-for 5'-TGACAGAGTCACCAGAAAAGGA-3' or 3-Prot-Mus-Moc-for 5'-TGACGAAGTCACCAGAAAAGGA-3'. The polymerase chain reaction (PCR) amplification profile for the first pair (V1aR 3'rep-ProtF/R) comprised an initial denaturation step at 95 °C for 2 min, followed by 40 cycles of denaturation at 95 °C for 1 min, annealing at 55 °C for 1 min and extension at 72 °C for 90 s. A final extension step of 72 °C for 10 min was added. For the other primer combinations, PCR conditions were identical except for annealing at 50 °C. Amplified fragments were controlled for size on a 1.5% agarose gel by comparing them with a 100 bp ladder (Invitrogen). After cleaning with GenElute PCR clean-up kit (Sigma) and dissolving products in 50  $\mu$ L bi-distilled water, sequencing reactions were carried out in a 10  $\mu$ L reaction volume using Terminator Ready Reaction Mix "Big Dye" Version 3.1 from Applied Biosystems. Both strands were sequenced using the following PCR conditions: An initial step of denaturation at 96 °C for 10 s, followed by 30 cycles of denaturation at 96 °C for 10 s, annealing at 55 °C for 10 s, and extension at 72 °C for 4 min 30 s. The products were cleaned using a DyeEx 96 spin kit (Qiagen) and were separated and detected on an ABI Prism 3100 Genetic Analyzer from Applied Biosystems. For the two nuclear sequence data sets, PCR products of individuals with multiple heterozygous sites were cloned as described in Fink et al. (2007) to resolve sequence haplotypes.

### AFLP Analyses

AFLP marker analyses were performed by adopting the protocol established by Vos et al. (1995). Six microliters of DNA (100 ng/ $\mu$ L) were digested for 2 h at 37 °C with 0.10  $\mu$ L *EcoRI* (100 U/ $\mu$ L; New England Biolabs, Ipswich, MA) and 0.02  $\mu$ L *MseI* (50 U/ $\mu$ L, New England Biolabs), whereas in the same reaction 1  $\mu$ L of each adapter pair (*EcoRI* adaptors 5 pmol/ $\mu$ L; *MseI* 50 pmol/ $\mu$ L) were ligated to cutting sites by adding 0.05  $\mu$ L T4 DNA ligase (2,000,000 U/mL; New England Biolabs). The following components were added to the reaction: 0.55  $\mu$ L BSA (1 mg/mL; New England Biolabs),

TABLE 1. Overview of taxonomical and geographical coverage of species, origins of samples and sequence accession numbers

<i>Microtus species</i>	Subgenus <sup>a</sup>	Continent	Country	Place of capture	Tissue number <sup>b</sup>	Label <sup>c</sup>	Accession numbers		AFLP <sup>d</sup> analysis
							avpr1a-exon	cytochrome <i>b</i>	
<i>arvalis</i>	<i>Microtus</i>	Europe	Switzerland	Gurnigel	CMPG	<i>arvalis_01_EU</i>	GU954331 GU954332	DQ663659	X
<i>arvalis</i>	<i>Microtus</i>	Europe	Switzerland	Gurnigel	CMPG	<i>arvalis_02_EU</i>	GU954402	GU954313	X
<i>arvalis</i>	<i>Microtus</i>	Europe	Switzerland	Belp	CMPG	<i>arvalis_03_EU</i>	EU175970 EU175971	GU954315 GU954314	X
<i>arvalis</i>	<i>Microtus</i>	Europe	Switzerland	Villa	CMPG	<i>arvalis_04_EU</i>	GU954339	DQ663658	
<i>arvalis</i>	<i>Microtus</i>	Europe	Switzerland	Ticino	CMPG	<i>arvalis_05_EU</i>	GU954333		
<i>agrestis</i>	<i>Agricola</i>	Europe	Switzerland	Bern	CMPG	<i>agrestis_01_EU</i>	GU954334		
<i>agrestis</i>	<i>Agricola</i>	Europe	Finland	Lapua	CMPG	<i>agrestis_02_EU</i>	EU175968 EU175969	GU954408 GU954409	
<i>agrestis</i>	<i>Agricola</i>	Europe	Germany	Waldbeck	CMPG	<i>agrestis_03_EU</i>	GU954333 GU954334		
<i>agrestis</i>	<i>Agricola</i>	Europe	Switzerland	St.Léonards	CMPG	<i>agrestis_04_EU</i>	GU954337	GU954405	
<i>agrestis</i>	<i>Agricola</i>	Europe	Sweden	Nättraby	CMPG	<i>agrestis_05_EU</i>	GU954406	GU954407	
<i>agrestis</i>	<i>Agricola</i>	Europe	Sweden	Norje Boke	CMPG	<i>agrestis_06_EU</i>	GU954407		
<i>tatricus</i>	<i>Terricola</i>	Europe	Slovakia	High Tatra Mountains	CMPG	<i>tatricus_01_EU</i>	EU175005	DQ663666	X
<i>tatricus</i>	<i>Terricola</i>	Europe	Slovakia	High Tatra Mountains	CMPG	<i>tatricus_02_EU</i>	GU954373		
<i>tatricus</i>	<i>Terricola</i>	Europe	Slovakia	High Tatra Mountains	CMPG	<i>tatricus_03_EU</i>	GU954370		
<i>rossiaemeridionalis</i>	<i>Microtus</i>	Europe	Macedonia	Gradsko	MVZ-159219	<i>rossiaemeridionalis_01_EU</i>	GU954371 GU954374	GU987117	X
<i>rossiaemeridionalis</i>	<i>Microtus</i>	Europe	Bulgaria	Obzor	CMPG	<i>rossiaemeridionalis_02_EU</i>	EU175973		
<i>rossiaemeridionalis</i>	<i>Microtus</i>	Europe	Romania	Maramures	CMPG	<i>rossiaemeridionalis_03_EU</i>	EU175972	DQ663664	X
<i>multiplex</i>	<i>Terricola</i>	Europe	Switzerland	Ticino	CMPG	<i>multiplex_01_EU</i>	EU175974	DQ663663	X
<i>nivalis</i>	<i>Chionomys</i>	Europe	Spain	Avila	CMPG	<i>nivalis_01_EU</i>	EU175975	GU987115	X
<i>nivalis</i>	<i>Chionomys</i>	Europe	Switzerland	Derborence	CMPG	<i>nivalis_02_EU</i>	GU954345	GU954317	
<i>nivalis</i>	<i>Chionomys</i>	Europe	Switzerland	Derborence	CMPG	<i>nivalis_03_EU</i>	GU954341 GU954342	GU954317	
<i>nivalis</i>	<i>Chionomys</i>	Europe	Switzerland	Derborence	CMPG	<i>nivalis_04_EU</i>	GU954343 GU954344	GU954316	
<i>felteni</i>	<i>Terricola</i>	Europe	Greece	Thessalia	CMPG	<i>felteni_01_EU</i>	EU175976	DQ663661	X
<i>thomasi</i>	<i>Terricola</i>	Europe	Greece	Nomos Arkadia	CMPG	<i>thomasi_01_EU</i>	EU175977	DQ663667	X
<i>thomasi</i>	<i>Terricola</i>	Europe	Greece	Kyllini	CMPG	<i>thomasi_02_EU</i>	GU954346	GU954320	
<i>thomasi</i>	<i>Terricola</i>	Europe	Greece	Kyllini	CMPG	<i>thomasi_03_EU</i>	GU954347 GU954348	GU954321	
<i>thomasi</i>	<i>Terricola</i>	Europe	Greece	Kyllini	CMPG	<i>thomasi_04_EU</i>	GU954349		
<i>cabraeae</i>	<i>Agricola</i>	Europe	Portugal	Cauda	CMPG	<i>cabraeae_01_EU</i>	EU175978	DQ663660	X
<i>cabraeae</i>	<i>Agricola</i>	Europe	Portugal	Cauda	CMPG	<i>cabraeae_02_EU</i>	GU954380	GU954322	
<i>socialis</i>	<i>Microtus</i>	Europe	Azerbaijan	Stepanakert	MVZ-199147	<i>socialis_01_EU</i>	EU175981	GU987118	X
<i>socialis</i>	<i>Microtus</i>	Europe	Iran	Tar Lake	MVZ-198799	<i>socialis_02_EU</i>	EU175982 GU954358 GU954359	GU954326	

Continued

TABLE 1. Continued

Microtus species	Subgenus <sup>a</sup>	Continent	Country	Place of capture	Tissue number <sup>b</sup>	Label <sup>c</sup>	Accession numbers		AFLP analysis <sup>d</sup>	
							avpr1a-exon	cypr1a-upstream		
<i>schelkovnikovi</i>	<i>Terricola</i>	Europe	Azerbaijan	Talysh	MVZ-179136	<i>schelkovnikovi_01_EU</i>	EU175979 EU175980 EU176006	GU954400	GU954311	X
<i>oeconomus</i>	<i>Alexandromys/ Pallasiinus</i>	Holarctic	unknown	unknown	CMPG	<i>oeconomus_01_HOL</i>	GU954382		GU987116	X
<i>oeconomus</i>	<i>Alexandromys/ Pallasiinus</i>	Holarctic	Poland	Bialowieza	CMPG	<i>oeconomus_02_HOL</i>	GU954350		GU987116	
<i>oeconomus</i>	<i>Alexandromys/ Pallasiinus</i>	Holarctic	Poland	Bialowieza	CMPG	<i>oeconomus_03_HOL</i>	GU954350		GU954319	
<i>oeconomus</i>	<i>Alexandromys/ Pallasiinus</i>	Holarctic	Poland	Bialowieza	CMPG	<i>oeconomus_04_HOL</i>	GU954351		GU987116	
<i>ochrogaster</i>	<i>Pedomys</i>	North America	USA	Kansas	CMPG	<i>ochrogaster_01_NA</i>	EU175983	GU954383 GU954384	DQ663671	X
<i>ochrogaster</i>	<i>Pedomys</i>	North America	USA	Illinois	CMPG	<i>ochrogaster_02_NA</i>	EU175984	GU954385	GU954318	X
<i>montanus</i>	<i>Mymomes</i>	North America	USA	Missoula	CMPG	<i>montanus_01_NA</i>	EU175986		DQ663650	
<i>montanus</i>	<i>Mymomes</i>	North America	USA	Missoula	CMPG	<i>montanus_02_NA</i>	EU175987			
<i>montanus</i>	<i>Mymomes</i>	North America	USA	Missoula	CMPG	<i>montanus_03_NA</i>	EU175988			
<i>pinetorum</i>	<i>Pitymys</i>	North America	USA	Calloway	CMPG	<i>pinetorum_01_NA</i>	EU175985	GU954414	DQ663654	X
<i>californicus</i>	<i>Mymomes</i>	North America	USA	Stanislaus	MVZ-207428	<i>californicus_01_NA</i>	EU175986	GU954386		
<i>californicus</i>	<i>Mymomes</i>	North America	USA	Ventura	MVZ-200053	<i>californicus_02_NA</i>	EU175987	GU954387	DQ663651	X
<i>californicus</i>	<i>Mymomes</i>	North America	USA	San Bernardino	MVZ-198778	<i>californicus_03_NA</i>	EU175988	GU954388		
<i>californicus</i>	<i>Mymomes</i>	North America	USA	Siskiyou	MVZ-199494	<i>californicus_04_NA</i>	GU954356	GU954418	GU954325	
<i>chrotorrhinus</i>	<i>Aulacomys</i>	North America	USA	Minnesota	CMPG	<i>chrotorrhinus_01_NA</i>	GU954357		GU954323	X
<i>chrotorrhinus</i>	<i>Aulacomys</i>	North America	USA	Minnesota	CMPG	<i>chrotorrhinus_02_NA</i>	GU954353	GU954415		
<i>richardsoni</i>	<i>Pitymys/ Aulacomys</i>	North America	USA	Probably Oregon	CMPG	<i>richardsoni_01_NA</i>	GU954354	GU954416	DQ663652	X
<i>richardsoni</i>	<i>Pitymys/ Aulacomys</i>	North America	USA	Probably Oregon	CMPG	<i>richardsoni_02_NA</i>	GU954355	GU954417	GU954324	X
<i>abbreviatus</i>	<i>Stenocranius/ Alexandromys</i>	North America	USA	Hall Island	MVZ-199498	<i>abbreviatus_01_NA</i>	EU175989	GU954390	DQ663652	X
<i>abbreviatus</i>	<i>Stenocranius/ Alexandromys</i>	North America	USA	Hall Island	MVZ-199500	<i>abbreviatus_02_NA</i>	EU175990	GU954391		
<i>oregoni</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216761	<i>oregoni_01_NA</i>	EU175991	GU954421	GU954327	X
<i>townsendii</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216762	<i>townsendii_01_NA</i>	EU175989	GU954422	DQ663655	X
<i>townsendii</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216766	<i>townsendii_02_NA</i>	EU175991	GU954392		
<i>richardsoni</i>	<i>Pitymys/ Aulacomys</i>	North America	USA	Probably Oregon	CMPG	<i>richardsoni_02_NA</i>	GU954361	GU954393		
<i>abbreviatus</i>	<i>Stenocranius/ Alexandromys</i>	North America	USA	Hall Island	MVZ-199498	<i>abbreviatus_01_NA</i>	GU954362	GU954394	GU954307	X
<i>abbreviatus</i>	<i>Stenocranius/ Alexandromys</i>	North America	USA	Hall Island	MVZ-199500	<i>abbreviatus_02_NA</i>	EU175994	GU954394	GU954330	X
<i>oregoni</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216761	<i>oregoni_01_NA</i>	GU954363	GU954394		
<i>townsendii</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216762	<i>townsendii_01_NA</i>	EU175995	GU954395	GU954308	X
<i>townsendii</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216766	<i>townsendii_02_NA</i>	EU175996	GU954401	GU954312	X
<i>townsendii</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216766	<i>townsendii_02_NA</i>	EU175997	GU954401		
<i>townsendii</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216766	<i>townsendii_02_NA</i>	EU175998	GU954366		
<i>townsendii</i>	<i>Mymomes</i>	North America	USA	Oregon	MVZ-216766	<i>townsendii_02_NA</i>	GU954367	GU954367		

Continued

TABLE 1. Continued

<i>Microtus</i> species	Subgenus <sup>a</sup>	Continent	Country	Place of capture	Tissue number <sup>b</sup>	Label <sup>c</sup>	Accession numbers		AFLP analysis <sup>d</sup>	
							avpr1a-exon	avpr1a-upstream		
<i>longicaudus</i>	<i>Aulacomys/Chionomys/Microtus</i>	North America	USA	California	CMPG	<i>longicaudus_01_NA</i>	GU954335	GU954399	DQ663656	X
<i>longicaudus</i>	<i>Aulacomys/Chionomys/Microtus</i>	North America	Canada	Yukon	CMPG	<i>longicaudus_02_NA</i>	EU175992	GU954419		X
<i>longicaudus</i>	<i>Aulacomys/Chionomys/Microtus</i>	North America	USA	California	MVZ-198761	<i>longicaudus_03_NA</i>	EU175993	GU954420	DQ663656	
<i>montebelli</i>	<i>Alexandromys/Pallasinus</i>	Asia	Japan	Takatsuki	CMPG	<i>montebelli_01_AS</i>	EU176000	GU954396	DQ663653	X
<i>montebelli</i>	<i>Alexandromys/Pallasinus</i>	Asia	Japan	Tottori lity	CMPG	<i>montebelli_02_AS</i>	EU175999	GU954423	GU954328	X
<i>kikuchii</i>	<i>Alexandromys/Pallasinus</i>	Asia	Taiwan	Tao-Yuan	MVZ-180804	<i>kikuchii_01_AS</i>	EU176000 EU176001	GU954424 GU954397	GU954309	X
<i>kikuchii</i>	<i>Alexandromys/Pallasinus</i>	Asia	Taiwan	Alsihan	MVZ-180746	<i>kikuchii_02_AS</i>	GU954364		GU954329	
<i>Arvicola terrestris</i>		Europe	Switzerland	Bern	CMPG	<i>A.terrestris-EU</i>	GU954365 EU176002	GU954398	GU954310	X

<sup>a</sup>Subgenus assigned according to Wilson and Reeder (2005) and webpage <http://www.bucknell.edu/msw3/>. For unclear taxonomic status, all possible subgenus names are given.

<sup>b</sup>Collection and tissue number of samples: CPMG = collection of Computational and Molecular Population Genetics lab (University of Bern, Switzerland); MVZ = Museum of Vertebrate Zoology of the University of California, USA.

<sup>c</sup>At least one individual per species was analyzed for all four different markers used in this study. Up to four individuals per marker and species were analyzed. In case of heterozygous individuals for nucDNA sequences, chromosome sequences determined by cloning were arbitrarily labeled with 1 and 2 following the sample number.

<sup>d</sup>Individuals analyzed for AFLPs are indicated in the last column with X.

1.10  $\mu$ L NaCl (0.5 M) and 1.10  $\mu$ L T4 DNA Ligase buffer (New England Biolabs).

Ligated samples were diluted 10 times before selective preamplification was performed with primers complementary to the adapters with an additional base (E01 = 5'-GACTGCGTACCAATTC-A-3' and M02 = 5'-GATGAGTCCTGAGTAA-C-3'). Each PCR was performed in a 20  $\mu$ L reaction volume using 4  $\mu$ L diluted and ligated product, 0.5  $\mu$ L of each primer (5  $\mu$ M), 1.2  $\mu$ L MgCl<sub>2</sub> (25 mM), 2  $\mu$ L dNTP (2.5 mM), 2  $\mu$ L 10 $\times$  PCR buffer (Qiagen), and 0.16  $\mu$ L *Taq* DNA polymerase (5,000 U/ml, Qiagen). Thermocycling parameters were 72  $^{\circ}$ C for 2 min, followed by 30 cycles of denaturation at 94  $^{\circ}$ C for 20 s, annealing at 56  $^{\circ}$ C for 30 s, and extension at 72  $^{\circ}$ C for 2 min, followed by a final extension step at 60  $^{\circ}$ C for 30 min. The quality and quantity of the preamplified products were determined on 1.5% agarose gels and with a spectrophotometer (NanoDrop ND-1000 Spectrophotometer; NanoDrop Technologies, Wilmington, DE).

The selective amplification was performed with eight combinations of unique primers (Table 2). Primer combinations are referred to using the last two selective bases, for example, the combination E01-AAC/M02-CAG is shortened to ACag. Each PCR was performed in a 20  $\mu$ L reaction volume using 4  $\mu$ L diluted (1:9 with ddH<sub>2</sub>O) preamplification product, 0.5  $\mu$ L fluorescence-labeled (10  $\mu$ M; 6-FAM, VIC, NED, PET) E-primer, 0.5  $\mu$ L M-primer (10  $\mu$ M), 1.3  $\mu$ L MgCl<sub>2</sub> (25 mM), 4  $\mu$ L dNTP (2.5 mM), 2  $\mu$ L 10 $\times$  PCR buffer (Qiagen) and 0.16  $\mu$ L *Taq* DNA polymerase (5 U/ $\mu$ L; Qiagen).

Touchdown PCR was used, starting with 94  $^{\circ}$ C for 2 min, followed by 10 cycles of 94  $^{\circ}$ C for 20 s, 65  $^{\circ}$ C for 30 s, 72  $^{\circ}$ C for 2 min, where the annealing temperature dropped 1.0  $^{\circ}$ C with each repetition, followed by 26 cycles of 94  $^{\circ}$ C for 20 s, 56  $^{\circ}$ C for 30 s, 72  $^{\circ}$ C for 2 min, with a final 60 min temperature hold at 60  $^{\circ}$ C. From each PCR, 2  $\mu$ L PCR product was mixed with 11.85  $\mu$ L formamide and 0.15  $\mu$ L GeneScan-500 (LIZ) size standard (Applied Biosystems), and run on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems).

Special care was taken to ensure reproducibility of AFLP marker analyses. One individual of each species

was independently reprocessed and reanalyzed for all eight primer combinations (starting with the AFLP enzyme digestion) to check consistency of individual results. These replicates were analyzed independently in different processing runs, and the well position of the samples on the 96-well plates was altered to account for potential position effects on the plate of the sequencer. Furthermore, up to three individuals per species were analyzed to examine variation among individuals.

### Statistical Analyses

*Sequence analyses.*—Sequences from each individual were aligned using the Clustal W algorithm (Thompson et al. 1997) implemented in the program BioEdit 5.0.9 (Hall 1999) and revised manually to ensure conservative alignments (see in Creer et al. 2006). Haplotypes were identified with the program Arlequin 3.1 (Excoffier et al. 2005), and nucleotide diversity was estimated using DNAsp (Rozas et al. 2003).

Phylogenetic relationships among sequence haplotypes were reconstructed by obtaining neighbor-joining (NJ; Saitou and Nei 1987) and maximum-likelihood (ML) trees rooted with *A. terrestris* with 10,000 bootstrap replicates in Mega 3.1 (Kumar et al. 2004) and Paup 4.0b (Swofford 1999). For the ML analysis, Modeltest 3.06 (Posada and Crandall 1998) implemented in Paup 4.0b was used to estimate the most suitable model of DNA substitution. This program uses hierarchical likelihood ratio tests and the Akaike information criterion (Akaike 1974) to compare the log-likelihood scores of 56 models based on a NJ tree with Jukes-Cantor (Jukes and Cantor 1969) corrected distances to identify the model with the best fit. We additionally used likelihood ratio tests to evaluate whether substitution models violated molecular clock assumptions. The topologies inferred with these methods were compared with Bayesian analyses (BA) trees obtained with MrBayes 3.0 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003), which runs a multiple-chain Metropolis-coupled Markov Chain Monte Carlo analysis. We applied this Bayesian method to the various sequence markers as well as to the AFLP analyses by running two independent runs with four chains (one cold and three hot chains) for 1,000,000 generations, with a sample frequency of 100 and a burn-in of 25% (see in Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). We used the web-based program AWTY (Nylander et al. 2008), which identifies lack of convergence through graphical exploration of output trees from MrBayes. The post burn-in consensus phylogenies were drawn in the program TreeView (Page 1996).

Similarities among phylogenies obtained from sequence data and AFLPs were evaluated with the software TOPD/FMTS (Puigbo et al. 2007) that uses randomization of quartets (Estabrook et al. 1985) to evaluate whether two trees are more similar than expected by chance. We used the option "random" to generate 100 trees with the same taxa as in the original tree but

TABLE 2. Selective primer combinations and their respective extensions chosen (marked with X) for use in the AFLP assay<sup>a</sup>

<i>Eco</i> RI	<i>Mse</i> I							
	CAA	CAC	CAG	CAT	CTA	CTC	CTG	CTT
AAC, NED			X					
AAG, NED							X	
ACA, 6-FAM					X			
ACC, VIC		X						
ACG, PET								X
ACT, 6-FAM	X							
AGC, VIC				X				
AGG, PET						X		

<sup>a</sup>Rows contain selective trinucleotide extensions attached to the 3' end of the fluorescently labeled *Eco*RI primer 5'-GACTGCGTACCAATTC-NNN-3'. Columns contain selective tri-nucleotide extensions attached to the 3' end of the *Mse*I primer 5'-GATGAGTCCTGACCGA-NNN-3'.

random changes in the topology and consequently the relationship of the taxa. For the comparison of AFLPs and *cytb*, we compared the 10 topologies each with the highest posterior probability and calculated the average dissimilarity ( $\pm$ standard deviation [SD]) weighted by the posterior probability of the different trees. For the nuclear sequences (EXON1 and UPSTREAM) we compared the full diploid topologies (50 end branches) to cover diversity within individuals. For the comparison of EXON1 and UPSTREAM with effectively haploid phylogenies (AFLPs; *cytb*), we arbitrarily fixed the gametic phase for nuclear sequences and calculated the average dissimilarity ( $\pm$ SD) of the two resulting haploid trees with AFLPs and *cytb*.

We conducted a Bayesian Concordance Analysis (BCA) of phylogenetic trees with the program BUCKy 1.2 (Ane et al. 2007). BCA estimates the concordance factor of the tree topologies as well as of the splits based on analyses with MrBayes. This method can account for biological processes like hybridization, incomplete lineage sorting, or lateral gene transfer, which may result in different loci to have different genealogies. The a priori level of discordance among loci is controlled by the parameter  $\alpha$ . For BCA, we used a flat prior ( $\alpha = 1.0$ ), three hot and one cold chain, and a burn-in of 100,000 simulations followed by 1,000,000 simulations. Analyses were restricted to the comparison of *cytb* and AFLP tree topologies because additional intraindividual variation in the nuclear loci EXON1 and UPSTREAM cannot be taken into account by BUCKy 1.2 and the number of terminal branches is currently restricted to a maximum of 32.

Phylogenies based on single loci were compared with a combined marker phylogeny based on the same individual per species analyzed for UPSTREAM, EXON1, and *cytb*. Combined phylogenies were reconstructed with the program MrBayes, which allows variable rates for each marker-specific partition. For EXON1, analyses including all codon positions were compared with analyses with the selectively neutral third codon position only in order to avoid confounding effects of selection on this locus (see in Fink et al. 2007; Heckel and Fink 2008). Additionally, we used published data to reconstruct combined sequence trees based on all nuclear information available on GenBank for *Microtus*. This analysis included data from the growth hormone receptor gene (GHR) from eight *Microtus* species (*M. nivalis*, *M. oeconomus*, *M. arvalis*, *M. kikuchii*, *M. ochrogaster*, *M. richardsoni*, *M. chrotorrhinus*, and *M. longicaudus*) and *A. terrestris* (Galewski et al. 2006).

Biogeographical hypotheses of the colonization of Europe and North America by *Microtus* were explicitly tested by assessing the proportion of compatible gene trees following Carstens et al. (2005). We tested the single-wave hypothesis (SWH) that states that the continents were colonized by a single wave each (Fig. 1, Brunet-Lecomte and Chaline 1991; Chaline et al. 1999) and multiple-wave hypotheses (MWH) that specify several waves for the colonization of each continent. We used a burn-in of 25% for the Bayesian runs and com-

bined trees then in a posterior distribution of 15,000 trees. We then used the tree filters in Paup 4.0b to assess the proportion of the trees in the distribution that were consistent with the SWH or MWH. The proportion of trees consistent with the topology predicted by the hypothesis is taken as the posterior probability for the hypothesis given the priors and data (Carstens et al. 2005).

For the nuclear sequence data sets, we tested for recombination within marker regions using Recombination Detection Program (RDP; Martin and Rybicki 2000), Geneconv (Padidam et al. 1999), Bootscanning (Salminen et al. 1995), Maximum Chi Square (MaxChi2; Maynard Smith 1992; Posada and Crandall 2001), Chimaera (Posada and Crandall 2001), and Sister scanning (Gibbs et al. 2000) of the RDP2 package (Martin et al. 2005). Because methods using substitution-based approaches might outperform analyses using phylogenetic information to estimate recombination (Posada and Crandall 2001; Posada 2002), we applied various types of methods (phylogenetic: RDP, Bootscanning, sister scanning; substitution: Geneconv, MaxChi2, Chimera). For the RDP method, we used internal and external references (to avoid potential rate heterogeneity problems) and a window size of 10 bp. For the Geneconv method, triplets of sequences were investigated and each indel was treated as a polymorphic site, whereas the G-scale was set to zero, which allows for a high mismatch penalty and is thus useful to detect recent recombination events. The following settings were used for the Bootscanning method: window size 100 bp, step size 20 bp, 1000 bootstrap replicates and the Jukes and Cantor (1969) model of substitutions. We scanned sequence triplets with the MaxChi2 method by using a variable window size with 0.1% fraction of variable sites to ensure an adjusted window size when considering closely and distantly related sequences simultaneously. The same settings were applied for the Chimaera method. When using the sister scanning method, we applied a window size of 100 bp and a step size of 20 bp. To focus the analysis on sites differing between the sequences in a triplet, we used the setting "1/2/3 variable positions" and as the fourth sequence we chose the nearest outlier to the triplet investigated.

*AFLP analyses.*—AFLP fragment scoring was done with GeneMapper software version 3.7 (Applied Biosystems). The multilocus AFLP profiles were scored as presence (1) or absence (0) of a particular band to create binary matrices. To avoid artifacts, only bands with intensity above a preset amplitude (100), and only markers that occurred in at least two individuals or in two independent replicates of one individual were scored. Fragments in the range of 50–500 bp length were analyzed in comparison with the internal GeneScan-500 (LIZ) size standard.

Genetic distances from the binary AFLP matrices were calculated with the Restdist program of the Phylip package (Felsenstein 2005) using the modified restriction

sites distance method of Nei and Li (1979). NJ trees based on the genetic distances were generated using the program Neighbor in the Phylip package (Felsenstein 2005) and bootstrapped 10,000 times. Bootstrap values were calculated by the Phylip programs Seqboot and Consense (Felsenstein 2005). Additionally, BA phylogenetic inferences were performed with MrBayes by adding the option that characters absent in all taxa (all 0) could not be scored (see in Felsenstein 1992).

We evaluated the relative gain in resolving the single individual phylogeny with increasing number of AFLP primer pairs and numbers of markers. The bootstrap support resulting from the analysis of the markers from each single primer pair was calculated separately. Furthermore, the binary data matrices of the last 4 primer pairs were permuted (not all eight due to calculation limitations) to evaluate the variation in bootstrap support of a single primer pair conditional on the data from other primers. The mean bootstrap support of each permutation was calculated for 5, 6, 7, and 8 primer pairs to evaluate primer pair-specific effects. For each new permuted primer pair, the phylogenetic tree and the mean bootstrap support were calculated with the Phylip package (see above, Felsenstein 2005). From the resulting trees, the overall mean and standard deviation ( $\pm$ SD) of the mean bootstrap support and the number of nodes with bootstrap support higher than 50% and 70% were counted and plotted against the number of primer pairs used.

## RESULTS

### Mitochondrial *cytb* Analyses

Sequencing of 48 individuals for 930 bp of the mitochondrial *cytb* gene and inclusion of 18 sequences from GenBank resulted in 62 different sequences (data submitted to TreeBASE, accession number TB2:S10754). Shared haplotypes were found among conspecific individuals (e.g. in *M. oeconomus*, *M. nivalis*) but not among species. Haplotype alignments revealed 40% variable positions and a nucleotide diversity of 0.1179.

Phylogenies based on NJ, ML and BA methods revealed the same tree topologies (Figs. 2a and 3a). Following Modeltest results, the ML analyses of single individuals per species were based on the General Time Reversible model with a gamma distribution and a proportion of invariable sites (GTR+G+I; Lanave et al. 1984; Tavaré 1986; Rodriguez et al. 1990), and a molecular clock with the following parameters: base frequencies of A = 0.3253, C = 0.3461, G = 0.1067, T = 0.2219; a substitution rate matrix of A $\leftrightarrow$ C = 0.8944, A $\leftrightarrow$ G = 8.0313 and A $\leftrightarrow$ T = 1.9282, C $\leftrightarrow$ T = 0.0217, C $\leftrightarrow$ G = 16.3736, G $\leftrightarrow$ T = 1.0000, and a proportion of invariable sites (I) = 0.5701. The gamma shape parameter of 0.8233 indicated low rate heterogeneity and thus no confounding phylogenetic signals would be expected. For the analyses with additional intraspecific data (Fig. 3a), the following model was chosen: Transversion model (TVM; Posada and Crandall 1998) with base frequency of A = 0.3420,

C = 0.3388, G = 0.0870, T = 0.2323; a substitution rate matrix of A $\leftrightarrow$ C = 0.6233, A $\leftrightarrow$ G and C $\leftrightarrow$ T = 8.8327, A $\leftrightarrow$ T = 1.1083, C $\leftrightarrow$ G = 0.4045, G $\leftrightarrow$ T = 1.0000; I = 0.5659, and a gamma shape parameter of 0.9894, again indicating low rate heterogeneity. The molecular clock assumption was rejected in this case.

### Nuclear *avpr1a* Analyses

The 703 bp analyzed for the noncoding UPSTREAM region of the first exon of the nuclear *avpr1a* gene showed 28% variable sites with additional 19% of pure insertion-deletion sites without additional substitutions at these positions. Overall nucleotide diversity was estimated as 0.0378. Among the 44 individuals analyzed, 59 sequence types were detected (data submitted to TreeBASE, accession number TB2:S10754). Identical sequences occurred only among individuals of the same species. Seventeen of the 44 individuals were heterozygous (39%).

All methods (NJ, ML, and BA) revealed similar topologies (Figs. 2b and 3b). The substitution model identified by Modeltest for the single individual trees (Fig. 2b) was the transition model (TIM; Posada and Crandall 1998) without molecular clock assumptions. The base frequencies were estimated as follows: A = 0.2744, C = 0.2154, G = 0.3141, T = 0.1961, and the substitution rate matrix was A $\leftrightarrow$ C and G $\leftrightarrow$ T = 1.0000, A $\leftrightarrow$ G = 1.6773, A $\leftrightarrow$ T and C $\leftrightarrow$ G = 0.5382, C $\leftrightarrow$ T = 2.4596. The gamma shape parameter was estimated as 0.7749, indicating low rate heterogeneity, which should thus not lead to confounding phylogenetic signals.

For the phylogeny with several individuals per species (Fig. 2b), the General Time Reversible model with a gamma distribution (GTR+G; Lanave et al. 1984; Tavaré 1986; Rodriguez et al. 1990) was used without molecular clock assumptions with the following parameters for the ML analyses: base frequency of A = 0.2762, C = 0.2237, G = 0.3128, T = 0.1874; a substitution rate matrix of A $\leftrightarrow$ C = 0.7717, A $\leftrightarrow$ G = 1.4736, and A $\leftrightarrow$ T = 0.6839, C $\leftrightarrow$ T = 2.2943, C $\leftrightarrow$ G = 0.3671, G $\leftrightarrow$ T = 1.0000; a gamma shape parameter of 0.7641. The relatively low rate heterogeneity together with no detection of recombination for the single- and the multiple individual data sets with many different methods indicates only weak confounding phylogenetic signals for this marker.

The analyses of 783 bp of the coding nuclear EXON1 of the *avpr1a* gene revealed 74 sequence types among the 54 individuals that we sequenced and the two sequences from GenBank (data submitted to TreeBASE, accession number TB2:S10754). Identical sequence types were detected within species (*M. agrestis*, *M. chrotorrhinus*, *M. montebelli*, *M. oeconomus*) but also between species (*M. oeconomus* and *M. taticus*). Twenty-six individuals were heterozygotes (48%); 18% of variable sites and a nucleotide diversity of 0.0161 were estimated.

Tree topologies obtained with different methods (NJ, ML, and BA) were identical for single and multiple individual analyses (Figs. 2c and 3c). The most likely

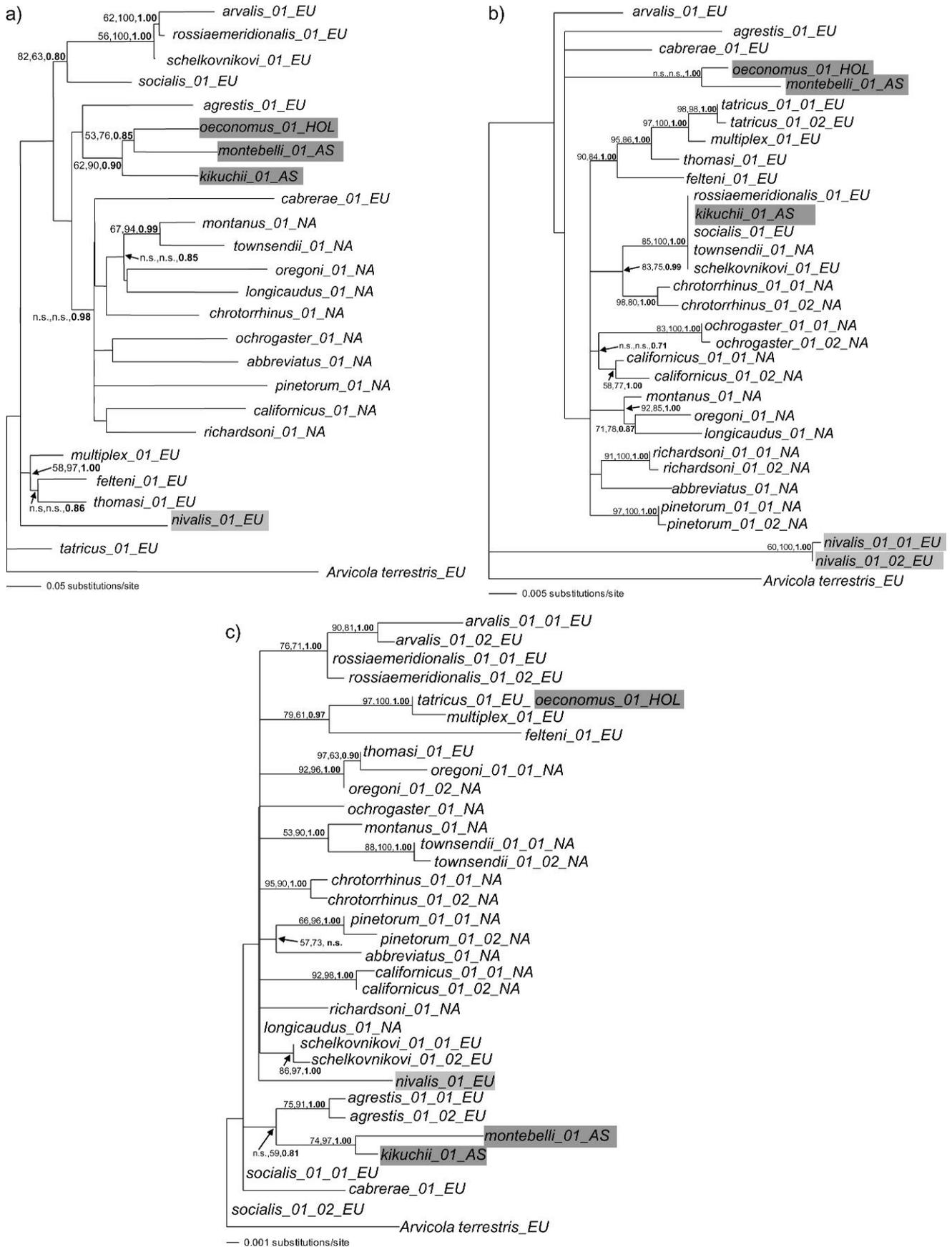


FIG 2.

substitution model for the single individual analyses (Fig. 2c) was the TVM model (Posada and Crandall 1998), without molecular clock assumptions with base frequencies of A = 0.1823, C = 0.2963, G = 0.2945, T = 0.2269; substitution rate matrix of A↔C = 2.7917, A↔G and C↔T = 8.9405, A↔T = 1.1664, C↔G = 0.8767, G↔T = 1.0000; gamma shape parameter of 0.2229. The relatively high rate heterogeneity together with potential selection on the first two codon positions (Fink et al. 2007) could possibly lead to confounding phylogenetic signals. We therefore reconstructed phylogenies based on the third codon position only, which revealed topologies identical to the NJ, ML, and BA trees based on all three codon positions.

The three substitution type model (K81; Kimura 1981) without molecular clock assumptions was identified as the best substitution model for the multiple individual data set (Fig. 2c). The following parameters were estimated: base frequencies: A = 0.1869, C = 0.2963, G = 0.2916, T = 0.2253; substitution rate matrix: A↔C and G↔T = 1.0000, A↔G and C↔T = 4.7254, A↔T and C↔G = 0.4801; proportion of invariable sites = 0.5222; gamma shape parameter = 1.0155. Contrary to the single individual analyses, the gamma shape parameter of the multiple individual data set indicates low-rate heterogeneity, and because none of the applied methods detected recombination, confounding signals in the phylogeny should be rare, except for the possible influence of selection.

#### AFLP Analyses

The AFLP data set consists of 1869 characters derived from eight independent selective primer pair combinations (data submitted to TreeBASE, accession number TB2:S10754). Of these, 1835 (98%) were polymorphic among the 25 taxa studied (Figs. 4 and 5), with an average of 229 polymorphic markers per primer combination across all individuals (Table 3). We scored on average 77 polymorphic markers per primer combination and individual. An average of 619 polymorphic markers were positively scored per species with a range from 542 (*M. schelkownikovi*) to 674 markers (*M. montanus*).

Mean bootstrap support of phylogenetic trees based on AFLPs improved with increasing number of primer combinations (Figs. 4 and 6). Nevertheless, the slope flattened after the first two or three primer combina-

tions (50% threshold and 70% threshold, respectively). The mean bootstrap support of the phylogeny depends to some extent on the specific combination of primer pairs, but this is apparently independent of the bootstrap support provided by a primer pair on its own (Table 3 and Fig. 6). For example, the primer pair GGtc, which provided the lowest bootstrap support when analyzed on its own (32%, Table 3), considerably increased the mean bootstrap support of the phylogeny in combination with other primer pairs. Contrary to that, the primer pair CGtt provided per se an average bootstrap support of 38% (Table 3), but did not increase the mean bootstrap support in combination with other primer combinations.

#### Comparison of Phylogenies Obtained by Different Markers

The resulting phylogenetic trees of the *Microtus* genus from all markers separately or combined were all star shaped with short basal distances among nodes and long terminal branches indicating a burst of speciation. Several clusters of sequences from species from the same continent can be identified, but the basal splits between vole species from different continents were only moderately supported by sequence data sets (Figs. 2–5, 7, and 8). Comparatively, the AFLP phylogeny showed considerably higher bootstrap values for the main clades (Fig. 4).

Statistical evaluation of tree topologies based on the same individual per species with TOPD/FMTS (quartets method) showed the highest similarity between AFLPs and *cytb* among all comparisons. Dissimilarity between AFLP and *cytb* topologies was estimated as  $0.2001 \pm 0.0179$  compared with a random dissimilarity between trees of  $0.8599 \pm 0.0627$ . The dissimilarity between the AFLP and UPSTREAM tree topology was  $0.6685 \pm 0.0004$  (random:  $0.8572 \pm 0.0593$ ), AFLP and EXON1  $0.6869 \pm 0.0018$  (random:  $0.8517 \pm 0.0659$ ), *cytb* and UPSTREAM  $0.5442 \pm 0.0014$  (random:  $0.8612 \pm 0.0630$ ), and *cytb* and EXON1  $0.7544 \pm 0.0001$  (random:  $0.8582 \pm 0.0683$ ). Dissimilarity was relatively high between the two diploid nuclear markers UPSTREAM and EXON1 with  $0.7527$  (random:  $0.8710 \pm 0.0201$ ) despite their physical linkage.

Further evaluation of agreement between AFLP and *cytb* with BCA rejected full concordance between the entire topologies but indicated high concordance in some splits of the two markers (Table 4). Sample-wide posterior probabilities for concordance were particularly

←

FIGURE 2. Maximum-likelihood trees of the mitochondrial cytochrome *b* gene (a), the UPSTREAM region of the first exon (EXON1) of the nuclear *avpr1a* gene (b) and the EXON1 (c) of 24 *Microtus* species and *Arvicola terrestris* for a single individual per species. Bootstrap values >50 of main clusters are shown for ML, NJ, and BA (in bold) methods. *Microtus* species are labeled according to Table 1. Different chromosomes of heterozygous individuals are indicated with 01 and 02 following the species name. Classification of distribution ranges: EU = European; NA = North American; AS = Asian; HOL = Holarctic. a) Cytochrome *b* provides a geographic structure in the phylogeny with a separate clade comprised of two Asian (*M. kikuchii* and *M. montbelli*) and one Holarctic (*M. oeconomus* species, dark gray). *Microtus nivalis* (light gray) forms a basal branch. b) The group of Holarctic and Asian species is not present in the UPSTREAM analysis. *Microtus nivalis* still forms a basal branch. c) The phylogeny based on EXON1 is less geographically structured than that of *cytb* and UPSTREAM given that the Asian and Holarctic group split up. *Microtus nivalis* does not form a basal branch. The two chromosomes within an individual are not always most closely related (*M. oregoni*).

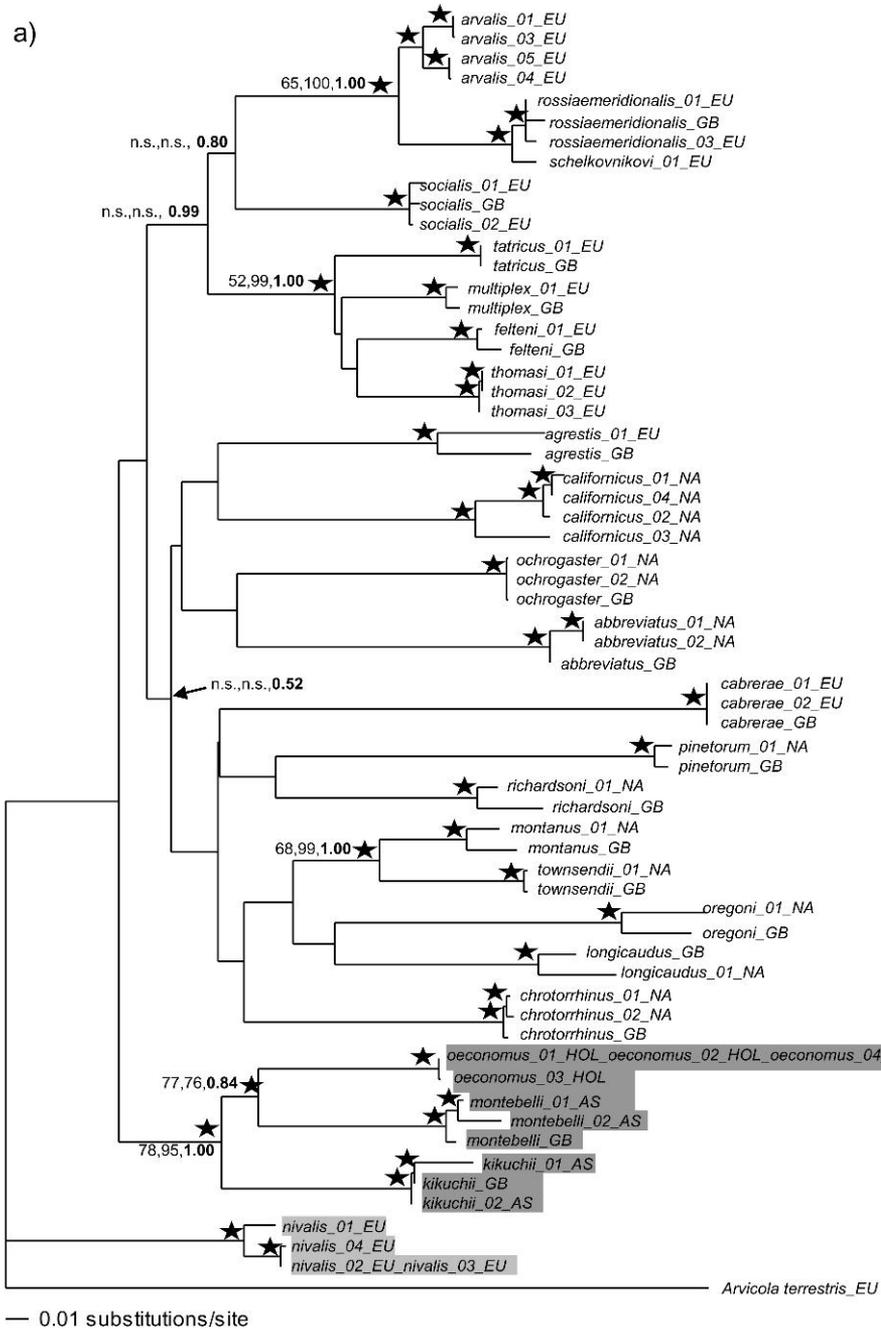


FIG 3. Continued

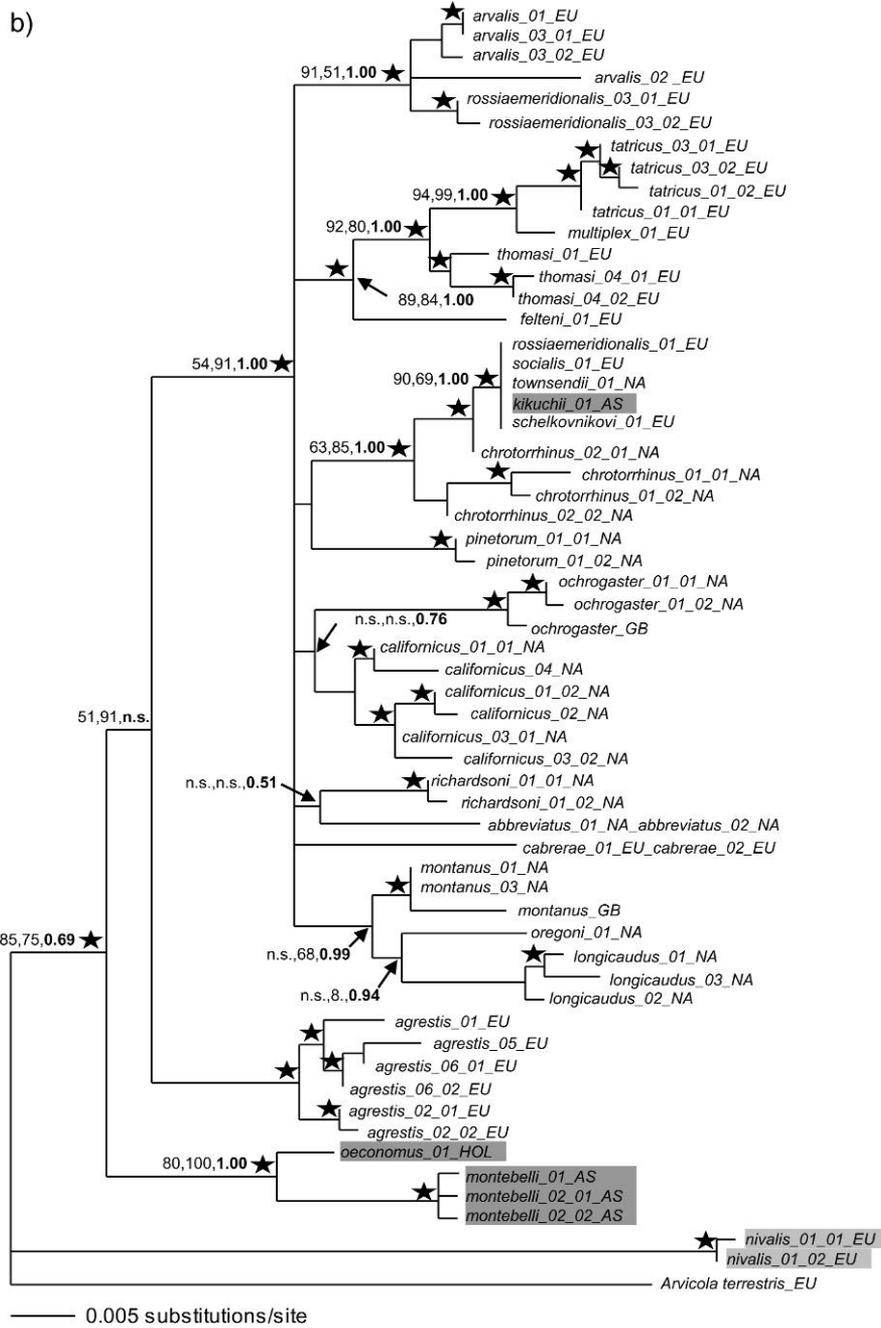


FIG 3. Continued

FIG. Continued

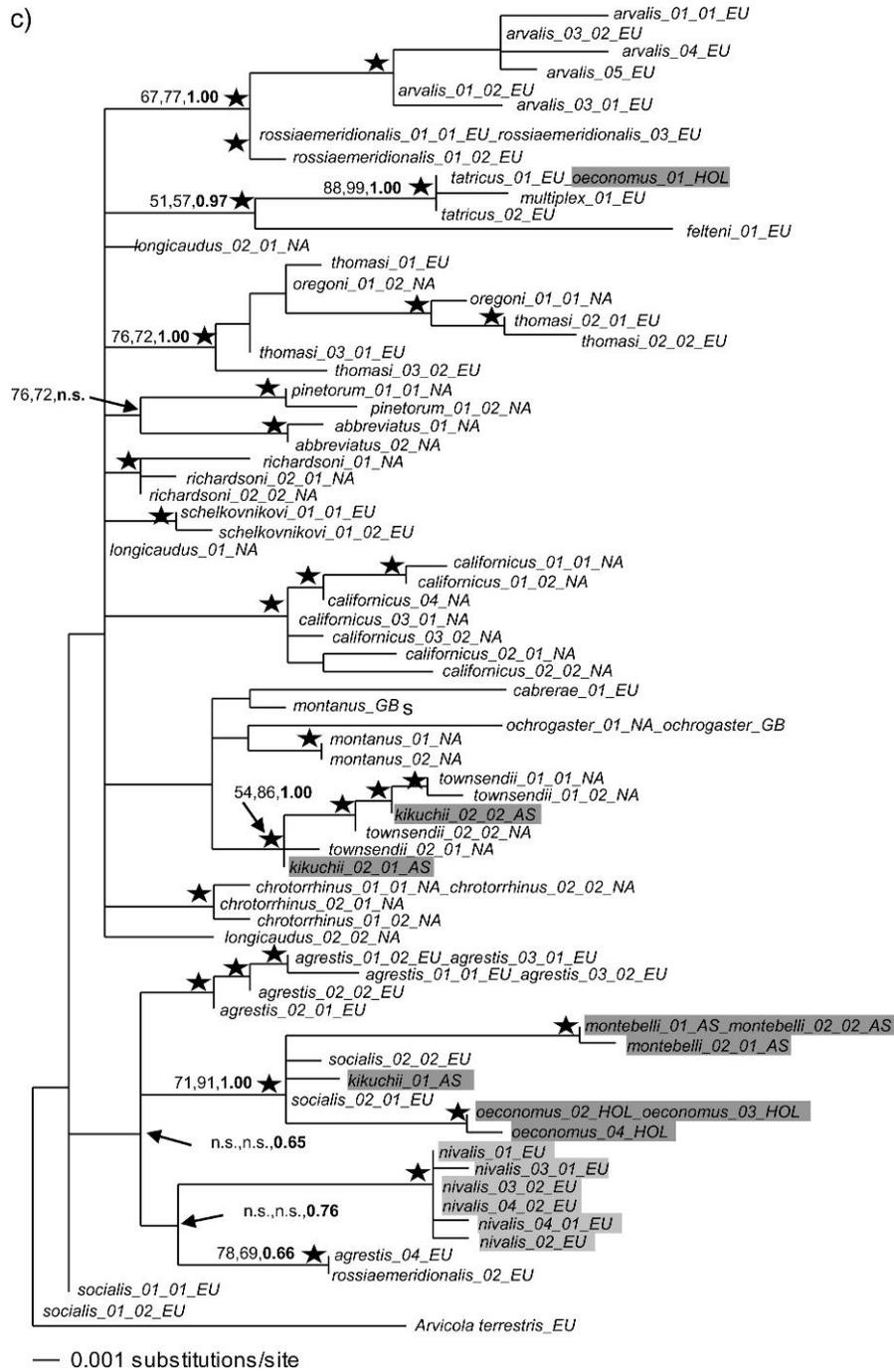


FIGURE 3. Maximum-likelihood trees of the mitochondrial cytochrome *b* gene (a), the UPSTREAM region of the nuclear *avpr1a* gene (b) and its first exon (EXON1) (c) of 24 *Microtus* species and *Arvicola terrestris* with additional samples to assess intraspecific diversity. Bootstrap support >50 for both NJ and ML methods as well as clade credibility intervals >0.5 for the Bayesian method are indicated with a star. Additionally, values of clusters of different species are shown for ML, NJ, and BA methods (in bold) on branches. n.s. indicates that one method does not show a bootstrap support >50. *Microtus* species are labeled according to Table 1. Different chromosomes of heterozygous individuals are indicated with 1 and 2 following the label. GenBank sequences are labeled with the suffix GB. a) Individuals of the same species cluster together for *cytb*, but there is no supported separation according to the geographic origin of the samples, except for a separate clade containing Asian (*M. kikuchii* and *M. montebelli*) and the Holarctic species (*M. oeconomus*, dark grey). The European species *M. agrestis* and *M. cabrerar* do not cluster with other European species. *Microtus nivalis* (light gray) forms a separate cluster toward the root of the phylogeny. b) Individuals belonging to the same species cluster together with the exception of *M. rossiaemeridionalis* and *M. chrotorrhinus* individuals. The group of Holarctic and Asian species is no longer present. *Microtus nivalis* is still a basal branch. c) No general clustering of sequences from the same species, and even the two chromosomes within one individual may cluster apart (see *M. oregoni*: *oregoni.01.01.NA* and *oregoni.01.02.NA*).

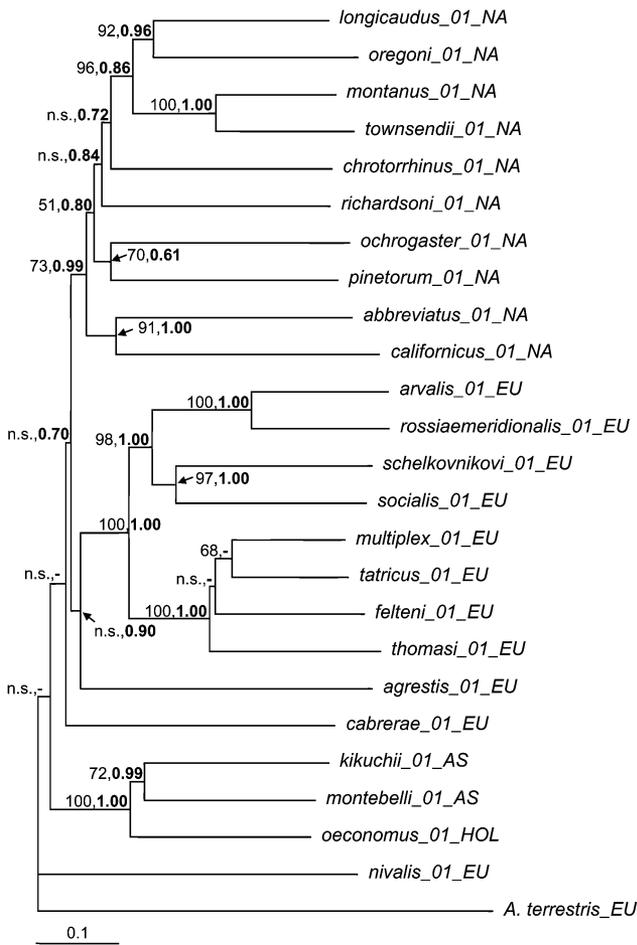


FIGURE 4. Phylogeny obtained from AFLP markers for one individual per species for *Microtus*, with *Arvicola terrestris* as outgroup. NJ bootstrap values (>50) are shown first, followed by the clade credibility values of BA in bold. Support values <50 are indicated with n.s. and missing nodes that appear only in one topology and are absent in the other are indicated with a hyphen. The Holarctic, Asian, and American species form separate, supported (>70) clusters, whereas the majority of the species from Europe form a supported subcluster (bootstrap 100) with two species (*M. agrestis*, *M. cabreræ*) lying outside. These two species might reflect evolutionarily older lineages similar to *M. nivalis* (see Discussion section).

high for larger clades (e.g., subgenera *Terricola* and *Microtus*) and still moderate for smaller clades or less well-sampled clades (e.g., Asian voles). Further runs with the use of a nonflat prior for the level of discordance among trees showed that a higher probability for concordance ( $\alpha = 0.1$ ) had no influence on the posterior probabilities.

Clustering of most species was largely comparable between markers despite differences in overall topologies. Differences were, for example, due to the clustering of *M. nivalis*, which formed a basal branch for *cytb*, UPSTREAM, and AFLPs (Figs. 2 and 4) but grouped among other *Microtus* species for EXON1 (Fig. 2). The Holarctic *M. oeconomus* and the Asian *M. montebelli* and *M. kikuchii* form a separate group in *cytb* and AFLPs but not for the two nuclear sequence markers. UPSTREAM has lower transition to transversion ratios than EXON1

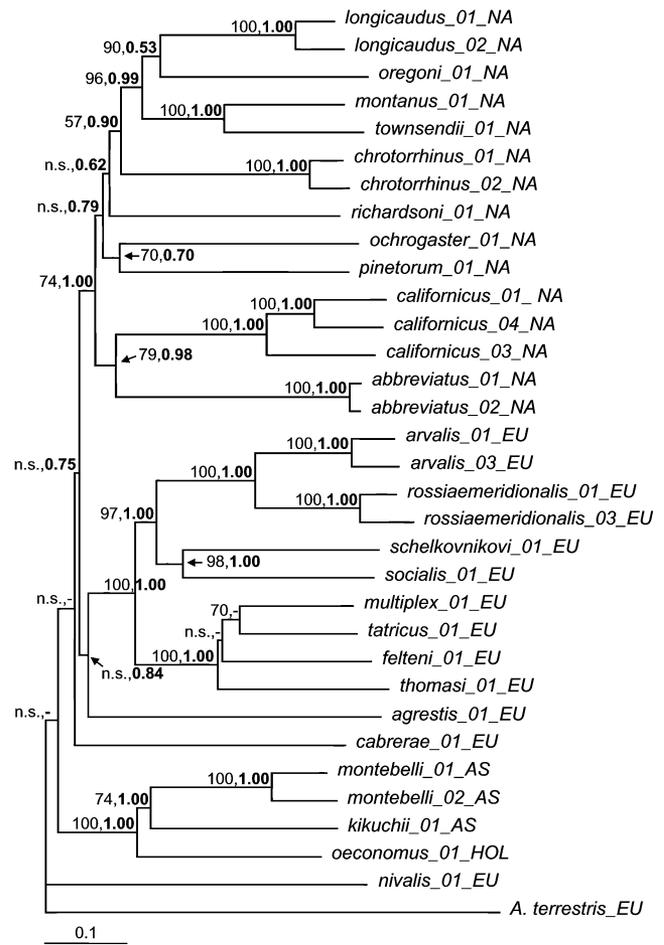


FIGURE 5. Phylogeny obtained from AFLP markers with additional samples to assess intraspecific diversity for *Microtus*, with *Arvicola terrestris* as outgroup. NJ bootstrap values (>50) are shown first, followed by the clade credibility values of BA in bold. If a method does not show bootstrap support >50, it is indicated with n.s. Missing nodes (that appear only in one topology and are absent in the other) are indicated with a hyphen. Individuals of the same species always cluster together with 100 bootstrap support. The topology and the separation according to the geographic origin of the samples is the same as in Figure 4.

and *cytb* indicating more power to detect phylogenetic patterns. UPSTREAM showed increased phylogenetic structure and support for some nodes of groups of American and European species compared with *cytb* (e.g., cluster *M. tatricus*, *M. multiplex*, *M. thomasi*, and *M. felteni*). Nevertheless, there is also a mixed cluster with individuals from three different continents (Asia: *M. kikuchii*; North America: *M. townsendii*; Europe: *M. schelkovnikovii*, *M. socialis*, *M. rossiaemeridionalis*) in the UPSTREAM phylogeny, which may indicate the lack of a clear phylogenetic signal for some sequences for this marker.

Contrary to the lack of clearly supported species clusters in the sequence analyses, the AFLP markers resolved the phylogeny with higher bootstrap and clade credibility support. The AFLP tree shows a supported clade with the American species and a supported group

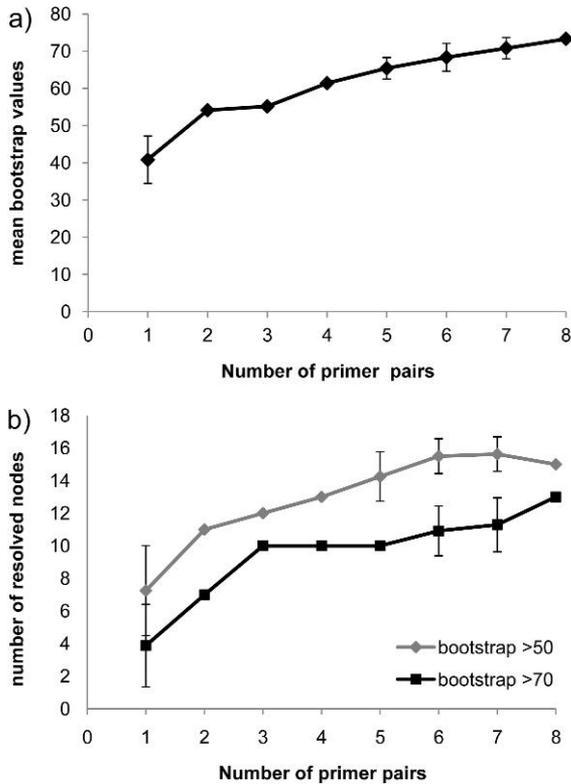


FIGURE 6. Relationships between numbers of primer pairs analyzed for AFLPs and support of the phylogenetic tree. a) Mean bootstrap values versus number of selective AFLP primer pairs scored for all nodes. Error bars give the standard deviation resulting from the permutation of the first or the last 4 primer pairs respectively (see text). b) Number of resolved nodes (bootstrap >50 or >70) versus the number of selective AFLP primer pairs used.

with the two Asian species and the Holarctic *M. oeconomus* (Fig. 4). The European species are not monophyletic because three species (*M. agrestis*, *M. cabreriae* and *M. nivalis*) do not cluster within the main clade. The two morphologically distinguished subgenera *Microtus* and *Terricola* can be seen as subclades within the main European clade. The topologies of the NJ and BA trees based on AFLPs are nearly identical, resolving the same main nodes. There is a slight difference in the American clade: *M. abbreviatus* and *M. californicus* are basal in the clade in the NJ tree, whereas the BA phylogeny reveals *M. ochrogaster* and *M. pinetorum* as basal species. Additionally, we found some incongruence in the European *Terricola* subgenus, where *M. multiplex* and *M. tatricus*

form an evolutionarily young sister group in the NJ tree, whereas they are basal to *M. thomasi* and *M. felteni* in the BA analyses.

The analysis of multiple individuals per species (Figs. 3 and 5) revealed generally similar tree topologies compared with the single individual trees (Fig. 2 and 4). Again, AFLPs showed larger bootstrap clade support (Fig. 5), and the European voles *M. nivalis*, *M. cabreriae* and *M. agrestis* remain still outside of the supported main European clade. However, there were relatively strong discrepancies among markers in regard to the resolution of individuals at the species level. UPSTREAM, *cytb*, and AFLP data clustered individuals according to species with all reconstruction algorithms except for *M. rossiaemeridionalis* and *M. chrotorrhinus* in UPSTREAM (see Fig. 3). EXON1 provided no obvious geographic signal and much less separation between species, and many individuals of the same species clustered in different groups than their conspecifics (e.g. *M. rossiaemeridionalis*, *M. kikuchii*, *M. agrestis*, and *M. oeconomus*). Even the two different chromosomes of one individual can appear in highly supported separate clusters (e.g., *M. oregoni* and *M. kikuchii*; Fig. 3c).

#### Combinations of Different Sequence Markers

The combined data set of *cytb*, UPSTREAM and EXON1 for the same individual per species showed a topology very similar to those obtained from *cytb* and AFLPs (Fig. 7). The grouping of the different species is highly supported by BA methods, although the topology was not identical when all codon positions of EXON1 were used compared with the analysis based on the third codon position only. *Microtus agrestis* forms a separate lineage when using all codon positions, but clusters with the North American *M. ochrogaster* in third codon analyses. There is no supported split between species from the North American and European continents, but the Asian and the Holarctic species form a monophyletic clade at the basal end of the topology, close to *M. nivalis* and the root.

Similar to these findings, a tree based on the combined nuclear sequences from UPSTREAM, EXON1, and GHR showed *M. nivalis* as a basal offshoot of the other *Microtus* species (Fig. 8). The topologies are identical with some highly supported clades, irrespective of whether all codon positions or only the third codon position of EXON1 was included. No clear separation was detected between European and Asian species or a

TABLE 3. Number of polymorphic and monomorphic AFLP markers resulting from the different primer pairs across 25 Arvicoline species, and mean bootstrap support of the single-individual phylogeny provided by each primer pair

Primer pair	CCac	CTaa	CAta	GCat	ACag	GGtc	AGtg	CGtt	Sum	Mean	±SD
Polymorphic	221	247	225	233	230	217	233	229	1835	229.4	9.1
Monomorphic	2	4	6	7	5	0	9	1	34	4.25	3.1
Sum	223	251	231	240	235	217	242	230	1869	233.6	11
Mean bootstrap support (%)	37.1	36.7	40.8	42.8	52.4	32.2	46.6	38.0		40.83	6.4

Note: SD = standard deviation.

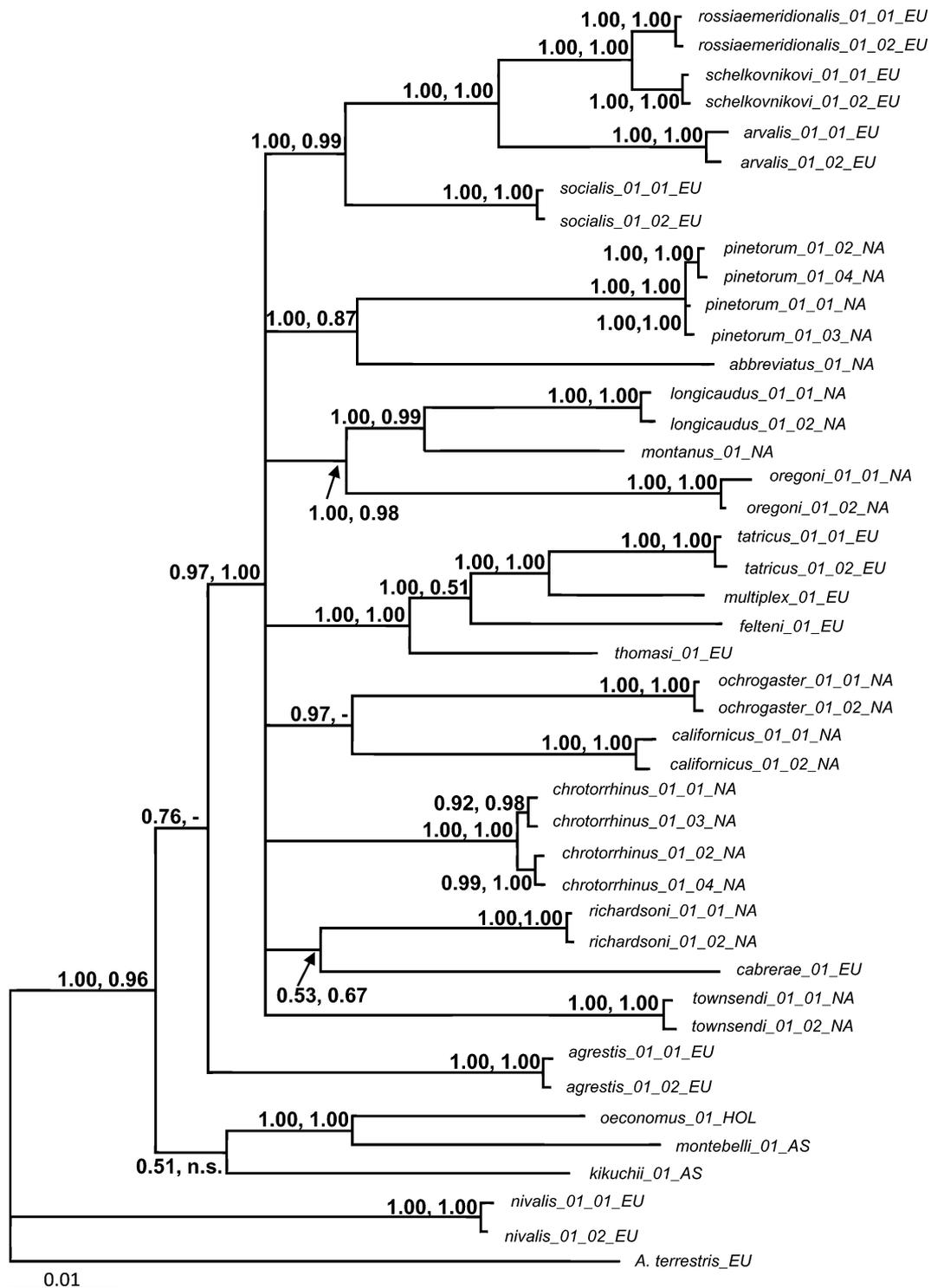


FIGURE 7. Combined Bayesian tree based on *cytb*, UPSTREAM and EXON1. Bayesian clade credibility values (50% majority rule) are given for the analysis with the whole EXON1 first, followed by values obtained by including only the third codon position of EXON1 in the analysis (see text). Every species is represented by one individual. As heterozygous sequences exist for UPSTREAM as well as for EXON1, all possible combinations of sequences are given (labeled 01–04; *cytb* was kept constant). Topologies obtained by including all codon positions of EXON1 and by using the third codon position only were very similar. Only the European *M. agrestis* clustered with the North American *M. ochrogaster* when only the selectively neutral third position was used. Because *M. agrestis* forms a separate lineage in the phylogeny based on all positions (as seen in the figure), whereas it clusters with *M. ochrogaster* in the phylogeny on the third codon position only, the missing nodes (and missing bootstrap support) are indicated with a hyphen.

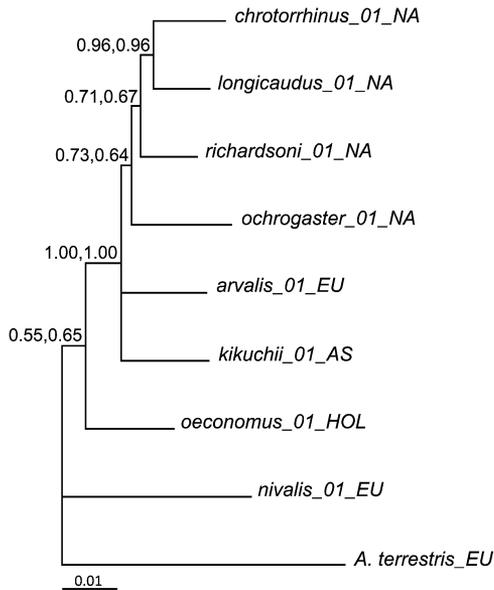


FIGURE 8. Combined tree of three nuclear markers, UPSTREAM and EXON1 of the *avpr1a* gene, and part of the exon11 of the GHR gene from Galewski et al. (2006) with one chromosomal sequence per species. Clade credibility values of BA (50% majority rule) are given on branches, where the first number represents the support for nodes when using all three codon positions of EXON1, whereas the second number indicates the bootstrap support if only the third codon position of EXON1 was included. Both analyses revealed similar topologies.

supported Holarctic and Asian cluster as in the analyses without GHR.

*Biogeographical Hypothesis Testing*

None of the markers or analysis methods clustered species and individuals strictly according to the continent

of origin of the samples. Biogeographical hypotheses for *Microtus* were explicitly tested based on the topologies of Bayesian AFLP trees as these showed highest support for different clusters in the genus. The SWH for each of the European and American continents (Fig. 1; Brunet-Lecomte and Chaline 1991; Chaline et al. 1999) had a Bayesian posterior probability of <0.01. The MWH for two colonizations of America (one by the North American main clade and one by *M. oeconomus*) had a posterior probability >0.99. The MWH tested for the European continent—one colonization by *M. oeconomus* and one by all the other European voles (including *M. agrestis* and *M. cabreræ*)—had a moderate posterior probability of 0.36. When we tested a simpler backbone tree (excluding *M. cabreræ*) we obtained a posterior probability of 0.96. This suggests that the European continent was colonized in at least two, potentially even three (*M. cabreræ*) or more colonization waves.

DISCUSSION

Single and combined mtDNA and nucDNA sequence markers and the genome scan approach demonstrate the complexity of phylogenetic relationships in the *Microtus* genus. Basal relationships among species from different continents are only weakly resolved in sequence analyses even though most markers show generally a clear separation between species. Contrary to this, the genome-wide approach with AFLP markers substantially increases resolution of the shallow phylogeny and reveals supported clades of species belonging to different continents. This new microtine phylogeny with better support for the main clades justifies a re-interpretation of the phylogenetic and biogeographic history of this fast-radiating genus.

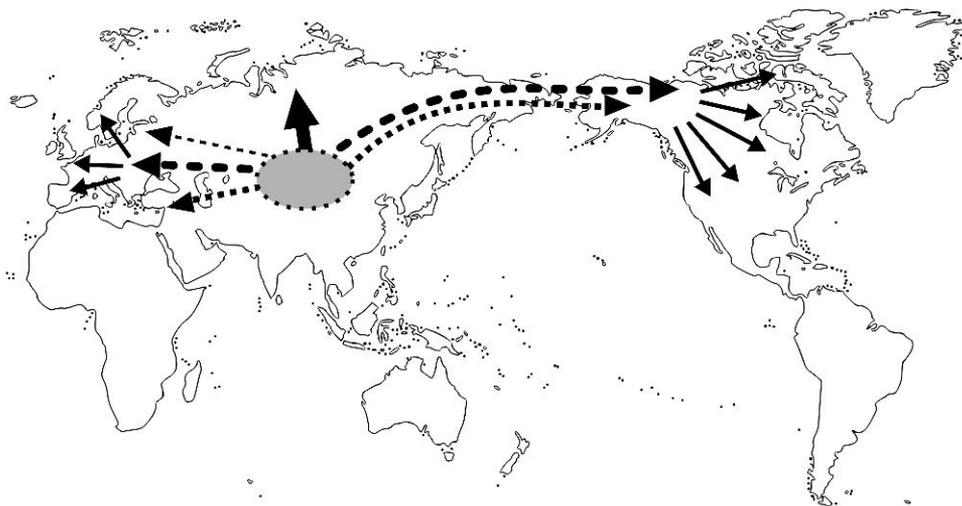


FIGURE 9. New colonization scenario for the *Microtus* genus supported by the AFLP tree topology and biogeographical hypothesis testing. The putative origin in South Asia is represented as a grey oval from which several colonization waves to Europe and North America (black dashed arrows) occurred with subsequent radiations on each continent (see text). Data for North Asia is too limited to test for a single or repeated colonization waves.

TABLE 4. Sample-wide mean concordance factor (CF) of splits (supported clades) in the Primary Concordance Tree of AFLPs and *cytb* based on the analyses with BUCKY (Ane et al. 2007)

Splits in the primary concordance tree	Sample-wide mean CF (95% credibility)
Subgenus: <i>Terricola</i>	1.000 (1.0,1.0)
Subgenus: <i>Microtus</i> + <i>Terricola</i>	0.999 (1.0,1.0)
<i>M. montanus</i> + <i>townsendii</i>	0.994 (1.0,1.0)
Asian + <i>M. oeconomus</i>	0.949 (0.5,1.0)
Subgenus: <i>Microtus</i>	0.895 (0.5,1.0)
<i>M. montanus</i> + <i>townsendii</i> + <i>oregoni</i> + <i>longicaudus</i>	0.861 (0.5,1.0)
<i>M. felteni</i> + <i>tatricus</i>	0.711 (0.5,1.0)
<i>M. oregoni</i> + <i>longicaudus</i>	0.611 (0.0,1.0)
Asian voles	0.522 (0.5,1.0)

#### *Phylogenies from Mitochondrial and Nuclear Sequences versus AFLP Markers*

Different patterns appear in the tree topologies of the two nuclear sequence markers and the mitochondrial gene. The *cytb* trees show clearly different haplotypes for all species and a strong clustering of individuals belonging to the same species, whereas the nuclear EXON1 and UPSTREAM phylogenies show some polyphyletic species groups, as well as shared sequence types between species. This pattern might have easily been missed by analyses of a single individual per species because for UPSTREAM only two species were detected in which not all individuals clustered together. The comparison of the tree topologies from single versus multiple individuals per species showed that additional intraspecific data do not necessarily change the conclusions on the phylogenetic relationships (see also Peters et al. 2005). However, given the high intraspecific diversity in *Microtus*, it still appears useful to assess its influence on the phylogeny, and to examine it for patterns of potentially confounding effects such as selection or recombination in nuclear markers. Our analyses provide no evidence for recombination events or sites in the two nuclear markers. There is some evidence for selection on EXON1 (Fink et al. 2007; Heckel and Fink 2008) and this may not only explain topological differences with *cytb* or AFLPs, but it might also influence phylogenetic patterns in UPSTREAM because of the physical linkage between the two loci. Nevertheless, phylogenetic trees based on these linked nuclear loci show almost the lowest similarity of all comparisons. This is not only due to differences in bipartitions within clusters of species originating from the same continent but also differences in basal nodes separating whole-continent clusters. The disagreement between the phylogenies of the physically closely linked markers thus does not simply reflect the accumulation of more mutations in the noncoding than in the coding sequence. Given this low similarity, variability in the noncoding locus with many insertions and deletions does not seem affected by selective pressure on the nearby gene. UPSTREAM provides a phylogenetic signal that is overall more similar to *cytb* and AFLPs.

Direct statistical testing for incongruence among tree topologies derived from these different data sources was complicated by the intraindividual variation in nuclear sequences. Current approaches to incongruence testing do not allow for unequal numbers of terminal branches (see Planet 2006), which prevents direct comparisons of nuclear, mitochondrial, and AFLP trees without eliminating considerable variation. Current phylogenetic studies of nuclear loci often ignore heterozygosity in DNA sequences, but heterozygosity is common and large nuclear data sets will be increasingly accessible in the near future (Levy et al. 2007). Analytical developments in incongruence testing that can accommodate for such additional levels of variation are therefore likely to gain high importance in future years particularly for the study of closely related taxa.

Generally, incongruent tree phylogenies of nuclear and mitochondrial markers are more common at the genus or species level (DeSalle and Val Giddings 1986; Seehausen 2004; Canestrelli et al. 2007) than in deep-level phylogenies (Reyes et al. 2004). Divergent phylogenies can be caused by differences in the mutation rates of the loci or in the effective population sizes due to different modes of inheritance. The latter is also a likely reason for the apparent incomplete lineage sorting in EXON1 and UPSTREAM compared with *cytb*. Additionally, the substitution rates per base differ between the markers: *cytb* and EXON1 show an excess of transitions over transversions, whereas UPSTREAM shows a low transitions to transversions bias, which might prevent misleading phylogenetic signals because of the accumulation of noninformative (mainly transitions) relative to informative (mainly transversions) mutations for basal branches (Goldstein and Pollock 1994; Hillis et al. 1996). In contrast, *cytb* shows high transitions to transversions ratios and additional signals of saturation and homoplasy, which could contribute to the low bootstrap support of the clades.

Another possible cause of cytonuclear discordance in adaptively radiated genera is introgressive hybridization (reviewed in Seehausen 2004). Ancient hybridization events at the burst of the microtine radiation should no longer interfere with the phylogenetic signal given that sufficient time has passed for the accumulation of mutations to form clearly distinct lineages consistent with species definitions. Moreover, the results from our nuclear sequence analyses would only be compatible with very old hybridization events between species occupying different continents after initial divergence, which is very unlikely for a small rodent with low dispersal capacities (see Hamilton et al. 2005; Schweizer et al. 2007; Braaker and Heckel 2009). Recent hybridization events should mainly occur between closely related species and should therefore not lead to basal differences in the topology, as seen between the different markers in this study. Additionally, for the genus *Microtus*, not many hybridizations between species have been reported (e.g., between *M. rossiaemerdionalis* and *M. arvalis* but  $F_1$  hybrids are sterile in this case; Meier et al. 1985). Introgressive hybridization is thus not a

likely explanation for incongruence of tree phylogenies between mitochondrial and nuclear markers in voles.

Given the detected partial incongruence, particular caution should be exerted when attempting to resolve the phylogeny of *Microtus* and related rodents with more sequence loci as previously suggested (Galewski et al. 2006). As expected from partially conflicting information, the joint phylogeny of the mitochondrial and the two nuclear sequence markers did not substantially improve phylogenetic resolution, although the topology was similar to the one obtained with *cytb*. This is probably because of the relative similarity of the noncoding UPSTREAM and *cytb*, which could have masked the topological changes of the coding EXON1 sequence. Partial incongruence may also explain why our analyses of a subset of species with increased length of purely nuclear sequences with additional data from GHR revealed no clear phylogenetic patterns or general increase in bootstrap support for groups of species clustering according to continents of origin. Furthermore, the short internal branches caused by rapid radiations are comparatively likely to lead to incorrect trees when loci are concatenated (Kubatko and Degnan 2007). Nuclear sequences from a very large number of loci or full-genome sequences would be needed to allow statistical examination of the contribution of these factors for difficulties in resolving the *Microtus* radiation (see Excoffier and Heckel 2006).

#### *The Utility of AFLP Markers for Resolving Phylogenetic Signals in Radiations*

Contrary to the combined sequence markers, the genome-wide approach revealed relatively high bootstrap support for major nodes in the microtine phylogeny. The AFLP phylogeny agreed with the grouping suggested by the mitochondrial *cytb* analyses, indicated by the highest similarity between the two topologies and high probability of concordance for larger clades (Table 4). Congruence of phylogenetic patterns between these different marker types is not unique to the *Microtus* radiation (Creer et al. 2004), and although the strict selective neutrality of mitochondrial genes is sometimes challenged and mutational saturation can lead to unresolved phylogenies (Ballard and Whitlock 2004), it seems parsimonious to assume that *cytb* reflects a closer approximation of the underlying species tree of *Microtus* than the nuclear sequence markers used in this study. Nevertheless, *cytb* did not provide bootstrap support for deeper splits that were supported with AFLP scans. This suggests that AFLPs may be the markers of choice for resolving relatively shallow phylogenies associated with clusters of rapid speciation events compared with single and multiple sequence loci. Because the AFLP method allows one to investigate thousands of markers across a genome, confounding effects related to the particular evolutionary history of single loci—such as selection, incomplete lineage sorting, recombina-

tion, or mutational saturation—are likely to be balanced (Albertson et al. 1999).

Specific simulations would be necessary to test whether the combination of AFLP markers could lead to anomalous gene trees (AGT) analogous to the concatenation of sequence data when internal branches of the species tree are short (see Kubatko and Degnan 2007). Under certain conditions, even a single short branch deep in the tree can lead to an AGT that may nevertheless be topologically close to the species tree (Kubatko and Degnan 2007). In our study, identical topologies of trees with single and multiple individuals per species indicate that the AGT phenomenon is unlikely to have affected AFLP results (see Degnan and Rosenberg 2006; Maddison and Knowles 2006; Kubatko and Degnan 2007).

An important consideration for the application of the genomic AFLP approach to phylogenetic questions is that evaluating homology of AFLP fragments may not always be straightforward (McKinnon et al. 2008). Computer simulations have shown that the analysis of highly divergent taxa in a phylogenetic framework may lead to the scoring of evolutionarily unrelated fragments (i.e., homoplasies, Koopman 2005; Althoff et al. 2007; McKinnon et al. 2008). Scoring of homoplasies is relatively unlikely in cases where species are closely related as in the fast radiating *Microtus* genus (see Koopman 2005). It has also been shown in silico that larger genomes are likely to yield more AFLP fragments, and consequently may lead to a higher probability of overlap particularly among shorter fragments that can result in homoplasious scoring (Althoff et al. 2007). Experimental procedures that strive to avoid homoplasy include the use of stringent primer pairs (e.g., three or four selective nucleotides) and the exclusion of fragments of very small size (Althoff et al. 2007). In our study, the exclusion of fragments shorter than an arbitrary threshold of 125 bp (15% of all markers for our data set; see also in Althoff et al. 2007) resulted in very similar phylogenies compared with the total data (results not shown). Fewer selective nucleotides in primers generally result in more amplified fragments, but practically it seems not necessary to use more than three selective nucleotides (Blears et al. 1998). Four selective nucleotides might be more likely to increase the possibility for mismatches during the amplification step than to increase selectivity of primers (Vos et al. 1995; Blears et al. 1998; own data from initial tests). In any case, initial tests of primers for AFLP experiments are necessary to ensure that they produce an appropriate number of fragments for the study question and organism. However, this step is similar to choosing the most informative locus for sequence analyses.

For larger phylogenetic studies, it seems advisable to balance AFLP primers as much as possible. For example, we used only combinations of unique primers (all eight *EcoRI* and eight *MseI* primers only used once; Table 2) to avoid potential primer-specific correlations of markers that might lead, for example, to scoring of the same fragment multiple times. Additionally, loci linked to genes under selection may not be randomly

distributed among primer combinations (Campbell and Bernatchez 2004). Therefore, our analyses of unique primer combinations should reduce potentially confounding effects arising from specific primer pairs under selection. Moreover, we used primer pairs balanced for GC or AT content in order to ensure a random coverage of the whole genome given that the mammalian genome is heterogeneous with respect to base composition in coding and noncoding regions (Galtier et al. 2001; Duret et al. 2002). Further assurance for the robustness of our AFLP results stems from the analyses of several individuals per species, which always showed monophyletic relationships for each *Microtus* species (Fig. 5).

It is important to note here that some additional technical issues are essential to consider when using AFLPs. Strict quality standards, low-tolerance laboratory procedures, and control measures (repetition of positive and negative control samples, batch processing, etc.) are more important for these markers than for standard sequencing techniques (Blears et al. 1998; Müller and LaReesa Wolfenbarger 1999). Although general reproducibility of AFLP fragments across different laboratories has been shown (Jones et al. 1997), this issue should be considered for every new study. To ensure reproducibility of our data, one individual of each species was independently reprocessed and reanalyzed for all eight primer combinations (starting from the DNA solution). The average reproducibility of these reprocessed individuals was 99.7% (each primer pair separately: 99.32–99.96%), calculated from the automatically scored output-file (binary data matrix) from GeneMapper software. The deviations from an identical match between two independent analyses of the same sample were mainly because of marginal technical variation in fragment detection that prevented automatic scoring of peaks (e.g., masking of fragments close to fragments with very strong fluorescence emission). In any case, it is advisable to check automatically scored AFLP data thoroughly before further analyses and either correct errors manually or exclude technically problematic fragments—not unlike quality control needed in sequence analyses (Bensch and Akesson 2005).

Polymorphism in our AFLP data set is relatively high compared with intraspecific analyses or some radiating species (e.g., cichlids of Lake Malawi; Albertson et al. 1999). However, other organisms also show high intra- and interspecies AFLP polymorphism (e.g., anolis lizards, 93% polymorphic markers with 3 primer pairs; Ogden and Thorpe 2002), and high polymorphism levels in *Microtus* are reflected at other markers (e.g., microsatellites; Heckel et al. 2005; Schweizer et al. 2007; Walser and Heckel 2008). Bootstrap support for the resulting AFLP phylogeny depends to some extent on the number of primer combinations investigated (see also Albertson et al. 1999). In our analyses, bootstrap support of phylogenetic trees based on AFLPs improved with increasing number of primer combinations, and the number of supported nodes has not fully reached a plateau with 8 primer pairs. This trend suggests that

resolution may increase further with more markers. In agreement with our observations, the main nodes of the Lake Malawi cichlids phylogeny were resolved with the first seven primer combinations (700 informative characters), but mean bootstrap values continued to increase as more characters were scored (up to 11 primer combinations investigated; Albertson et al. 1999). Taken together, our analyses are consistent with other studies (Albertson et al. 1999; Buntjer et al. 2002; Ogden and Thorpe 2002) that AFLPs are able to overcome some of the caveats linked to the use of sequence data at single or a few loci, particularly in the case of highly related species (see also in Althoff et al. 2007). Given proper experimentation, AFLPs are likely the markers of choice for resolving with moderate expense shallow phylogenies of closely related species that proved hitherto difficult to resolve with classical sequence markers.

#### *Phylogeographic Patterns and Biogeographic History*

The AFLP phylogeny reveals a distinct North American clade, a joint Asian and Holarctic clade, and to a much lesser extent a European clade. Although the European clade shows some internal structure, which is consistent with previously recognized subgenera (*Microtus* and *Terricola*; Jaarola and Searle 2004), some species do not cluster with the main group (*M. cabreræ*, *M. agrestis*, *M. nivalis*). The assignment of *M. nivalis* to the genus *Microtus* is still an ongoing controversy. It was originally assigned to *Microtus* (Niethammer and Krapp 1982), but it can be differentiated by morphological characteristics such as body size, skin color, molar pattern, and habitat type from other *Microtus* species (Niethammer and Krapp 1982; Musser and Carleton 1993; Mitchell-Jones et al. 1999). Fossil evidence (see in Jaarola et al. 2004) additionally sparked the discussion whether it belongs to the genus *Chionomys*. Our analyses and a previous study (Jaarola et al. 2004) both show a basal clustering of *M. nivalis* in the *Microtus* genus, but data on additional *Chionomys* species are needed to resolve the taxonomic uncertainty about this subgenus or genus. Still, branch lengths in the AFLP phylogeny are comparable for all species and in particular similar for *M. nivalis*, *M. agrestis*, and *M. cabreræ*. Contrary to *M. nivalis*, *M. agrestis*, and *M. cabreræ* were always incorporated into the *Microtus* genus. They seem to derive from very old splits having occurred early in the history of *Microtus* because they do not cluster with other European species into one main clade (see also Jaarola et al. 2004). They may thus be the trace of a first radiation in Europe, whereas the species in the subgenera *Terricola* and *Microtus* seem to derive from independent and evolutionarily younger radiations (see below).

The relationship of taxa within subgenera and especially the exact relationship of sister species and their closest relative is not unambiguously resolved. Although the *cytb* and AFLP phylogenies show relatively high similarity for major splits (Table 4), the main differences between the two topologies occur within

continent clusters. The European subgenus *Terricola* (*M. tatricus*, *M. multiplex*, *M. felteni*, and *M. thomasi*) is monophyletic in both *cytb* and AFLP phylogenies, but the sister taxa relationships within the subgenus are not resolved (e.g., *M. tatricus* is basal in the *cytb* topology but not in the AFLP phylogeny; see Figs. 3 and 5). Additionally, results for the AFLP data obtained by different phylogenetic methods do not unambiguously resolve the relationships within these groups (see unsupported nodes in Fig. 5). Similar to species clusters from Europe, closer relationships among American species are also not fully resolved. The AFLP phylogeny reveals a supported monophyletic North American cluster, but the substructure within this clade is not supported by all phylogenetic inference methods, and species do not cluster according to subgenus designation (see Fig. 5 and Table 1). Additional AFLP primer pairs would probably increase the overall support (see trend in Fig. 6a) and consequently also the support of deeper nodes. Entire mitochondrial genomes and additional nuclear markers might also provide further resolution at the subgenus level if the potential analytical problems discussed above can be solved. Further research is necessary at this systematic level.

Divergence within species might partially reflect either ongoing speciation or cryptic species in the genus. The multi-individual phylogenies show for *cytb* sequences and AFLP data clear differences among conspecifics (in *M. arvalis*, *M. agrestis*, *M. abbreviatus*, *M. longicaudus*, *M. oregoni*, *M. montebelli*, etc.). Phylogeographic analyses have shown that genetically divergent evolutionary lineages exist within several species (e.g., *M. oeconomus*, Brunhoff et al. 2003; *M. arvalis*, Fink et al. 2004; Galbreath and Cook 2004; *M. agrestis*, Jaarola and Searle 2004; Hamilton et al. 2005), and these lineages can be partially seen in our analyses. Nevertheless, some species show almost no genetic divergence between individuals despite large geographic distances between sampling sites (e.g., *M. socialis*, *M. ochrogaster*). These species may have undergone very recent range expansions and divergence might still be ongoing or too recent to be visible. More detailed analyses at the within-species levels are needed to investigate this further.

None of the three single-gene phylogenies as well as none of the earlier analyses of *cytb* (Conroy and Cook 2000a; Jaarola et al. 2004) revealed a supported main split between American and European species, but such a split is detected with AFLPs. Such a basal split would be predicted for the genus *Microtus*, given that it is derived from the fossil genus *Allophaiomys*. *Allophaiomys* is thought to have spread from southern Asia in three independent colonization processes to Central Asia, to Northern Eurasia, and to North America via the Beringian landbridge 1.2–2 Ma (Brunet-Lecomte and Chaline 1991; Chaline et al. 1999). Basal splits corresponding to different continents and colonization events have been found in other widespread mammal genera (e.g., humans, Hey 2005; hares of the genus *Lepus*, Ben Slimen et al. 2006), sometimes with evidence

for multiple colonization events (e.g., bats of the genus *Myotis*, Ruedi and Mayer 2001). Such multiple colonization events can lead to a reticulation of phylogenetic signals (Parent and Crespi 2006), making it impossible to resolve bifurcating basal splits.

Our analyses of AFLPs, in line with the other markers, clearly show that the evolutionary history of *Microtus* is better characterized by a scenario of repeated colonization events (MWH; Fig. 9) than by the hypothesis of single expansions into the western Palearctic and eastward into the Nearctic (SWH; Fig. 1). This is particularly evident for the species currently inhabiting North America. The American endemic species form a separate cluster, whereas the Holarctic species *M. oeconomus* clusters with the Asian species. Similar patterns can be seen in *cytb*, only without bootstrap support. Periodically low sea levels and land bridges between Asia and North America allowing for multiple crossings (see in Galbreath and Cook 2004) are consistent with American *M. oeconomus* populations originating from a colonization event out of Asia that is distinct from the one that gave rise to the other North American species. Additional support for this scenario comes from intraspecific analyses of mitochondrial and nuclear markers in *M. oeconomus* that revealed in the Great Beringian region 2 of 4 evolutionary lineages that may reflect different founder sources or time points of colonization (Brunhoff et al. 2003; Galbreath and Cook 2004).

We further suggest multiple independent colonization waves into Europe as the most likely evolutionary scenario for the *Microtus* species in this continent (Fig. 9). The co-occurrence of young and old lineages together with a high probability in biogeographical hypothesis testing for at least two colonization waves into Europe (*M. oeconomus* and the European main clade including *M. agrestis*), three waves (*M. oeconomus*, *M. cabreræ*, and the European main clade), or even more (*M. agrestis*) support this scenario. The exact routes over which these proceeded from the putative origin of *Microtus* in Southern Asia are currently unknown, as there is generally relatively little information about phylogeographic patterns within Asia. Dedicated phylogenetic and phylogeographic studies of Asian species incorporating fossil evidence wherever possible are needed to evaluate the number and timing of *Microtus* colonizations across this continent. An extended coverage of this region may not only provide deeper insights into the onset and mode of the explosive radiation of the *Microtus* genus but also more precise information on its place of origin.

## CONCLUSIONS

Single sequence marker phylogenies showed low bootstrap support for deeper nodes in the *Microtus* genus and some incongruence, possibly because of different modes of evolution and marker-specific properties. Neither the combination of different markers nor longer nuclear sequences have been able to resolve the complex evolutionary history of this genus.

In contrast, the genome-wide analysis of a large number of AFLP markers led to much higher support for major clades providing strong evidence for several independent colonization waves of the small *Microtus* rodents into North America as well as into Europe. We conclude that carefully applied AFLP markers can be very powerful tools for phylogenetic analyses particularly at lower taxonomic levels. The AFLP approach should therefore find wider application in the future to resolve complex evolutionary patterns in closely related species.

#### SUPPLEMENTARY MATERIAL

Supplementary material can be found at <http://www.sysbio.oxfordjournals.org/>.

#### FUNDING

Support for our study was from the University of Bern, the Berne University Research Foundation and Swiss NSF grant 3100A0-112072 and 3100-126074 to L.E. and G.H.

#### ACKNOWLEDGMENTS

We thank I. Keller, J. Sullivan, B. Riddle, S. Creer, and an anonymous reviewer for very helpful comments on the manuscript, and S. Braaker, S. Tellenbach, and D. Wegmann for technical assistance. We also thank W. Durka, C. Thiel-Egenter, F. Gugerli, and E. Bezaul for helpful advice for the establishment of AFLPs. We are grateful to the following people for providing access to samples: A. Bannikova, S. Braaker, R. Burri, F. Catzeflis, C. Conroy, B. Cushing, T. Derting, M. Jaarola, T. Maddalena, N. Martinkova, J. Michaux, J.-P. Müller, M. Pfunder, R. Pita, J. Suchomel, J. Robovsky, J. Runge, L. Vinciguerra, P. Vogel, and the Museum of Vertebrate Zoology of the University of California.

#### REFERENCES

- Akaike H. 1974. A new look at the statistical model identification. *IEEE T. Automat. Contr.* 19:716–723.
- Albertson R.C., Markert J.A., Danley P.D., Kocher T.D. 1999. Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa. *Proc. Natl. Acad. Sci. U S A.* 96:5107–5110.
- Althoff D.M., Gitzendanner M.A., Segraves K.A. 2007. The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. *Syst. Biol.* 56:477–484.
- Ane C., Larget B., Baum D.A., Smith S.D., Rokas A. 2007. Bayesian estimation of concordance among gene trees. *Mol. Biol. Evol.* 24:412–426.
- Ballard J.W.O., Whitlock M.C. 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13:729–744.
- Bensch S., Akesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Mol. Ecol.* 14:2899–2914.
- Ben Slimen H., Suchentrunk F., Memmi A., Sert H., Kryger U., Alves P.C., Ben Ammar Elgaaied A. 2006. Evolutionary relationships among hares from North Africa (*Lepus* sp. or *Lepus* spp.), cape hares (*L. capensis*) from South Africa and brown hares (*L. europaeus*), as inferred from mtDNA, PCR-RFLP and allozyme data. *J. Zool. Syst. Evol. Res.* 44:88–99.
- Bleas M., De Grandis S., Lee H., Trevor J. 1998. Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *J. Ind. Microbiol. Biot.* 21:99–114.
- Braaker S., Heckel G. 2009. Transalpine colonisation and partial phylogeographic erosion by dispersal in the common vole *Microtus arvalis*. *Mol. Ecol.* 18:2518–2531.
- Brunet-Lecomte P., Chaline J. 1991. Morphological evolution and phylogenetic relationships of the European ground voles (Arvicolinae, rodentia). *Lethaia.* 24:45–53.
- Brunhoff C., Galbreath K.E., Fedorov V.B., Cook J.A., Jaarola M. 2003. Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late Quaternary biogeography of high latitudes. *Mol. Ecol.* 12:957–968.
- Buntjer J.B., Otsen M., Nijman I.J., Kuiper M.T.R., Lenstra J.A. 2002. Phylogeny of bovine species based on AFLP fingerprinting. *Heredity.* 88:46–51.
- Campbell, D., Bernatchez L. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Mol. Biol. Evol.* 21: 945–956.
- Canestrelli D., Verardi A., Nascetti G. 2007. Genetic differentiation and history of populations of the Italian treefrog *Hyla intermedia*: lack of concordance between mitochondrial and nuclear markers. *Genetica.* 130:241–255.
- Carstens B.C., Brunsfeld S.J., Demboski J.R., Good J.M., Sullivan J. 2005. Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. *Evolution* 59:1639–1652.
- Chaline J., Brunet-Lecomte P., Montuire S., Viriot L., Courant F. 1999. Anatomy of the arvicoline radiation (Rodentia): palaeogeographical, palaeoecological history and evolutionary data. *Ann. Zool. Fenn.* 36:239–267.
- Conroy C.J., Cook J.A. 2000a. Molecular systematics of a Holarctic rodent (*Microtus*: Muridae). *J. Mammal.* 81:344–359.
- Conroy C.J., Cook J.A. 2000b. Phylogeography of a post-glacial colonizer: *Microtus longicaudus* (Rodentia: Muridae). *Mol. Ecol.* 9:165–175.
- Creer S., Thorpe R.S., Malhotra A., Chou W.-H., Stenson A.G. 2004. The utility of AFLPs for supporting mitochondrial DNA phylogeographical analyses in the Taiwanese bamboo viper, *Trimeresurus stejnegeri*. *J. Evol. Biol.* 17:100–107.
- Creer S., Pook C.E., Malhotra A., Thorpe R.S. 2006. Optimal intron analyses in the trimeresurus radiation of Asian pitvipers. *Syst. Biol.* 55:57–72.
- Danley P.D., Kocher T.D. 2001. Speciation in rapidly diverging systems: lessons from Lake Malawi. *Mol. Ecol.* 10:1075–1086.
- Degnan J.H., Rosenberg N.A. 2006. Discordance of species trees with their most likely gene trees. *PLoS Genet.* 2:762–768.
- DeSalle R., Val Giddings L. 1986. Discordance of nuclear and mitochondrial phylogenies in Hawaiian *Drosophila*. *Proc. Natl. Acad. Sci. U S A.* 83:6902–6906.
- Duret L., Semon M., Piganeau G., Mouchiroud D., Galtier N. 2002. Vanishing GC-rich isochores in mammalian genomes. *Genetics.* 162:1837–1847.
- Estabrook G.F., McMorris F.R., Meacham C.A. 1985. Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. *Syst. Zool.* 34:193–200.
- Excoffier L., Heckel G. 2006. Computer programs for population genetics data analysis: a survival guide. *Nat. Rev. Genet.* 7:745–758.
- Excoffier L., Laval G., Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinf. Online* 1:47–50.
- Felsenstein, J. 1992. Phylogenies from restriction sites: a maximum-likelihood approach. *Evolution.* 46:159–173.
- Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Seattle (WA): Department of Genome Sciences, University of Washington.
- 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. 2006. Mammalian monogamy is not controlled by a single gene. *Proc. Natl. Acad. Sci. U S A.* 103:10956–10960.

- Fink S., Excoffier L., Heckel G. 2007. High variability and non-neutral evolution of the mammalian *avpr1a* gene. *BMC Evol. Biol.* 7:176.
- Galbreath K.E., Cook J.A. 2004. Genetic consequences of Pleistocene glaciations for the tundra vole (*Microtus oeconomus*) in Beringia. *Mol. Ecol.* 13:135–148.
- Galewski T., Tilak M.-K., Sanchez S., Chevret P., Paradis E., Douzery E. 2006. The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies. *BMC Evol. Biol.* 6:80.
- Galtier N., Piganeau G., Mouchiroud D., Duret L. 2001. GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. *Genetics.* 159:907–911.
- Giannasi N., Thorpe R.S., Malhotra A. 2001. The use of amplified fragment length polymorphism in determining species trees at fine taxonomic levels: analysis of a medically important snake, *Trimeresurus albolabris*. *Mol. Ecol.* 10:419–426.
- Giannini N.P., Simmons N.B. 2003. A phylogeny of megachiropteran bats (Mammalia: Chiroptera: Pteropodidae) based on direct optimization analysis of one nuclear and four mitochondrial genes. *Cladistics.* 19:496–511.
- Gibbs M.J., Armstrong J.S., Gibbs A.J. 2000. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics.* 16.
- Goldstein D.B., Pollock D.D. 1994. Least squares estimation of molecular distance-noise abatement in phylogenetic reconstruction. *Theor. Popul. Biol.* 45:219–226.
- Griffiths A.J.F., Miller J.H., Suzuki D.T., Lewontin R.C., Gelbart W.M. 2000. An introduction to genetic analysis. 7 ed. New York: Freeman.
- Hall T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41:95–98.
- Hamilton G., Currat M., Ray N., Heckel G., Beaumont M., Excoffier L. 2005. Bayesian estimation of recent migration rates after a spatial expansion. *Genetics.* 170:409–417.
- Heckel G., Burri R., Fink S., Desmet J.-F., Excoffier L. 2005. Genetic structure and colonization processes in European populations of the common vole *Microtus arvalis*. *Evolution.* 59:2231–2242.
- Heckel G., Fink S. 2008. Evolution of the arginine vasopressin 1a receptor and implications for mammalian social behavior. In: Landgraf R., Neumann I.D. editors. *Advances in vasopressin and oxytocin—from gene to behavior*. Oxford: Elsevier B. V., . *Prog. Brain Res.* 170:321–330.
- Hellborg L., Gunduz I., Jaarola M. 2005. Analysis of sex-linked sequences supports a new mammal species in Europe. *Mol. Ecol.* 14:2025–2031.
- Hey J. 2005. On the number of new world founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol.* 3:965–975.
- Hillis D.M., Moritz C., Mable B.K. 1996. *Molecular Systematics*. Sunderland (MA): Sinauer Associates.
- Huelsenbeck J.P., Ronquist F. 2001. MRBAYES: Bayesian phylogenetic inference. *Bioinformatics.* 17.
- Jaarola M., Martinkova N., Gunduz I., Brunhoff C., Zima J., Nadachowski A., Amori G., Bulatova N.S., Chondropoulos B., Fragedakis-Tsolis S., Gonzalez-Esteban J., Lopez-Fuster M.J., Kandaurov A.S., Kefelioglu 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. 2004. A highly divergent mitochondrial DNA lineage of *Microtus agrestis* in southern Europe. *Heredity.* 92:228–234.
- Jansa S.A., Barker F.K., Heaney L.R. 2006. The pattern and timing of diversification of Philippine endemic rodents: evidence from mitochondrial and nuclear gene sequences. *Syst. Biol.* 55:73–88.
- Jones C.J., Edwards K.J., Castaglione S., Winfield M.O., Sala F., van de Wiel C., Bredemeijer G., Vosman B., Matthes M., Daly A., Bretschneider R., Bettini P., Buiatti M., Maestri E., Malcevski A., Marmioli N., Aert R., Volckaert G., Rueda J., Linacero R., Vazquez A., Karp A. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol. Breed.* 3:381–390.
- Jordan S., Simon C., Polhemus D. 2003. Molecular systematics and adaptive radiation of Hawaii's endemic damselfly genus *Megalagrion* (Odonata: Coenagrionidae). *Syst. Biol.* 52:89–109.
- Jukes T., Cantor C. 1969. Evolution of protein molecules. In: Munro, H. N. editors. *Mammalian protein metabolism*. New York: Academic Press. p. 21–132.
- Kimura M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci. U S A.* 78:454–458.
- Kingston S.E., Rosel P.E. 2004. Genetic differentiation among recently diverged delphinid taxa determined using AFLP markers. *J. Hered.* 95:1–10.
- Koopman W.J.M. 2005. Phylogenetic signal in AFLP data sets. *Syst. Biol.* 54:197–217.
- Kubatko L.S., Degnan J.H. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* 56:17–24.
- Kumar S., Tamura K., Nei M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinformatics.* 5:150–163.
- Lanave C., Preparata G., Saccone C., Serio G. 1984. A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* 20:86–93.
- Levy S., Sutton G., Ng P.C., Feuk L., Halpern A.L., Walenz B.P., Axelrod N., Huang J., Kirkness E.F., Denisov G., Lin Y., MacDonald J.R., Pang A.W.C., Shago M., Stockwell T.B., Tsiamouri A., Bafna V., Bansal V., Kravitz S.A., Busam D.A., Beeson K.Y., McIntosh T.C., Remington K.A., Abril J.F., Gill J., Borman J., Rogers Y.-H., Frazier M.E., Scherer S.W., Strausberg R.L., Venter J.C. 2007. The diploid genome sequence of an individual human. *PLoS Biol.* 5:e254.
- Maddison W.P., Knowles L.L. 2006. Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* 55:21–30.
- Martin D., Rybicki E. 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics.* 16:562–563.
- Martin D.P., Williamson C., Posada D. 2005. RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics.* 21:260–262.
- Maynard Smith, J. 1992. Analyzing the mosaic structure of genes. *J. Mol. Evol.* 34:126–129.
- McKinnon G.E., Vaillancourt R.E., Steane D.A., Potts B.M. 2008. An AFLP marker approach to lower-level systematics in *Eucalyptus* (Myrtaceae). *Am. J. Bot.* 95:368–380.
- Meier M.N., Radjabli S.I., Bulatova N.S., Golenishchev F.N. 1985. Karyological peculiarities and probable relations of common voles of the group *arvalis* (Rodentia, Cricetidae, *Microtus*). *Zool. Zhurnal* 64:417–428.
- Mitchell-Jones A.J., Amori G., Bogdanowicz W., Krystufek B., Reijnders P.J.H., Spitzenberger F., Stubbe M., Thissen J.B.M., Vohralik V., Zima J. 1999. *The atlas of European mammals*. London: T. Poyser, A. D. Poyser.
- Modi W.S. 1987. Phylogenetic analyses of chromosomal banding patterns among the nearctic arvicolidae (Mammalia: Rodentia). *Syst. Zool.* 36:109–136.
- Musser G.G., Carleton M.D. 1993. Family Muridae. In: Wilson D.E., Reeder D.M., editors. *Mammal species of the world: a taxonomic and geographic reference*. Washington DC: Smithsonian Institution Press. p. 501–756.
- Müller, U.G., LaReesa Wolfenbarger L. 1999. AFLP genotyping and fingerprinting. *Trends Ecol. Evol.* 14:389–394.
- Nei M., Li W.-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U S A.* 76:5269–5273.
- Nichols R. 2001. Gene trees and species trees are not the same. *Trends Ecol. Evol.* 16:358–364.
- Niethammer J., Krapp F. 1982. *Handbuch der Säugetiere Europas*. Wiesbaden (Germany): Akademische Verlagsgesellschaft.
- Nowak R.M. 1991. *Walker's mammals of the world*. 5th ed. Baltimore (MD): Johns Hopkins University Press.
- Nylander J.A.A., Wilgenbusch J.C., Warren D.L., Swofford D.L. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics.* 24:581–583.
- Ogden R., Thorpe R.S. 2002. The usefulness of amplified fragment length polymorphism markers for taxon discrimination across graduated fine evolutionary levels in Caribbean *Anolis* lizards. *Mol. Ecol.* 11:437–445.

- Padidam M., Sawyer S., Fauquet C.M. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology*. 265:218–225.
- Page R.D.M. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357–358.
- Parent C.E., Crespi B.J. 2006. Sequential colonization and diversification of Galapagos endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora). *Evolution*. 60:2311–2328.
- Peters J.L., McCracken K.G., Zhuravlev Y.N., Lu Y., Wilson R.E., Johnson K.P., Omland K.E. 2005. Phylogenetics of wigeons and allies (Anatidae: *Anas*): the importance of sampling multiple loci and multiple individuals. *Mol. Phylogenet. Evol.* 35:209–244.
- Planet P.J. 2006. Tree disagreement: measuring and testing incongruence in phylogenies. *J. Biomed. Inform.* 39:86–102.
- Posada D. 2002. Evaluation of methods for detecting recombination in DNA sequences: empirical data. *Mol. Biol. Evol.* 19:708–717.
- Posada D., Crandall K.A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics*. 14:817–818.
- Posada D., Crandall K.A. 2001. Evaluation of methods for detecting recombination in DNA sequences: computer simulations. *Proc. Natl. Acad. Sci. U S A*. 98:13757–13762.
- Puigbo P., Garcia-Vallve S., McInerney J.O. 2007. TOPD/FMTS: a new software to compare phylogenetic trees. *Bioinformatics*. 23:1556–1558.
- Reyes A., Gissi C., Catzeflis F.M., Nevo E., Pesole G., Saccone C. 2004. Congruent mammalian trees from mitochondrial and nuclear genes using Bayesian methods. *Mol. Biol. Evol.* 21:397–403.
- Rodriguez F., Oliver J.L., Marin A., Medina J.R. 1990. The general stochastic model of nucleotide substitution. *J. Theor. Biol.* 142:485–501.
- Ronquist F., Huelsenbeck J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 19:1572–1574.
- Rozas J., Sanchez-DelBarrio J.C., Messeguer X., Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*. 19:2496–2497.
- Ruedi M., Mayer F. 2001. Molecular systematics of bats of the genus *Myotis* (Vespertilionidae) suggests deterministic ecomorphological convergences. *Mol. Phylogenet. Evol.* 21:436–448.
- Saitou N., Nei M. 1987. The neighbour joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- Salminen M.O., Carr J.K., Burke D.S., McCutchan F.E. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by Bootscanning. *AIDS Res. Hum. Retroviruses*. 11:1423–1425.
- Sambrook J., Fritsch E.F., Maniatis T. 1989. *Molecular cloning: a laboratory manual*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
- Schmidt-Chanasit J., Essbauer S.S., Petraityte R., Yoshimatsu K., Tackmann K., Conraths F.J., Sasnauskas K., Arikawa J., Thomas A., Pfeffer M., Scharninghausen J.J., Spletstosser W., Wenk M., Heckel G., Ulrich R. 2010. Extensive host sharing of Central European Tula virus. *J. Virol.* 84:459–474.
- Schweizer M., Excoffier L., Heckel G. 2007. Fine-scale genetic structure and dispersal patterns in the common vole *Microtus arvalis*. *Mol. Ecol.* 16:2463–2473.
- Sato A., Tichy H., O’Hugin C., Grant P.R., Grant B.R., Klein J. 2001. On the origin of Darwin’s finches. *Mol. Biol. Evol.* 18:299–311.
- Seehausen O. 2004. Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19:198–207.
- Springer M.S., Stanhope M.J., Madsen O., de Jong W.W. 2004. Molecules consolidate the placental mammal tree. *Trends Ecol. Evol.* 19:430–438.
- Sullivan J.P., Lavoué S., Arnegard M.E., Hopkins C.D. 2004. AFLPs resolve phylogeny and reveal mitochondrial introgression within a species flock of African electric fish (Mormyroidea: Teleostei). *Evolution*. 58:825–841.
- Suzuki H., Shimada T., Terashima M., Tsuchiya K., Aplin K. 2004. Temporal, spatial, and ecological modes of evolution of Eurasian *Mus* based on mitochondrial and nuclear gene sequences. *Mol. Phylogenet. Evol.* 33:626–646.
- Swofford D.L. 1999. PAUP\*. Phylogenetic analysis using parsimony (\*and other methods). version 4. Sunderland (MA): Sinauer Associates.
- Tamarin R.H. 1985. *Biology of New World Microtus*. Shippensburg (PA): American Society of Mammalogists.
- Tavaré S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. *Lect. Math. Life Sci.* 17:57–86.
- Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876–4882.
- Vos P., Hogers R., Bleeker M., Reijans M., Vandeele T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. 1995. AFLP—a new technique for DNA-fingerprinting. *Nucleic Acids Res.* 23:4407–4414.
- Walser B., Heckel G. 2008. Microsatellite markers for the common vole (*Microtus arvalis*) and their cross-species utility. *Cons. Genet.* 9:479–481.
- Wilson D.E., Reeder D.-A.M. editors. 2005. *Mammal Species of the World. A Taxonomic and Geographic Reference*. Baltimore (MA): Johns Hopkins University Press.
- Young L.J., Nilsen R., Waymire K.G., MacGregor G.R., Insel T.R. 1999. Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. *Nature*. 400:766–768.