



# Out of Himalaya: the impact of past Asian environmental changes on the evolutionary and biogeographical history of Dipodoidea (Rodentia)

Julie Pisano<sup>1,2\*</sup>, Fabien L. Condamine<sup>3</sup>, Vladimir Lebedev<sup>4</sup>, Anna Bannikova<sup>5</sup>, Jean-Pierre Quéré<sup>2</sup>, Gregory I. Shenbrot<sup>6</sup>, Marie Pagès<sup>1,2†</sup> and Johan R. Michaux<sup>1†</sup>

<sup>1</sup>Laboratory of Conservation Genetics, Institute of Botany (B22), University of Liège, 4000 Liège (Sart-Tilman), Belgium, <sup>2</sup>INRA, UMR 1062 CBGP (INRA/IRD/CIRAD/Montpellier SupAgro), 34988 Montferrier-sur-Lez, France, <sup>3</sup>CNRS, UMR 7641 Centre de Mathématiques Appliquées (Ecole Polytechnique), 91128 Palaiseau, France, <sup>4</sup>Zoological Museum of Moscow State University, 125009 Moscow, Russia, <sup>5</sup>Lomonosov Moscow State University, 119992 Moscow, Russia, <sup>6</sup>Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of Negev, 84990 Midreshet Ben-Gurion, Israel

## ABSTRACT

**Aim** We assessed the influence of past environmental changes, notably the importance of palaeogeographical and climatic drivers, in shaping the distribution patterns of Dipodoidea (Rodentia), the superfamily most closely related to the large species-rich superfamily Muroidea (c. 1300–1500 species). Dipodoids are suitable for testing several biogeographical hypotheses because of their disjunct distribution patterns in the Northern Hemisphere and the numerous species distributed in Asian deserts.

**Location** Holarctic.

**Methods** We inferred molecular phylogenetic relationships for Dipodoidea (34 out of 51 species and 15 out of 16 genera) based on five coding genes. A time-calibrated phylogeny was estimated using a Bayesian relaxed molecular clock with four fossil calibrations. A cross-validation procedure was adopted to examine the impact of each fossil on our estimates. The ancestral area of origin and biogeographical scenarios were reconstructed using time-stratified dispersal–extinction–cladogenesis models.

**Results** Phylogenetic analyses recovered a well-resolved and supported topology. The divergence between Dipodoidea and Muroidea occurred in the late Palaeocene (c. 57.72 Ma) and modern Dipodoidea diversified during the middle Eocene (c. 40.62 Ma). Similar results were found with each calibration strategy used with the cross-validation procedure. The reconstruction of ancestral areas and biogeographical events indicated that modern Dipodoidea originated in the Himalaya–Tibetan and Central Asian region.

**Main conclusions** At the time when Dipodoidea diversified (middle Eocene), the Central Asia and Himalaya–Tibetan Plateau region experienced major uplift episodes due to the collision of India with Asia, which also induced diversification events in many other groups. Other important diversification events (e.g. divergence between Zapodidae and Dipodidae in Central Asia) took place during the Eocene–Oligocene transition when the global temperature decreased significantly and rodent/lagomorph-dominant faunas replaced Eocene perissodactyl-dominant faunas. All of these climatic and geological disruptions in the Central Asia and Himalaya–Tibetan Plateau region modified landscapes and offered new habitats that favoured diversification events, thus triggering the evolutionary history of Dipodoidea.

## Keywords

Asian deserts, biogeography, Bering land bridge, Dipodidae, dispersal–extinction–cladogenesis, Holarctic, Himalayan uplift, rodent phylogeny.

\*Correspondence: Julie Pisano, Laboratory of Conservation Genetics, Institute of Botany (B22), University of Liège, 4000 Liège (Sart-Tilman), Belgium.

E-mail: pisano.julie@gmail.com

†Co-senior authors.

## INTRODUCTION

Historical biogeographical studies aim to assess the influence of past environmental changes on the distribution and the evolution of organisms. Generally, global environmental changes have left their footprint on the evolutionary history of living organisms (e.g. Fabre *et al.*, 2012). Following the recent development of genetic tools (e.g. DNA sequences), the impact of such factors on species evolution can be assessed using phylogenetic approaches. Since the collision of the Indian plate with the Asian plate 40 million years ago (Ma), the uplifts of the Himalayan mountains and Tibetan plateau have thoroughly modified Asian environments and have already been suggested as one of the main driving forces behind long-term Asian Cenozoic climate changes (Bouilhol *et al.*, 2013). Both are the cause of many important biogeographical barriers that led to diversification events in many taxa (e.g. Zhang *et al.*, 2006). Understanding which past Asian environmental changes best explain diversification events can shed light on current distribution patterns and can help to improve our prediction of future range shifts of living organisms.

To better understand the impact of these environmental changes on the diversification of Asian – but also, more generally, of Holarctic – organisms, we propose to study the evolutionary history of the superfamily Dipodoidea (Rodentia). Dipodoidea includes 16 genera and 51 species distributed throughout the Holarctic (Holden & Musser, 2005). Based on morphological and molecular data, three families are now recognized: Sminthidae (syn. Sicistinae, 13 species), Zapodidae (five species), and Dipodidae, which includes Cardiocraniinae (seven species), Euchoreutinae (one species), Alactaginae (16 species) and Dipodinae (nine species) (see the taxonomic revision proposed by Lebedev *et al.*, 2012). Birch mice (Sminthidae) – mainly found in the subalpine meadows and the boreal and alpine forests of Europe, Russia and Central and Eastern Asia – are essentially an arboreal-adapted group, yet some species also occur in steppes or semi-deserts (e.g. *Sicista subtilis*, *S. severtzovi*). Jumping mice (Zapodidae) typically inhabit riparian or wooded areas and marshlands within coniferous forests in North America and, more anecdotally, in China. Jerboas (Dipodidae) are distributed in the deserts, semi-deserts and steppes of North Africa and Eurasia (Holden & Musser, 2005; Shenbrot *et al.*, 2008; IUCN, 2012). As dipodoids exhibit disjoint distribution patterns in the Holarctic, with many species found in different remote arid habitats, they are particularly suitable for testing biogeographical scenarios.

The first occurrences of Dipodoidea in the fossil record are from North America with *Elymys* (?Zapodidae, middle Eocene) and *Simimys* (Simimyidae, middle to late Eocene). In Asia, the oldest dipodoid representatives correspond to other genera: *Heosminthus* (Zapodidae or Dipodidae depending on the studies, middle Eocene to late Oligocene) and *Sinosminthus* (Zapodidae, middle Eocene to middle Miocene) (Wang, 1985; Emry & Korth, 1989; Kelly, 1992; Tong, 1997;

Daxner-Höck, 2001). Whether these genera belong to extant taxa, or represent extinct sister groups of Dipodoidea has yet to be determined. Based on a single-nuclear marker tree of 16 dipodoid species, Zhang *et al.* (2012) proposed that the diversification of modern dipodoids took place during the middle Eocene. Combining a time-calibrated phylogeny with a compilation of the fossil record, they further suggested that diversification events and range expansions were mostly influenced by new ecological opportunities triggered by an increasing aridity and the development of open habitats. Indeed, birch mice (Sminthidae) were shown to diversify during the warming period of the Oligocene–Miocene (24 Ma), while jumping mice (Zapodidae) and jerboas (Dipodidae) are assumed to have radiated during the global cooling following the mid-Miocene climatic optimum (15 Ma). In another study based on nine fragments of nuclear genes, Wu *et al.* (2012) estimated the origin of modern Dipodoidea during the early Oligocene, which provides a different evolutionary history. Overall, despite these sound studies, the timeframe and biogeography of Dipodoidea are unclear and the direct effects of the Himalayan uplift remain untested. Further taxonomic coverage and phylogenetic data are needed to assess the centre of origin, test the effect of major environmental drivers (e.g. climatic oscillations, vegetation changes, Himalayan uplift), and unravel the putative colonization routes that spurred the diversification of the group through the Holarctic region. In addition, Lebedev *et al.* (2012) investigated the relationships of Dipodoidea based on 15 out of the 51 described species while comparing the morphological- and molecular-based phylogenetic trees. They provided the first molecular classification of Dipodoidea. However, their taxonomy and systematics are not yet fully understood, which hampers efforts to unravel the evolutionary and biogeographical history of Dipodoidea.

We tackled this challenge by reconstructing the most complete species-level phylogeny for Dipodoidea based on one mitochondrial and four nuclear coding genes for 34 out of the 51 dipodoid species belonging to 15 out of the 16 genera. This new phylogenetic framework was then used to reconstruct the temporal and biogeographical origins of the group with: (1) estimates of divergence times using a Bayesian relaxed fossil-calibrated molecular clock; and (2) inferences of the biogeographical and evolutionary history using the dispersal–extinction–cladogenesis model.

## MATERIALS AND METHODS

### Taxon sampling and DNA sequence acquisition

Fifty new dipodoid vouchers corresponding to 18 species were sampled. Additional DNA sequences from 20 specimens corresponding to 17 dipodoid species analysed by Lebedev *et al.* (2012) were also added to our dataset. Consequently, our sampling encompassed 34 out of the 51 species belonging to 15 out of the 16 genera of Dipodoidea described in *Mammal Species of the World* (Holden & Musser, 2005). The

single missing genus in our dataset was the Pakistani *Salpingotulus* [note that based on morphological characters, Lebedev *et al.* (2012) demonstrated that *Salpingotulus* should be included into the genus *Salpingotus*]. Twelve outgroup species belonging to Muroidea subfamilies (Gerbillinae, Murinae, Glirinae, Leithiinae), and to Sciuridae and Aplodontiidae families were also included in the sampling. Outgroups were selected to recover specific nodes in the phylogeny so as to be able to use fossils as calibration points to constrain nodes (see below). Twenty-four out of the 53 outgroup DNA sequences were generated in this study, while the others were mined from GenBank. Voucher material descriptions and GenBank accession numbers are given in Appendix S1 in Supporting Information.

Total DNA was extracted and purified from ethanol- and dried-preserved tissues using the Qiagen DNEasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Five genes were selected based on rodent (Fan *et al.*, 2009; Pagès *et al.*, 2010; Lebedev *et al.*, 2012) or mammal (Huchon *et al.*, 1999; Poux & Douzery, 2004) phylogenies. We sequenced one mitochondrial gene, cytochrome *b* (*cytb*, 1.14 kilo base, kb), and four nuclear fragments: exon 1 of the interstitial retinoid-binding protein (*IRBP*, *c.* 1.2 kb), exon 10 of the growth hormone receptor (*GHR*, *c.* 0.9 kb), exon 11 of the breast cancer type 1 susceptibility protein (*BRCA1*, *c.* 0.7 kb), and a portion of the recombination activating gene 1 (*RAG1*, *c.* 1.0 kb). Primer sets used to amplify these different markers are listed in Appendix S2.

Amplifications were carried out in 25 µL reactions containing about 30 ng of extracted DNA, 1 unit of *Taq* DNA polymerase (Qiagen), 2.5 µL of 10× buffer, 0.5 mM of extra MgCl<sub>2</sub>, 100 µM of each dNTP, and 0.2 µM of each primer. Cycling conditions were as follows: one activation step at 94 °C for 4 min followed by 40 denaturation cycles at 94 °C for 30 s, annealing at 50–60 °C depending on the primers for 30 s (see Appendix S2 for temperature), elongation at 72 °C for 1 min or 1 min 30 s depending on the length of the target, and a final extension at 72 °C for 10 min. When amplifications with the Qiagen *Taq* polymerase failed, PCR reactions were performed in 20 µL reactions containing about 30 ng of extracted DNA, 0.4 µM of each primer and 10 µL of Qiagen Multiplex PCR Master Mix. The cycling conditions were similar to the previous ones except for the activation step at 95 °C for 15 min. PCR products were sequenced by Eurofins MWG Operons (Ebersberg, Germany) or Macrogen (Seoul, South Korea). Sequences were corrected using SeqScape (Applied BioSystems), aligned by eye and translated in amino acids using SEAVIEW (Galtier *et al.*, 1996) to ensure sequence orthology.

### Phylogenetic analyses

Phylogenetic trees were reconstructed using maximum likelihood (ML) and Bayesian inferences (BI). Two molecular datasets were used for phylogenetic reconstructions: the densely sampled matrix that contained several individuals per

species, and the species-level matrix consisting of one single individual per species. The appropriate subset partitions and their relative sequence evolution substitution models were determined using the 'greedy' algorithm and the corrected Akaike information criterion (AIC<sub>c</sub>) implemented in PARTITIONFINDER 1.1.1 (Lanfear *et al.*, 2012). Branch lengths were estimated independently for each subset by setting 'branch-lengths = unlinked'. The list of selected evolution models for each partition is available in Appendix S2.

### Maximum likelihood analyses

We first carried out ML analyses on each gene independently using PHYML 3.0 (Guindon *et al.*, 2010). For each analysis, the transition/transversion ratio, the number of substitution rate categories, the proportion of invariable sites and the gamma distribution parameter (if necessary; Appendix S2) were estimated and the starting tree was determined by BioNJ analysis of the dataset. Using optimization options, 1000 bootstrap replicates were performed. Gene tree congruence was checked by visual comparisons. As gene trees were congruent, all genes were concatenated into supermatrices to gain insight into the dipodoid species tree (Douzery *et al.*, 2010). As PHYML software does not allow partitioning of combined datasets, partitioned ML analyses of combined datasets were performed using RAXMLGUI 1.31 (Silvestro & Michalak, 2011). We carried out RAXMLGUI analyses with the following settings: (1) a GTR+GAMMA substitution model for each partition; and (2) robustness of the best tree assessed using the thorough bootstrap (BP) procedure with 1000 replications.

### Bayesian analyses

Bayesian inferences were performed on each gene independently and on the partitioned supermatrices using MRBAYES 3.2.2 (Ronquist *et al.*, 2012). The settings were as follows: (1) two independent runs with four Markov chain Monte Carlo (MCMC) algorithms; (2) 20 million generations; (3) trees sampled every 1000 generations; (4) appropriate independent evolution models for each partition (Appendix S2); and (5) reconstruction of the consensus tree using the 'all-compact' option. A burn-in period of 25% of total generations was determined graphically and the effective sample size (ESS) of the trace of each parameter was checked using TRACER 1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>). The branch supports were estimated using posterior probabilities (PP). The potential scale reduction factors (PSRF) were checked after the end of each analysis to ensure that runs converged (i.e. PSRF reaching 1).

### Assessing the confidence of the tree topology

Using MRBAYES, alternative topological hypotheses were assessed with Bayes factors (BF) (Kass & Raftery, 1995) using TRACER. Lebedev *et al.* (2012) discussed the trichotomy of

Euchoreutinae, Dipodinae and Allactaginae and the paraphyly of *Allactaga* species. We thus tested these hypotheses by grouping the single Euchoreutinae species (*Euchoreutes naso*) with (1) Allactaginae and (2) Dipodinae and (3) constraining the monophyly of *Allactaga*.

## Bayesian divergence time estimations

### Dating analyses

Molecular divergence dates were estimated using Bayesian relaxed clock (BRC) approaches that account for changes in evolutionary rate over time and across clades (Drummond *et al.*, 2006), which are implemented in BEAST 1.8.0 (Drummond *et al.*, 2012). To minimize the size of the parameter space, we enforced the topology obtained from MRBAYES in all dating analyses. Given the use of fossil calibrations, BRC analyses were performed using the species-level topology obtained using MRBAYES as the constrained tree because intraspecific mutation rates are characterized by higher values than interspecific substitution rates (Ho *et al.*, 2005). Using BEAUTY (Drummond *et al.*, 2012), settings were applied as follows: (1) *nucleotide substitution models* were specifically applied for each partition (Appendix S2); (2) the *clock model* was set to an uncorrelated lognormal relaxed clock; (3) *tree models* were set to a Yule or birth–death speciation process; and (4) the *MCMC parameters* were fixed to 50 million of generations, sampling every 5000 generations. The remaining parameters were left in their default settings. We designed 10 dating analyses that differed by the combination of settings and priors that we used for node calibrations.

### Fossil calibrations

Soft bounds were applied to take fossil date uncertainties into account (Ho & Phillips, 2009). The parameters of the four fossil calibration points were all set to lognormal distributions with the 95% interval bounded by the minimum age (2.5% quartile) of the geological interval where the fossil of each calibration point was found and a maximum age (97.5% quartile) of 54 Ma, corresponding to the geological interval where the oldest known fossil of Myodonta was found (i.e. *Erlianomys* from the Eocene Arshanto formation, Nei Mongol region, China) (Li & Meng, 2010). The standard error was set to 0.75. Among the Rodentia fossil records, we selected the following four fossil calibrations (FC):

FC1: *Douglassciurus jeffersoni* (36 Ma) is defined as the oldest known fossil of Sciuridae (McKenna & Bell, 1997). The divergence between Sciuridae and its sister group, Aplodontiidae, happened at least before this date. Hence, we assigned the oldest record of Sciuridae, *D. jeffersoni* at 36 Ma (McKenna & Bell, 1997) to the split between *Aplodontia rufa* (Aplodontiidae) and the monophyletic group composed of *Sciurus aestuans* and *Marmota marmota* (Sciuridae) [offset = 34.94; log(mean) = 1.48; 2.5% quartile = 35.95; 97.5% quartile = 54.05].

FC2: Following Stepan *et al.* (2004), we considered the fossil record from the Siwalik succession in Pakistan as an accurate depiction of the murine history. *Progonomys* has been described as either the most recent common ancestor (MRCA) of extant murines or a predecessor (8.1–12.3 Ma) (Jacobs & Flynn, 2005). We assigned the oldest record of *Progonomys* (Jacobs & Flynn, 2005) to the split between the basal tribe Phloemyini (*Batomys granti*) and the other tribes of Murinae (Apodemini, *Apodemus sylvaticus* and *A. mystacinus*; Rattini, *Rattus tanezumi* and *Maxomys surifer*) [offset = 7.95; log(mean) = 2.36; 2.5% quartile = 10.39; 97.5% quartile = 54.01].

FC3: Fossils of *Apodemus jeanteti* (7 Ma) and *Apodemus dominans* (7 Ma) are considered to be close to extant *A. mystacinus* and *A. sylvaticus*, respectively (Michaux *et al.*, 1997). Consequently, we assigned a minimum age of 7 Ma for the split between *A. mystacinus* and *A. sylvaticus* [offset = 4.376; log(mean) = 2.4345; 2.5% quartile = 7.0; 97.5% quartile = 54.0].

FC4: *Sicista primus* is the earliest known fossil attributed to the *Sicista* genus and was recovered from the 17 million-year-old deposits in Nei Mongol, China (Kimura, 2011). Following Zhang *et al.* (2012), we assumed that the radiation of modern Sminthidae happened at least 17 Ma. Consequently, we assigned the oldest record of *Sicista* at 17 Ma to the crown group of Sminthidae [offset = 14.95; log(mean) = 2.195; 2.5% quartile = 17.01; 97.5% quartile = 54.01].

### Cross-validation analyses

The analyses were performed by omitting, one by one, each of the fossil constraints in turn to identify putative inconsistencies (i.e. incongruence between the molecular credibility interval obtained for the omitted constraint and its palaeontological estimate; Table 1). The maximum clade credibility tree was generated using a burn-in period of 25% with TREEANNOTATOR 1.8.0 (included in the BEAST package). Finally, the BEAST output files were analysed with TRACER to check the convergence of runs and the ESS of the trace of each parameter, and to confirm the use of a relaxed molecular clock (using the standard deviation of the UCLD relaxed clock, 'uclid.stdev' parameter) (Drummond *et al.*, 2012). The best-fit calibration strategy was selected using Bayes factors (Kass & Raftery, 1995) implemented in TRACER.

## Biogeographical analyses

Ancestral area reconstructions for Dipodoidea were inferred using LAGRANGE v. 20130526 and the dispersal–extinction–cladogenesis model (DEC) (Ree & Smith, 2008). Updated Wallace's zoogeographical regions were used to determine the area boundaries and were further split into smaller biogeographical units. These units were delineated using: (1) palaeogeographical criteria (Scotese, 2004; Blakey, 2008); and (2) the revised distribution of extant dipodoid species (Hol-

**Table 1** Major dipodoid divergence times inferred using BEAST. Molecular dating estimations (Ma) of 10 analyses differing by the combination of fossil calibrations (FC) used and the speciation tree priors specified [Yule or BD (birth-death)]. Node numbers refer to those in Fig. 2. 95% highest posterior densities are indicated in parentheses. The text in bold indicates the estimated divergence times when, one by one, FC were removed. The underlined analysis indicates the chronogram that received the highest likelihood score (see Results).

Node identification	All FC		No FC1		No FC2		No FC3		No FC4	
	Yule	BD	Yule	BD	Yule	BD	Yule	BD	Yule	BD
1 Crown of Rodentia	64.84 (60–71.72)	64.85 (60.75–71.75)	64.15 (56.18–75.22)	63.97 (56.64–76.01)	64.8 (60.69–71.8)	64.65 (60.65–71.53)	64.4 (60.56–70.79)	64.46 (60.69–71.1)	64.11 (60.46–70.87)	64.18 (60.6–71.37)
3 FC1 – <i>Douglassciurus jeffersoni</i>	37.9 (35.49–42.09)	37.91 (35.51–41.94)	<b>37.5</b> ( <b>32.84–43.97</b> )	<b>37.39</b> ( <b>33.11–44.43</b> )	37.88 (35.48–41.97)	37.79 (35.46–41.81)	37.65 (35.4–41.38)	37.68 (35.48–41.56)	37.48 (35.34–41.43)	37.52 (35.42–41.72)
9 FC2 – <i>Progonomys</i> (crown of Murinae)	12.95 (12.12–14.38)	12.95 (12.13–14.33)	12.81 (11.22–15.02)	12.78 (11.31–15.18)	<b>12.94</b> ( <b>12.12–14.34</b> )	<b>12.91</b> ( <b>12.11–14.29</b> )	12.86 (12.1–14.14)	12.87 (12.12–14.2)	12.8 (12.07–14.15)	12.82 (12.1–14.25)
12 FC3 – <i>Apodemus jeanteti</i> & <i>Apodemus dominans</i>	7.37 (6.9–8.18)	7.37 (6.9–8.16)	7.29 (6.39–8.55)	7.27 (6.44–8.64)	7.36 (6.9–8.16)	7.35 (6.89–8.13)	<b>7.32</b> ( <b>6.88–8.05</b> )	<b>7.33</b> ( <b>6.9–8.08</b> )	7.29 (6.87–8.05)	7.29 (6.89–8.11)
14 FC4 – <i>Sciasta primus</i> (crown of Sminthidae)	17.9 (16.76–19.87)	17.9 (16.77–19.81)	17.71 (15.51–20.76)	17.66 (15.63–20.98)	17.89 (16.75–19.82)	17.85 (16.74–19.75)	17.78 (16.72–19.54)	17.79 (16.75–19.63)	<b>17.7</b> ( <b>16.69–19.56</b> )	<b>17.72</b> ( <b>16.73–19.7</b> )
6 Split Dipodoidea / Muroidea	58.51 (54.78–64.97)	58.51 (54.82–64.74)	57.89 (50.7–67.88)	57.72 (51.11–68.59)	58.47 (54.76–64.79)	58.34 (54.73–64.55)	58.11 (54.65–63.88)	58.16 (54.77–64.16)	57.85 (54.55–63.95)	57.91 (54.68–64.4)
13 Radiation of Dipodoidea	41.18 (38.55–45.72)	41.18 (38.58–45.56)	40.74 (35.68–47.77)	40.62 (35.97–48.27)	41.15 (38.54–45.6)	41.05 (38.52–45.43)	40.9 (38.46–44.95)	40.93 (38.54–45.15)	40.71 (38.39–45)	40.76 (38.48–45.32)
21 Divergence Zapodidae	34.99 (32.76–38.85)	34.99 (32.78–38.72)	34.62 (30.32–40.59)	34.52 (30.56–41.02)	34.96 (32.75–38.75)	34.89 (32.73–38.6)	34.75 (32.68–38.2)	34.78 (32.75–38.37)	34.6 (32.62–38.24)	34.63 (32.7–38.51)
26 Divergence Cardiocraninae	27.18 (25.45–30.18)	27.19 (25.47–30.08)	26.89 (16.44–22.01)	26.82 (23.74–31.87)	27.16 (25.44–30.1)	27.1 (25.43–29.99)	27 (25.39–29.68)	27.02 (25.44–29.81)	26.88 (25.35–29.71)	26.91 (25.41–29.92)
28 Divergence Euchoreutinae	22.51 (21.08–25)	22.52 (21.09–24.91)	22.28 (19.51–26.12)	22.21 (19.67–26.39)	22.5 (21.07–24.93)	22.45 (21.06–24.84)	22.36 (21.03–24.58)	22.38 (21.08–24.69)	22.26 (20.99–24.61)	22.29 (21.04–24.78)
29 Split Allactaginae / Dipodinae	22.09 (20.69–24.53)	22.1 (20.7–24.45)	21.86 (19.14–25.63)	21.8 (19.3–25.9)	22.08 (20.68–24.47)	22.03 (20.67–24.37)	21.94 (20.64–24.12)	21.96 (20.68–24.23)	21.85 (20.6–24.15)	21.87 (20.65–24.32)
22 Radiation of modern Zapodidae	20.52 (19.21–22.79)	20.52 (19.22–22.71)	20.3 (17.78–23.8)	20.24 (17.92–24.06)	20.51 (19.21–22.72)	20.46 (19.2–22.64)	20.38 (19.17–22.4)	20.4 (19.21–22.5)	20.29 (19.13–22.43)	20.31 (19.18–22.59)
27 Radiation of modern Cardiocraninae	18.97 (17.76–21.07)	18.98 (17.78–21)	18.77 (16.44–22.01)	18.72 (16.57–22.24)	18.96 (17.76–21.01)	18.92 (17.75–20.93)	18.84 (17.72–20.71)	18.86 (17.76–20.81)	18.76 (17.69–20.74)	18.78 (17.73–20.88)
30 Radiation of modern Allactaginae	8.87 (8.3–9.85)	8.87 (8.31–9.81)	8.77 (7.68–10.29)	8.75 (7.75–10.4)	8.86 (8.3–9.82)	8.84 (8.29–9.78)	8.81 (8.28–9.68)	8.81 (8.3–9.72)	8.77 (8.27–9.69)	8.78 (8.29–9.76)
38 Radiation of modern Dipodinae	16.33 (15.29–18.13)	16.33 (15.3–18.07)	16.16 (14.15–18.94)	16.11 (14.26–19.14)	16.32 (15.28–18.08)	16.28 (15.28–18.02)	16.22 (15.25–17.83)	16.23 (15.29–17.91)	16.15 (15.23–17.85)	16.16 (15.26–17.98)

den & Musser, 2005; Buerki *et al.*, 2011; IUCN, 2012). Finally, the dipodoid range, which extends over the entire Holarctic, was divided into nine areas: A, Nearctic (North America); B, West Palaearctic (from Western Europe to Ural Mountains, without North Africa); C, Siberia (from Ural Mountains to Bering Sea); D, Central Asia (Turkmenistan, Uzbekistan and Kazakhstan); E, Mongolia and South-East Russia (Altai Mountains, Mongolian steppe and Yablonoi Mountains); F, Turkey, Iran, Georgia, Azerbaijan and Armenia (Persian plateau, Anatolian region and Caucasus, Iranian plateau); G, Himalaya and Tibetan Plateau; H, Gobi and Taklamakan deserts; I, North Africa and Arabia (Arabian peninsula and Sahara region) (see Appendix S3). The species distributions were defined by presence or absence coding for each area. Species ranges were refined to better fit their present-day distributions, as the distributions available on the IUCN website (IUCN, 2012) or in *Mammal Species of the World* (Holden & Musser, 2005) appear to be inaccurate (G.I. Shenbrot was used as the reference authority; see Appendix S3). Marginal distributions or human introduction events were excluded. The number of area subsets was constrained by setting the 'maxareas' parameter to four, given the widest dipodoid range (*Dipus sagitta*). All ranges or combination of ranges were allowed in the analysis.

We added temporal constraints on dispersal rates between areas according to palaeogeographical reconstructions of the Earth (Scotese, 2004; Blakey, 2008). Specific constraints on dispersal rates were set for a series of five time slices (TS): TS1, Quaternary and Pliocene (0–5.3 Ma); TS2, late and middle Miocene (5.3–16 Ma); TS3, early Miocene (16–23 Ma); TS4, Oligocene (23–34 Ma); and TS5, Eocene (34–56 Ma). The TS boundaries fit with pulses of species diversification identified from the maximum clade credibility tree and assumed to coincide with past key environmental events (Buerki *et al.*, 2011). We tested three types of matrix to assess the impact of dispersal rates on the results (Appendix S3). For each time slice, a Q matrix was defined in which transition rates were dependent on the geographical connectivity between areas (Buerki *et al.*, 2011). For the null hypothesis M0, all dispersal rates were set to 1, which implies no barrier between distinct areas. For the first alternative hypothesis M1, dispersal rates were set between 0 and 1, whereas in the second alternative hypothesis M2, dispersal rates were set between 0 and 0.5. In the absence of barriers (adjacent areas), the dispersal rate was fixed to 1 for M1 and to 0.5 for M2 (e.g. the B and C areas, Appendix S3). When a geographical barrier had to be crossed (e.g. Caucasus Mountains), a dispersal rate of 0.7 was specified for M1 and 0.25 for M2 (e.g. between areas D and F in TS1). Whenever a substantial barrier had to be overcome (e.g. Bering Strait), a dispersal rate of 0.5 for M1 and 0.125 for M2 was attributed (e.g. between areas A and C in TS1). Long-distance dispersal was set to 0.1 in M1 and 0.01 in M2 (e.g. between areas A and B, or G and I, in TS1).

All DEC analyses were carried out using the maximum clade credibility tree that produced the highest likelihood

score compared with the other cross-validation procedure analyses. Outgroups were removed for biogeographical analyses because of their distant phylogenetic relationships with the ingroup. To decrease basal node uncertainties, several range constraints on the root were tested (combination of one to four areas). Their global likelihood scores were compared to determine the most likely ancestral area. For all nodes of the chronogram (including the root), a given distribution area was treated as significantly supported when its score was greater than or equal to two log-likelihood units compared with the scores of other tested analyses (Ree & Smith, 2008).

## RESULTS

### Phylogenies and rare genomic changes

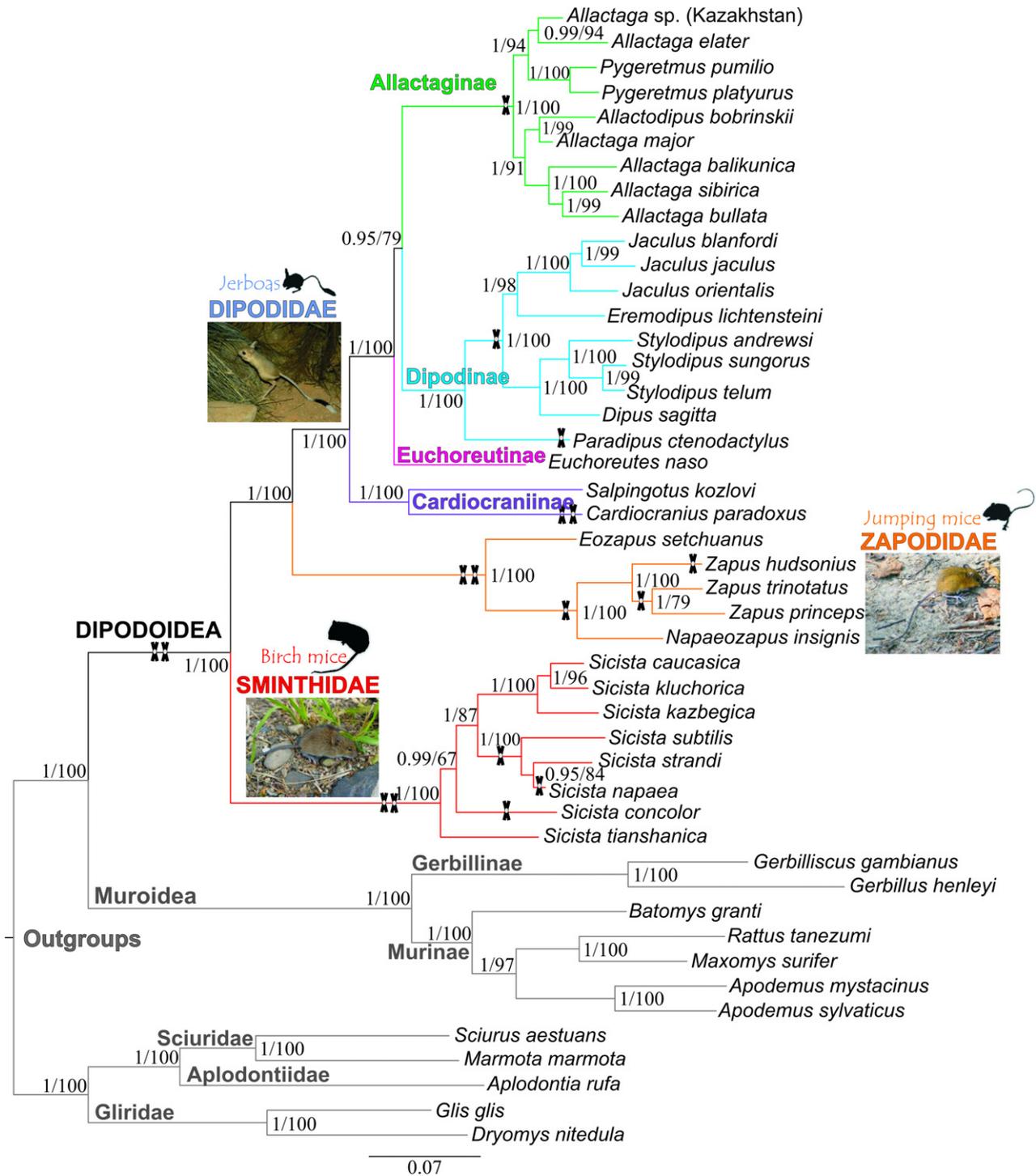
#### *Phylogenetic inference and topological hypotheses*

Maximum likelihood and Bayesian inference analyses based on each gene independently yielded congruent topologies. Accordingly, all genes were concatenated in a single supermatrix. The final supermatrices (4973 nucleotides) consisted of 46 species, 34 of which belong to Dipodoidea. ML and BI combined analyses recovered a similar well-resolved and supported topology. Phylogenetic results based on the species-level matrix are discussed below and presented in Fig. 1, while those based on the densely sampled matrix are shown in Appendix S2. All nodes have PP  $\geq$  0.95, and 82% of branches have BP values  $>$  95%. Sequences were deposited in GenBank under accession numbers KM397124 to KM397347 (Appendix S1).

Bayes factors showed significant differences between our best tree and the alternative topological hypotheses (BF<sub>No constraint tree vs. H(1,2,3)</sub>  $>$  10) (Appendix S2). These results confirm the paraphyly of the *Allactaga* genus and the sister grouping between *Euchoreutes naso* (Euchoreutinae) and the clade including Allactaginae and Dipodinae (PP/BP = 0.95/79). The monophyly of Dipodoidea, Sminthidae, Zapodidae, Dipodidae and all dipodid subfamilies was confirmed with maximum support (1/100).

#### *Rare genomic changes*

We observed 17 rare genomic changes (RGC), corresponding to indels of three or multiples of three nucleotides (Springer *et al.*, 2004) in *BRCA1*, *IRBP* and *GHR* sequences (Fig. 1, Appendix S2). RGC strengthened the obtained topology by independently confirming: the monophyly of Dipodoidea, Sminthidae, Zapodidae and Allactaginae; the basal branching of *Paradipus ctenodactylus* (Dipodinae) and the monophyly of the remaining Dipodinae; the monophyly of the American Zapodidae genera (*Napaeozapus* and *Zapus*); and the sister grouping between *Zapus princeps* and *Z. trinotatus* (Zapodidae) and between *Sicista napaea*, *S. strandi* and *S. subtilis* (Sminthidae), respectively.

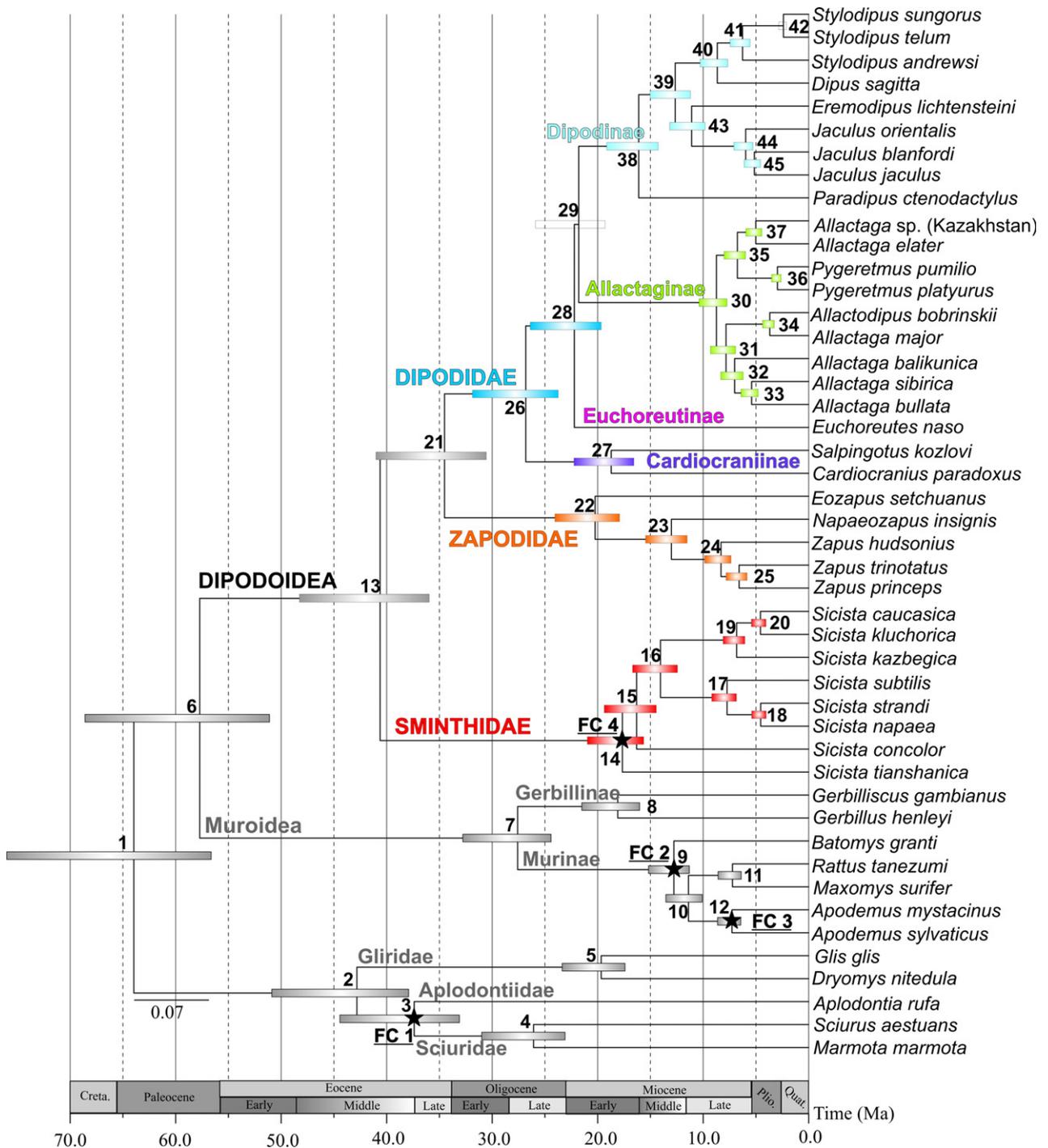


**Figure 1** Phylogenetic relationships among 34 species of Dipodoidea obtained using the species-level matrix (Bayesian inference tree). Analyses were performed using the partitioned dataset of the combined *cytb*, *IRBP*, *GHR*, *BRCA1* and *RAG1* genes. Bayesian inferences and maximum likelihood analyses gave an identical topology. Numbers above branches reflect node supports obtained using MRBAYES and RAXML: posterior probability (PP)/ bootstrap (BP) values. Black crosses on branches indicate the presence of rare genomic changes in our alignment, which strengthened the obtained topology and confirmed independently the monophyly of specific groups. Colours refer to clades. The black shades refer to the three distinct dipodoid morphotypes. Pictures © Wikimedia Commons.

### Divergence time estimates

All analyses gave similar results whatever the calibration strategy used during the cross-validation procedure

(Table 1). Our fossil calibration constraints were thus validated. None of the 10 dating strategies was significantly better than the others (see Appendix S2 for BF and likelihood scores). Consequently, we selected the analysis that



**Figure 2** Dated phylogeny of Dipodoidea. The figure shows the maximum clade credibility tree with median ages from the Bayesian uncorrelated lognormal method that is discussed in further details in this study. Black stars indicate fossil calibrations of node. Numbers at nodes refer to those in Table 1 (see ‘NoFCI’ analysis) and Appendix S2. Coloured rectangles at nodes refer to the 95% highest posterior density (95% HPD) of estimated divergence times (see Appendix S2 for all detailed values). In the geological time-scale, ‘Quat.’, ‘Plio.’ and ‘Creta.’ refer to the Quaternary, Pliocene and Cretaceous, respectively. Colours refer to those in Fig. 1.

produced the highest likelihood score, i.e. the chronogram obtained without the use of the fossil *Douglassciurus jeffer-soni* and using the birth–death model of speciation (‘No FCI’ in Table 1). This chronogram is presented in Fig. 2.

Estimated node ages and the 95% highest posterior density (95% HPD) for the main nodes are detailed in Table 1 (see Appendix S2 for all node estimations and their 95% HPD).

## Historical biogeography

Analyses performed using the M2 hypothesis with no geographical constraint on the root showed the highest likelihood score ( $\log L_{\text{no-constraint}} (M2) = -103.4$ ) compared with the M0 and M1 analyses (Table 2). Given that the combination of the geographical area inferred for the root of Dipodoidea at the time of their origin was biologically unlikely (e.g. 'ABDG'), we constrained the root of Dipodoidea. When constraining the root with one area, the analysis with Central Asia (Area 'D') as root constraint provided the best likelihood of all analyses constrained with one area ( $\log L_{\text{D}} (M2) = -107.7$ ). We then added a second area to constrain the root. The analysis with Central Asia and the Himalaya-Tibetan plateau (areas 'DG') as root constraint provided a better likelihood ( $\log L_{\text{DG}} (M2) = -107.4$ ). Increasing the number of constrained areas at the root (i.e. using three areas to constrain the root) failed to improve the likelihood. We thus selected the analyses constrained with the geographical area 'DG' as the most likely biogeographical scenario. Analyses with the 'DG' root constraint obtained using the M0, M1 and M2 stratified DEC models yielded highly con-

**Table 2** Results of biogeographical analyses of Dipodoidea. The table shows likelihood scores of dispersal–extinction–cladogenesis (DEC) analyses constrained with biogeographical zones of the stratified model. 'No constraint on the root' refers to null hypotheses assuming no geographical constraint on the root of Dipodoidea. M0, M1 and M2 refer to stratified DEC models. The alphabet code refers to the nine areas of the biogeographical model and is the same as the one in Fig. 3. The analysis in bold and underlined indicates the biogeographical scenario that received the highest likelihood score and that is discussed in further detail in this study.

Likelihood scores for biogeographical analyses using DEC and stratified models			
	M0	M1	M2
No constraint on the root	-109.4	-104.9	-103.4
Root A	-120	-117.1	-115.2
Root B	-116.2	-113.7	-112.5
Root C	-117	-113.3	-111.9
Root D	-113	-108.9	-107.7
Root E	-117.9	-113.6	-111.7
Root F	-117.2	-114.7	-113.3
Root G	-114.6	-111	-109.8
Root H	-115.9	-112.1	-110.3
Root I	-122	-123.1	-122.6
Root DH	-113.5	-109.4	-108
Root DE	-114.6	110.4	108.8
<b><u>Root DG</u></b>	<b>-112.8</b>	<b>-108.6</b>	<b>-107.4</b>
Root HG	-114	-110.4	-108.9
Root DF	-114.7	-110.8	-109.4
Root DGH	-112.8	-108.8	-107.4
Root FGH	-113.9	110.5	-109.4
Root GHE	-115	-110.7	-109.1
Root DEGH	-114.1	-110	-108.5
Root DFGH	-113.8	-110	-108.7

gruent results (Appendix S2). The ancestral areas and biogeographical processes (vicariance, dispersal and colonization routes) reconstructed using the M2 matrix of dispersal rates and the 'DG' constraint on the root are shown in Fig. 3.

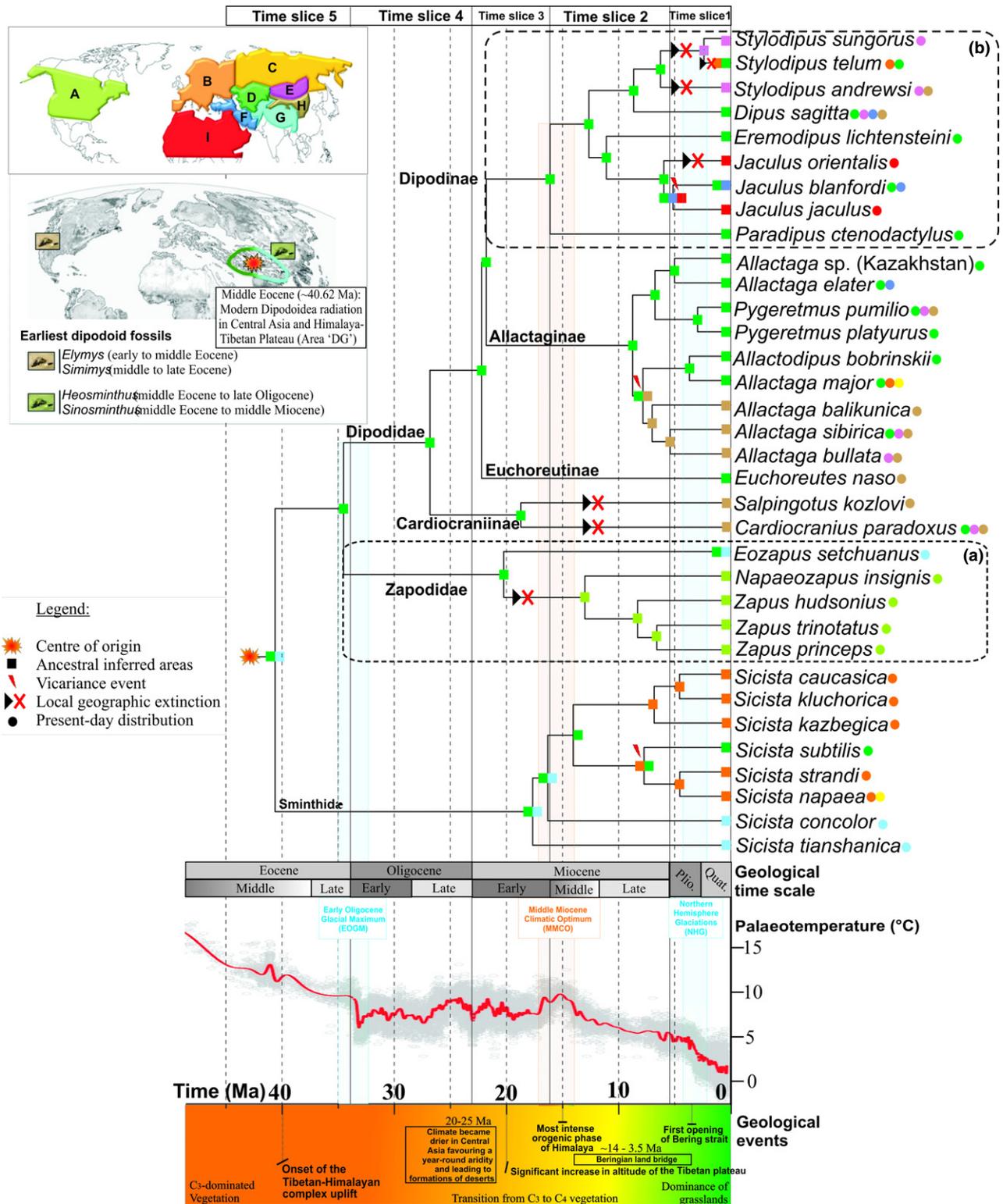
## DISCUSSION

Inferring the impact of historical events on the evolution of faunas is particularly difficult. It is especially true when dispersals and/or local extinctions occurred between biogeographical regions, making them difficult to tease apart. It is also not trivial to connect records that lie within rocks and fossils with records captured into DNA sequences. One way to sort the information contained in palaeontological and molecular data is thus to use biogeographical events as connectors to infer the biogeographical history of living organisms.

### Origin and evolutionary history of Dipodoidea

By including 34 out of the 51 described Dipodoidea species, this study investigated the evolutionary history of Dipodoidea in further detail. The phylogenetic results were congruent with those of previous dipodoid studies (Fan *et al.*, 2009; Lebedev *et al.*, 2012) and confirmed the paraphyly of the *Allactaga* genus and the phylogenetic position of *Euchoreutes naso* (Euchoreutinae). These systematic results were required to understand their evolutionary history. The dating estimates were congruent with those of Meredith *et al.* (2011) and Zhang *et al.* (2012) but differed from those of Wu *et al.* (2012), who estimated younger node ages. This incongruence was probably because of a larger taxonomic sampling (focus on major clades of Rodentia in Wu *et al.*), calibration strategies that relied on distinct calibration constraints (a single calibration point in common out of the seven selected by Wu *et al.*), but also on a different interpretation of the fossil record (i.e. interpretation of *Progonomys*, see justification in the 'Fossil calibrations' section). In addition, fitting the best partitioning schemes and the best molecular evolution models to nucleotide alignments allowed us to better estimate the branch lengths of our trees and thus to better estimate the node ages.

The oldest dipodoid fossil, *Elymys complexus* (described as '?Zapodidae') found in the early Bridgerian of North America, Nevada (46.2–50.3 Ma) (Emry & Korth, 1989), suggests an American origin of the early Dipodoidea (i.e. stem lineage). It consolidates our molecular estimation of the split between Dipodoidea and Muroidea during the late Palaeocene (c. 57.72 Ma; 95% HPD: 51.11–68.59 Ma). But, depending on our interpretation of the phylogenetic position of the fossil *Elymys*, our findings could conflict with the palaeontological evidence if we consider the American *Elymys* as belonging to Zapodidae (i.e. within modern Dipodoidea). Molecular and palaeontological data provide independent ways to estimate when and where clades appeared and evolved, but neither approach can be considered straightforward. Dating the time of origin of taxa is complicated and is confounded by both preservation biases of the fossil record



and inaccuracies of molecular clock estimation. On the contrary, the oldest known fossils of Dipodoidea discovered in Asia, *Heosminthus* and *Sinosminthus*, found from the middle Eocene of Central Asia (Wang, 1985; Tong, 1997; Daxner-Höck, 2001) rather support our estimation of the diversification of modern Dipodoidea during the middle Eocene (c.

40.62 Ma; 95% HPD: 35.97–48.27 Ma) in the Himalaya-Tibetan plateau and Central Asia regions (areas 'DG', Fig. 3).

Over 50–40 Ma, the Himalaya-Tibetan plateau and Central Asia regions encountered geological disturbances (i.e. uplift episodes) as a result of the collision of India with Asia (Guo et al., 2002; Bouilhol et al., 2013). This geological event

**Figure 3** Temporal and geographical history of Dipodoidea based on results of the dispersal–extinction–cladogenesis (DEC) analysis for which the root of Dipodoidea was constrained with areas ‘DG’ and inferred using the M2 stratified model. The maximum clade credibility tree with the highest likelihood was used for biogeographical analyses of dipodoid lineages (outgroups removed). Names of major clades are indicated in bold above branches. The top left corner rectangular map represents the geographical model, which was divided into nine biogeographical areas (A–I). Coloured areas on the rectangular map correspond to coloured squares of nodes, which represent the most likely inferred ancestral area(s). The black and white map is a representation of the Earth during middle Eocene, and indicates where modern Dipodoidea radiated and where the oldest dipodoid fossils have been found. Coloured circles at tips represent dipodoid present-day distributions. Red crosses preceded by black arrows represent local geographical extinctions in the previous area. Grey dotted boxes (‘a’ and ‘b’) refer to clades, on which we particularly focused. The red curve representing palaeotemperatures and vertical blue and orange bars indicating cooling and warming Cenozoic climatic events are represented according to Zachos *et al.* (2008). A 5-Ma geological time-scale is at the bottom of the figure. Major geological events are indicated inside the coloured rectangle that indicates the transition from C<sub>3</sub> to C<sub>4</sub> grasses (Cerling *et al.*, 1997).

already known to have induced vicariance events in many vertebrate groups [e.g. glyptosternoid fishes (He *et al.*, 2001) and warblers (Johansson *et al.*, 2007)] could have also favoured the radiation of modern Dipodoidea (i.e. crown lineage). Our analyses support an ‘out-of-Himalaya’ origin for the dipodoids because most of their early diversifications have occurred in (or close to) the proto-Himalaya during the Eocene and Oligocene. In the Miocene, the geographical evolution has been influenced by climatic and geological events that were induced by the rise of Himalaya (e.g. aridification). We propose that Asian climatic and geological disruptions that modified landscapes and offered new habitats favoured the early diversification events of many regional clades. Besides, Zapodidae together with Ctenodactylidae became the dominant groups during the Tabenbulukian (i.e. Asian land mammal ages from late Oligocene) (Wang *et al.*, 2007). Our study suggests that the split between Zapodidae and Dipodidae in Central Asia occurred during the Eocene–Oligocene transition (c. 34.52 Ma; 30.56–41.02 Ma), while climatic conditions were declining (Zachos *et al.*, 2008) and Eocene perissodactyl-dominant faunas were replaced by rodent/lagomorph-dominant faunas (i.e. members of Dipodoidea, Cricetidae or other rodent taxa) (Wang *et al.*, 2007; Fabre *et al.*, 2012).

Our sampling was not exhaustive and missing (living or extinct) species may introduce biases in biogeographical reconstructions (Mao *et al.*, 2012). The missing species in our study are distributed in regions inferred as ancestral areas (Appendix S3), so there was likely to have been little bias. As the sampling of Zapodinae and Dipodinae was exhaustive in our study, we preferred to focus on the evolutionary history of these two groups from this point forward.

### Colonization of the New World and Zapodidae diversification

During the early Miocene, the Himalaya–Tibetan plateau experienced major uplift episodes, promoting the aridification of the region (Zhisheng *et al.*, 2001; Guo *et al.*, 2002). Induced disruptions in climatic and environmental conditions also favour changes in the ecological niches, which might have affected many clades. Indeed, our biogeographical reconstructions suggest that modern Zapodidae have responded to these changes and radiated in Central Asia dur-

ing the early Miocene (c. 20.24 Ma; 17.92–24.06 Ma). Thereafter, jumping mice underwent local geographical extinction in their original ancestral area and expanded their range towards North America (see Box ‘A’ in Figs 3 & 4a).

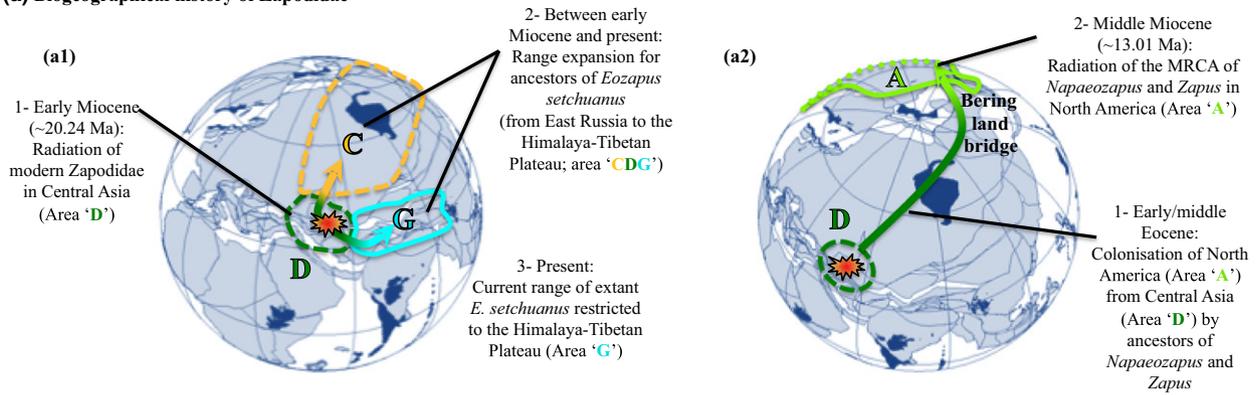
Between c. 14 and 3.5 Ma, the Bering land bridge (BLB) was covered by a continuous boreal coniferous forest belt. This region extensively used as a trans-Beringian connector between Asia and North America promoted faunal exchanges between both continents (Sanmartín *et al.*, 2001). In the fossil record, it is shown that the extinct genus *Megasminthus* from the middle Miocene of North America constitutes the first occurrence of Zapodidae in the Nearctic region (Zhang *et al.*, 2012). It is interesting to note that our biogeographical reconstructions are in agreement with both the fossil record and studies concerning the BLB. Indeed, our results suggest that the ancestors of *Zapus* and *Napaeozapus* colonized North America by the BLB between the early (c. 20.24 Ma, split with the basal zapodid *Eozapus*) and middle Miocene (c. 13.01 Ma, 11.52–15.46 Ma; diversification of the American genera) (Fig. 4b). Zapodidae are now the only dipodoid representatives still inhabiting North America.

### Expansion through Eurasia, the conquest of Africa, and Dipodinae diversification

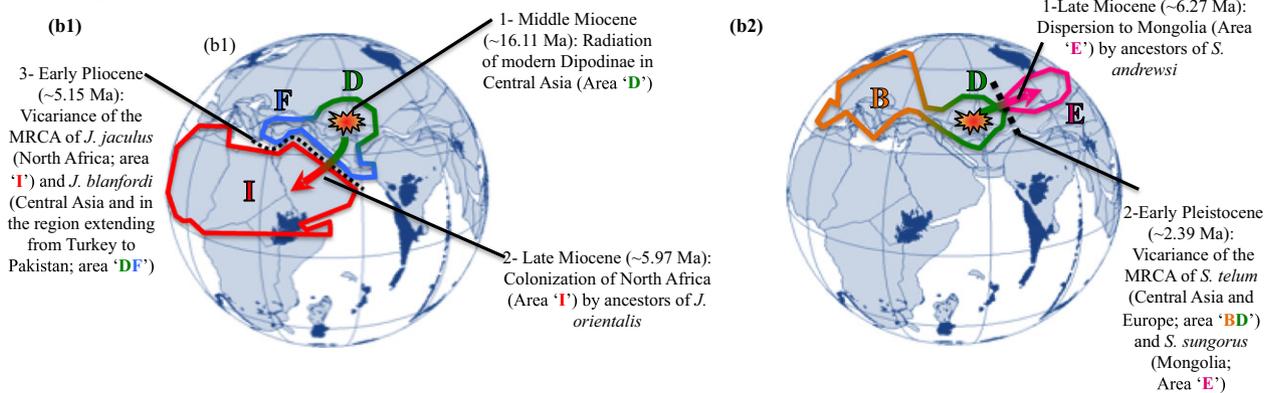
During the early–middle Miocene boundary, the Himalayan Mountains underwent an important and rapid uplift phase, which, coupled with the period of considerable warming called the mid-Miocene climatic optimum, induced strong modifications in climatic and thus, environmental conditions in Central Asia (Tangelder, 1988; Harrison *et al.*, 1992; Zachos *et al.*, 2008). Favourable towards the emergence of new taxa, our results show that modern Dipodinae took advantage of these changes by diversifying in Central Asia during this early–middle Miocene boundary (c. 16.11 Ma; 14.26–19.14 Ma) (see Box ‘B’ in Figs 3 & 4b).

In the late Miocene, the Mediterranean Sea dried up (i.e. the Messinian Salinity Crisis). It is assumed that this event promoted faunal exchanges between Africa and adjacent region. Indeed, thanks to the earliest occurrence of *Mus* in Kenya from 4.5 Ma (Winkler, 2002), it is suggested that the colonization of Africa was already occurring. Besides, based on molecular data, Lecompte *et al.* (2008) demonstrated

**(a) Biogeographical history of Zapodidae**



**(b) Biogeographical history of Dipodinae**



**Figure 4** Biogeographical scenarios for the distribution patterns of (a) Zapodidae and (b) Dipodinae, with specific palaeogeographical maps. Concerning areas, the colours and the alphabet codes are the same as those in Fig. 3. Dotted lines refer to ancestral areas. Red splashes refer to the centre of origin of clades. (a1) During the early Miocene occurred the radiation of modern Zapodidae in Central Asia. Ancestors of the Asian *Eozapus setchuanus* expanded their range across East Russia and the Himalaya-Tibetan Plateau. Nowadays, *E. setchuanus* is exclusively distributed in the Himalaya-Tibetan Plateau. (a2) Between the early and middle Miocene North America was colonized by the most recent common ancestor (MRCA) of *Napaeozapus* and *Zapus*, where they then diversified. (b1) Modern Dipodinae originated in Central Asia during the middle Miocene. The dispersal to North Africa would first have happened by ancestors of *Jaculus orientalis*. The divergence between *J. jaculus* and *J. blanfordi* was promoted by a vicariance event in the region separating North Africa and Asia. (b2) While ancestors of *Stylodipus andrewsi* dispersed to Mongolia during the late Miocene, the MRCA of *S. telum* and *S. sungorus* diversified by vicariance in the region between Mongolia and Central Asia. Palaeogeographical maps have been modified from Blakey (2008).

that *Mus* actually colonized North Africa around 6.6 and 4.0 Ma. Other examples of late Cenozoic murine dispersals between Asia and Africa are also provided by the fossil record, with African sites of the latest Miocene–Pliocene age displaying several ‘Indian’ genera (e.g. *Millardia* and *Golunda*) (WoldeGabriel *et al.*, 1994; Benammi *et al.*, 1996; Wynn *et al.*, 2006). We thus hypothesized that after the radiation of *Jaculus* species in Central Asia at the end of the late Miocene (5.97 Ma, 5.29–7.09 Ma; Fig. 4b1), the MRCA of *J. jaculus* and *J. blanfordi* first colonized a wide region from Central Asia to North Africa (Area ‘DFI’; Fig. 4b1). During the early Pliocene (5.15 Ma; 4.56–6.12 Ma), it gave rise by vicariance to *J. jaculus* in North Africa and to *J. blanfordi* in the region encompassing Central Asia and the area that extends from Turkey to Pakistan (Area ‘DFI’ split into areas ‘DF’ and ‘I’).

In addition, the global cooling of the late Miocene promoted grasslands and arid habitats in Europe and Central

Asia. Woodland-adapted mammals were then replaced by more open-habitat representatives (Cerling *et al.*, 1997). Our biogeographical analyses show that Dipodinae, species adapted to open and arid habitats, would have responded to these changes by expanding their distribution area (e.g. *D. sagitta* is currently found in Mongolia, Gobi-Taklamakan deserts, Central Asia, or northern Iran; area ‘DEFH’).

**CONCLUSIONS**

Exhaustive taxonomic sampling for Dipodoidea is laborious. Some dipodoid species are only known from the type specimens (e.g. *Salpingotus thomasi*) (Holden & Musser, 2005), while others are hard to trap because of difficulties in accessing their range (e.g. Taklamakan desert), or because they are elusive (e.g. *Sicista pseudonapaea* is listed as data deficient; IUCN, 2012). In this study, we collected two-thirds of the dipodoid diversity. Based on this sam-

pling, we have inferred the biogeographical history of the superfamily, in particular for Zapodidae and Dipodinae. Since the middle Eocene, the evolutionary history of Dipodoidea has been influenced by geological and climatic upheavals that occurred in Central Asia, especially the uplift of the Himalayan-Tibetan Mountains, which promoted the development of new habitats, in turn favouring the diversification of several Dipodoidea clades. Accordingly, this study highlighted the importance of such palaeontological and palaeoclimatic events for the diversification of Palaearctic mammals.

## ACKNOWLEDGEMENTS

We thank the editors and the anonymous referees who provided constructive comments. We are particularly grateful to F. Catzeffis, G. Dobigny J.-M. Duplantier, W. Fuwen, L. Granjon, G. Musser, the Burke Museum (Seattle, USA), and the Muséum national d'Histoire naturelle (MNHN; Paris, France) for donations of dipodoid tissues. Analyses were performed at the CBGP HPC computational platform, maintained by A. Dehne-García. J.P. is financed by an 'aspirant FNRS' scholarship also granted by the FRS-FNRS. M.P. and J.M. are supported by a Belgian research fellowship from the FRS-FNRS (respectively, 'mandat chargé de recherches' and 'mandat maître de recherches'). F.L.C. is grateful for support from the French National Agency for Research (ANR ECO-EVOBIO-CHEX2011 grant awarded to H. Morlon). The research of A.B. and V.L. was partly supported by RFBR no. 14-04-00034a. This research was sponsored by financial grants from the Belgian FNRS.

## REFERENCES

- Benammi, M., Calvo, M., Prevot, M. & Jaeger, J.J. (1996) Magnetostratigraphy and paleontology of Ait Kandoula Basin (High Atlas, Morocco) and the African-European late Miocene terrestrial fauna exchanges. *Earth and Planetary Science Letters*, **145**, 15–29.
- Blakey, R.C. (2008) Gondwana paleogeography from assembly to breakup – a 500 million year odyssey. *Resolving the late Paleozoic ice age in time and space* (ed. by C.R. Fielding, T.D. Frank and J.L. Isbell), pp. 1–28. The Geological Society of America Special Paper 441, Boulder, CO.
- Bouilhoul, P., Jagoutz, O., Hanchar, J.M. & Dudas, F.O. (2013) Dating the India–Eurasia collision through arc magmatic records. *Earth and Planetary Science Letters*, **366**, 163–175.
- Buerki, S., Forest, F., Alvarez, N., Nylander, J.A.A., Arrigo, N. & Sanmartín, I. (2011) An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *Journal of Biogeography*, **38**, 531–550.
- Cerling, T.E., Harris, J.M., MacFadden, B.J., Leakey, M.G., Quade, J., Eisenmann, V. & Ehleringer, J.R. (1997) Global vegetation change through the Miocene/Pliocene boundary. *Nature*, **389**, 153–158.
- Daxner-Höck, G. (2001) New zapodids (Rodentia) from Oligocene-Miocene deposits in Mongolia. Part 1. *Senckenbergiana lethaea*, **81**, 359–389.
- Douzery, E.J., Blanquart, S., Criscuolo, A., Delsuc, F., Douady, C., Lartillot, N., Philippe, H. & Ranwez, V. (2010) Phylogénie moléculaire. *Biologie évolutive* (ed. by F. Thomas, T. Lefèvre and M. Raymond), pp. 183–243. De Boeck, Bruxelles.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.
- Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
- Emry, R.J. & Korth, W.W. (1989) *Rodents of the Bridgerian (Middle Eocene) Elderberry Canyon local fauna of Eastern Nevada*. Smithsonian Institution Press, Washington, DC.
- Fabre, P.H., Irestedt, M., Fjeldså, J., Bristol, R., Groombridge, J.J., Irham, M. & Jönsson, K.A. (2012) Dynamic colonization exchanges between continents and islands drive diversification in paradise-flycatchers (Terpsiphone, Monarchidae). *Journal of Biogeography*, **39**, 1900–1918.
- Fan, Z., Liu, S., Liu, Y., Zeng, B., Zhang, X., Guo, C. & Yue, B. (2009) Molecular phylogeny and taxonomic reconsideration of the subfamily Zapodinae (Rodentia: Dipodidae), with an emphasis on Chinese species. *Molecular Phylogenetics and Evolution*, **51**, 447–453.
- Galtier, N., Gouy, M. & Gautier, C. (1996) SEAVIEW and PHYLO\_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Cabios*, **12**, 543–548.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, **59**, 307–321.
- Guo, Z.T., Ruddiman, W.F., Hao, Q.Z., Wu, H.B., Qiao, Y.S., Zhu, R.X., Peng, S.Z., Wei, J.J., Yuan, B.Y. & Liu, T.S. (2002) Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. *Nature*, **416**, 159–163.
- Harrison, T.M., Copeland, P., Kidd, W.S. & Yin, A. (1992) Raising Tibet. *Science*, **255**, 1663–1670.
- He, S., Cao, W. & Chen, Y. (2001) The uplift of Qinghai-Xizang (Tibet) Plateau and the variance speciation of glyptosternoid fishes (Siluriformes: Sisoridae). *Science in China*, **44**, 650–651.
- Ho, S.Y.W. & Phillips, M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367–380.
- Ho, S.Y.W., Phillips, M.J., Cooper, A. & Drummond, A.J. (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, **22**, 1561–1568.
- Holden, M.E. & Musser, G.G. (2005) Superfamily Dipodoidea. *Mammal species of the world: a taxonomic and geo-*

- graphic reference, 3rd edn (ed. by D.E. Wilson and D.M. Reeder), pp. 871–893. Johns Hopkins University Press, Baltimore, MD.
- Huchon, D., Catzeflis, F.M. & Douzery, E.J.P. (1999) Molecular evolution of the nuclear von Willebrand factor gene in mammals and the phylogeny of rodents. *Molecular Biology and Evolution*, **16**, 577–589.
- IUCN (2012) *IUCN Red List of Threatened Species (Version 2012.2)*. Available at: <http://www.iucnredlist.org> (accessed 10 January 2014).
- Jacobs, L.L. & Flynn, L.J. (2005) Of mice... again: the Siwalik rodent record, murine distribution, and molecular clocks. *Interpreting the past: essays on human, primate, and mammal evolution* (ed. by D.E. Lieberman, R.J. Smith and J. Kelley), pp. 63–80. Brill Academic Publishers, Boston, MA.
- Johansson, U.S., Alstrom, P., Olsson, U., Ericson, P.G., Sundberg, P. & Price, T.D. (2007) Build-up of the Himalayan avifauna through immigration: a biogeographical analysis of the *Phylloscopus* and *Seicercus* warblers. *Evolution*, **61**, 324–33.
- Kass, R.E. & Raftery, A.E. (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.
- Kelly, T.S. (1992) New Uintan and Duchesnean (Middle and Late Eocene) rodents from the Sespe formation, Simi Valley, California. *Bulletin Southern California Academy of Sciences*, **91**, 97–120.
- Kimura, Y. (2011) The earliest record of birch mice from the Early Miocene Nei Mongol, China. *Naturwissenschaften*, **98**, 87–95.
- Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–701.
- Lebedev, V.S., Bannikova, A.A., Pagès, M., Pisano, J., Michaux, J.R. & Shenbrot, G.I. (2012) Molecular phylogeny and systematics of Dipodoidea: a test of morphology-based hypotheses. *Zoologica Scripta*, **42**, 231–249.
- Lecompte, E., Aplin, K., Denys, C., Catzeflis, F., Chades, M. & Chevret, P. (2008) Phylogeny and biogeography of African Murinae based on mitochondrial and nuclear gene sequences, with a new tribal classification of the subfamily. *BMC Evolutionary Biology*, **8**, 199.
- Li, Q. & Meng, J. (2010) *Erlanomys combinatus*, a primitive myodont rodent from the Eocene Arshanto Formation, Nuhetingboerhe, Nei Mongol, China. *Vertebrata Palasiatica*, **48**, 133–144.
- Mao, K., Milne, R.I., Zhang, L., Peng, Y., Liu, J., Thomas, P., Mill, R.R. & Renner, S.S. (2012) Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National Academy of Sciences USA*, **109**, 7793–7798.
- McKenna, M.C. & Bell, S.K. (1997) *Classification of mammals above the species level*. Columbia University Press, New York.
- Meredith, R.W., Janecka, J.E., Gatesy, J. et al. (2011) Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science*, **334**, 521–524.
- Michaux, J., Aguilar, J.P., Montuire, S., Wolff, A. & Legendre, S. (1997) Les Murinae (Rodentia, Mammalia) néogènes du Sud de la France: évolution et paléoenvironnements. *Geobios*, **20**, 379–385.
- Pagès, M., Chaval, Y., Herbretreau, V., Waengsothorn, S., Cosson, J.F., Hugot, J.P., Morand, S. & Michaux, J. (2010) Revisiting the taxonomy of the Rattini tribe: a phylogeny-based delimitation of species boundaries. *BMC Evolutionary Biology*, **10**, 184.
- Poux, C. & Douzery, E.J.P. (2004) Primate phylogeny, evolutionary rate variations, and divergence times: a contribution from the nuclear gene IRBP. *American Journal of Physical Anthropology*, **124**, 1–16.
- Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.
- Sanmartín, I., Enghoff, H. & Ronquist, F. (2001) Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biological Journal of the Linnean Society*, **73**, 345–390.
- Scotese, C.R. (2004) A continental drift flipbook. *The Journal of Geology*, **112**, 729–741.
- Shenbrot, G.I., Sokolov, V.E., Heptner, V.G. & Koval'skaya, Y.M. (2008) *Jerboas: mammals of Russia and adjacent regions*. Science Publishers Inc, Enfield, NH.
- Silvestro, D. & Michalak, I. (2011) raxmlGUI: a graphical front-end for RAXML. *Organisms Diversity & Evolution*, **12**, 335–337.
- Springer, M.S., Stanhope, M.J., Madsen, O. & de Jong, W.W. (2004) Molecules consolidate the placental mammal tree. *Trends in Ecology and Evolution*, **19**, 430–8.
- Steppan, S.J., Adkins, R.M. & Anderson, J. (2004) Phylogeny and divergence-date estimates of rapid radiations in murid rodents based on multiple nuclear genes. *Systematic Biology*, **53**, 533–553.
- Tangelder, I.R.M. (1988) The biogeography of the Holarctic *Nephrotoma dorsalis* species-group (Diptera, Tipulidae). *Beaufortia*, **38**, 1–35.
- Tong, Y. (1997) Middle Eocene small mammals from Liguangqiao basin of Henan province and Yuanqu basin of Shanxi province. *Palaeontologica Sinica*, **26**, 1–256.
- Wang, B. (1985) Zapodidae (Rodentia, Mammalia) from the Lower Oligocene of Qujing, Yunnan, China. *Mainzer Geowissenschaftliche Mitteilungen*, **14**, 345–367.
- Wang, Y., Meng, J., Ni, X. & Li, C. (2007) Major events of Paleogene mammal radiation in China. *Geological Journal*, **42**, 415–430.

- Winkler, A.J. (2002) Neogene paleobiogeography and East African paleoenvironments: contributions from the Tugen Hills rodents and lagomorphs. *Journal of Human Evolution*, **42**, 237–256.
- WoldeGabriel, G., White, T.D., Suwa, G., Renne, P., de Heinzelin, J., Hart, W.K. & Heiken, G. (1994) Ecological and temporal placement of early Pliocene hominids at Aramis, Ethiopia. *Nature*, **371**, 330–333.
- Wu, S., Wu, W., Zhang, F., Ye, J., Ni, X., Sun, J., Edwards, S.V., Meng, J. & Organ, C.L. (2012) Molecular and paleontological evidence for a post-Cretaceous origin of rodents. *PLoS ONE*, **7**, e46445.
- Wynn, J.G., Alemseged, Z., Bobe, R., Geraads, D., Reed, D. & Roman, D.C. (2006) Geological and palaeontological context of a Pliocene juvenile hominin at Dikika, Ethiopia. *Nature*, **443**, 332–336.
- Zachos, J.C., Dickens, G.R. & Zeebe, R.E. (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, **451**, 279–283.
- Zhang, P., Chen, Y.Q., Zhou, H., Liu, Y.F., Wang, X.L., Papenfuss, T.J., Wake, D.B. & Qu, L.H. (2006) Phylogeny, evolution, and biogeography of Asiatic salamanders (Hynobiidae). *Proceedings of the National Academy of Sciences USA*, **103**, 7360–5.
- Zhang, Q., Xia, L., Kimura, Y., Shenbrot, G., Zhang, Z., Ge, D. & Yang, Q. (2012) Tracing the origin and diversification of Dipodoidea (Order: Rodentia): evidence from fossil record and molecular phylogeny. *Evolutionary Biology*, **40**, 32–44.
- Zhisheng, A., Kutzbach, J.E., Prell, W.L. & Porter, S.C. (2001) Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. *Nature*, **411**, 62–66.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Taxon sampling for Dipodoidea species and outgroups used in this study.

**Appendix S2** Supplementary information for phylogenetic and dating analyses.

**Appendix S3** Supplementary information for biogeographical analyses.

## BIOSKETCHES

**Julie Pisano** is a PhD student of the Conservation Genetics Unit of the University of Liège (Belgium) headed by **Johan Michaux**, and is conducting her study at the *Centre de Biologie pour la Gestion des Populations* (UMR CBGP, Montferrier-sur-Lez, France). She is interested in the evolutionary history of rodents and also investigates the genetic structure of rodents in hybrid zones.

The research groups of **Johan Michaux** and **Marie Pagès** focus on documenting biodiversity and understanding the origin, evolution and conservation of diverse mammal groups.

Author contributions: J.P., F.L.C., M.P. and J.R.M. conceived the ideas; J.P., J.R.M., V.L., J.-P.Q., and G.I.S. collected the data; J.P., M.P., A.B. and V.L. achieved the molecular work; J.P., F.L.C. and M.P. analysed the data; and J.P. led the writing with revisions of all co-authors.

---

Editor: Brett Riddle