Talpid Mole Phylogeny Unites Shrew Moles and Illuminates Overlooked Cryptic Species Diversity

Kai He,[‡,†,1,2] Akio Shinohara,[†,3] Kristofer M. Helgen,[4] Mark S. Springer,[5] Xue-Long Jiang,[*,1] and Kevin L. Campbell[*,2]

[1]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

[2]Department of Biological Sciences, University of Manitoba, Winnipeg, MN , Canada

[3]Department of Bio-resources, Division of Biotechnology, Frontier Science Research Center, University of Miyazaki, Miyazaki, Japan

[4]National Museum of Natural History Smithsonian Institution, Washington, DC

[5]Department of Biology, University of California, Riverside, CA

[‡]Present address: The Kyoto University Museum, Kyoto University, Kyoto, Japan

[†]These authors contributed equally to this work.

***Corresponding authors:** E-mails: jiangxl@mail.kiz.ac.cn; kevin.campbell@umanitoba.ca

Associate editor: Emma Teeling

Abstract

The mammalian family Talpidae (moles, shrew moles, desmans) is characterized by diverse ecomorphologies associated with terrestrial, semi-aquatic, semi-fossorial, fossorial, and aquatic-fossorial lifestyles. Prominent specializations involved with these different lifestyles, and the transitions between them, pose outstanding questions regarding the evolutionary history within the family, not only for living but also for fossil taxa. Here, we investigate the phylogenetic relationships, divergence times, and biogeographic history of the family using 19 nuclear and 2 mitochondrial genes (~16 kb) from ~60% of described species representing all 17 genera. Our phylogenetic analyses help settle classical questions in the evolution of moles, identify an ancient (mid-Miocene) split within the monotypic genus *Scaptonyx*, and indicate that talpid species richness may be nearly 30% higher than previously recognized. Our results also uniformly support the monophyly of long-tailed moles with the two shrew mole tribes and confirm that the Gansu mole is the sole living Asian member of an otherwise North American radiation. Finally, we provide evidence that aquatic specializations within the tribes Condylurini and Desmanini evolved along different morphological trajectories, though we were unable to statistically reject monophyly of the strictly fossorial tribes Talpini and Scalopini.

Key words: **Talpidae, tree of life, cryptic species, aquatic, fossorial.**

Article

Introduction

Members of the mammalian family Talpidae—the moles, shrew moles, and desmans—include some of the least studied mammals on Earth. The 43 recognized living species reflect a surprisingly diverse set of lifestyles including terrestrial, semi-fossorial, semi-aquatic, fossorial, and aquatic-fossorial (Nowak 1999; Hutterer 2005). The evolution of adaptive specializations to these diverse environments produced substantial morphological variation, for example in skeletal anatomy (Freeman 1886) and sensory systems (Catania 2000). However, increasingly derived "stepping-stone" phenotypic adaptations can be identified across a morphological gradient from the terrestrial shrew-like moles (Uropsilinae), to the semi-fossorial shrew moles/long-tailed moles (Neurotrichini, Urotrichini, and Scaptonychini), and further to the fossorial "true moles" (Talpini and Scalopini). Conversely, the more aquatic members of the family, the desmans (Desmanini) and the aquatic-fossorial star-nosed mole (Condylurini), are among the most unique and morphologically isolated of living mammals.

Prominent locomotory adaptations involved with the varied talpid lifestyles have attracted considerable interest from natural historians and evolutionary biologists for more than a century (Shimer 1903; Edwards 1937; Reed 1951; Hutchison 1974; Yates and Moore 1990; Whidden 2000; Piras et al. 2012; Meier et al. 2013). Biologists have especially been fascinated by: (i) whether the aquatic forms (Condylurini and Desmanini) derived from semi-fossorial/fully fossorial moles, or *vice versa*; and (ii) whether the striking morphological similarity of the two strictly fossorial "true mole" tribes (Talpini and Scalopini) reflects synapomorphy or evolved convergently from an ancestral shrew-like morphology. This latter dispute is particularly topical in light of a recent molecular hypothesis (Bannikova, Zemlemerova, Lebedev, et al. 2015) that strongly refutes the sister-taxa association consistently recovered between the two fossorial tribes in anatomical studies (Whidden 2000; Motokawa 2004; Sanchez-Villagra et al. 2006; Schwermann and Thompson 2015). Additional outstanding systematic questions that have attracted vigorous debate involve the phylogenetic placement of key

taxonomic and morphological lineages, including whether the Chinese endemic Gansu mole (*Scapanulus*) is a shrew mole (Van Valen 1967) or a member of the North American true mole tribe Scalopini (Hutchison 1968; Kawada, Li, et al. 2008; Bannikova, Zemlemerova, Lebedev, et al. 2015; Schwermann and Thompson 2015), and whether the Southeast Asian endemic long-tailed mole (*Scaptonyx*) is more closely related to fossorial moles or to the trans-Pacific shrew mole tribes (Hutchison 1968). Numerous phylogenetic hypotheses have been proposed for the family based on different character sets, including osteological, myological, and molecular data (Hutchison 1976; Yates and Moore 1990; Corbet and Hill 1992; Whidden 2000; Motokawa 2004; Shinohara et al. 2004; Cabria et al. 2006; Sanchez-Villagra et al. 2006; Bannikova, Zemlemerova, Lebedev, et al. 2015; Schwermann and Thompson 2015), but results from these disparate comparisons have many conflicts, with subfamily and inter-tribal relationships in particular lacking strong statistical support.

Conflicting taxonomic classifications regarding the establishment and composition of subfamilies and tribes have been proposed for the talpid family (Hutchison 1968; Yates 1984; McKenna and Bell 1997; Hutterer 2005), with substantial confusion also surrounding the phylogenetic placement and taxonomic assignment of talpids in the fossil record (Ziegler 2003, 2012; Klietmann et al. 2015). Because of the robustness of the forelimb in some talpids, relatively plentiful numbers of humeri are preserved as fossils for some lineages (e.g., Talpini, Scalopini, Urotrichini), and can provide insights into locomotory adaptations through geological time. However, the resolution of locomotory adaptations through time requires a well-resolved phylogeny, which is currently not available.

Herein, we investigated the phylogenetic relationships, divergence times, and biogeographic history of the family Talpidae using comprehensive molecular approaches that incorporated data from 19 nuclear (~14 kb) and 2 mitochondrial genes (~2 kb) from 52 specimens, representing all 17 genera and ~60% (25 of 43) of recognized talpid species (supplementary table S1, Supplementary Material online). To further access potential overlooked species diversity (Wan et al. 2013; He et al. 2014; Bannikova, Zemlemerova, Colangelo, et al. 2015; Feuda et al. 2015), we also constructed a maximum-likelihood tree using an expanded supermatrix that included published sequence data from 15 additional described talpid species. Our results provide a robust scaffold for systematic studies of living and fossil species, reveal the tempo of talpid diversification in the context of the Paleocene extinction events, identify up to 12 undescribed cryptic species across 6 genera, and offer key insights to major questions regarding the evolutionary history, morphological transitions, and systematics of talpid moles.

Results

Phylogenetic Relationships and Species Diversity

Our concatenated gene tree (21 genes) indicates that the basal split in Talpidae is between the subfamily Uropsilinae and the seven recognized mole tribes (PP = 1.0, BS = 100) (fig. 1). Although the fossorial tribe Scalopini is strongly supported as sister to the remaining six tribes in our Bayesian analysis (PP = 1.0), this relationship was only weakly supported in our ML tree (BS = 65). Similar discrepancies between these analyses were recovered regarding the monophyly of the aquatic desmans (tribe Desmanini) with the aquatic-fossorial Condylurini (PP = 0.97, BS = 25), and their grouping with the strictly fossorial tribe Talpini (PP = 0.99, BS = 47). While the latter clade was recovered when the mitochondrial genes were removed from the dataset (PP = 0.89), the monophyly of Condylurini and Desmanini was not supported (supplementary fig. S1A, Supplementary Material online). Conversely, all concatenated analyses provided strong support (PP = 1.0, BS = 100) for: (i) a sister-taxa relationship between the Gansu mole *Scapanulus oweni* and *Parascalops breweri* within the North American Scalopini, (ii) a monophyletic clade that unites long-tailed moles (tribe Scaptonychini) with the shrew mole tribes Neurotrichini and Urotrichini, and (iii) the placement of *Euroscaptor mizura* sister to the genus *Mogera*. Notably, both analyses also identified two long-tailed mole (*Scaptonyx fusicaudus*) lineages with strikingly long branches (fig. 1, Supplementary fig. S1A, Supplementary Material online).

The coalescent species tree estimations were highly congruent with the concatenated trees at the tribal level (fig. 2 and supplementary fig. S1*B* and *C*, Supplementary Material online), though most tribal interrelationships, including the positions of the star-nosed mole and the desmans, were not resolved. In addition, while *Scapanulus* was unambiguously placed within Scalopini, its sister-relationship with *Parascalops* only received modest support (BS = 49–72). Conflicts among the coalescent trees included the interrelationships among Southeast Asian *Euroscaptor*, and between long-tailed moles and the two shrew-mole tribes, though support scores for these alternative nodes were low (BS = 49–74). SVDquartets further placed *Euroscaptor parvidens* sister to a clade containing other Southeast Asian *Euroscaptor* plus *Parascaptor* + *Scaptochirus* (fig. 2B), resulting in a polyphyletic *Euroscaptor* (BS = 31).

As bootstrap support scores for inter-tribal nodes were often <50%, we conducted three separate parsimony analyses to statistically test the monophyly versus diphyly of Scalopini + Talpini. These tests revealed no significant differences between the two hypotheses. Six trees at 18,707 steps were recovered for the diphyly hypothesis and are only four steps shorter than three trees at 18,711 steps that were recovered when Scalopini + Talpini monophyly was enforced ($P > 0.53$; supplementary table S2, Supplementary Material online).

Species Diversification through Time and Space

The species delimitation analyses provided strong support for undescribed cryptic species in the genera *Uropsilus*, *Euroscaptor*, and *Parascaptor*, and identified 33 distinct lineages (confidence interval [CI] = 30–37) in our concatenated 21-gene dataset. Moreover, our analysis with an expanded supermatrix, which included ~93% of accepted species diversity, suggests that talpid species richness may be nearly 30%

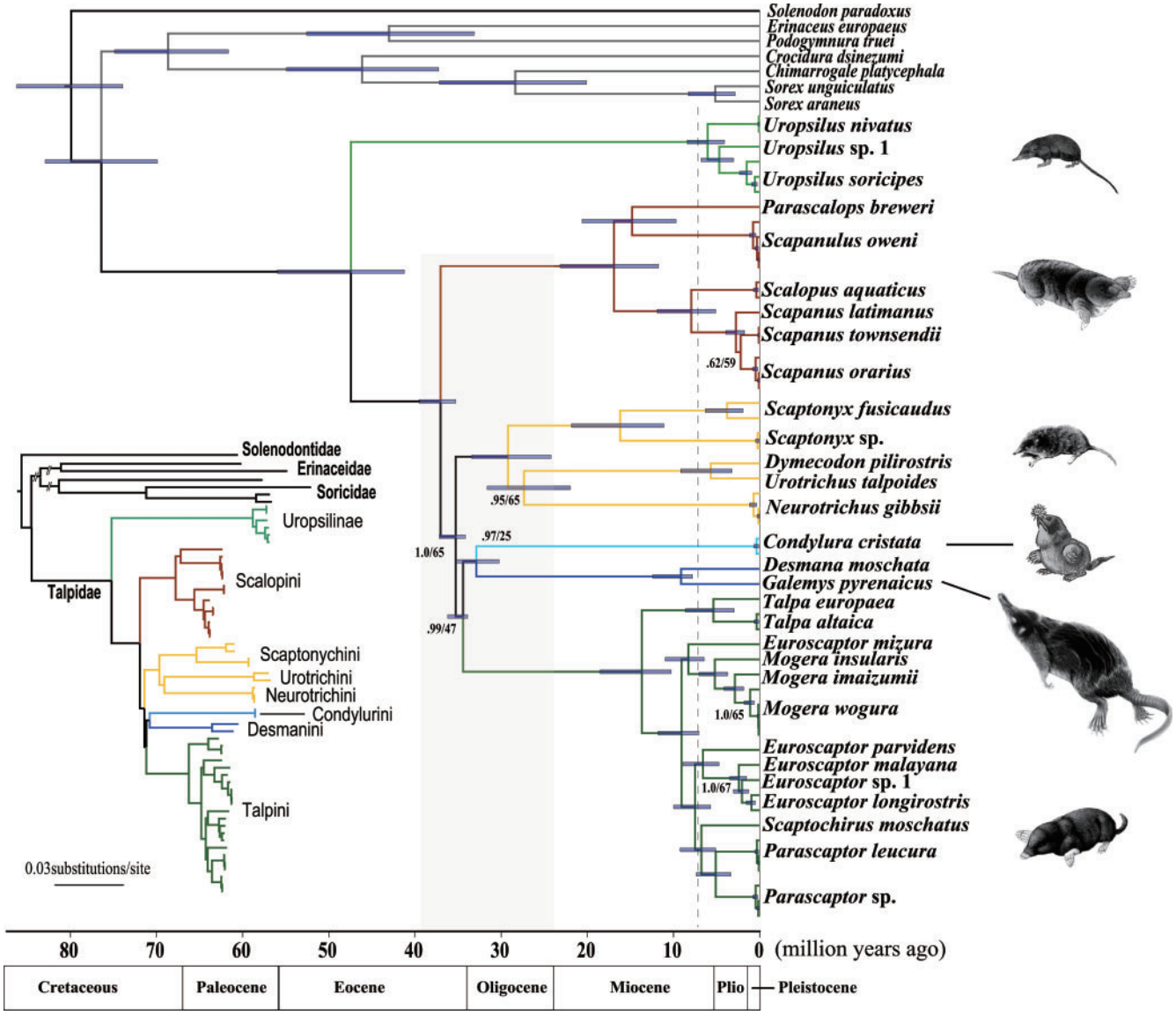

FIG. 1. Time calibrated Bayesian species tree of the family Talpidae based on a concatenated alignment of 19 nuclear and 2 mitochondrial genes. Branch lengths represent time and node bars indicate the 95% CI for each clade age. All relationships are highly supported (posterior probabilities [PP] = 1.00 and bootstrap values [BS] $\geq$ 95), otherwise support values are provided. Inset: Branch lengths represent substitutions per site.

higher than previously recognized (i.e., 55 vs. 43 species) as 12 uncharacterized lineages across 6 genera were identified (fig. 3; supplementary table S1, Supplementary Material online).

Divergence time estimates suggest that the most recent common ancestor of the family dates from the Eocene boundary at ca. 47 million years ago (Ma) with a 95% CI of 57–42 Ma (fig. 1). Tribal divergences commenced during the Priabonian Stage (39.5–33.9 Ma) of the late Eocene when Scalopini diverged from other tribes at 37.0 Ma (95% CI = 39.5–35.3 Ma), and continued through the Rupelian (33.9–28.1 Ma) and Chattian (28.1–23.03 Ma) stages of the Oligocene when the three shrew mole tribes (Scaptonychini, Urotrichini, Neurotrichini) diverged from each other between 30 and 27 Ma (95% CI = 34–23 Ma). Divergences between Chinese and North American scalopins, between European and Asian talpins, and between the two long-tailed mole lineages all occurred in the mid-Miocene (ca. 16 Ma; 95% CI = 22–10 Ma). Finally, divergences between sister genera generally occurred between 18 and 7 My (95% CI = 26–4 Ma). The lineage-through-time (LTT) plot (supplementary fig. S2A, Supplementary Material online) showed an increasing diversification rate beginning ca. 10 Ma. However, the Pybus and Harvey's γ-value was insignificantly negative (−0.70; $P = 0.24$), while the likelihood-based test found pure-birth as the best constant rate model (AIC = 68.67) and the yule-3-rate as the best rate variable model (AIC = 67.94), suggesting that speciation rate did not shift significantly through time. The fossil-diversity plot (supplementary fig. S2B, Supplementary Material online) demonstrated that the number of semi-fossorial and nonfossorial species have substantially declined in the last four million years.

The results of biogeographic inference analyses with and without fossil taxa were discordant (fig. 4A and B). For example, the extant-species-only analysis supported an Asian origination for Talpidae, while the origin of the family is uncertain when fossil taxa are incorporated. When incorporating fossils,

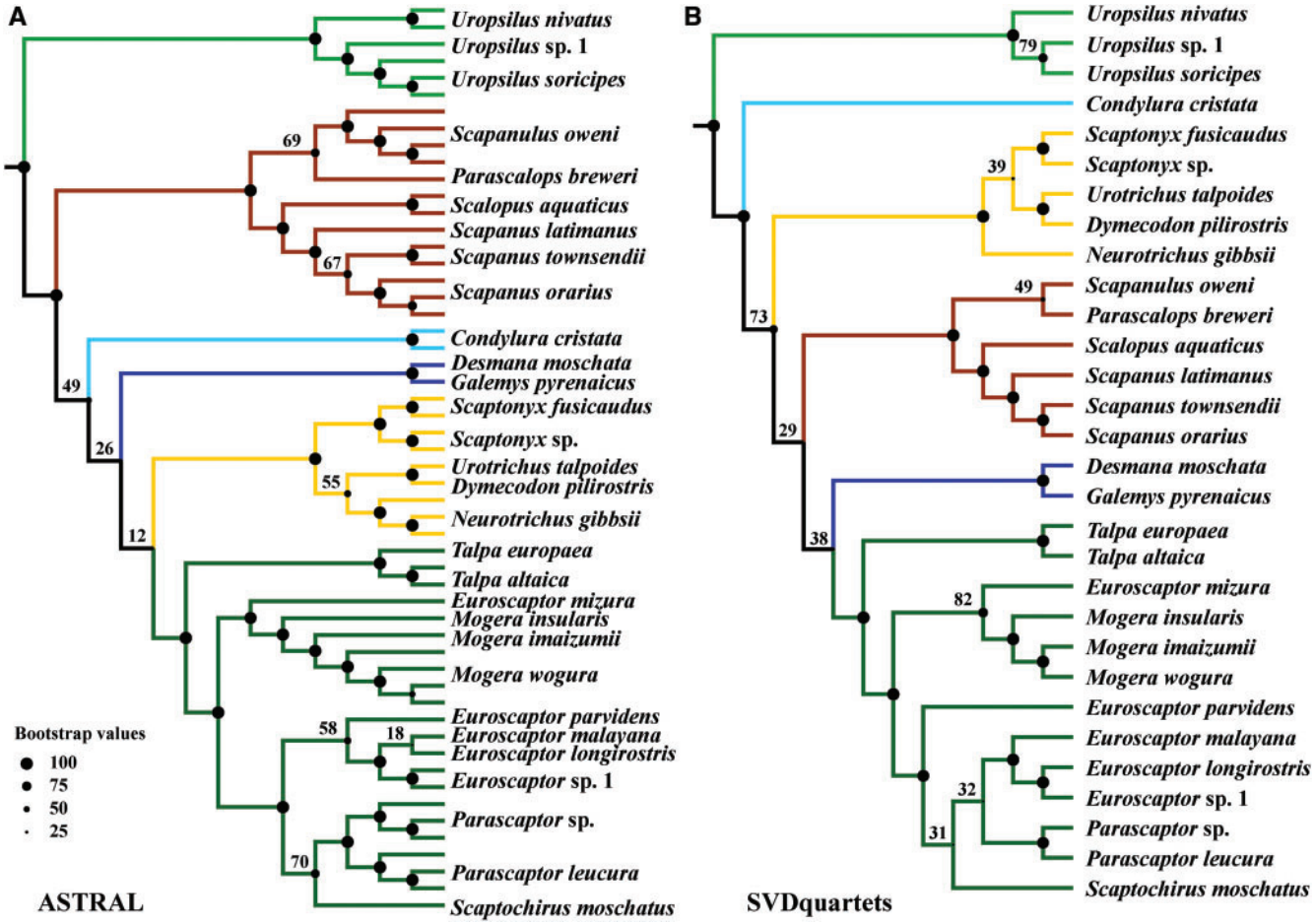

FIG. 2. Bootstrap coalescent species trees of the family Talpidae estimated using (*A*) ASTRAL-II, and (*B*) SVDquartets. Support values are provided for all non-highly supported nodes (posterior probabilities [PP] <0.95 or bootstrap values [BS] <90).

the analysis supported a North American origin of Scalopini, and European origins of Talpini as well as the ancestor of Condylurini + Desmanini + Talpini. The extant-species-only analysis did not strongly support biogeographic origins for any of these clades. The two analyses also supported different origination scenarios for the ancestor of the long-tailed mole and the shrew moles.

Morphology-Based Trees

We also explored phylogenetic relationships based on 175 morphological characters (Schwermann and Thompson 2015), and recovered nine most parsimonious (MP) trees with tree lengths of 569 steps (table 1; supplementary file S1, Supplementary Material online). These trees broadly conflicted with molecular phylogenetic relationships, in that they reject shrew mole monophyly, place desmans as sister to the other non-*Uropsilus* tribes, and recovered *Condylura* as the sister group of Talpini + Scalopini (supplementary fig. S3, Supplementary Material online). Eighteen transformation series (TS) supported the sister relationship of Talpini + Scalopini, and another 12 TS supported the clade uniting Condylurini with the fossorial moles, with most pertaining to the clavicle, humerus, manubrium, scapula, and sternum (supplementary text S1, Supplementary Material online). When characters were optimized on our molecular phylogenies, single trees of 621–622 steps (fully constrained) or 5 trees of 581 steps (partly constrained) were recovered (table 1; supplementary file S1, Supplementary Material online). When the monophyly of *Condylura* + Desmanini was constrained, five MP trees with tree lengths of 582 steps were recovered, with only five TS supporting this semi-aquatic clade (supplementary text S2, Supplementary Material online); notably, these characters do not appear to be specializations for aquatic locomotion or lifestyles.

Discussion

Accurate species trees are the lifeblood of many questions in evolutionary biology, though have proven to be challenging to elucidate in practice due to ancient rapid radiation events and the confounding effects of recombination, hybridization, and incomplete lineage sorting (ILS). While numerous multi-gene phylogenetic reconstruction approaches (e.g., concatenation and coalescent estimation) have been developed to overcome the influence of these forces, the optimal analytical approach remains contested as each method is subject to different tradeoffs, drawbacks, and biases (Chou et al. 2015; Mirarab and Warnow 2015; Springer and Gatesy 2016). Despite their different limitations, concatenation, ASTRAL-II, and SVDquartets provide 100% bootstrap support for many of the same clades including a basal split between Uropsilinae

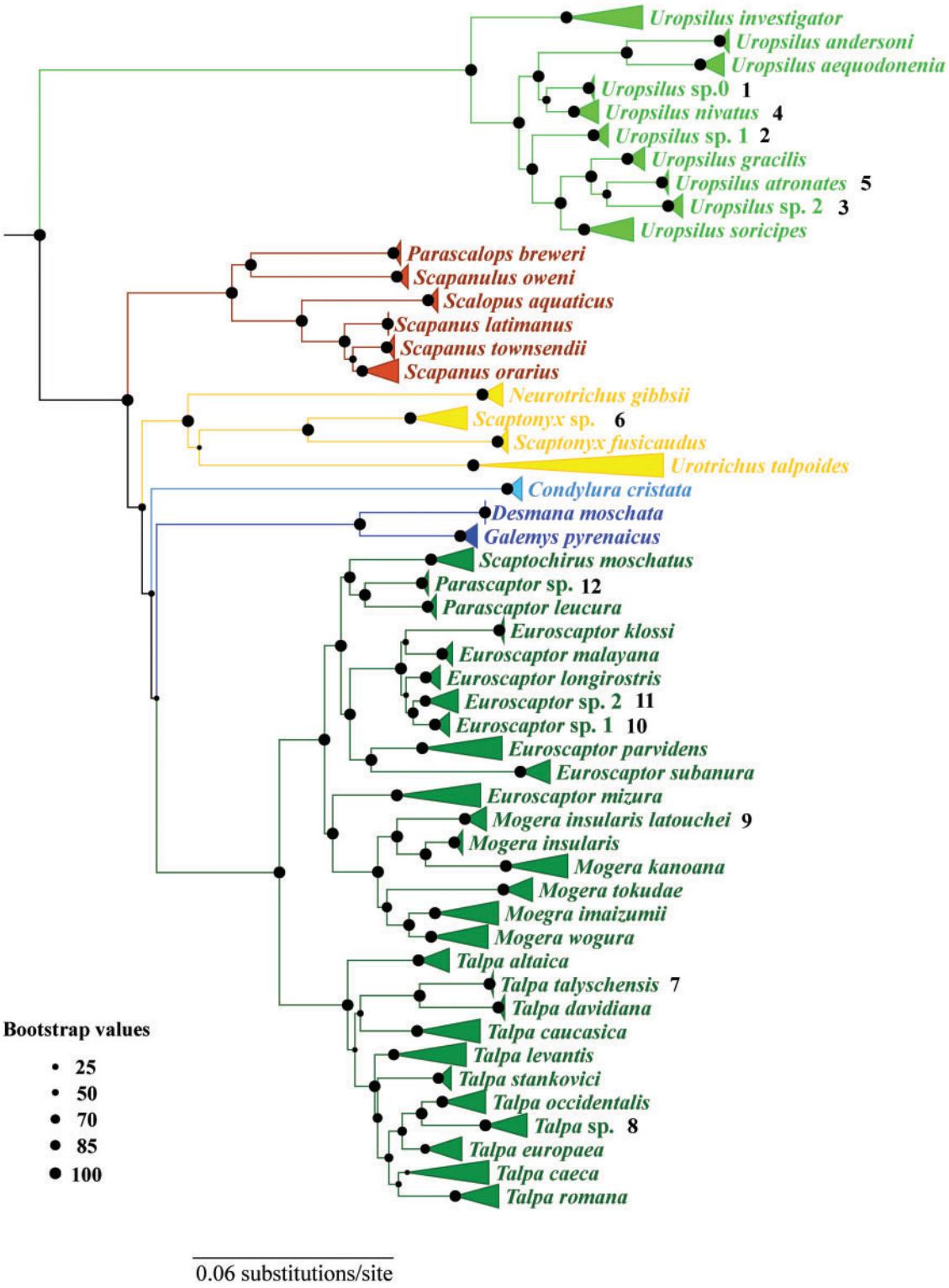

FIG. 3. Maximum likelihood tree (-ln likelihood = 138472) of the family Talpidae based on an expanded supermatrix for 19 nuclear and 2 mitochondrial genes from 782 specimens (2,836 gene sequences; supplementary file S2, Supplementary Material online). Currently undescribed, potential cryptic species are numbered 1–12 (see supplementary table S1, Supplementary Material online). The tree was constructed using RAxML, with tree branch lengths representing substitutions per site.

and other talpids, inclusion of the Gansu mole (*S. oweni*) within the otherwise New World tribe Scalopini, monophyly of Scaptonychini + Urotrichini + Neurotrichini, and monophyly of each of the remaining tribes that include more than one species (Desmanini, Talpini). While conflicts were apparent among the species trees, in all cases these differences were not robustly supported by two or more competing methods.

In general, bootstrap support percentages are higher for concatenation with 21 genes (mean bootstrap

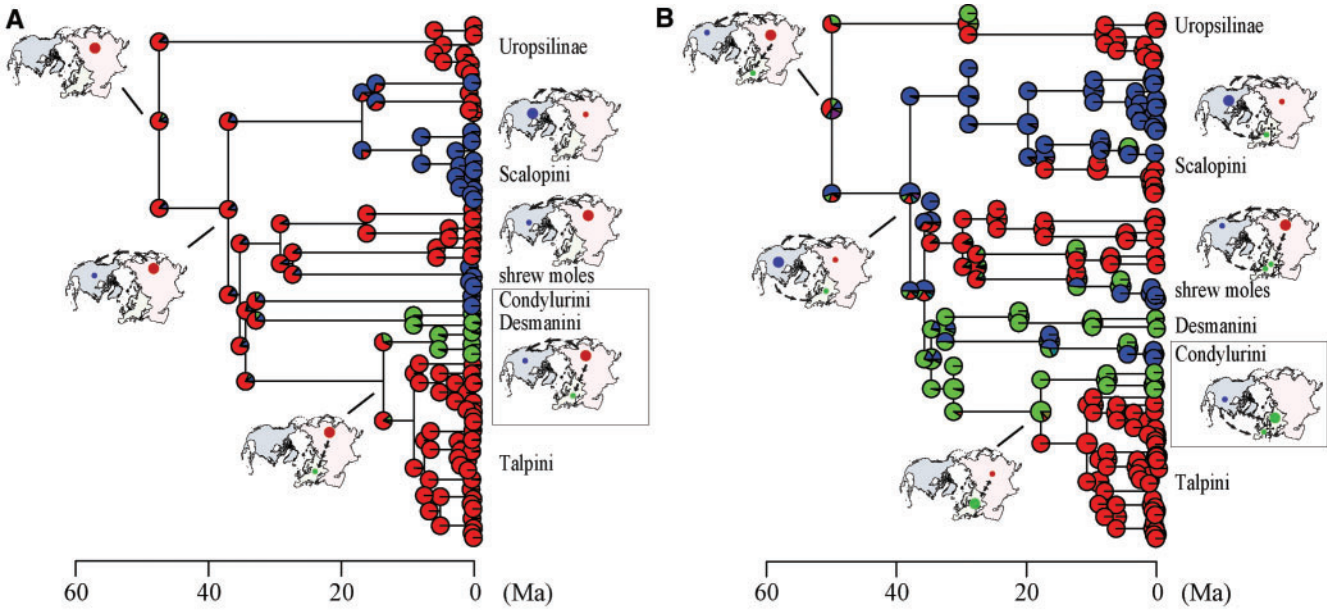

Fig. 4. Biogeographical reconstruction of ancestral ranges of the family Talpidae estimated using the DEC*+J model with only living species (*A*), and with 14 fossil taxa included (*B*). Pie charts represent the probabilities of the likely ancestral ranges at each node. Map insets show the most likely trans-continental migration scenarios at ancestral nodes or for each tribe. Crosses indicate extinction after migration.

Table 1. Summary of the Results of Morphological Analyses.

	# of best trees	# steps	CI	RI	RC
Maximum parsimony trees[a]	9	569	0.424	0.676	0.286
Full constraint (21 genes/concatenation)[b]	1	621	0.388	0.625	0.242
Full constraint (21 genes/ASTRAL-II)	1	622	0.387	0.624	0.242
Full constraint (21 genes/SVDquartets)	1	622	0.387	0.624	0.242
Partial constraint (molecular consensus)[c]	1	581	0.415	0.664	0.275
***Condylura* + Desmanini constrained**	5	582	0.414	0.663	0.275

Note.—all input datasets and output results are given in the supplementary file S1, Supplementary Material online.
[a]Provided in supplementary fig. S3, Supplementary Material online.
[b]The maximum clade credibility tree estimated using the concatenated 21-gene dataset implemented in BEAST.
[c]Only those relationships that were strongly supported (PP $\geq$0.95 and BS $\geq$ 85) in concatenated and coalescent species trees using 21 genes were constrained.

support = 94%) than for ASTRAL-II (90%) and SVDquartets (87%). This finding is consistent with other studies (e.g., Prum et al. 2015) that have reported higher bootstrap percentages for methods when individual gene segments are not binned through local or global concatalescence (Gatesy and Springer 2013) to mitigate gene tree reconstruction errors. Differences between concatenation and coalescence species trees may also reflect the distinction between ILS ignorant (concatenation) and ILS aware (coalescence) methods, but evaluations of empirical datasets suggest that problems with gene tree inaccuracy are more problematic for shortcut coalescent methods than ILS is for concatenation for deep-level problems in the Tree of Life (Gatesy and Springer 2014; Springer and Gatesy 2016). Also, species trees based on two coalescence methods (ASTRAL-II, SVDquartets) exhibit topological differences with each other, but these differences cannot be explained by ILS. The strongly supported basal relationships among the five nonuropsiline clades in the concatenated 21-gene Bayesian analysis (PP > 0.95 in all cases) is similarly suspect and may reflect model mis-specification (Waddell and Shelley 2003) as placements of these lineages received low support scores in all other analyses conducted both with and without inclusion of the two mitochondrial loci. The low branch supports we recovered for the clade containing the six non-Scalopini tribes in our non-Bayesian analyses are particularly surprising given that this clade was highly supported in ML analyses (BS = 92) conducted on a much smaller (six loci; ~5.8 kb) dataset (Bannikova, Zemlemerova, Lebedev, et al. 2015). Nonetheless, our inability to statistically refute Talpini + Scalopini monophyly re-opens the debate regarding the single versus convergent evolution of fossorial adaptations in these tribes. This lack of resolution and the discordant placement of tribes among our concatenated and coalescent trees are consistent with a series of rapid radiation events, and calls for additional phylogenomic data and alternative analyses (e.g., retroposon insertions) to fully resolve difficult nodes.

As noted above, however, even with these differences there are no topological incongruences that are strongly supported by both concatenation and coalescence species trees, that is, all methods resulted in a congruent set of strongly supported clades, and thus close several century-long debates about the taxonomy and phylogeny of moles. For example, all

analyses strongly supported inclusion of the long-tailed moles (Scaptonychini) into a monophyletic clade containing the two shrew mole tribes, a point that has previously generated considerable debate, mostly disputing their monophyly (Van Valen 1967; Ziegler 1971; Hutchison 1976; Yates and Moore 1990; Corbet and Hill 1992; McKenna and Bell 1997; Shinohara et al. 2003; Motokawa 2004; Cabria et al. 2006; Sanchez-Villagra et al. 2006; Kawada, Li, et al. 2008; Schwermann and Thompson 2015). The Gansu mole (*S. oweni*) is also unambiguously placed with the North American scalopins (Hutchison 1968; Bannikova, Zemlemerova, Lebedev, et al. 2015; Schwermann and Thompson 2015) rather than with Asian scaptonychins (Van Valen 1967). Moreover, crown Scalopini is equally or slightly older than crown Talpini (fig. 1), suggesting that the Gansu mole is a relict scalopin in Asia that has survived a long history of climatic changes and competition with radiations of Old World moles as well as various subterranean rodent lineages (Nevo 1979). In addition, the surprisingly ancient divergence (~18 Ma) within the monotypic long-tailed mole genus *Scaptonyx*, and paraphyletic or polyphyletic *Euroscaptor* indicate the presence of uncharacterized genus-level talpid lineages, with strong evidence for cryptic species also found in *Euroscaptor*, *Parascaptor*, and *Uropsilus*, as suggested in previous studies (Wan et al. 2013; He et al. 2014). Remarkably, our expanded supermatrix analysis indicates that up to 12 genetically distinct talpid lineages remain to be characterized (fig. 3), indicating that talpid diversity may be nearly a third higher than previously recognized (i.e., 55 vs. 43 species). Most of these putative species are from the mountains of southwest China, implying that the "sky island" landscape has had significant effects on speciation (Lei 2012; He and Jiang 2014). Extensive sampling in this and adjacent areas is likely to reveal additional undescribed lineages, and is required to elucidate the full species diversity of the region.

Timing and Tempo of Talpid Diversification

Divergence time estimates indicate that all tribes originated during the Eocene–Oligocene transition, which may be attributed to global cooling and desiccation, and the consequent development of non-forest flora (Nevo 1979). Although most extant talpids have very limited distributions and dispersal abilities, there is strong evidence for transcontinental and oversea dispersals following these radiation events (fig. 4; supplementary text S3, Supplementary Material online). For example, the fossil history of Uropsilinae (including *Asthenoscapter* and *Desmanella*) extends much further back in Europe than it does in Asia (Qiu and Storch 2005). Similarly, members of the North American tribes Neurotrichini, Condylurini, and Scalopini were also distributed in Europe during the Oligocene to Pleistocene (supplementary fig. S2*C*, Supplementary Material online), but have since been extirpated (van den Hoek Ostende et al. 2005). This decline of non-fossorial talpids was clearly evident on the fossil-diversity plot (supplementary fig. S2*B*, Supplementary Material online), but not accounted for on the LTT plot using living species only (supplementary fig. S2*A*, Supplementary Material online). Consequently, biogeographical reconstructions in the absence of fossil evidence (see fig. 4) may be highly misleading, though it is important to note that they may also be strongly biased by both missing and incorrectly placed fossil taxa, as well as by uncertain phylogenetic relationships among living taxa. These constraints call for studies addressing the definitive phylogenetic placements of all known fossil species to better assess the biogeographic origins of the various major lineages within Talpidae, and for the family itself, as morphological character matrices are currently only available for a few fossil European species (Schwermann and Thompson 2015; Hooker 2016).

Ecomorphological Evolution

The adaptation of talpid moles to fossorial and aquatic environments has generated substantial research attention. Mapping morphological characters on our molecular phylogenetic trees resulted in an increase in tree length from 569 to at least 581 steps (table 1). When morphological characters primarily associated with digging are mapped on the most parsimonious morphological tree, a pattern of increasing robustness of the clavicle, manubrium, scapula, sternum, and humerus are apparent from the shrew moles through to the star-nosed moles, and finally to the true (Scalopini + Talpini) mole tribes (supplementary fig S3 and supplementary text S1, Supplementary Material online). As a sister-taxa relationship between these latter two tribes was neither supported nor rejected by molecular evidence, however, questions regarding evolution of their fossorial specializations remain unresolved (Schwermann and Thompson 2015). Conversely, the five uniquely shared anatomical characters in the tribes Condylurini and Desmanini are unlinked to semi-aquatic lifestyles (supplementary text S2, Supplementary Material online), supporting the contention that star-nosed moles and desmans evolved aquatic specializations along independent morphological trajectories from a common shrew-mole-like ancestor.

Conclusions

We provide a phylogenetic hypothesis that robustly unites long-tailed moles with shrew moles, identifies a deep (mid-Miocene) split within the monotypic genus *Scaptonyx*, and illuminates parallel adaptive routes to semi-aquatic habits. Moreover, our phylogeny challenges numerous taxonomic classifications pertaining to the mole family and reopens the Talpini + Scalopini monophyly versus diphyly debate. Finally, our findings suggest that talpid species richness is underestimated by nearly 30%, with most of these undescribed lineages inhabiting the meagerly sampled mountains of southwest China—a known biodiversity hotspot that is likely to harbor additional taxonomic diversity.

Materials and Methods

Taxon Sampling and Data Collection

Forty-three talpid species, including four recently described species (Kawada et al. 2007; Kawada, Yasuda, et al. 2008; Kawada, Son, et al. 2012; Liu et al. 2013), are recognized worldwide. Our taxon sampling included representatives from 25

accepted species and 5 putative cryptic lineages identified in previous studies (Wan et al. 2013; He et al. 2014) from all 17 talpid genera (52 specimens in total), together with 3 shrew species (supplementary table S3, Supplementary Material online). Standard phenol–chloroform extraction protocol or Qiagen DNeasy Blood and Tissue kits were used to extract total DNA.

Two mitochondrial (*12S*, *CYTB*) and segments of 19 nuclear genes (*ADORA3*, *ADRB2*, *APOB*, *APP*, *ATP7A*, *BCHE*, *BDNF*, *BMI1*, *BRCA1*, *BRCA2*, *CREM*, *DMP1*, *ENAM*, *GHR*, *PLCB4*, *RAG1*, *RAG2*, *TTN2*, *VWF*) were amplified following (Meredith et al. 2011). Primer sequences are provided in supplementary table S4, Supplementary Material online. Corresponding sequences of two moles, two shrews, one hedgehog, one gymnure, and one solenodon were downloaded from GenBank and included in our dataset. Sequences from 54 ingroups and 7 outgroups were assembled using SeqMan (Lasergene v7.1) and aligned using MUSCLE (Edgar 2004). Patterns of variation including mean genetic distances and informative sites for each gene are given in supplementary table S5, Supplementary Material online.

Gene Trees and Divergence Time

We first constructed 20 maximum likelihood gene trees (one for each nuclear locus plus a combined mitochondrial gene tree) rooted with *Solenodon* and enforcing the monophyly of each family. As preliminary analyses using RAxML revealed high percentages of arbitrarily resolved relationships with branch lengths $<1 \times 10^{-8}$ for all loci, we instead implemented the default settings in GARLI v2.0.1 (Bazinet et al. 2014) and collapsed branches with lengths $<1 \times 10^{-8}$ into polytomies. The partitioning schemes and evolutionary models of each gene/codon position were simultaneously estimated using PartitionFinder v1.1.1 (Lanfear et al. 2012) (supplementary table S6, Supplementary Material online). GARLI was run twice for each gene alignment to estimate 100 bootstrap trees and to determine the best scoring ML tree.

Two concatenated datasets were compiled: one solely based on the 19 nuclear markers (13,992 bp) and the other incorporating both mitochondrial and nuclear loci (15,986 bp). Bayesian and ML gene trees were estimated from these datasets using BEAST v1.8.0 (Drummond et al. 2012) and RAxML v8.2 (Stamatakis 2014), respectively. The partitioning schemes and evolutionary models were estimated using PartitionFinder (supplementary tables S7 and S8, Supplementary Material online). The concatenated BEAST analyses consisted of 100 million generations, with chains being sampled every 10,000 generations, and the nuclear genes and mitochondrial genes given independent relaxed molecular clocks. Each analysis used a random starting tree, birth–death tree prior. Ten calibrations were used for divergence time estimation (supplementary text S4, Supplementary Material online). Convergence of parameters was accessed using Tracer v1.6. We ran RAxML analyses of the partitioned data with random starting trees with GTR + Γ model, performed on CIPRES (Miller et al. 2015). Supports for relationships were estimated with a rapid Bootstrap algorithm and 500 replicates were used to estimate the Bootstrap supports.

Species Delimitation, Species Tree Reconstruction, and Biogeographic Reconstruction

Species delimitation analyses based on the phylogenetic species concept were calculated using a generalized mixed Yule-coalescent model (GMYC) implemented with the R package splits (Pons et al. 2006).

As our dataset contained gene trees with bifurcations connected by branches with effective lengths of 0 (see above), species trees were estimated using a recently devised and statistically consistent summary coalescent approach—ASTRAL-II—that accommodates polytomies (Mirarab and Warnow 2015). In brief, we ran ASTRAL-II using the best scoring ML trees estimated by GARLI, with 100 bootstrap trees of each gene used for multi-locus bootstrapping. We also estimated species trees with SVDquartets as implemented in PAUP* 4.0a. This SNP summary approach infers quartet trees for all subsets of four species from single-site data, and then combines the set of quartet trees into a species tree, thereby avoiding potential problems arising from gene tree discordance (Chifman and Kubatko 2014). We evaluated 100,000 random quartets with 100 bootstrap replicates. For both the ASTRAL-II and SVDquartet analyses, species trees were estimated from both the nuclear and combined nuclear/mitochondrial datasets.

Because not all phylogenetic relationships in our concatenated tree were well supported by species tree estimations, we further tested the monophyly versus diphyly of Scalopini + Talpini using three parsimony-based methods—Kishino–Hasegawa (Kishino and Hasegawa 1989), Templeton, and winning site tests (Templeton 1983)—as implemented in PAUP* 4.0a and CONSEL v0.1, respectively.

We estimated ancestral distributions by applying a modified likelihood-based Dispersal–Extinction–Cladogenesis (DEC) model—DEC*+J (Massana et al. 2015)—which allowed dispersal events (J) and prohibited transitions into the null range in the transition rate matrix (*). We used our 21-gene ultrametric tree to assign living species to their contemporary distributions. The talpid distributions were divided by continents: Asia, Europe, and North America, and we allowed a maximum of two areas at each node, assuming such small animals could not be distributed through three continents at the same time. Analyses were implemented in the R package BioGeoBEARS (Matzke 2014). We also included fossil taxa though adopted a more conservative approach than Piras et al. (2012). In brief, we placed 14 fossil taxa whose phylogenetic positions have been strongly inferred (see supplementary text S5, Supplementary Material online, for justifications) in previous studies on our concatenated gene tree, and re-ran the analysis. All fossil taxa were treated as terminal taxa instead of direct ancestors, thus accounting for dispersal, extinction, and cladogenesis in the analyses.

Expanded Supermatrix Analysis

To assess potential overlooked genetic diversity in the talpid family, we assembled an expanded molecular supermatrix

(supplementary file S2, Supplementary Material online) containing 2,856 sequences from 782 specimens (~2.38 mbps) that included representatives from 40 of the 43 recognized talpid species, plus 7 outgroup species using SequenceMatrix v1.7 (Vaidya et al. 2011). In brief, we searched GenBank for sequences that were orthologous to those in our 21-gene dataset, >300 bp in length, and that included voucher or isolate numbers. New sequences were aligned with the above 21-gene dataset using MUSCLE. Phylogenetic relationships were estimated using RAxML v8.2, as this program is able to handle short sequences and place them on the tree using an MP-based evolutionary placement algorithm (Berger et al. 2011). The alignment was partitioned by genes and codon positions. We employed a GTR model of sequence evolution for searching the best tree and GTR + Γ model for bootstrapping. The mitochondrial *CYT B* sequences were separated into three partitions by codon positions, while all nuclear coding genes were separated into two partitions ($1^{st} + 2^{nd}$ codon positions were in the same partition).

Tracing Morphological Evolution

We removed fossil talpid species from a recently published morphological matrix containing 175 characters (Schwermann and Thompson 2015), and used this dataset to construct the most parsimonious (MP) trees using PAUP* 4.0a (Swofford 2002). We performed heuristic searches with 10,000 random addition replicates using the TBR branch-swapping algorithm. The characters were optimized using "delayed transformation" on the trees in memory. Tree lengths were obtained and apomorphy lists were generated. We then re-ran this analysis (i) using constrained topologies according to our concatenated and coalescent species trees, (ii) using a consensus tree whereby only those relationships that were strongly supported (i.e., PP ≥0.95 and BS ≥ 85) in our concatenated and coalescent species trees estimated using 21 genes were constrained, and (iii) with only the sister-relationship of *Condylura* and Desmanini constrained.

Supplementary Material

Supplementary tables S1–S8, figures S1–S3, supplementary texts and files S1 and S2 are available at *Molecular Biology and Evolution* online.

Acknowledgments

We appreciate the three anonymous reviewers for constructive comments and suggestions, and thank Achim Schwermann and Marcelo Sanchez-Villagra for their advice on the morphological transition series. We also thank Sharon Birks at the Burke Museum of Natural History and Culture, University of Washington, and Patricia Gegick at the New Mexico Museum of Natural History for providing tissue samples. New sequences obtained in this study were submitted to GenBank (accession numbers: KX754466-KX755245). We thank Nancy Halliday for allowing us to use her illustration of *Scalopus aquaticus*. This study was supported by the National Natural Science Foundation of China (no. 31301869 to K.H.), by the Ministry of Education, Culture, Sports, Science and Technology, Japan (no. 15770060 to A.S.), by Natural Sciences and Engineering Research Council of Canada Discovery and Discovery Accelerator Supplement grants (NSERC 238838 and 412336, respectively, to K.L.C.), by the National Science Foundation (NSF DEB-1457735 to M. S), and by Smithsonian Institution (to K.M.H.).

References

Bannikova AA, Zemlemerova ED, Colangelo P, Sözen M, Sevindik M, Kidov AA, Dzuev RI, Kryštufek B, Lebedev VS. 2015. An underground burst of diversity—a new look at the phylogeny and taxonomy of the genus *Talpa* Linnaeus, 1758 (Mammalia: Talpidae) as revealed by nuclear and mitochondrial genes. *Zool J Linnean Soc*. 175:930–948.

Bannikova AA, Zemlemerova ED, Lebedev VS, Aleksandrov DY, Fang Y, Sheftel BI. 2015. Phylogenetic position of the Gansu mole *Scapanulus oweni* Thomas, 1912 and the relationships between strictly fossorial tribes of the family talpidae. *Dokl Biol Sci*. 464:230–234.

Bazinet AL, Zwickl DJ, Cummings MP. 2014. A gateway for phylogenetic analysis powered by grid computing featuring GARLI 2.0. *Syst Biol*. 63:812–818.

Berger SA, Krompass D, Stamatakis A. 2011. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst Biol*. 60:291–302.

Cabria MT, Rubines J, Gomez-Moliner B, Zardoya R. 2006. On the phylogenetic position of a rare Iberian endemic mammal, the Pyrenean desman (*Galemys pyrenaicus*). *Gene* 375:1–13.

Catania KC. 2000. Epidermal sensory organs of moles, shrew-moles, and desmans: a study of the family talpidae with comments on the function and evolution of Eimer's organ. *Brain Behav Evol*. 56:146–174.

Chifman J, Kubatko L. 2014. Quartet Inference from SNP data under the coalescent model. *Bioinformatics* 30:3317–3324.

Chou J, Gupta A, Yaduvanshi S, Davidson R, Nute M, Mirarab S, Warnow T. 2015. A comparative study of SVD quartets and other coalescent-based species tree estimation methods. *BMC Genomics* 16:S2.

Corbet GB, Hill JE. 1992. The mammals of the Indomalayan region: a systematic review. London: Natural History Museum.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 29:1969–1973.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792–1797.

Edwards LF. 1937. Morphology of the forelimb of the mole (*Scalops aquaticus*, L.) in relation to its fossorial habits. *Ohio J Sci*. 37:20–41.

Feuda R, Bannikova AA, Zemlemerova ED, Di Febbraro M, Loy A, Hutterer R, Aloise G, Zykov AE, Annesi F, Colangelo P. 2015. Tracing the evolutionary history of the mole, *Talpa europaea*, through mitochondrial DNA phylogeography and species distribution modelling. *Biol J Linnean Soc*. 114:495–512.

Freeman RA. 1886. The anatomy of the shoulder and upper arm of the mole (*Talpa europaea*). *J Anat Physiol*. 20:201–219.

Gatesy J, Springer MS. 2013. Concatenation versus coalescence versus "concatalescence". *Proc Natl Acad Sci U S A*. 110:E1179–E1179.

Gatesy J, Springer MS. 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. *Mol Phylogenet Evol*. 80:231–266.

He K, Jiang X. 2014. Sky islands of southwest China. I: an overview of phylogeographic patterns. Chinese Science Bulletin 59:585–597.

He K, Shinohara A, Jiang X-L, Campbell KL. 2014. Multilocus phylogeny of talpine moles (Talpini, Talpidae, Eulipotyphla) and its implications for systematics. *Mol Phylogenet Evol*. 70:513–521.

Hooker JJ. 2016. Skeletal adaptations and phylogeny of the oldest mole *Eotalpa* (Talpidae, Lipotyphla, Mammalia) from the UK Eocene: the beginning of fossoriality in moles. *Palaeontology* 59:195–216.

Hutchison JH. 1968. Fossil talpidae (Insectivora, Mammalia) from the later Tertiary of Oregon. *Bull Mus Nat Hist Univ Oregon* 11:1–117.

Hutchison JH. 1974. Notes on type specimens of European Miocene Talpidae and a tentative classification of old world Tertiary Talpidae (Insectivora: Mammalia). *Geobios* 7:211–256.
Hutchison JH. 1976. The Talpidae (Insectivora, Mammalia): evolution, phylogeny, and classification. [PhD dissertation]: University of California, Berkeley.
Hutterer R. 2005. Order Soricomorpha. In: Wilson DE, Reeder DA, editors. Mammal species of the world: a taxonomic and geographic reference. Baltimore: John Hopkins University Press. p. 220–311.
Kawada S-I, Li S, Wang Y-X, Mock OB, Oda S-I, Campbell KL. 2008. Karyotype evolution of shrew moles (Soricomorpha: Talpidae). *J Mammal.* 89:1428–1434.
Kawada S-I, Son NT, Ngoc Can D. 2012. A new species of mole of the genus *Euroscaptor* (Soricomorpha, Talpidae) from northern Vietnam. *J Mammal.* 93:839–850.
Kawada S-I, Yasuda M, Shinohara A, Lim BL. 2008. Redescription of the Malaysian mole as to be a true species, *Euroscaptor malayana* (Insectivora, Talpidae). *Mem Natl Mus Nat Sci*. 45:65–74.
Kawada S, Shinohara A, Kobayashi S, Harada M, Oda S, Lin LK. 2007. Revision of the mole genus *Mogera* (Mammalia: Lipotyphla: Talpidae) from Taiwan. *Syst Biodivers.* 5:223–240.
Kishino H, Hasegawa M. 1989. Evaluation of the maximum-likelihood estimate of the evolutionary tree topologies from DNA-sequence data, and the branching order in Hominoidea. *J Mol Evol.* 29:170–179.
Klietmann J, Nagel D, Rummel M, van den Hoek Ostende LW. 2015. A gap in digging: the Talpidae of Petersbuch 28 (Germany, Early Miocene). *Paläontologische Zeitschrift* 89:563–592.
Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol.* 29:1695–1701.
Lei FM. 2012. Global endemism needs spatial integration. *Science* 335:284–285.
Liu Y, Liu SY, Sun ZY, Guo P, Fan ZX, Murphy RW. 2013. A new species of *Uropsilus* (Talpidae: Uropsilinae) from Sichuan, China. *Acta Theriologica Sinica* 33:113–122.
Massana KA, Beaulieu JM, Matzke NJ, O'Meara BC. 2015. Non-null effects of the null range in biogeographic models: exploring parameter estimation in the DEC Model. *bioRxiv*. doi: http://dx.doi.org/10.1101/026914.
Matzke NJ. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in Island Clades. *Syst Biol* 63:951–970.
McKenna MC, Bell SK. 1997. Classification of mammals: above the species level. New York: Columbia University Press.
Meier PS, Bickelmann C, Scheyer TM, Koyabu D, Sanchez-Villagra MR. 2013. Evolution of bone compactness in extant and extinct moles (Talpidae): exploring humeral microstructure in small fossorial mammals. *BMC Evol Biol.* 13:10.
Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TL, Stadler T, et al. 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521–524.
Miller MA, Schwartz T, Pickett BE, He S, Klem EB, Scheuermann RH, Passarotti M, Kaufman S, O'Leary MA. 2015. A RESTful API for Access to Phylogenetic Tools via the CIPRES science gateway. *Evol Bioinform*. 11:43–48.
Mirarab S, Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31:44–52.
Motokawa M. 2004. Phylogenetic relationships within the family Talpidae (Mammalia: Insectivora). *J Zool*. 263:147–157.
Nevo E. 1979. Adaptive convergence and divergence of subterranean mammals. *Annu Rev Ecol Syst.* 10:269–308.
Nowak RM. 1999. Walker's mammals of the world. Baltimore: Johns Hopkins University Press.
Piras P, Sansalone G, Teresi L, Kotsakis T, Colangelo P, Loy A. 2012. Testing convergent and parallel adaptations in talpids humeral mechanical performance by means of geometric morphometrics and finite element analysis. *J Morphol.* 273:696–711.
Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol.* 55:595–609.
Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526:569–573.
Qiu Z, Storch G. 2005. The fossil record of the Eurasian Neogene insectivores (Erinaceomorpha, Soricomorpha, Mammalia), Part I: China. *Scripta Geologica. Special Issue* 5:37–50.
Reed CA. 1951. Locomotion and appendicular anatomy in three soricoid insectivores. *Am Midland Nat.* 45:513–671.
Sanchez-Villagra MR, Horovitz I, Motokawa M. 2006. A comprehensive morphological analysis of talpid moles (Mammalia) phylogenetic relationships. *Cladistics* 22:59–88.
Schwermann AH, Thompson RS. 2015. Extraordinarily preserved talpids (Mammalia, Lipotyphla) and the evolution of fossoriality. *J Vert Paleontol.* 35:e934828.
Shimer HW. 1903. Adaptations to aquatic, arboreal, fossorial and cursorial habits in Mammals. III. Fossorial adaptations. *Am Nat.* 37:819–825.
Shinohara A, Campbell KL, Suzuki H. 2003. Molecular phylogenetic relationships of moles, shrew moles, and desmans from the new and old worlds. *Mol Phylogenet Evol.* 27:247–258.
Shinohara A, Suzuki H, Tsuchiya K, Zhang YP, Luo J, Jiang XL, Wang YX, Campbell KL. 2004. Evolution and biogeography of talpid moles from continental East Asia and the Japanese islands inferred from mitochondrial and nuclear gene sequences. *Zool Sci.* 21:1177–1185.
Springer MS, Gatesy J. 2016. The gene tree delusion. *Mol Phylogenet Evol.* 94:1–33.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
Swofford DL. 2002. PAUP*: phylogenetic analysis using parsimony (* and other methods). Version 4.0b.
Templeton AR. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37:221–244.
Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27:171–180.
van den Hoek Ostende LW, Doukas CS, Reumer JWF. 2005. The fossil record of the Eurasian Neogene insectivores (Erinaceomorpha, Soricomorpha, Mammalia), Part 1. *Scripta Geologica Special Issue* 5:1–300.
Van Valen LM. 1967. New Paleocene insectivores and insectivore classification. *Bull Am Mus Nat Hist.* 135:217–284.
Waddell PJ, Shelley S. 2003. Evaluating placental inter-ordinal phylogenies with novel sequences including RAG1, gamma-fibrinogen, ND6, and mt-tRNA, plus MCMC-driven nucleotide, amino acid, and codon models. *Mol Phylogenet Evol.* 28:197–224.
Wan T, He K, Jiang X-L. 2013. Multilocus phylogeny and cryptic diversity in Asian shrew-like moles (*Uropsilus*, Talpidae): implications for taxonomy and conservation. *BMC Evol Biol.* 13:232.
Whidden HP. 2000. Comparative myology of moles and the phylogeny of the Talpidae (Mammalia, Lipotyphla). *Am Mus Novit.* 3294:1–53.
Yates TL. 1984. Insectivores, elephant shrews, tree shrews and dermopterans. In: Anderson S, Jones JK, Jr., editors. Orders and families of recent mammals of the world. New York: Wiley. p. 117–144.
Yates TL, Moore DW. 1990. Speciation and evolution in the family Talpidae (Mammalia: Insectivora). *Prog Clin Biol Res.* 335:1–22.
Ziegler AC. 1971. Dental homologies and possible relationships of recent Talpidae. *J Mammal.* 52:50–68.
Ziegler R. 2003. Moles (Talpidae) from the late Middle Miocene of South Germany. *Acta Palaeontologica Polonica* 48:617–648.
Ziegler R. 2012. Moles (Talpidae, Mammalia) from Early Oligocene karstic fissure fillings in South Germany. *Geobios* 45:501–513.