Histoires évolutives de rongeurs Holarctiques

Approche micro- & macroévolutive

Julie Pisano

fnrs LA LIBERTÉ DE CHERCHER
CBGP
Université de Liège ULg
EDENext

THESE DE DOCTORAT
DE L'UNIVERSITE DE LIEGE

Spécialité
Biologie évolutive moléculaire

Présentée par
Julie PISANO

Pour l'obtention du diplôme de Docteur
en Sciences de l'Université de Liège (ULg)

HISTOIRE EVOLUTIVE DE RONGEURS HOLARCTIQUES:
Approche micro- & macroévolutive

Numéro EDENext : EDENext369

Devant le jury composé de :

Jean-Pierre Thomé	Professeur, ULg	Président
Denis Baurain	Professeur, ULg	Secrétaire
Johan Michaux	Directeur de recherche FNRS	Promoteur
Gauthier Dobigny	Chargé de recherche (CR1), IRD (France)	Examinateur
Olivier Hardy	Maître de recherche FNRS, ULB	Examinateur
Patrick Mardulyn	Professeur, ULB	Examinateur
Claire Remacle	Professeur, ULg	Examinatrice
Alain Vanderpoorten	Professeur, ULg	Examinateur

Résumé

La biodiversité n'est pas stable dans le temps et l'espace. Elle évolue en réponse à différents facteurs. A l'échelle macroévolutive, les moteurs de diversité sont essentiellement les changements tectoniques majeurs, climatiques globaux et environnementaux. Ils sont connus pour avoir façonné les patrons évolutifs de groupes d'espèces sur de grandes échelles spatiales et temporelles. A l'échelle microévolutive, les moteurs de diversité sont majoritairement liés à des forces évolutives telles que la mutation, la dérive génétique, la sélection ou la dispersion. Ils rythment l'évolution de la biodiversité populationnelle à une plus petite échelle spatiale et temporelle. Dans le cadre de cette thèse, le but a été de construire un cadre évolutif stable permettant de nous éclairer sur les processus évolutifs et/ou les facteurs qui ont rythmé l'histoire évolutive d'espèces et de populations de rongeurs.

Pour étudier l'évolution de la biodiversité à l'échelle macroévolutive, nous avons pris comme modèle biologique la superfamille des Dipodoidea (Rongeurs : Myodonta). Groupe frère des Muroidea, la superfamille des Dipodoidea comprend trois grands groupes d'organismes : les sicistes (Sicistinae), les souris-sauteuses (Zapodinae) et les gerboises (Allactaginae, Cardiocraniinae, Dipodinae et Euchoreutinae). Dans la littérature, la superfamille des Dipodoidea comprend 51 espèces réparties dans 16 genres de six sous-familles, toutes de la famille des Dipodidae mais cette classification basée essentiellement sur des données morphologiques est très controversée. Avant cette thèse, aucune phylogénie moléculaire des Dipodoidea n'avait été reconstruite. De plus, les Dipodoidea sont particulièrement intéressants pour tester divers scénarios biogéographiques étant donné certaines distributions disjointes dans l'Holarctique et les nombreuses espèces réparties dans les déserts d'Asie et d'Afrique. Il est donc intéressant de comprendre comment ces patrons de distribution disjoints sur l'Holarctique (e.g. Afrique du Nord, Amérique du Nord) ont été mis en place et 'quand et où' ces différents groupes sont apparus.

Lors de cette thèse, pour la première fois, une phylogénie moléculaire comprenant 20 des 51 espèces de Dipodoidea a été reconstruite à partir de quatre gènes nucléaires (BRCA1, GHR, IRBP, RAG1). Cette phylogénie moléculaire a ensuite été comparée à une phylogénie morphologique reconstruite sur base des caractères de la dentition, de la bulle auditive, du gland du pénis et des glandes reproductives accessoires. Cela a permis de comprendre que les nombreuses controverses autour de la taxonomie et de la systématique des Dipodoidea étaient dues à des homologies qui brouillaient le signal phylogénétique. Ainsi, une nouvelle taxonomie des Dipodoidea a pu être proposée. La superfamille des Dipodoidea est dorénavant constituée de 3 familles (Sminthidae, Zapodidae, Dipodidae) et de 19 genres. Ensuite, pour étudier l'histoire évolutive biogéographique de la superfamille des Dipodoidea, l'échantillonnage taxonomique a été augmenté. La phylogénie moléculaire la plus complète à ce jour incluant 34 espèces de Dipodoidea a pu ainsi être reconstruite sur base du gène mitochondrial du cytochrome *b* et des mêmes gènes nucléaires utilisés précédemment. Lors de cette seconde étude, nous avons pu montrer que la radiation des Dipodoidea modernes a

eu lieu au Paléocène supérieur dans la région d'Asie Centrale et de l'Himalaya-Plateau Tibétain et que, de façon générale, leur histoire évolutive a été rythmée par les grands bouleversements climatiques et environnementaux engendrés par la surrection de l'Himalaya et du Plateau Tibétain.

Pour étudier l'évolution de la biodiversité à l'échelle microévolutive, nous avons pris comme modèle biologique le campagnol roussâtre (*Myodes glareolus*). Les populations de campagnol roussâtre sont réparties en plusieurs lignées mitochondriales distribuées sur une large zone de la région Paléarctique. L'une d'elles, caractérisée par le génome mitochondrial du campagnol de la Taïga (*Myodes rutilus*), se distribue de la moitié supérieure de la Suède à travers la Finlande jusqu'au centre de la Russie. En Finlande, cette lignée introgressée (mitotype RUT) vient au contact d'une autre lignée du campagnol roussâtre (mitotype GLA). Il a été proposé que cette zone de contact en Finlande soit le résultat d'un contact secondaire. Cependant, étant donné qu'aucune différenciation nucléaire n'a été observée entre les mitotypes GLA et RUT, il n'est pas clair si cette zone de contact résulte bien d'un contact secondaire (deux évènements de recolonisation de la Finlande). Une autre hypothèse suggérant un seul événement de recolonisation de la Finlande pourrait également expliquer ce patron de discordance mito-nucléaire.

Lors de cette thèse, nous avons étudié la zone de contact entre les mitotypes GLA et RUT située au centre de Finlande sur base de 17 marqueurs microsatellites et du cytochrome *b*. Notre but était d'estimer si la Finlande a connu un ou deux évènements de recolonisation postglaciaire et donc, de mieux comprendre si la zone de contact résulte ou non d'un contact secondaire entre les deux mitotypes. Les approches classiques de génétique des populations et de « clustering » ne nous ont pas permis de valider l'une ou l'autre des hypothèses du fait que la dispersion était limitée dans l'espace et que la différenciation génétique nucléaire entre les campagnols de Finlande était faible. Par conséquent, pour valider définitivement une des deux hypothèses, nous avons dû utiliser des analyses de clines de fréquences alléliques de marqueurs neutres. Celle-ci a montré que le cytb et 16 des 17 microsatellites présentaient des changements de fréquences alléliques entre les mitotypes GLA et RUT et que, par conséquent, la zone de contact entre les mitotypes GLA et RUT correspondait bien à une zone de contact secondaire résultant de deux évènements de recolonisation indépendants.

En conclusion, cette thèse m'a permis de mieux comprendre comment la biodiversité évolue en réponse à différents facteurs. Etudier la biodiversité en utilisant des approches macroévolutives et microévolutives est très intéressant car cela permet d'avoir un regard large sur la manière avec laquelle les espèces, les populations et leurs génomes évoluent.

Abstract

Biodiversity is unevenly distributed across space and time and evolves in response to a variety of factors. On a macroevolutionary scale, drivers of biodiversity are mostly: climate, tectonic events, and environmental changes. They shape patterns at large spatial and temporal scales. On a microevolutionary scale, drivers known to operate at smaller time and spatial scales are mutation, genetic drift, selection and dispersion. In this thesis, we aim at constructing an evolutionary and stable framework enabling to study evolutionary processes and/or factors that have influenced the evolutionary history of species and populations of rodents.

Our macroevolutionary studies were performed on the superfamily Dipodoidea (Rodentia: Myodonta). This group – known as the sister-group of the superfamily Muroidea – includes: birch mice (Sicistinae), jumping mice (Zapodinae) and jerboas (Allactaginae, Cardiocraniinae, Dipodinae et Euchoreutinae). In the literature, this superfamily contains one single family (Dipodidae), 16 genera and 51 species. However, this taxonomy based on morphological markers is controversial. Actually, before this thesis, the molecular phylogeny of the Dipodoidea had never been reconstructed. Moreover, Dipodoidea are particularly attractive for testing biogeographical hypotheses because of their disjunct distribution patterns in the Northern Hemisphere and the numerous species distributed in Asian and African deserts. It is interesting to investigate how these disjunct patterns of distribution were established, and 'when and where' groups appeared.

In this thesis, the molecular phylogeny of 20 out of 51 dipodoid species was reconstructed for the first time based on four nuclear genes (BRCA1, GHR, IRBP, RAG1). This molecular phylogeny was then compared with the morphological one reconstructed based on characters of dentition, auditory bulla, glans penis, and accessory reproductive glands. This has enabled to understand that controversies over the taxonomy and systematic of Dipodoidea were caused by homologies. The new taxonomy proposed herein is divided in three families (Sminthidae, Zapodidae, Dipodidae) and includes 19 genera. In addition, for the biogeographical evolutionary study of Dipodoidea, our taxonomic sampling was improved. Based on one mitochondrial (cytb) and four nuclear genes (BRCA1, GHR, IRBP, RAG1), we reconstructed the most exhaustive molecular phylogeny of Dipodoidea (34 dipodoid species). We showed that the radiation of modern Dipodoidea occurred during the late Palaeocene in the region of Central Asia and Himalaya-Tibetan Plateau, and that, generally, the evolutionary history of Dipodoidea has mostly been influenced by climatic and environmental changes caused by the emergence of Himalaya and Tibetan Plateau.

Our microevolutionary study was performed on populations of the bank vole (*Myodes glareolus*). This rodent has a wide range in the Palearctic, stretching from the British Isles through continental Europe and Russia and extending in the North beyond the Arctic Circle. A peculiar mitochondrial lineage of *M. glareolus* ranges from central Sweden, through northern Finland and up to central and southern Russia; it carries the mtDNA of a northern

neighbour and closely related species, the red vole, *Myodes rutilus*. In Central Finland, this introgressed mitochondrial lineage (mitotype RUT) comes in contact with another lineage (mitotype GLA). It has been suggested that this contact zone between GLA and RUT results from a secondary contact but no nuclear differentiation has been observed between the two mitotypes. It was thus not clear if this contact zone is a result of a secondary contact. Another hypothesis could explain this pattern of mito-nuclear discordance.

In this thesis, we studied the Finnish contact zone between mitotypes GLA and RUT based on 17 microsatellite loci and the cytochrome *b* gene. Our goal was to estimate if Finland encountered one or two postglacial recolonisation events and thus, to better understand if Central Finland is a region of secondary contact between mitotypes GLA and RUT. Classical population genetics and clustering methods did not enable to estimate if one or two recolonisations occurred because of isolation by distance and low intraspecific differentiation among bank voles. Consequently, to solve this issue, we used neutral clines analyses. The subsequent results showed that the cytochrome *b* and 16 out of the 17 microsatellites displayed breaks in allelic frequencies between mitotypes GLA and RUT. Thus, the Finnish contact zone between mitotypes GLA and RUT still shows signs of a past contact between two differentiated lineages that colonised Finland independently.

This thesis enables to better understand how biodiversity responds to distinct factors. ‘Studying biodiversity using macro- and microevolutionary approaches’ is thus interesting given that it provides a broader view about the way species, populations and genomes can evolve.

Remerciements

"On ne peut rien enseigner à autrui. On ne peut que l'aider à le découvrir lui-même."

- Galilée -

Une thèse, c'est un peu comme un long marathon. Au début, c'est facile… puis ce qui te permet de tenir et d'atteindre la lignée d'arrivée, c'est tout d'abord l'endurance physique et morale qui t'anime, mais c'est aussi l'aide et le soutien de personnes spéciales que tu vas avoir sur ton chemin qui vont t'aider à garder le cap.

Tout a commencé en 2011 lors d'une proposition de thèse… dès lors, merci Johan Michaux pour m'avoir accordée votre confiance et m'avoir permis de réaliser cette thèse.

Fabien Condamine, que dire, tu as été un mentor, un exemple, une aide précieuse que je ne remercierai jamais assez. Tu m'as transmis ta passion pour la biogéographie. Ce fut un vrai régale de travailler avec toi ! Merci pour tout ce que tu m'as apportée. « You rock ! »

Nathalie Charbonnel, de la perfection scientifique au soutien morale, tu as été mon pilier, ma référence dans chaque moment. Tu m'as aidée à dépasser mes limites, tu m'as donné envie de toujours faire mieux et du mieux que je peux. Je ne sais pas comment te dire merci… Merci d'avoir été là sur ma route, merci de m'avoir soutenue, merci pour ton sourire, merci pour tout ! « Longue vie à la *Happye Family* !!! »

Flavie, Jean-Marc, Jean-Pierre et Laurent, merci de m'avoir accueillie au CBGP pendant toutes ces années. Votre institut représente une véritable mine d'or aussi bien scientifiquement qu'humainement ! J'ai de la chance d'avoir pu évoluer ici. Merci pour tout ce que vous avez fait pour moi. Je n'aurais qu'un seul regret en partant… ne pas avoir pu terminer 'en beauté' en soutenant ma thèse dans la grande salle de réunion et puis en fêtant ça, comme il se doit, à la cafèt'. Encore une fois, Mille Merci !

Victor, un collègue de bureau, un pote brésilien… mon AMI ! Un ami cher, comme on en a peu, un qui restera dorénavant pour toujours !!! Tu fais partie des 'amis de la main' ☺, n'oublie pas ! Merci pour ton sourire, pour ta joie de vivre, pour toutes nos conversations qui nous ont fait rêvé de 'jour meilleur' ! Le B1.22, le 'bureau des stars' restera gravé pour toujours ! « Muito obrigado meu amigo ! »

Raphaël, Merci pour avoir accepté de faire partie de mon comité, pour m'avoir emmené sur le chemin de la génét' des pop' ! Je ne dirai pas qu'elle n'a plus de secret pour moi, mais en tout cas, tes lumières m'ont permis d'y voir plus clair ! Merci pour tout Raph' !

Carine B., merci énormément pour son aide précieuse qui a su me redonner courage quand j'en avais le plus besoin !

Audrey, merci pour ton soutien quotidien, pour ton amitié, pour tes conseils ! Alex, maître du cluster, ami à temps partiel, merci pour tout ! Max, expert de la BM, merci pour ton aide et pour ta gentillesse ! Nath V. merci pour cette joie de vivre, cette 'patate' que tu transmets au quotidien et accessoirement aussi pour ton aide indispensable concernant la logistique computer !

Martin, Andrea, Maria, Ivhino, Carole K., Denise, Ghais, Nadine, Allan, Sigrid, Guillaume C., Adélaïde, Jorian, Marie-stéphane, Caro… Amis et collègues CBGPiens en générale, je vous remercie pour les conversations échangées autour d'un bon café ou d'une petite clope, pour les échanges et le soutien que vous m'avez apporté au cours de ces quatre années. Merci pour ces bons moments passés ensemble !

Ma poulettttte, mon AMIE, mon trésor de Montpellier, Mathilde G. je te remercie pour toutes ces soirées, pour toutes ces discussions partagées, pour tes conseils et ton soutien. Une thèse, c'est aussi ça, faire la fête et s'amuser ! T'y as fortement participé alors, merci mon bichon pour m'avoir soutenue comme tu l'as fait ! T'es une amie en or <3

Venons-en aux Belges ^^ ! Vous qui étiez là – sans être là – pendant ces quatre années. The best ones… celle qui fait partie de moi et qui m'a soutenue de façon hebdomadaire, qui a toujours cru en moi, qui a toujours su m'écouter et me conseiller. Merci ma July ! Ici je peux le dire haut et fort, t'es pas nulle, t'assures trop ! Merci ma sœur ! Et mon Piet, t'as été d'un soutien moral exceptionnel pendant cette thèse, merci ! Merci pour ton amitié !

Thomas & Magali, je n'oublierai jamais le jour où vous êtes venus sonné chez moi à 23h… pour venir m'apporter votre soutien en cette fin de thèse. Improbable, inattendu et pourtant ! Les liens de sang, c'est plus fort que ces kilomètres qui nous séparent. Merci d'avoir cru en moi ! Vincent & Léti' merci à vous aussi pour vos visites, pour avoir amené près de moi ma 'super-copine' ! Elle me redonne du courage et me pousse à devenir meilleure ! Merci à vous, pour tout !

Bien entendu, je ne peux arriver à la fin de ces remerciements, sans remercier mes parents ! C'est grâce à vous si je suis là aujourd'hui. Merci de m'avoir toujours pousser à réaliser mes rêves et à croire en mes idées et mes convictions ! De tout mon cœur, merci <3

Enfin, je voudrais terminer en remerciant l'homme qui est à mes côtés. Lorenzo, merci d'avoir toujours cru en moi, de m'avoir toujours soutenue, de m'avoir rassurée quand j'en avais besoin. Une fin de thèse, c'est pas facile à supporter – surtout pour le conjoint – alors merci pour ta patience et ton amour.

Un grand Merci à tous,
Julie

Abréviations

ADNmt	Acide désoxyribonucléique mitochondrial
ADRAB2	Gène du récepteur adrénergique a-2B
BRCA1	Gène de susceptibilité du cancer du sein de type 1
c.	Du latin *circa* qui signifie 'environ'
cf.	Du latin *confer* qui signifie 'se reporter à'
cytb	Gène du cytochrome *b*
e.g.	Du latin *exempli gratia* qui signifie 'par exemple'
Fig.	Figure
G6pd	Gène du glucose-6-phosphate déshydrogénase
GHR	Gène du récepteur de l'hormone de croissance
i.e.	Du latin *id est* qui signifie communément 'c'est-à-dire'
IRBP	Gène de la protéine de liaison de l'inter-photorécepteur rétinoïde
LCAT	Gène de la lécithine-cholestérol acyltransférase
Ma	Du latin *mega annum* qui signifie 'million d'années'
MSW	Mammal Species of the World
PCR	Abréviation pour 'réaction de polymérisation en chaine'
RAG1	Gène activant la recombinaison 1
sp.	Espèce
spp.	Espèces
Univ.	Université
ULB	Université Libre de Bruxelles
vs	Versus
vWF	Gène du facteur de vonWillebrand

SOMMAIRE

Chapitre 1:

INTRODUCTION GENERALE

1.1 Préambule

Rien n'a de sens en biologie, si ce n'est à la lumière de l'évolution.
- T. Dobzhansky

Figure 1 Bref aperçu de la biodiversité des vertébrés.

Ce célèbre aphorisme résume en peu de mots l'importance qu'a prise la biologie évolutive dans le monde moderne de la recherche en seulement un peu plus d'un siècle et demi. La biologie de l'évolution est en effet une science jeune qu'on peut dater de l'époque de Charles Darwin (1859 – date de parution du célèbre ouvrage '*L'origine des espèces*') et de Johann Gregory Mendel (1865 – date de publication des travaux sur les lois de l'hérédité) dont les théories ont induit une avancée majeure sur l'étude de la diversité biologique (syn. biodiversité). La biodiversité représente le large réseau de la diversité du vivant (Fig. 1) (Tilman, 2000). A l'heure actuelle, une accélération de l'extinction des espèces est observée ~100 à 1000 fois plus rapidement qu'attendu en situation naturelle. Différents facteurs peuvent jouer un rôle dans ce phénomène tels que l'urbanisation, la destruction ou la transformation des habitats pour l'agriculture et les entreprises, les changements climatiques, ou encore la pollution. Dans ce contexte, la préservation de la biodiversité devient une priorité pour les sociétés humaines car la biodiversité représente un enjeu économique (e.g. les ressources de l'humanité), éthique (e.g. la disparition des espèces) et social (i.e. le lien entre l'Homme et son environnement) important (Parmesan, 2006; IUCN, 2015). En 2010, à l'occasion de l'*Année de la Biodiversité*, le professeur R. Barbault (ex-directeur du département 'Ecologie et Gestion de la Biodiversité' du MNHN, et ex-président du programme *Man and Biosphere* de l'Unesco) soutenait le fait que désormais '*la biodiversité, ce n'est plus simplement envisager l'Homme et la nature mais plutôt considérer l'Homme avec la nature'*.

1.2 Etude de la biodiversité

Le terme de biodiversité est souvent employé en référence à un nombre d'espèces. Il y aurait environ ~8.7 millions d'espèces vivantes sur la Terre (May, 1988; Mora *et al.*, 2011). Parmi cette incroyable diversité, certains groupes – plus facilement observables que d'autres étant donné leur grande taille – sont mieux connus et répertoriés (e.g. les vertébrés constitués de ~47,000 espèces connues sur ~50,000 estimées, vs les invertébrés représentés par ~1 million d'espèces connues sur ~3 millions estimées) (Blaxter, 2003). Obtenir une estimation précise de la richesse spécifique d'un groupe peut être laborieux voire utopique. En effet, de nombreuses espèces restent encore à découvrir notamment à cause du fait qu'elles vivent dans des milieux difficilement accessibles et/ou que certaines espèces sont difficilement différenciables (e.g. les espèces cryptiques) (Bickford *et al.*, 2007; Petit & Excoffier, 2009). En biologie évolutive, il est donc essentiel de procéder régulièrement à des révisions systématiques et taxonomiques pour établir pertinemment les liens de parentés entre les espèces et de façon plus générale, entre les groupes d'organismes.

La biodiversité n'est pas stable dans le temps et l'espace (Fig. 2). Elle évolue en réponse à différents facteurs. Les facteurs abiotiques sont liés à l'action du non-vivant sur le vivant tels que le climat, l'environnement ou la tectonique des plaques. Ils sont des déterminants évolutifs extrinsèques qui façonnent généralement les processus évolutifs à grandes échelles spatiales (i.e. globales et régionales) et temporelles (i.e. sur des millions d'années). Les facteurs biotiques sont rattachés à l'action du vivant sur un autre être vivant tels que la compétition ou le parasitisme. Ils sont des déterminants évolutifs intrinsèques qui auront un impact à une échelle locale et agissent souvent sur de courtes périodes de temps. Pour étudier l'évolution de la biodiversité et déterminer quels sont les facteurs impliqués, deux types d'approches peuvent être utilisées : la 'macroévolution' qui vise à mieux comprendre les processus évolutifs à une échelle interspécifique, et la 'microévolution' qui étudie la variabilité de la diversité biologique à un niveau intraspécifique.

1.2.1 Les moteurs de diversité à l'échelle macroévolutive

Comme nous le démontre le registre fossile, la biodiversité peut augmenter et diminuer au cours du temps (Fig. 2). La diversité des espèces a notamment évolué en réponse aux grands bouleversements abiotiques que la Terre a connu (Quental & Marshall, 2010). Les changements tectoniques majeurs (e.g. les mouvements des plaques, la surrection d'une montagne), les changements climatiques globaux (e.g. les réchauffements) ou encore les gros changements environnementaux (e.g. la désertification) ont effectivement eu un impact considérable sur la diversification du vivant au cours du temps (Zachos *et al.*, 2001; Zachos *et al.*, 2008; Lomolino *et al.*, 2010). Ces perturbations ont aussi bien conduit à des augmentations qu'à des diminutions de la richesse spécifique du vivant au cours du temps. Par exemple, alors que pendant des millions d'années (de l'Oligocène inférieur au Miocène moyen, i.e. de -30 à -12 millions d'années (Ma)) la richesse spécifique des cétacés n'a fait qu'augmenter, l'augmentation importante des températures globales lors de l'optimum climatique du Miocène moyen a engendré une diminution significative de la diversité spécifique des cétacés (Fig. 2) (Quental & Marshall, 2010; Morlon *et al.*, 2011).

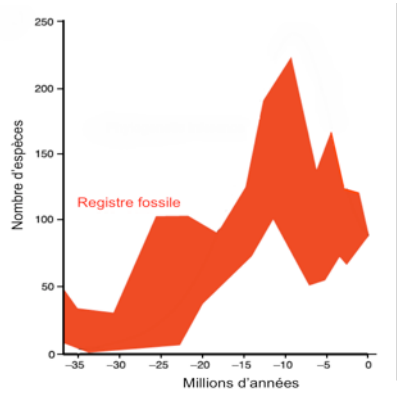


Figure 2 Courbe de diversité des Cétacés estimée à partir du registre fossile (Figure extraite et traduite de Morlon *et al.* (2011)).

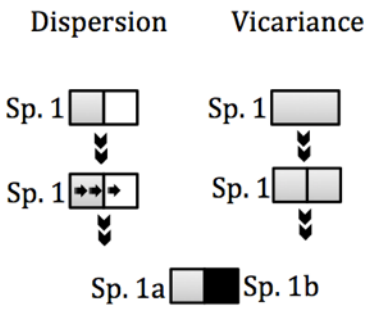


Figure 3 Processus de spéciation à l'échelle macroévolutive. (A gauche) La dispersion : l'espèce Sp.1 est restreinte dans un premier temps à l'aire grisée. Ensuite, elle va se disperser (du moins, une partie de la population de Sp.1) dans l'aire blanche. (A droite) La vicariance : l'aire d'origine de l'espèce Sp.1 va se scinder en deux. Ces deux processus de spéciation favorisent la différentiation d'une espèce en deux nouvelles espèces distinctes (Sp.1a & Sp.1b).

Le processus conduisant un ensemble d'individus généalogiquement interconnectés à se séparer en deux ensembles divergents est la spéciation. De façon générale, la naissance à partir d'une seule espèce de deux espèces filles distinctes peut résulter de divers mécanismes d'isolement reproductif prézygotiques (e.g. isolement écologique, éthologique, temporel) ou postzygotiques (e.g. viabilité réduite des hybrides, fécondité réduite des hybrides). En

biologie macroévolutive, les biologistes évolutionnistes s'intéressent principalement à deux modes de spéciation : la vicariance et la dispersion (Fig. 3) (Blondel, 1995). La spéciation par vicariance constitue un processus de spéciation allopatrique *in situ* qui est induit par la scission d'une aire de distribution en deux ou plus et où chacune des nouvelles espèces occupent une portion différente de l'aire de distribution de l'espèce-mère. Cela ne nécessite pas forcément l'émergence d'une barrière physique. La spéciation par dispersion est induite par la colonisation d'une nouvelle aire de distribution d'une population souvent périphérique à l'aire de distribution « de l'espèce-mère ». Les remaniements chromosomiques et l'effet fondateur vont à terme favoriser l'émergence de deux nouvelles espèces (Mayr, 1963).

1.2.2 Les moteurs de diversité à l'échelle microévolutive

Une population est un ensemble d'individus d'une même espèce qui échangent continuellement du matériel génétique et dont on pense qu'il fonctionne comme une unité du point de vue évolutif et écologique (Roughgarden, 1979). La structure allélique d'une population peut évoluer en fonction de différentes forces évolutives : la mutation, la dérive génétique, la sélection et la dispersion (ou migration).

La mutation est un changement ponctuel du matériel génétique (i.e. une transformation d'un allèle en un autre) dont les taux varient d'un locus à l'autre, d'une espèce à l'autre et en fonction du milieu (Harry, 2001). Par exemple, les taux de mutation des marqueurs microsatellites peuvent varier de 10^{-6} à 10^{-2} par génération selon les espèces (Oliveri, 2010). Il existe trois types de mutations ponctuelles : (1) la mutation par substitution, il y changement d'un nucléotide (d'une base) par un autre nucléotide, (2) la mutation par insertion, il y a ajout d'un ou plusieurs nucléotides, (3) la mutation par délétion, il y a perte d'un ou plusieurs nucléotides.

La dérive génétique est une modification aléatoire des fréquences alléliques dans une population causée par un échantillonnage au hasard des individus dans une population de petite taille efficace qui a pour conséquence, la perte ou la fixation de l'un ou l'autre allèle (Harry, 2001). L'effet fondateur est un exemple de cette force évolutive où un petit nombre d'individus d'une population vont être isolés, n'emmenant avec eux qu'une portion du bagage génétique de la 'population-mère'.

La sélection est un processus par lequel des variations phénotypiques/génétiques vont avoir tendance à augmenter ou diminuer en fréquence au sein d'une population. Il existe différents types de sélection (e.g. directionnelle, stabilisatrice, diversifiante). Par exemple, la sélection directionnelle positive est une force qui permet de fixer une mutation qui sera favorable à la survie de la population et/ou de l'espèce (e.g. le mélanisme industriel de la phalène du bouleau, *Biston betularia* (Grant *et al.*, 1998)).

Enfin, la dispersion désigne le mouvement des individus et donc des gènes entre les populations (Henry & Gouyon, 1998; Carlson, 2004; Baer *et al.*, 2007). Il existe différents types de modèles de dispersion (e.g. le modèle en île, le 'stepping-stone', 'l'isolement par la distance'). Lorsque la dispersion est limitée dans l'espace, les individus ont plus tendance à se reproduire avec ceux qui leur sont géographiquement proches qu'avec ceux qui sont plus éloignés (Fig. 4, 'isolement par la distance') (Rousset, 1997). Cela entraine une corrélation positive entre la distance géographique et la différentiation génétique entre individus.

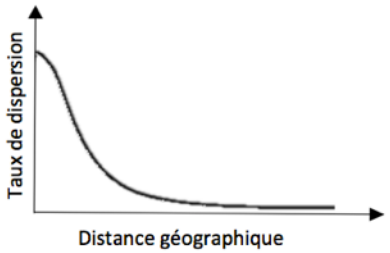


Figure 4 Graphe du taux de dispersion en fonction de la distance géographique en cas d'isolement par la distance entre individus ou populations.

Dans le cadre de cette thèse, je me suis intéressée à la diversité génétique et spatiale des organismes à une échelle macroévolutive et microévolutive dans le but d'identifier des facteurs qui peuvent influencer l'histoire évolutive des organismes vivants.

1.2.3 Des approches basées sur l'étude de la diversité génétique à l'échelle macroévolutive et microévolutive

1.2.3.1 Qui est plus proche parent de qui ?

Pour étudier l'histoire macroévolutive d'un groupe d'espèces, l'une des premières questions qu'il convient de se poser est : « Qui est plus proche parent de qui ? Quel(le)s espèces/genres/clades sont phylogénétiquement proches ? ». Pour répondre à ces questions, la meilleure solution est de reconstruire un arbre retraçant les relations phylogénétiques du groupe taxonomique étudié. A ce propos, les phylogénies moléculaires reconstruites à partir d'approches probabilistes constituent un outil de recherche avec un pouvoir exploratoire considérable permettant notamment d'établir les liens de parenté entre les espèces (voir Encadré 1, p. 8) (Delsuc & Douzery, 2004a, b; Douzery *et al.*, 2010).

Dans l'absolu, l'idéal serait d'avoir un arbre phylogénétique complet qui retracerait les relations de parenté entre toutes les espèces (vivantes et éteintes). Cet arbre constituerait alors un cadre d'évidence 'parfait' pour étudier l'évolution des groupes depuis leurs origines, jusqu'au présent. Or, ces données n'existeront sans doute jamais. Nous avons majoritairement aujourd'hui à notre disposition des données morphologiques et/ou moléculaires d'espèces vivantes.

Chaque clade (i.e. groupe monophylétique) possède une histoire évolutive qui lui est propre (e.g. Almeida *et al.* (2012); Condamine *et al.* (2013b)). Dès lors, une des premières étapes à accomplir pour mieux comprendre l'histoire évolutive d'un clade est de reconstruire une phylogénie robuste la plus exhaustive possible. Celle-ci peut ensuite être replacée sur une échelle de temps géologique de manière à pouvoir dater les évènements de divergence. En étudiant les distributions spatiales passées et présentes des organismes à l'aide des cartes de distribution des espèces et du registre fossile, il est ensuite possible de reconstruire l'histoire évolutive du clade étudié en lien avec les changements planétaires majeurs (climatiques, géologiques ou environnementaux).

Encadré 1 La Phylogénie Moléculaire

La phylogénie est une discipline qui tend à comprendre l'histoire évolutive des espèces à travers l'identification des liens de parenté entre les êtres vivants sur base de caractères morphologiques (e.g. la dent) et/ou moléculaire (e.g. des séquences d'ADN).

1. L'arbre phylogénétique

L'arbre phylogénétique est un objet géométrique qui modélise l'histoire évolutive du vivant (Douzery *et al.*, 2010). Il en existe différents types : (1) le *cladogramme* qui représente les relations de parenté (les longueurs de branches n'ont aucune signification évolutive) (Fig. E1.1 - A), (2) le *phylogramme* qui possède des longueurs de branches différentes en fonction des degrés de divergence des taxa (les espèces ne sont pas toutes équidistantes de la racine) (Fig. E1.1 - B), (3) le *chronogramme* qui replace l'arbre phylogénétique dans un cadre temporel, où chaque espèce est équidistante de la racine et où les longueurs de branches sont proportionnelles aux degrés de changements évolutifs par unité de temps (Fig. E1.1 - A).

Les branches de l'arbre

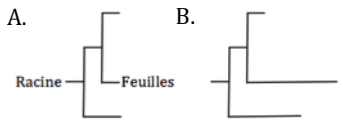


Figure E1.1 Les arbres phylogénétiques. (A) un cladogramme ou un chronogramme, (B) un phylogramme.

sont les représentations graphiques des liens de parenté. Aux extrémités des branches se trouvent les espèces. En remontant le long des branches (des feuilles jusqu'au tronc), on tombe sur des intersections autrement appelés les 'nœuds' de l'arbre. Ils représentent les ancêtres communs hypothétiques – qui en réalité, correspondent plutôt à un ensemble de caractères ancestraux qu'à un seul et même organisme (Douzery *et al.*, 2010). En phylogénie moléculaire, enraciner un arbre revient à utiliser des séquences de gène d'un groupe 'externe', i.e. de taxa extérieurs phylogénétiquement au groupe interne étudié (Lecointre & Le Guyader, 2001).

2. Les méthodes probabilistes

Les méthodes de reconstruction souvent utilisées en phylogénie moléculaire sont les méthodes probabilistes (e.g Springer *et al.* (2004); Steppan *et al.* (2004); Schenk *et al.* (2013)). Elles ont l'avantage d'être performantes et peu sensibles à certains artéfacts de reconstruction (Philippe *et al.*, 2005). En effet, l'un des problèmes fréquemment rencontrés, par exemple en maximum de parcimonie, est le phénomène d'attraction des longues branches : certaines séquences avec des taux d'évolution plus élevés qu'observés normalement au sein du clade font que des taxa non-apparentés se retrouvent plus proches phylogénétiquement l'un de l'autre à cause de leurs branches anormalement longues qui sont le résultat de convergences qui sont en fait interprétées comme des synapomorphies (Bergsten, 2005; Douzery *et al.*, 2010). Les méthodes probabilistes s'affranchissent de ce genre de problème. Elles font intervenir des modèles complexes d'évolution des séquences qui permettent de bénéficier d'un cadre statistique robuste. Ces méthodes probabilistes sont basées sur le concept de vraisemblance qui, en phylogénie, peut s'exprimer comme étant 'la probabilité conditionnelle d'observer les données (un alignement de séquences) sous l'hypothèse évolutive (un arbre phylogénétique) de modèles d'évolution particuliers' (Delsuc & Douzery, 2004a).

La méthode du *Maximum de Vraisemblance* ('ML' pour *Maximum Likelihood*) consiste à estimer les paramètres qui maximisent la vraisemblance d'une topologie donnée et ensuite, à sélectionner l'arbre ayant la plus forte probabilité d'avoir conduit aux données observées – les séquences comparées – étant donné le modèle d'évolution sélectionné (Delsuc & Douzery, 2004a).

En *Inférence Bayésienne* ('BI' pour *Bayesian Inference*), un grand nombre de topologies avec une forte vraisemblance vont être générées. Le meilleur arbre correspondra à un consensus de toutes ces topologies obtenues sur la base des données *a priori* (Delsuc & Douzery, 2004b). Cette méthode utilise simultanément plusieurs chaînes de Markov Monte-Carlo (MCMC) qui explorent les paramètres durant un nombre de générations donné. Contrairement à la méthode de ML, les différentes MCMC ne gardent pas qu'une seule topologie de référence en mémoire mais échantillonnent un grand nombre d'arbres au cours de leur recherche (Delsuc & Douzery, 2004b). De plus, en BI, les risques d'optimums locaux sont réduits grâce à la capacité des MCMC à échanger leurs états respectifs en communiquant entre elles ; celles-ci peuvent d'ailleurs également

explorer des topologies dont la vraisemblance est moins forte (Delsuc & Douzery, 2004b). Les premières étapes (générations) d'uneanalyse par BI constituent une étape de chauffe – appelé *burn-in* – qu'il est nécessaire d'éliminer en fin d'analyse. Après convergence des MCMC (phase stationnaire), les générations ayant atteint une valeur plateau de vraisemblance seront conservées pour établir l'arbre consensus.

3. La notion de robustesse

En ML, la méthode du *bootstrap* est la plus fréquemment utilisée pour tester la fiabilité des branches (Felsenstein, 1985). Son principe consiste à perturber la matrice de données initiales et à reconstruire un nouvel arbre pour estimer dans quelle mesure ce dernier est proche de l'arbre initialement inféré. Les N caractères de la matrice de départ sont tirés aléatoirement N fois pour produire une nouvelle matrice (le tirage au sort se fait avec remise). A partir de cette nouvelle matrice, un arbre est reconstruit de la même manière que précédemment. Ces étapes sont répétées un certain nombre de fois (e.g. 100 fois). La valeur de bootstrap indiquée à chaque nœud représente le pourcentage de fois où un nœud est apparu au sein des différents réplicats. Si l'information phylogénétique soutenant un nœud est robuste, la matrice de départ aura beau être perturbée, il y aura toujours suffisamment de support pour que le nœud apparaisse dans la phylogénie reconstruite à partir de la matrice perturbée (la valeur de bootstrap sera forte). Toutefois, si l'information phylogénétique soutenant un nœud est insuffisante, la probabilité que celle-ci disparaisse quand on perturbera la matrice est forte. Le nœud n'apparaîtra donc pas dans le nouvel arbre reconstruit à partir de la matrice perturbée (la valeur de bootstrap sera faible).

En BI, un grand nombre d'arbres sont produits au cours d'une même analyse bayésienne de phylogénie et l'arbre final correspond au consensus majoritaire de l'ensemble de ces représentations phylogénétiques. Le soutien de chaque nœud est alors apporté par les probabilités postérieures estimées a posteriori à partir de la fréquence d'apparition d'une topologie durant l'analyse, i.e. la fréquence avec laquelle les différents nœuds apparaissent dans les arbres visités (Delsuc & Douzery, 2004b). Les probabilités postérieures correspondent donc à la probabilité qu'un clade donné soit vrai étant donné les modèles d'évolution, les probabilités *a priori*, et les données considérées (Huelsenbeck *et al.*, 2001).

4. Le choix des marqueurs

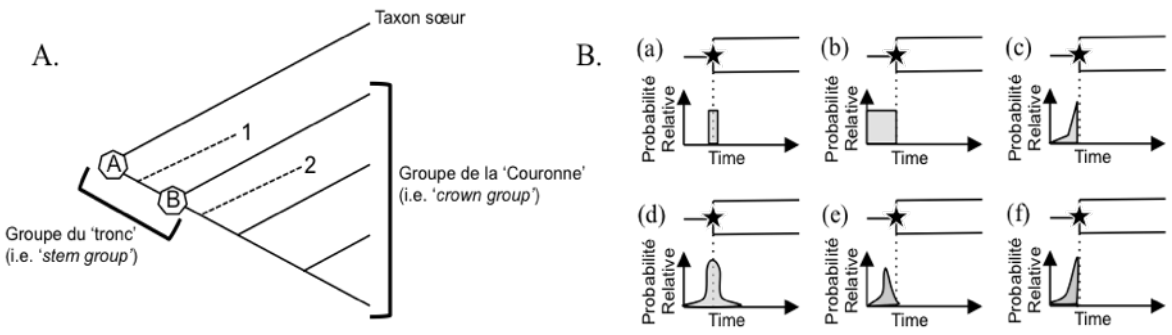
Les gènes nucléaires possèdent une vitesse d'évolution lente. Ils permettent principalement de résoudre les nœuds profonds (Wan *et al.*, 2004). Ils ont un mode de transmission biparental et autorisent donc les évènements de recombinaison, par opposition aux gènes mitochondriaux qui sont de transmission quasi exclusivement maternelle. Les gènes mitochondriaux ont des taux d'évolution rapide qui permettent de résoudre majoritairement les nœuds terminaux (Hudson & Coyne, 2002; Wan *et al.*, 2004).

Etant donné qu'un gène ne constitue qu'une infime portion du génome, il est utopique d'imaginer qu'un seul puisse inférer une phylogénie soutenue et correctement résolue qui correspondrait à « l'arbre vrai ». De plus, l'histoire évolutive des gènes n'équivaut pas toujours à l'histoire des espèces. Les marqueurs ne donnent pas nécessairement la même topologie (principe de congruence des gènes) comme par exemple : (1) dans le cas de discordances mito-nucléaires causées par la capture du génome mitochondrial d'une autre espèce ; (2) lorsqu'un allèle n'est pas fixé avant qu'une nouvelle lignée émerge (tri incomplet des lignées), deux taxa qui ne sont pas directement apparentés présentent deux copies plus proches d'un gène (Douzery *et al.*, 2010). Il est donc plus prudent de combiner plusieurs gènes congruents – préférentiellement mitochondriaux et nucléaires – de manière à renforcer le signal phylogénétique de la racine jusqu'aux feuilles.

1.2.3.2 Comment dater un arbre du vivant?

Pour dater les évènements de divergence d'un arbre phylogénétique (i.e. pour que les longueurs des branches soient mesurées en unité de temps) à partir de données moléculaires, il existe deux méthodes : '*l'horloge moléculaire'* et '*l'horloge moléculaire* dite *relâchée'*. L'horloge moléculaire considère un taux de mutation constant dans le temps et entre taxa. La méthode des *horloges moléculaires* dites *relâchées* autorise des taux de mutation hétérogènes dans le temps et entre les taxa (Zuckerkandl & Pauling, 1965; Bromham & Penny, 2003; Drummond *et al.*, 2006). Dans la plupart des cas, il existe des variations de taux de mutation entre les lignées ; par conséquent, la seconde méthode est souvent préférentiellement appliquée (Smith & Peterson, 2002). Pour procéder à des analyses de datation moléculaire, des données paléontologiques (i.e. des fossiles) sont souvent utilisées comme points de calibration (ou étalonnage) afin d'estimer des âges absolus de divergence entre les organismes car les fossiles témoignent de l'existence d'un taxon ou d'une lignée à une période donnée (Donoghue & Benton, 2007). Il faut cependant faire attention car un fossile ne représente qu'une preuve ponctuelle de la présence d'un clade à un moment donné. Par conséquent, il est conseillé de se servir du fossile uniquement pour contraindre l'âge minimum du nœud de l'arbre car le taxa auquel le fossile se réfère aurait pu apparaître avant l'âge attribué au fossile (Fig. 5) (Ho & Phillips, 2009). Il est de plus préférable que les calibrations fossiles soient spécifiées comme des intervalles de temps (par exemple en

Figure 5 Utilisation de fossiles pour calibrer des nœuds dans un chronogramme. (A) Arbre illustrant l'utilisation de fossiles pour calibrer les nœuds. Le groupe de la 'couronne' comprend l'ensemble des taxa vivants et éteints d'un clade ainsi que leur ancêtre commun et leurs descendances. Le groupe du 'tronc' constitue la branche qui joint le groupe de la couronne avec son groupe sœur vivant. '1' & '2' se réfèrent au taxa éteints, i.e. les fossiles. Le fossile '1' partage certaines synapomorphies avec le groupe de la couronne ; il peut être utilisé pour poser une contrainte minimum de l'âge du nœud « A ». Le fossile '2 ' possède toutes les synapomophies du groupe de la couronne ainsi que certaines apomorphies qui permettent de l'identifier sans ambiguïté comme un ancêtre d'une lignée du groupe de la couronne ; il peut ainsi servir d'âge minimum pour contraindre le nœud « B ». Les fossiles rattachés au taxon sœur peuvent également servir d'âge minimum pour contraindre « A ». (B) Méthodes d'incorporation des calibrations fossiles lors d'une analyse de datation moléculaire : (a) Point de calibration, (b) contrainte 'stricte' (i.e. '*hard bound*') sur l'âge minimum du nœud, (c) contrainte 'souple' sur l'âge minimum du nœud, (d) distribution normale, (e) distribution log-normale, (f) distribution exponentielle.(Figure redessinée et traduite de Ho & Phillips (2009)).



utilisant des contraintes souples, Fig. 5) plutôt que des points fixes car cela permet de prendre en compte l'incertitude liée à l'estimation géologique de l'âge du fossile (Ho & Phillips, 2009). Aussi, introduire un point de calibration fossile dans un arbre ne nécessite pas seulement de le placer correctement dans le temps mais aussi, de le placer correctement dans la phylogénie. En effet, un fossile ne peut être utilisé comme point de calibration que lorsqu'il représente le plus vieil ancêtre commun connu de deux lignées présentes dans l'arbre (Donoghue & Benton, 2007; Ho & Phillips, 2009). Enfin, la théorie neutraliste de l'évolution selon laquelle des mutations silencieuses (e.g. à la 3ème base du codon) apparaissent au cours de l'évolution explique pourquoi le taux de substitution entre des individus d'une même espèce devrait être égal au taux de substitution entre taxa d'espèces différentes. Néanmoins, le taux de substitution estimé en interspécifique est, en pratique, moins élevé que le taux de substitution en intraspécifique à cause de la sélection purifiante (Ho et al 2005). Il est donc conseiller de travailler avec une matrice ne présentant qu'un seul individu par espèce afin de ne pas mélanger des taux de substitution interspécifiques et intraspécifiques.

1.2.3.3 Où les groupes sont-ils apparus ?

Les aires de distribution des espèces représentent plus qu'une simple carte. En effet, en plus de définir comment les espèces se distribuent à la surface de la Terre, elles permettent aussi d'identifier les préférences écologiques qui caractérisent chacune des espèces et de comprendre comment les espèces évoluent dans le temps et dans l'espace (Lomolino *et al.*, 2010). Retracer les parcours géographiques des espèces selon leur histoire évolutive appartient au domaine de la biogéographie (voir Encadré 2, p. 12). En étudiant les aires de répartition actuelles des espèces (i.e. les cartes de distribution spatiale), les biogéographes tentent (1) d'inférer où les évènements de divergence qui ont marqué l'histoire évolutive d'un groupe d'espèces ont eu lieu, (2) de déterminer quelles sont les routes de colonisation qui ont été empruntées, (3) d'identifier les facteurs écologiques (e.g. surrection d'une montagne, aridification d'un milieu) responsables des patrons de distribution observés et enfin, (4) de proposer des scénarios qui permettent de décrire l'histoire évolutive du clade étudié (Blondel, 1986). En biogéographie, les fossiles sont très souvent utilisés. Ils sont des preuves ponctuelles de la répartition spatiale et temporelle de la biodiversité passée qui permettent de confirmer les scénarios qui ont été reconstruits, par exemple en validant l'aire et l'époque d'origine d'un clade (e.g. Zhang *et al.* (2012)).

Encadré 2 La Biogéographie Historique

La biogéographie constitue une discipline à la croisée des sciences de la Terre et des sciences de la Vie. Elle peut être définie comme l'étude spatio-temporelle des diversités biologiques, de leurs origines et de leur évolution dans des espaces hétérogènes et changeants (Blondel, 1986). La biogéographie historique évolutive vise à expliquer les distributions de grands groupes taxonomiques en relation avec les évènements climatiques et géologiques et environnementaux majeurs de la planète (Blondel, 1995).

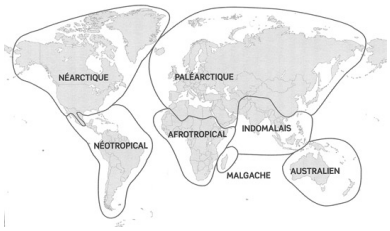


Figure E2.1 Les différentes régions biogéographiques de Sclater et Wallace.

1. Les régions biogéographiques

La biogéographie s'appuie sur les théories de Candolle (1855), Sclater (1858), Huxley (1868), Wallace (1876), Darwin (1859) et Humboldt & Bonpland (1807) qui subdivisent la Terre en plusieurs régions biogéographiques (Fig. E2.1) dont chacune est caractérisée par un ensemble d'espèces vivant dans un environnement écologique et géographique donné (Darlington, 1957). En plus de considérer les distributions des espèces, la subdivision de la Terre en régions biogéographiques prend en compte les grands mouvements tectoniques de l'écorce terrestre (théorie de la dérive des continents de Wegener (1924)) qui ont effectivement participé à la création des unités de diversité biologique (Blondel, 1995).

2. La biogéographie phylogénétique

Les phylogénies moléculaires ont une importance majeure en biogéographie car elles permettent d'établir le cadre d'étude spatio-temporel autour duquel les scénarios biogéographiques vont être reconstruits. Cette discipline qui couple 'phylogénie et biogéographie' est mieux connue sous le nom de 'biogéographie phylogénétique' (Fig. E2.2).

3. Méthode d'analyses

Les scénarios biogéograpiques peuvent être reconstruits en appliquant une méthode de maximum de vraisemblance (Lagrange) basée sur le modèle de dispersion-extinction-cladogenèse (DEC) (Ree & Smith, 2008). La première étape vise à établir un arbre phylogénétique le plus proche possible de l'arbre des espèces et ensuite de replacer les évènements de spéciation dans un cadre temporel (e.g. via des analyses de datation moléculaire). L'étape suivante vise à construire un modèle géographique qui subdivise les régions biogéographiques (Fig. E2.1) en fonction (1) des évènements paléogéographiques liés à la tectonique des plaques et (2) des aires de répartition des espèces actuelles. Ensuite, il faut identifier sur quelle(s) sous-région(s) les espèces sont présentes (cf. chapitre 3 – Appendix S3). La dernière étape avant de procéder aux analyses proprement-dites concerne la construction des matrices de taux de dispersion stratifiées au cours du temps (cette notion de stratification est liée au fait que la période de temps étudiée est divisée en plusieurs petites périodes pour chacune desquelles une matrice de taux de dispersion est établie). Cela consiste à établir la probabilité de disperser d'une aire à l'autre en fonction de la présence éventuelle de barrières géographiques (e.g. montagnes, mers). Pour définir ces probabilités de dispersion et identifier par la suite les facteurs responsables des évènements évolutifs, il est nécessaire de s'intéresser à l'évolution de la tectonique des plaques durant la période concernée par l'étude. Cela permet de définir les barrières physiques éventuelles qui existaient entre les différentes sous-régions du modèle biogéographique. Les analyses de biogéographie peuvent identifier deux modes de spéciation: la dispersion ou la vicariance (Fig. E2.2, et voir le point 1.2.1 pour l'explication de ces modes de spéciation) (Blondel, 1995). Ces deux processus reposent la règle suivante : l'aire d'origine d'un clade est toujours celle du clade ancestral.

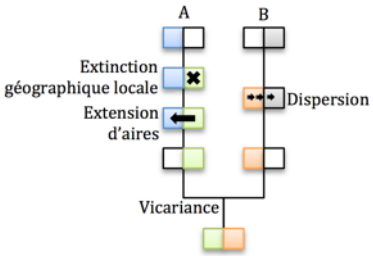


Figure E2.2 Arbre biogéographique. L'aire ancestrale d'origine (verte & orange) a été scindée en deux par vicariance. Dans la branche A, l'aire de distribution (vert) s'est étendue sur l'aire en bleu. Le clade a ensuite subit un événement d'extinction géographique locale sur l'aire verte. Dans la branche B, la dispersion a permis l'installation dans une nouvelle aire où le taxon n'était pas encore présent.

1.2.3.4 Comment étudier la diversité génétique à l'échelle des populations ?

Dans une population idéale de taille infinie où il n'y a pas de mutation/sélection/migration, la loi d'Hardy-Weinberg prédit que les fréquences alléliques ne seront pas modifiées d'une génération à l'autre. De plus, si les individus se reproduisent au hasard (principe de panmixie), les fréquences génotypiques se déduiront alors directement des fréquences alléliques. Or, dans la réalité, les populations sont de tailles finies et elles peuvent être soumises à diverses forces évolutives induisant des fluctuations des fréquences alléliques (cf. 1.2.2). Les modifications génétiques survenant au niveau intra- et inter-populationnel sont du domaine de la 'génétique des populations'. Cette discipline a pour but de mesurer la diversité génétique et de mieux comprendre comment cette diversité se transmet d'une génération à l'autre.

Pour mieux comprendre l'histoire évolutive spatiale et temporelle des populations d'une espèce (e.g. pour tenter d'établir un scénario de colonisation d'une aire géographique occupée par certaines populations d'une espèce), il faut faire appel à une autre discipline. La discipline qui vise à établir les relations de parenté entre les différentes populations géographiques d'une même espèce est la phylogéographie (Avise *et al.*, 1987). Elle intègre des notions de génétique des populations et de géographie (Hickerson *et al.*, 2010). Cette méthode était initialement basée sur l'analyse du polymorphisme du génome mitochondrial et de la distribution géographique des ses lignées. Aujourd'hui, les analyses de phylogéographie incluent souvent également des marqueurs nucléaires tels que, par exemple, les microsatellites dont les taux d'évolution sont assez rapides pour étudier l'histoire évolutive récente (i.e. de l'ordre de quelques milliers d'années) des organismes (Cutter, 2013).

Ainsi, pour mieux comprendre pourquoi les espèces et les populations sont réparties comme elles le sont actuellement, de nombreuses disciplines peuvent être utilisées. Le but principal de la présente thèse était de construire un cadre évolutif stable qui permettrait de nous éclairer sur les processus et les facteurs ayant façonné la biodiversité actuelle d'espèces et de populations de rongeurs.

1.3 Deux modèles de rongeurs pour étudier les distributions spatiales et/ou temporelles d'espèces ou de populations

1.3.1 La diversité des Rongeurs

Composé de ~2277 espèces de ~477 genres répartis dans plus de 33 familles, l'ordre des Rongeurs représente presque la moitié de la diversité mammalienne (Wilson & Reeder, 2005). De plus, chaque année de nouvelles espèces sont découvertes comme, par exemple, *Halmaheramys bokimekot* et *Laonastes aenigmamus* (Jenkins *et al.*, 2005; Fabre *et al.*, 2013). Les rongeurs occupent la quasi-totalité des habitats, excepté ceux situés en Antarctique et sur quelques îles océaniques. Ils occupent même les écosystèmes secs et chauds tels que les déserts, et les régions septentrionales plus froides (e.g. les forêts boréales) (Myers *et al.*, 2015).

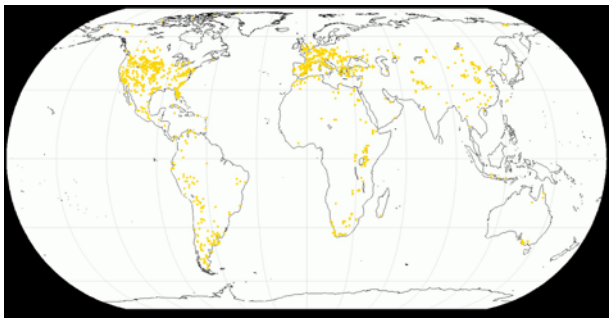


Figure 6 Carte de distributions des spécimens fossiles de Rongeurs (*Paleobiology Database*, août 2015).

Le registre fossile de l'ordre des Rongeurs est assez riche (i.e. à ce jour, ~4226 spécimens fossiles sont répertoriés, Fig. 6) (PaleobiologyDatabase, 2015). Le registre fossile et les analyses de datation moléculaire s'accordent à dire que l'ordre des Rongeurs s'est largement diversifié depuis le Cénozoïque (i.e. 65.5 Ma – Présent) (Fabre *et al.*, 2012b; PaleobiologyDatabase, 2015). En effet, les plus anciens fossiles connus de Rongeurs ont été découverts en Amérique, en Europe et en Asie et ont été datés de ~59.2 Ma (Stucky & McKenna, 1993). Cette incroyable diversité (présente et passée) fait des rongeurs des modèles intéressants pour comprendre les patrons de distribution temporels et spatiaux d'espèces ou de populations.

1.3.2 Les Dipodoidea, un modèle pour étudier l'évolution de la biodiversité à l'échelle macroévolutive

1.3.2.1 Les Dipodoidea (Rongeurs : Myodonta)

La superfamille des Dipodoidea constitue le groupe frère des Muroidea (Wilson & Reeder, 2005). Au sein des Dipodoidea, trois grands groupes d'organismes peuvent être distingués : les sicistes (Sicistinae, un genre), les souris-sauteuses (Zapodinae, trois genres) et les gerboises (Allactaginae, trois genres ; Cardiocraniinae, trois genres ; Dipodinae, cinq genres ; Euchoreutinae, un genre) (Holden & Musser, 2005). Chacun de ces groupes présente un morphotype différent : les sicistes – quadrupèdes – ont un morphotype 'souris', les souris-sauteuses – essentiellement quadrupèdes – présentent un allongement des membres postérieurs (sans toutefois présenter de disproportions de la taille des paires de pattes) et les gerboises – locomotion par saltation bipède – possèdent un morphotype 'kangourou' (i.e. des membres antérieurs de taille réduite, des pattes arrières significativement plus longues que les pattes avant, et une longue queue ornementée d'un plumeau) (Shahin, 2005; Myers *et al.*, 2015).

Actuellement, la majorité des espèces sont considérées comme '*préoccupation mineure*' mais 32% des espèces de Dipodoidea sont menacées d'extinction (IUCN, 2015). Les principales menaces sont pour elles : la destruction et la fragmentation de leurs habitats (e.g. déforestation) (IUCN, 2015). Selon la *liste rouge des espèces menacées de l'Union Internationale pour la Conservation de la Nature* (IUCN), la siciste du Caucase (*Sicista caucasica*) et la gerboise à quatre doigts (*Allactaga tetradactyla*) sont considérées '*vulnérables'*. La siciste de Klukhor (*Sicista kluchorica*), la gerboise de Vinogradov (*Allactaga vinogradovi*), la gerboise d'Euphrate (*Allactaga euphratica*), la gerboise à queue épaisse (*Pygeretmus zhitkovi*) sont '*quasi-menacées'*. La siciste d'Ossétie (*Sicista kazbegica*) et la siciste arménienne (*Sicista armenica*) sont désormais '*en danger'* d'extinction. De plus, pour de nombreuses espèces, nous ne disposons malheureusement pas de suffisamment de données et celles-ci sont donc classées dans la catégorie '*Données Insuffisantes'* (i.e. *Cardiocranius paradoxus, Sicista caudata, Sicista pseudonapaea, Salpingotulus michaelis, Salpingotus crassicauda, Salpingotus heptneri, Salpingotus pallidus, Allactaga firouzi*), ce qui revient à les considérer comme espèces menacées comme le signale l'IUCN.

1.3.2.2 Controverse à propos de leur systématique et taxonomie

Dans la littérature, la superfamille des Dipodoidea comprend 51 espèces réparties dans 16 genres de six sous-familles, toutes de la famille des Dipodidae (Holden & Musser, 2005). Cette classification établie sur la base de caractères morphologiques a amené à de nombreuses controverses à propos de la systématique et la taxonomie des Dipodoidea. Par exemple, de nombreuses discussions existent autour du nombre de familles qu'il convient d'établir au sein des Dipodoidea. Selon les auteurs et les traits morphologiques étudiés, la superfamille des Dipodoidea peut en effet être subdivisée en une (Dipodidae) (Winge, 1887; Thomas, 1896; Vinogradov, 1930; Vinogradov, 1937; Ellerman, 1940; McKenna & Bell, 1997; Holden & Musser, 2005), deux (e.g. Dipodidae et Zapodidae) (e.g. Lyon (1901); Stein (1990)), trois (e.g. Allactagidae, Dipodidae, et Zapodidae) (Zazhigin & Lopatin, 2000), quatre (Allactagidae, Dipodidae, Sicistidae, et Zapodidae) (Shenbrot, 1992), voire cinq familles (Sminthidae, Zapodidae, Dipodidae, Allactagidae, Euchoreutidae) (Shenbrot *et al.*, 2008). De façon générale, les études précédentes n'étaient pas d'accord sur les parentés évolutives et les statuts taxonomiques de plusieurs clades.

La classification proposée par Holden and Musser (2005) – qui peut être définie comme 'conservatrice' - a en réalité été appliquée en dépit des diverses taxonomies proposées durant le siècle dernier. Par conséquent, de nombreuses questions systématiques et taxonomiques restent toujours en suspens, comme par exemple celle en lien avec le clade basal des Dipodoidea. En fait, avant cette thèse, aucune phylogénie moléculaire des Dipodoidea n'avait encore été reconstruite et comparée aux données morphologiques pour établir pertinemment la systématique et la taxonomie de ces rongeurs.

1.3.2.3 Distributions et préférences écologiques

La majorité des espèces de Dipodoidea occupent la région Paléarctique (Afrique du Nord compris) mais les espèces appartenant aux genres *Zapus* et *Napaeozapus* se distribuent dans la région Néarctique. Les sicistes, les souris-sauteuses et les gerboises sont chacune inféodée à un habitat particulier (Fig. 7) : les sicistes sont essentiellement distribuées dans les prairies et les zones de toundra des forêts subalpines et boréales de Russie et d'Europe du Nord ; les souris-sauteuses peuplent quant à elles les aires lacustres, les marais et les milieux forestiers

du Canada, des Etats-Unis et de façon plus anecdotique, de Chine ; enfin, les gerboises sont assujetties à des environnements beaucoup plus secs et chauds tels les déserts, les semi-déserts et les steppes du Nord de l'Afrique et du Moyen-Orient (IUCN, 2015). Lorsqu'on examine indépendamment chaque clade, la superfamille des Dipodoidea est très intéressante d'un point de vue 'évolutif et reconstruction de scénarios biogéographiques' car certaines espèces occupent des habitats arides et isolés (e.g. déserts d'Afrique du Nord et steppes de Mongolie) (Fig. 7). Il est donc intéressant de comprendre comment ces patrons de distribution disjoints sur l'Holarctique (e.g. Afrique du Nord, Amérique du Nord) ont été mis en place.

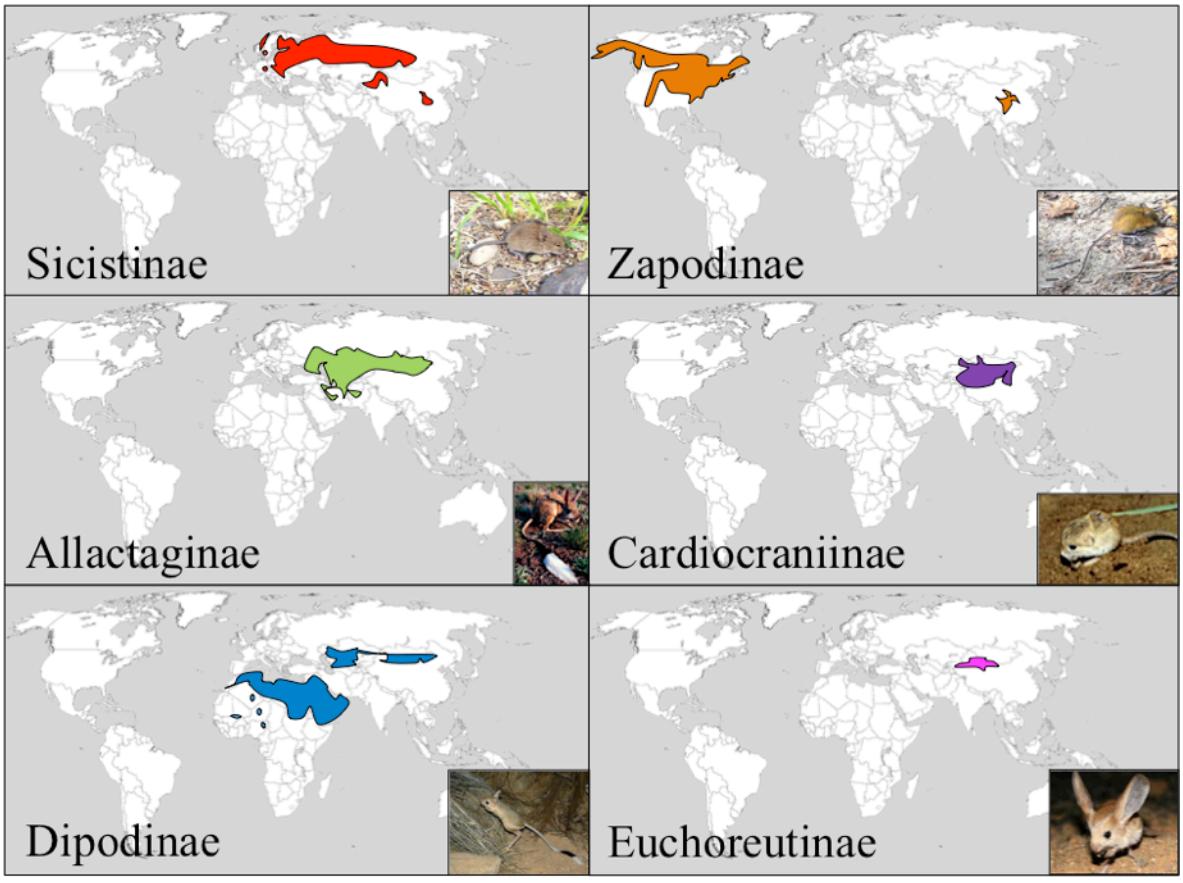
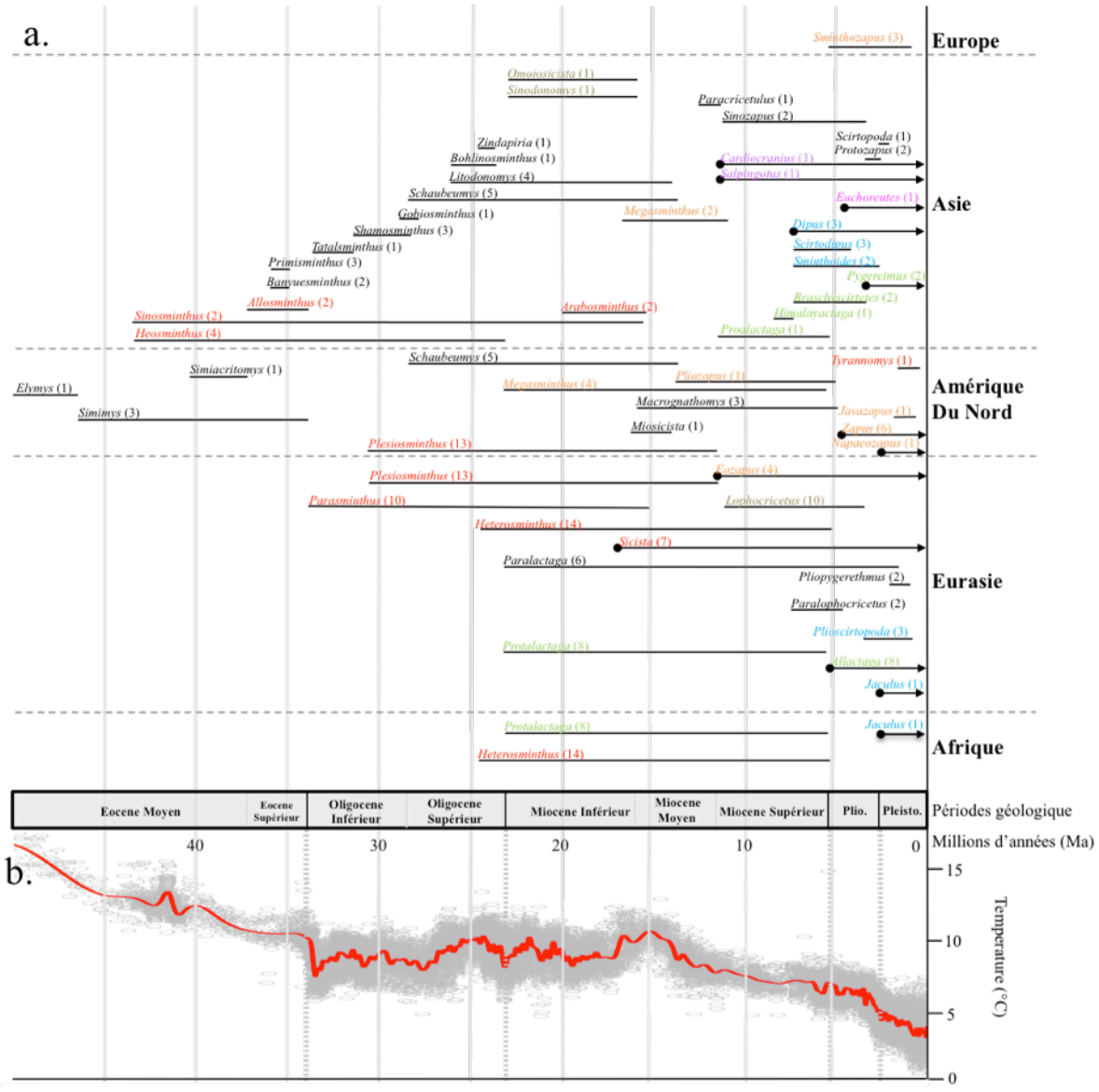


Figure 7 Cartes de distributions des sous-familles de Dipodoidea. Les aires de répartition par sous-famille ont été reconstruites à partir des cartes de distribution de chacune des espèces obtenues sur *The IUCN Red List of Threatened species* (IUCN 2015). Les photos associées aux sous-familles de Sicistinae, Zapodinae, Dipodinae, Allactaginae, Cardiocraniinae et Euchoreutinae correspondent respectivement aux espèces *Sicista betulina, Napaeozapus insignis, Jaculus orientalis, Allactaga major, Salpingotus crassicauda* et *Euchoreutes naso* (© Wikipédia Commons).

1.3.2.4 Le registre fossile des Dipodoidea

Le registre fossile des Dipodoidea est plutôt riche (Fig. 8 et Table S1). Les plus anciennes traces fossiles connues de Dipodoidea ont été découvertes en Amérique du Nord et datées de l'Eocène moyen. Elles ont été identifiées comme appartenant aux genres *Elymys* ('?Zapodidae', définit ainsi à cause de l'incertitude liée à sa position taxonomique) et *Simimys*

Figure 8 Registre fossile des Dipodoidea. (a) Graphe des genres fossiles de Dipodoidea en fonction de leur répartition géographique. Les couleurs se réfèrent aux différentes sous-familles : rouge – Sicistinae, orange – Zapodinae, violet – Cardiocraniinae, rose – Euchoreutinae, vert – Allactaginae, et bleu – Dipodinae. Les nombres entre parenthèses font référence au nombre d'espèces fossiles identifiées pour chaque genre. La ligne noire en dessous des noms de genre indique la période d'occurrence du genre (celles qui sont ornées d'une flèche indique les genres toujours présents aujourd'hui). (b) Courbe des paléo-températures selon Zachos *et al.* 2008.



(Simimyidae, famille éteinte des Dipodoidea). En Asie, les plus anciens représentants connus de Dipodoidea correspondent à *Heosminthus* (Zapodidae ou Dipodidae, genre présent de l'Eocène moyen à l'Oligocène supérieur) et *Sinosminthus* (Zapodidae, de l'Eocène moyen au Miocène moyen) (Wang, 1985; Emry & Korth, 1989; Kelly, 1992; Tong, 1997; Daxner-Höck, 2001). Sur base de ces données paléontologiques, les hypothèses les plus couramment admises ont été que les prémices de l'histoire des Dipodoidea ont eu lieu en Amérique du Nord et que rapidement, dans le courant de l'Eocène moyen, ils sont passés de la région Néarctique à la région Paléarctique.

1.3.3 Le campagnol roussâtre (*Myodes glareolus*), un modèle pour étudier l'évolution de la biodiversité à l'échelle microévolutive

1.3.3.1 Distribution spatiale et écologie de *Myodes glareolus* (Muroidea : Arvicolinae)

Le campagnol roussâtre possède une large distribution sur la région Paléarctique, s'étendant des îles britanniques à travers l'Europe et la Russie jusqu'au lac Baïkal et au-delà du Cercle Polaire dans la région de Fennoscandie (Fig. 9) (IUCN, 2015). Son habitat préférentiel se situe au sein des forêts de ligneux (surtout de type chênaie-hêtraie) à sous-bois (ronciers, broussailles et zones à fougères) moyennement dense ornés d'arbres morts et de branches cassées. Le campagnol roussâtre apprécie les sols couverts de mousse et d'une épaisse couche de litière (Quéré & Le Louarn, 2011; IUCN, 2015).

Figure 9 Carte de distribution du campagnol roussâtre (*Myodes glareolus*) (IUCN, 2015). Photographie © Roger Butterfield.

Son domaine vital est d'environ 800 m^2, mais il peut exceptionnellement atteindre 2000 m^2. Certains individus effectuent des déplacements à longue distance (environ 700 m). La densité peut varier de 6 à 12 individus par hectare, voire environ 50 en fonction de la réussite de reproduction (Quéré & Le Louarn, 2011).

1.3.3.2 Distributions des populations

En Europe, les populations de campagnol roussâtre se répartissent en plusieurs lignées mitochondriales (Fig. 10) (Deffontaine *et al.*, 2005). Parmi elles, une lignée particulière a été identifiée (Fig. 10, la lignée introgressée (bleu foncée)). Elle se distribue de la moitié supérieure de la Suède à travers la Finlande jusqu'au centre de la Russie. Elle est caractérisée par le génome mitochondrial d'une autre espèce, le campagnol de la Taïga (*Myodes rutilus*) (Tegelström, 1987; Deffontaine *et al.*, 2005; Potapov *et al.*, 2007; Abramson *et al.*, 2009a; Boratynski *et al.*, 2011). L'espèce *M. rutilus* a une répartition plutôt nordique par rapport à *M. glareolus*. Elle est caractéristique des habitats secs et froids des forêts boréales de conifères de l'Holarctique (Fig. 11). Au nord de l'Europe, cette lignée introgressée du campagnol roussâtre vient au contact de la lignée Ouest européenne (Fig. 10, lignée verte) au centre de la Suède et de la lignée Est européenne au centre de la Finlande (Fig. 10, lignée jaune), mettant ainsi en évidence deux zones de contact entre des lignées mitochondriales différenciées de campagnol roussâtre (Deffontaine *et al.*, 2005).

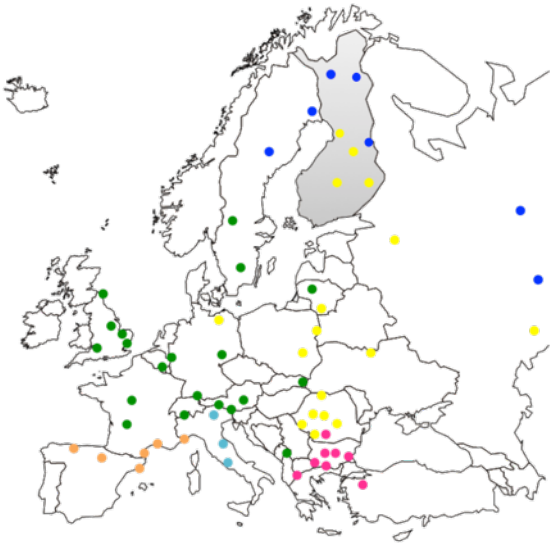


Figure 10 Carte de distribution des lignées mitochondriales de campagnol roussâtre (*M. glareolus*) en Europe. De gauche à droite et de bas en haut : orange – lignée espagnole, vert – lignée Ouest-européenne, bleu clair – lignée italienne, rose – lignée des Balkans, jaune – lignée Est européenne, bleu foncé – lignée introgressée. Le pays grisé est la Finlande. (Carte modifiée à partir de Boratynsky *et al.* 2011 & Deffontaine *et al.* 2005).

Figure 11 Carte de distribution de *Myodes glareolus* et *M. rutilus*. Les aires de répartitions sont indiquées en : rouge pour *M. glareolus*, vert pour *M. rutilus*, et bleu pour la zone de sympatrie entre les deux espèces. (Boratynski *et al.*, 2014).

La zone de contact finlandaise entre la lignée mitochondriale introgressée (mitotype RUT) et la lignée Est européenne (mitotype GLA) est particulièrement intéressante. Des études phylogénétiques à travers la zone de contact finlandaise ont montré que d'un point de vue mitochondrial, deux groupes monophylétiques étaient effectivement bien présents : celui au nord qui est proche des spécimens de *M.*

rutilus (le mitotype RUT), et l'autre au sud, proche des autres spécimens de *M. glareolus* (le mitotype GLA) (Deffontaine *et al.*, 2005; Boratynski *et al.*, 2011; Boratynski *et al.*, 2014) ; alors que concernant la variabilité au sein du génome nucléaire, les gènes (i.e. *BRCA1, G6pd, GHR, LCAT, vWF, ADRAB2*) n'ont montré aucune différentiation entre les spécimens GLA et RUT (Boratynski *et al.*, 2011; Boratynski *et al.*, 2014). Sur base d'un gène mitochondrial (le cytochrome *b,* cytb), Deffontaine *et al.* (2005) avaient proposé que cette zone de contact en Finlande soit le résultat d'un contact secondaire entre des populations de campagnol roussâtre qui auraient divergé en allopatrie pendant le dernier âge glaciaire et qui seraient revenues en contact lors de la recolonisation postglaciaire. Or, sous cette hypothèse, on se serait attendu à observer au moins une légère différentiation au sein du génome nucléaire entre les lignées mitochondriales GLA et RUT situées de part et d'autre de la zone de contact finlandaise. Dès lors, appuyer cette hypothèse de contact secondaire entre ces deux lignées mitochondriales différenciées nécessite de tester avec d'autres marqueurs moléculaires (tels que les microsatellites qui ont des taux d'évolution assez rapide (de l'ordre de 10^{-6} à 10^{-2}) pour étudier l'histoire évolutive récente des campagnols de Finlande), si la zone de contact finlandaise est associée à de la différenciation sur le génome nucléaire. Si ce n'est pas le cas, une hypothèse alternative serait que la zone de contact résulterait d'un événement unique de recolonisation et l'introgression mitochondriale par *M. rutilus* aurait alors eu lieu en Finlande au sein d'une population génétiquement homogène.

1.4 Questions & objectifs de la thèse

La question « Qu'est ce qui détermine les patrons de distribution et de diversité d'espèces et de populations dans le temps et l'espace ? » a donné matière à réflexion lors de cette thèse.

'Comprendre ce qui façonne la biodiversité et quel(s) facteur(s)/processus est(sont) impliqué(s) dans la distribution spatiale et temporelle du vivant' est un des thèmes de recherche les plus explorés de nos jours (Pennisi, 2005). Grâce aux avancées notoires dans le domaine de la biologie évolutive moléculaire, différentes approches permettent désormais d'aborder ce sujet dans le but de mettre en évidence les facteurs responsables des patrons évolutifs observés. Dans le cadre de cette thèse de doctorat, j'ai donc étudié divers patrons de distributions d'espèces et de populations à grandes et petites échelles spatio-temporelles à partir d'une approche moléculaire. Ce manuscrit de thèse est composé de trois chapitres reprenant chacun un thème de recherche ayant permis la rédaction d'un article en préparation ou publié dans un journal scientifique à comité de lecture. Ces trois sujets de recherche sont :

- **CHAPITRE 2 – Quand la phylogénie moléculaire aide à résoudre la taxonomie et la systématique d'un groupe : cas de la superfamille des Dipodoidea**

Lebedev V.S., Bannikova A.A., Pagès M., **Pisano J.,** Michaux J.R. & Shenbrot G.I. (2013) Molecular phylogeny and systematics of Dipodoidea: a test of morphology-based hypotheses. *Zoologica Scripta,* 42(3) : 231-249.

La phylogénie des Dipodoidea est restée longtemps sujette à controverses à cause d'hypothèses divergentes concernant les parentés phylogénétiques au sein de la superfamille des Dipodoidea, lesquelles étaient essentiellement basées sur des caractères morphologiques. Pour la première fois, les relations phylogénétiques entre les espèces et les clades de cette superfamille ont été reconstruites sur base d'une approche moléculaire. Cette phylogénie moléculaire a ensuite été comparée à une phylogénie reconstruite à partir de caractères morphologiques. Cela a permis de répondre à diverses questions de taxonomie et de systématique et au final, de réviser la taxonomie des Dipodoidea sur la base de caractères moléculaires et morphologiques. Mon implication dans le cadre de cette étude s'est principalement axée autour de la reconstruction de la phylogénie moléculaire des Dipodoidea à partir d'approches probabilistes.

- **CHAPITRE 3 – Impact des changements environnementaux, climatiques et géologiques passés sur la distribution et l'histoire évolutive des espèces : cas de la superfamille des Dipodoidea**

Pisano J., Condamine F.L., Lebedev V.S., Bannikova A.A., Quéré J.-P., Shenbrot G.I., Pagès M. & Michaux J.R. (2015) Out-of-Himalaya: the impact of past Asian environmental changes on the evolutionary and biogeographical history of Dipodoidea (Rodentia). *Journal of Biogeography,* 42(5) : 856-870.

Avant cette thèse, l'histoire évolutive des Dipodoidea était très peu connue, aussi bien au niveau phylogénétique que biogéographique. Une série de questions concernant leur histoire évolutive demeuraient ainsi sans réponse : quand et où les clades de Dipodoidea sont-ils apparus ? Quel est leur centre d'origine ? Quelles routes de colonisation les clades ont-ils empruntés ? Quels facteurs ont influencé leurs évènements de diversification ? Dans le cadre de cette étude biogéographique, l'idée a été de récolter un maximum d'espèces de Dipodoidea pour ensuite reconstruire la phylogénie moléculaire la plus complète de la superfamille à partir de laquelle il serait alors possible d'estimer les dates de divergences et les routes de colonisation permettant de retracer leur parcours évolutif depuis leurs origines jusqu'au Présent.

- **CHAPITRE 4 – Estimation de la différenciation nucléaire entre lignées mitochondriales et impact sur l'identification d'une zone de contact secondaire : cas des populations de campagnol roussâtre (*Myodes glareolus*) de Finlande**

Pisano J., Leblois R., Charbonnel N., Cosson J.-F., Piry S., Galan M., Huitu O., Henttonen H., Michaux J.R. (in prep.) Testing the presence of a barrier to nuclear gene flow between two distant mitochondrial lineages of the bank vole (*Myodes glareolus*) in central Finland. *Molecular Ecology.*

En Finlande, les patrons mitochondriaux de distribution du campagnol roussâtre sont déjà bien décrits. Deux lignées mitochondriales ont été identifiées : l'une introgressée (RUT) au Nord, et l'autre non-introgressée (GLA) au Sud, lesquelles se rencontrent au centre de la Finlande. Le flux génique nucléaire à travers la zone de contact mitochondrial entre GLA et RUT était cependant peu connu et l'hypothèse de zone de contact secondaire n'avait encore jamais été testée. Il restait donc à évaluer le statut de la zone de contact au niveau nucléaire. Le but a donc été d'estimer spécifiquement le flux de gènes nucléaire entre ces deux lignées mitochondriales de la même espèce pour estimer si la zone de contact constitue bien une zone de contact secondaire.

Ma thèse se termine enfin par un dernier chapitre de « discussion générale & perspectives ».

Chapitre 2 :

Quand la phylogénie moléculaire aide à résoudre la taxonomie et la systématique d'un groupe :

Cas de la superfamille des Dipodoidea

Mots-clés: Systématique, phylogénie moléculaire, révision taxonomique, caractères morphologiques, homologies

Au 19ème siècle, Darwin affirme que le vivant possède en lui la mémoire du passé à partir de laquelle il est possible d'identifier les liens de parenté entre les espèces (Lecointre & Le Guyader, 2001). Dans un premier temps, les études de systématique et taxonomie se sont intéressées à l'étude de caractères morphologiques mais après la découverte de l'ADN en 1953, une autre catégorie de caractères est née, les caractères moléculaires. Diverses méthodes ont alors pu être développées pour établir les liens de parenté entre les espèces : (1) la 'méthode cladistique de parcimonie' qui définit l'arbre le plus parcimonieux comme étant celui qui requiert le plus petit nombre d'évènements évolutifs pour les états de caractères étudiés, autrement-dit l'arbre 'le plus court' (Darlu & Tassy, 1993) ; (2) les 'méthodes phénétiques' (e.g. l'UPGMA ('*unweighted pair group method using averages*') et le 'neighbour-joining') qui diffèrent de la cladistique par le fait qu'elles quantifient la ressemblance générale entre les organismes en calculant un indice de similitude global entre deux taxa, i.e. une distance (e.g. le nombre de nucléotides différents entre deux espèces) (Lecointre & Le Guyader, 2001) ; (3) 'les approches probabilistes' qui utilisent des algorithmes de reconstruction d'arbre exploitant les probabilités de changement d'un état de caractère en un autre pour calculer la vraisemblance des données pour un arbre (Encadré 1, p. 8).

Quel que soit le type de marqueur (morphologique ou moléculaire), la systématique phylogénétique tend à établir les parentés évolutives sur base de caractères homologues, i.e. de synapomorphies (Lecointre & Le Guyader, 2001). Cependant, il se peut que des caractères non-hérités de l'ancêtre commun (i.e. des homoplasies) soient utilisés à tort pour reconstruire une phylogénie. Divers processus (e.g. convergence et parallélisme) conduisent alors à des reconstructions fausses des parentés phylogénétiques entre les clades et des conflits phylogénétiques peuvent alors être observés lorsque différents types de marqueurs sont comparés (Hillis, 1987). Par exemple, dues à des pressions biotiques et abiotiques sur certains traits morphologiques, des discordances entre les phylogénies moléculaire et morphologique ont été identifiées au sein d'espèces d'éponges (DeBiasse & Hellberg, 2015). De façon générale, les caractères morphologiques sont sujets à de l'homoplasie ; une phylogénie moléculaire robuste est donc souvent confrontée aux phylogénies morphologiques pour s'assurer de la validité de la reconstruction phylogénétique et établir les parentés selon l'histoire évolutive des clades (Sanderson & Hufford, 1996).

Résumé de l'article

La superfamille des Dipodoidea comprend les gerboises (espèces avec des pattes arrière significativement plus longues que les pattes avant, exclusivement bipèdes et avec un mode de locomotion exclusivement saltatoire), les souris-sauteuses (espèces principalement quadrupèdes ayant cependant la capacité d'effectuer de longs bonds grâce à leurs pattes arrière allongées) et les sicistes (espèces strictement quadrupèdes et non-saltatoires) (Fokin, 1978). Classiquement, on y distingue 51 espèces réparties dans 16 genres de six sous-familles (Allactaginae, Cardiocraniinae, Dipodinae, Euchoreutinae, Sicistinae et Zapodinae) toutes inclues dans la famille des Dipodidae (Holden & Musser, 2005). [Désormais, par souci de priorité, le clade des Sicistinae (Allen 1901) sera dénommé Sminthinae (Brandt, 1855).] Cette classification en six sous-familles est néanmoins controversée. De nombreuses études taxonomiques se sont en effet intéressées aux Dipodoidea sans établir explicitement les liens de parenté entre les clades majeurs de la superfamille. En fonction des caractères étudiés, le nombre de familles décrites peut varier de une à cinq et, les relations de parenté entre chacune des sous-familles peuvent également changer d'une classification à l'autre (Fig. 1, Lebedev *et al.* (2012), p. 34). L'objectif principal de cette étude a donc été dans un premier temps d'établir une phylogénie moléculaire robuste des Dipodoidea pour ensuite la confronter à des phylogénies reconstruites à partir de traits morphologiques. Dans un second temps, il a fallu répondre à diverses questions de taxonomie et systématique pour permettre une révision de la taxonomie des Dipodoidea sur la base de caractères moléculaires et morphologiques : (i) Quelle est la première sous-famille à avoir divergé ? (ii) Il a été proposé que les Cardiocraniinae formeraient un groupe monophylétique avec le genre *Paradipus* qui est communément classé au sein des Dipodinae ('Cardiocraniinae + *Paradipus'*) ou avec les Dipodinae ('Cardiocraniinae + Dipodinae'), et que *Euchoreutes naso* – la seule et unique espèce de la sous-famille des Euchoreutinae – serait phylogénétiquement proche du genre *Sicista* ('Euchoreutinae + Sminthinae') ou de la sous-famille des Allactaginae ('Euchoreutinae + Allactaginae') ou encore que *E. naso* formerait un groupe monophylétique avec les sous-familles de 'Cardiocraniinae + Dipodinae + Allactaginae'. Etant donné cela, quelle position ces taxa occupent-ils au sein de phylogénies moléculaire et morphologique ? (iii) Quelles sont les relations intergénériques au sein des sous-familles ? (iv) Les Sminthinae et les Zapodinae constituent-ils un groupe monophylétique ? (v) Le genre *Allactaga* (Allactaginae) est-il bien monophylétique ? (vi) Les phylogénies moléculaires et

morphologiques sont-elles congruentes ? (vii) Sur la base de caractères moléculaires, quelle serait l'origine de la bipédie chez les Dipodoidea ?

Au total, 20 espèces de 15 genres de Dipodoidea représentant les six sous-familles décrites dans *Mammal Species of the World* (Holden & Musser, 2005) ont été étudiées. En plus, 20 espèces de Rongeurs ont été utilisées en tant que groupes externes pour enraciner les phylogénies de Dipodoidea. Pour les analyses moléculaires, les régions de quatre gènes nucléaires codants ont été amplifiées et séquencées : les gènes du BRCA1 (breast cancer 1), du GHR (growth hormone receptor), de l'IRBP (interphotoreceptor binding protein) et du RAG1 (recombination activating gene 1). Les analyses phylogénétiques ont d'abord été réalisées sur chaque gène indépendamment puis sur le jeu de données concaténées. La congruence des gènes a été testée via des tests ILD (*Incongruence length difference test*). Les phylogénies moléculaires ont été reconstruites selon la méthode de maximum de vraisemblance (avec *Treefinder*), la méthode d'inférence bayésienne (avec *MrBayes*) et celle du maximum de parcimonie (avec *PAUP*). Des hypothèses concernant la position phylogénétique de certains taxa ont également été vérifiées en testant la topologie de l'arbre avec des AU tests (*approximate unbiased tests*). Pour les analyses morphologiques, des traits liés à la dentition, à la bulle auditive, au gland du pénis et aux glandes reproductrices accessoires ont été étudiés. Les phylogénies morphologiques ont été reconstruites sur base d'une approche de maximum de parcimonie (*PAUP*). Pour tester si les données morphologiques étaient compatibles avec la topologie moléculaire, les analyses phylogénétiques morphologiques ont été contraintes en fonction de la topologie moléculaire précédemment obtenue.

Les phylogénies obtenues sur chaque gène indépendamment n'ont pas donné des topologies exactement identiques ; cependant, les tests ILD et AU ont indiqué qu'il n'y avait pas d'incongruences entre les topologies de gènes ($P = 0.155$) et que la topologie obtenue à partir du jeu de données concaténées était significativement meilleure que celles obtenues sur chaque gène indépendamment ($P = 0.29, 0.27, 0.26, 0.28$ pour IRBP, GHR, BRCA1, RAG1 respectivement). Dès lors, nous nous focaliserons désormais sur les résultats obtenus à partir de la matrice de gènes combinés. Quelle que soit la méthode de reconstruction appliquée, la topologie obtenue était toujours identique. L'AU test a rejeté toutes les hypothèses testées sauf celle concernant la parenté de *E. naso* avec les Allactaginae ($P < 0.05$). Concernant les marqueurs morphologiques, les tests ILD ont montré une discordance significative entre les

différentes topologies obtenues sur la base de chaque marqueur indépendamment ($P < 0.01$). Normalement, les marqueurs morphologiques ne pouvaient donc pas être combinés. Cependant, il semblerait que le fait de rejeter H_0 ($P < 0.05$) lors d'un test ILD n'est pas forcément synonyme de combinaisons inadéquates entre marqueurs (Barker & Luzoni 2002, Hipp et al 2004). Dès lors, les phylogénies reconstruites à partir du jeu de données morphologiques combinées ont tout de même été reconstruites. En contraignant la même topologie que celle obtenue précédemment à partir des données moléculaires, la topologie de la phylogénie morphologique inférée était moins parcimonieuse que celle non-contrainte ($P < 0.05$).

Les données morphologiques ont ensuite été confrontées à des données moléculaires pour revoir la taxonomie de la superfamille. La phylogénie moléculaire obtenue à partir du jeu de données combinées s'est avérée bien résolue et supportée (Fig. 2, Lebedev *et al.* 2012, p. 42). La monophylie de chacune des sous-familles de gerboises a été confirmée. Le branchement situé à la base de l'arbre concerne les Sminthinae, qui se positionnent ainsi en groupe frère de toutes les autres sous-familles. Les Sminthinae et les Zapodinae ne forment pas un ensemble monophylétique. La proximité entre Cardiocraniinae et Dipodinae a été rejetée, alors que les caractères morphologiques soutenaient fortement cette monophylie. Au contraire, il a été démontré que les Cardiocraniinae représentent le groupe frère à toutes les autres sous-familles de gerboises (i.e. Allactaginae, Dipodinae, Euchoreutinae). A propos de la position d'*E. naso*, la potentielle relation de parenté avec le genre *Sicista* a été définitivement rejetée. Tandis que sa proximité phylogénétique avec les Allactaginae et Dipodinae s'est révélée monophylétique, sa position au sein de ce clade reste cependant toujours confuse (i.e. trichotomie non-résolue). Concernant les relations intergénériques, il est intéressant de noter que sur base de caractères morphologiques, *Allactaga, Allactodipus* et *Pygeretmus* constituent trois genres bien distincts, alors que la phylogénie moléculaire suggère plutôt une paraphylie du genre *Allactaga* (*A. elater* serait plus proche de *Pygeretmus* et l'espèce *A. major* du genre *Allactodipus*).

Les phylogénies moléculaires et morphologiques se sont avérées ne pas être congruentes entre elles. En réalité, les discordances observées sont surtout expliquées par des taux d'homoplasie élevés au sein des caractères morphologiques étudiés (notamment ceux liés à la dentition). En effet, par exemple, sur base des caractères moléculaires, la proximité

phylogénétique entre les Cardiocraniinae et les Dipodinae a été définitivement rejetée, alors que la phylogénie morphologique soutenait cette parenté (bootstrap = 93%) de façon robuste.

Pour estimer si l'acquisition de la bipédie a eu lieu une fois ou plusieurs fois indépendamment, les arbres phylogénétiques moléculaire et morphologique ont été analysés. La phylogénie moléculaire la plus parcimonieuse suggère que l'ancêtre commun des gerboises était bipède et qu'il y a eu une acquisition progressive de la bipédie : les sicistes basales sont quadrupèdes (Sminthinae), ensuite les souris-sauteuses (Zapodinae) sont plutôt quadrupèdes mais elles peuvent se déplacer sur leurs deux longues pattes arrières, les gerboises 'primitives' (Cardiocraniinae) ne se déplacent plus que par bonds sur leurs pattes arrière significativement plus longues que les pattes avant, et enfin le groupe 'avancé' des gerboises (Allactaginae, Dipodinae, Euchoreutinae) présente une fusion des métatarses en un os canon procurant plus de puissance lors du saut. Néanmoins, bien que les données laissent supposer une seule voie d'acquisition du caractère, il serait judicieux d'étudier plus en détail cette problématique avant de valider cette hypothèse car les caractères morphologiques sont bien connus pour être fortement sujets au parallélisme adaptatif.

Cette étude regroupant pour la première fois des phylogénies moléculaire et morphologique des Dipodoidea a donc permis de proposer une nouvelle taxonomie, au sein de laquelle trois familles sont dorénavant reconnues : les Sminthidae (les sicistes) – en groupe frère des autres familles –, les Zapodidae (les souris-sauteuses) et les Dipodidae (les gerboises). Vu la monophylie des taxa bipèdes, les sous-familles d'Allactaginae, de Cardiocraniinae, de Dipodinae et d'Euchoreutinae ont été regroupées dans la même famille, celle des Dipodidae. Cette structuration taxonomique fait référence à celle précédemment proposée par Vorontsov *et al.* (1971) qui était basée sur des caractères de type caryotypique. Au final, au terme de cette étude, deux questions restent en suspens: (1) bien que la monophylie des Allactaginae soit bien soutenue, des analyses complémentaires sont indispensables pour mieux cerner le statut paraphylétique d'*Allactaga* et statuer sur la parenté avec les espèces d'*Allactodipus* et *Pygeretmus,* (2) la trichotomie 'Allactaginae + Dipodinae + Euchoreutinae' reste non-résolue.

Article

Molecular phylogeny and systematics of Dipodoidea: a test of morphology-based hypotheses

Vladimir LEBEDEV[1], Anna BANNIKOVA[2], Marie PAGES[3, 4], Julie PISANO[3, 4], Johan R. MICHAUX[3, 4], Gregory I. SHENBROT[5]

Zoologica Scripta (2013), 42(3): 231-249, DOI: 10.1111/zsc.1200

[1] *Zoological Museum of Moscow State University, 125009 Moscow, Russia.*
[2] *Lomonosov Moscow State University, 119992 Moscow, Russia.*
[3] *INRA, UMR 1062 CBGP (INRA/IRD/CIRAD/Montpellier SupAgro), 34988 Montferrier-sur-Lez, France.*
[4] *Laboratory of Conservation Genetics, Institute of Botany (B22), University of Liège, 4000 Liège (Sart-Tilman), Belgium.*
[5] *Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of Negev, 84990 Midreshet Ben-Gurion, Israel.*

ABSTRACT

The superfamily Dipodoidea (Rodentia, Myomorpha) in its current interpretation contains a single family subdivided into six subfamilies. Four of them include morphologically specialized bipedal arid-dwelling jerboas (Dipodinae – three-toed jerboas, Allactaginae – five-toed jerboas, Cardiocraniinae – pygmy jerboas and Euchoreutinae – long-eared jerboas), the other two are represented by more generalized quadrupedal taxa (Zapodinae - jumping mice and Sminthinae – birch mice). Despite considerable effort from morphologists, the taxonomy as well as the phylogeny of the Dipodoidea remains controversial. Strikingly, molecular approach has never been envisaged to investigate these questions. In this study, the phylogenetic relationships among the main dipodoid lineages were reconstructed for the first time using DNA sequence data from four nuclear genes (*IRBP, GHR, BRCA1, RAG1*). No evidence of conflict among genes was revealed. The same robustly supported tree topology was inferred from the concatenated alignment whatever the phylogenetic methods used (maximum parsimony, maximum-likelihood and Bayesian phylogenetic methods). Sminthinae branches basally within the dipodoids followed by Zapodinae. Monophyletic Cardiocraniinae is sister to all other jerboas. Within the latter, the monophyly of both Dipodinae and Allactaginae is highly supported. The relationships between Dipodinae, Allactaginae and Euchoreutinae should be regarded as unresolved trichotomy. Morphological hypotheses were confronted to findings based on the presented molecular data. As a result, previously proposed sister group relationships between *Euchoreutes* and *Sicista*, *Paradipus* and Cardiocraniinae as well as the monophyly of Cardiocraniinae + Dipodinae were rejected. However, the latter association is consistently supported by most morphological analyses. The basis of the obvious conflict between genes and morphology remains unclear. Suggested modifications to the taxonomy of Dipodoidea imply recognition of three families: Sminthidae, Zapodidae and Dipodidae, the latter including Cardiocraniinae, Euchoreutinae, Allactaginae and Dipodinae as subfamilies.

INTRODUCTION

The superfamily of Dipodoidea is the sister group of Muroidea based either on current molecular (e.g. Michaux & Catzeflis 2000; Huchon *et al* 2002; DeBry 2003; Adkins *et al* 2003; Montgelard *et al.* 2008; Blanga-Kanfi *et al* 2009; Churakov *et al* 2010) or on morphological (e.g. Klingener 1964) data. Compared to the latter taxon, which is the most diverse group of mammals – no less than 6 families, 310 genera and more than 1518 species (Musser & Carleton 2005) - the diversity of dipodoids appears relatively modest - just 16 genera and 51 species (Holden & Musser 2005). However, the Dipodoidea demonstrate diverse ecological and morphological adaptations ranging from the forest- and meadow-dwelling mouse-like birch mice and jumping mice to the arid-dwelling saltatorial jerboas, which exhibit the most extreme specializations for bipedal locomotion among rodents (Fokin, 1978).

Although the superfamily has received much taxonomic attention (see bellow), its classification remains controversial (Fig. 1). Most researchers (e.g. Holden & Musser 2005) agree that dipodoids can be arranged into six major groups, which are usually assigned to subfamilial rank. These groups are as follows: Dipodinae (three-toed jerboas – 5 genera), Allactaginae (five-toed jerboas – 3 genera), Cardiocraniinae (pygmy jerboas – 2 genera), Euchoreutinae (long-eared jerboas – 1 genus), Zapodinae (jumping mice - 3 genera), Sminthinae (synonym Sicistinae, birch mice – 1 genus).

In contrast, family-level classification has long been a matter of debate, the number of recognized families ranging from one to five. This lack of consensus on dipodoid taxonomy is, to a large extent, accounted for by the fact that phylogenetic relationships among the main lineages have not been unambiguously established yet.

Some of the early precladistic classifications contrasted the bipedal arid-dwelling true jerboas (Dipodinae, Allactaginae, Cardiocraniinae and Euchoreutinae) to the forest and grassland quadrupedal birch mice (Sminthinae) and jumping mice (Zapodinae) considering these two groups as separate families: the Dipodidae and the Zapodidae (Miller & Gidley 1918; Vinogradov 1925; Simpson 1945). Obviously, the latter family was based upon plesiomorphic similarity between the two quadrupedal lineages and was never formally supported on phylogenetic grounds. According to the alternative view, all dipodoid groups should be placed in a single family (Vinogradov 1930; Vinogradov 1937; Ellerman 1940; Ognev 1948; Klingener 1984; Holden & Musser 2005). The rationale for this approach rests on the observation that bipedal Euchoreutinae and Cardiocraniinae share with Zapodinae and Sminthinae a number of unspecialized skeletal features (Vinogradov 1930; Vinogradov 1937). Based on equivocal and possibly intermediate position of Euchoreutinae and Cardiocraniinae it was concluded that simple dichotomy between bipedal and non-bipedal taxa is inadequate to accommodate significant morphological variation within the superfamily. The point is illustrated by an informal morphological phylogenetic tree of Vinogradov (1937), which supported a basal position of Zapodinae (here including also *Sicista*)

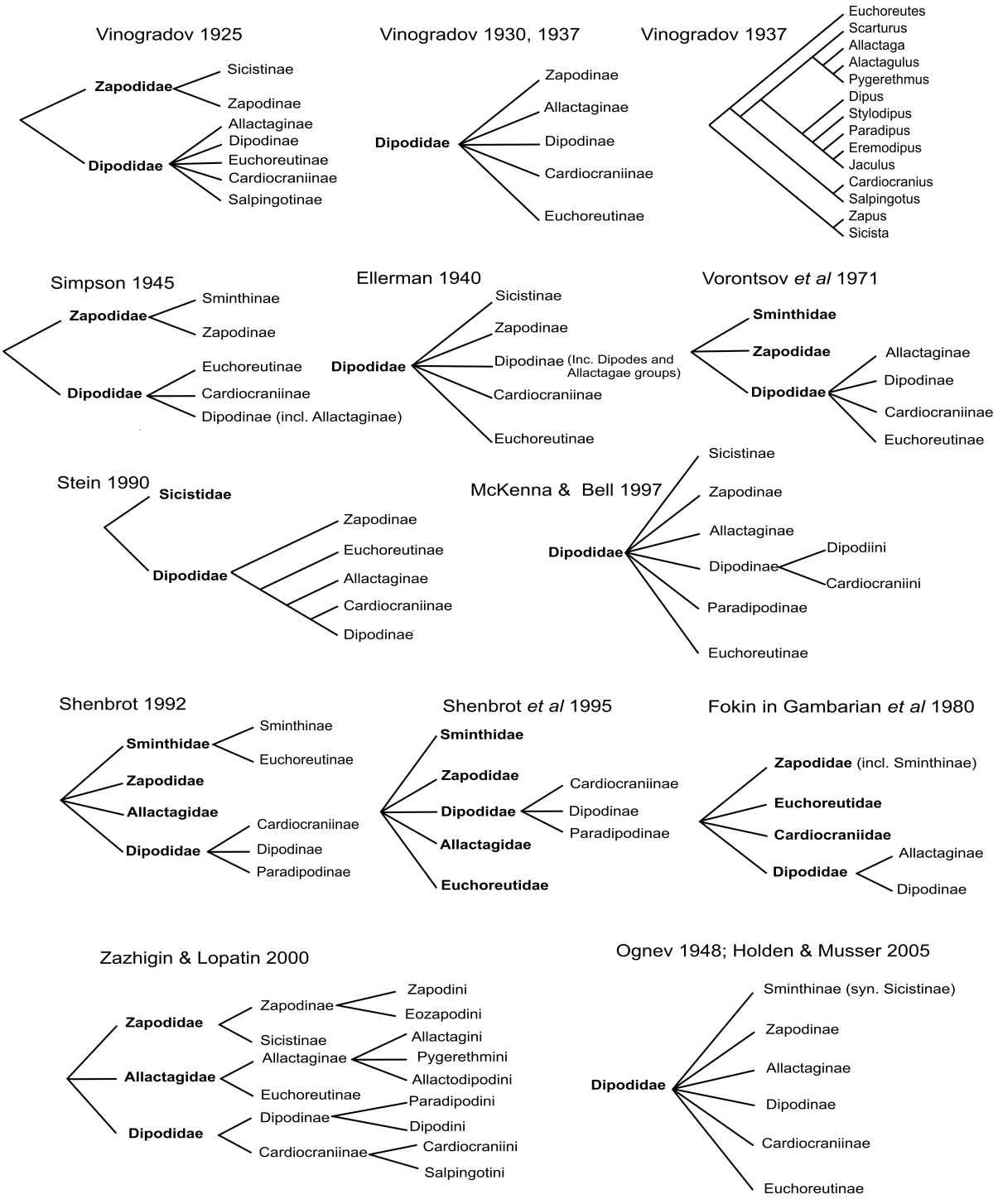


Figure 1 Graphic representation of previously suggested taxonomic systems and phylogenetic patterns in Dipodoidea.

with Cardiocraniinae and *Euchoreutes* branching deeper than Allactaginae and Dipodinae (Fig. 1).

A different taxonomic arrangement (Dipodidae, Zapodidae, Sminthidae as separate families) was proposed based on chromosome data (Vorontzov *et al* 1971). Zapodidae were considered karyotypically primitive as compared to jerboas and

birch-mice. No evident karyological similarity between any two of these well-defined groups could be traced.

The first explicitly phylogenetic system of the superfamily (Stein 1990) recognized only two families, Sicistidae (=Sminthidae) and Dipodidae, the latter including also Zapodinae. Contrary to Vinogradov's pattern, Cardiocraniinae is placed sister to the Dipodinae. This study was primarily based on cladistic analysis of limb myology.

To a large extent, the above morphology-based systems reflect the evolution of locomotory adaptations, with subfamilies (or families) corresponding to grades of evolutionary development from primitive quadrupedal to specialized bipedal locomotion. The least advanced stage is represented by birch mice (strictly quadrupedal, non-leaping), followed by jumping mice (predominantly quadrupedal but capable of making long leaps; hind foot elongated). All strictly bipedal forms (jerboas) show important modifications of hind limb skeletal and muscular morphology (see Fokin 1978 for a review). Among jerboa subfamilies, the highest degree of morphological specializations is achieved in fast-running allactagines and dipodines (maximum speed approximately 5.1-11.9 m/s), whereas the lowest is found in cardiocraniines (maximum speed <2.5 m/s; antipredator response is freezing rather than fleeing) (Fokin 1978). The problem is that, given the adaptive nature of morphological change and limitations on the number of pathways along which it can occur, one may expect that the evolutionary transformation of the locomotory system in dipodoids could be the subject of massive parallelisms, which are rather hard to differentiate from true synapomorphies. Therefore, it seems reasonable to focus a phylogenetic study upon traits that are not directly associated with locomotion. This approach was accepted in a cladistic analysis based on characters of dentition, male reproductive system, and auditory bulla (Shenbrot 1992). The analysis did not reveal any synapomorphies to support the monophyly of the bipedal taxa. On the contrary, a non-orthodox grouping inferred in this study (Fig. 1) suggested that long-eared jerboa (*Euchoreutes*) could be more related to *Sicista* rather than to Allactaginae or Dipodinae. This pattern would then imply at least two independent events of acquisition of bipedal locomotion in lineages leading to the long-eared jerboa and to Dipodinae - Allactaginae. In addition, this study is incongruent with traditional views in supporting a close relationship between Cardiocraniinae and *Paradipus,* which had been otherwise regarded as one of the dipodine genera. Shenbrot's arrangement concurs with that of Stein (1990) in placing Dipodinae and Cardiocraniinae as closest relatives; at the same time, the relationships among Zapodidae, Sminthidae, Allactagidae and Dipodidae remain unresolved. The resulting phylogeny suggested important changes in the dipodoid taxonomy. Some of Vinogradov` subfamilies were elevated to the familial rank dividing jerboas into four families: Allactagidae, Dipodidae (including Paradipodinae and Cardiocraniinae), Sminthidae (with Euchoreutinae) and Zapodidae.

Paleontological data as overviewed in Zazhigin & Lopatin (2000a, b; 2001) support this last classification in some aspects. Their system lists Zapodidae

(containing Sicistinae and Zapodinae), Allactagidae (containing Allactaginae and Euchoreutinae) and Dipodidae (which includes Cardiocraniinae, Dipodinae, and the extinct Lophocricetinae) (Zazhigin & Lopatin (2000b). The evolutionary scenario proposed therein is consistent with the independent origin of the two more specialized jerboa lineages suggesting that the roots of Allactagidae and Dipodidae are related to different groups of sminthine-like ancestors. The last two classifications are based on the view of dipodoid evolution as a complex process involving independent (and often parallel) locomotory, trophic and substrate adaptations.

Although far from complete, available molecular data consistently places *Sicista* as the sister group to other dipodoids with zapodines branching off next, while the three-toed and the five-toed jerboas (represented by single exemplars of *Dipus* or *Jaculus* and *Allactaga* respectively) appear as sister-taxa (Jansa & Weksler 2004; Steppan *et al* 2004; Montgelard *et al* 2008; Jansa *et al* 2009; Fan *et al* 2009). However, important taxa such as Cardiocraniinae and *Euchoreutes* remain unstudied.

Given all these contradictions and gaps in our knowledge of dipodoid phylogeny, it is not surprising that many taxonomists accept a conservative classification retaining a single family with six subfamilies. Actually, this uninformative system was first introduced by Ognev (1948) and now is reproduced in the third edition of Mammal Species of the World (MSW3, Holden & Musser 2005). In the present paper, we also follow the Holden and Musser's classification, with minor nomenclatural changes. In particular, we accept that Sminthinae Brandt, 1855 has priority over Sicistinae Allen, 1901 (see Data S1).

The goal of the present study is to examine the relationships between major branches of Dipodoidea based on molecular data (four nuclear exons), an approach that proved its effectiveness in similar studies in Muroidea (e. g. Michaux *et al* 2001; Steppan *et al* 2004; Jansa *et al* 2009). Available morphological hypotheses will be tested against genetic evidence. Specifically, we will address the following questions: (i) what is the basal branching order within Dipodoidea, (ii) what is the position of phylogenetically controversial taxa (Cardiocraniinae, *Euchoreutes* and *Paradipus*), (iii) what are the intergeneric relationships within Allactaginae and Dipodinae, (iv) is there significant discordance between molecular and morphological data, (v) do molecular data support independent origin of bipedal locomotion in several jerboa lineages. Finally, the taxonomy will be revised.

MATERIAL AND METHODS

Specimens examined

The original material consists of 26 specimens of 15 species of jerboas, two species of jumping mice (Zapodinae) and three birch mice (Sminthinae). Most of voucher specimens of animals used in this study are deposited in the Zoological Museum of Moscow Lomonosov State University (ZMMU). For phylogenetic analysis, 12 sequences of different genes of Dipodoidea were retrieved from GenBank (Data S1). The total matrix

Table 1 Characterization of the original material.

Species	Specimen code (Fig. 1-2 ref.)	Museum catalog number, tissue or field code	Collecting locality
Allactaga major	1	ZMMU 01/08_32	Russia, Dagestan, Chakanny
	2	t.c. Ama DNA donated by D. Kramerov	Locality unknown
Allactaga bullata	1	ZMMU S179572	Mongolia, Dundgovi aymag, Luus-somon 10 km W
	2	ZMMU S188089	Mongol ia, Ömnögovi aymag, Manlai–Mandakh road
Allactaga sibirica	2	ZMMU S181016	Mongolia, Khovd aymag, S shore Durgen-nur Lake
Allactaga elater		uncatalogized (coll. L.Khlyap)	Russia, Dagestan, Chakanny
Allactodipus bobrinskii		t.c. Abo DNA donated by D. Kramerov	Turkmenistan, Chardzhou region, left bank of Amur-Daryi, 6 km W Lebap vill.
Pygeretmus pumilio	1	t.c. Apy DNA donated by D. Kramerov	Kazakstan, Kzyl-Ordy region, N Kyzylkum, ~80 km WSW Kzyl-Ordy
	2	ZMMU S181007	Mongolia, Govi-Altai aymag, W shore Beger-nur Lake
Eremodipus lichtensteini		t.c. Eli DNA donated by D. Kramerov	Kazakstan, Kzyl-Ordy region, N Kyzylkum, ~80 km WSW Kzyl-Ordy
Dipus sagitta	1	ZMMU S179627	Mongolia, Bayankhongor aymag, S shore Bon-Tzagan-nur Lake
	2	ZMMU S188090	M ongolia, Ömnögovi aymag, Manlai – Mandakh road
Stylodipus telum		uncatalogized (coll. L. Savinetskaya)	Russia, Kalmykia, Chernye zemli
Jaculus jaculus		Moscow Zoo	Locality unknown
Jaculus blanfordi		San-Diego, t.c. 2006 ZIN 96242	Uzbekistan, C Qyzylqum, Mynbulak 40 km SW, Djarakuduk ibid
Paradipus ctenodactylys		ZMMU S143250 ZMMU S112338 (D3049)	Turkmenistan, Kazandjik 25 km NE Uzbekistan, W Qyzylqum, Takhtakupyr 100 km E, Kempir-Tyube
Cardiocranius paradoxus	1 2	ZMMU S188830 ZMMU S187344	China, Inner Mongolia Mongolia, Govi-Altai aymag, Buutsagaan ~50km SW
Salpingotus kozlovi		ZMMU S183051	Mongolia, Govi-Altai aymag, Nogon-Davon
Euchoreutes naso		ZMMU S179636	Mongolia, Bayankhongor aymag, Ekhiin-gol 47 km NE
Eozapus setchuanus		D3040 Donated by J.-P. Quéré	China, Sichuan
Napaeozapus insignis		ZMMU S183269	Canada, Ontario, vic. of Deepriver
Sicista ex. gr. subtilis		YuMK 1559	Russia, Volgograd region, Kamyshin 20 km S
Sicista ex. gr. caucasica		D764 Donated by F. Catzeflis	Russia, NW Caucasus

ZMMU, Zoological Museum of Moscow State University, Moscow; ZIN, Zoological Institute RAS, St.-Petersburg.

contains 33 specimens of Dipodoidea representing all six subfamilies, 15 genera (out of recognized 16) and 20 species (out of recognized 51). The data set includes 20 outgroup OTUs (members of Spalacidae, Muridae, Nesomyidae, Cricetidae, Calomyscidae, Geomyidae, Heteromyidae, Castoridae, Anomaluridae, Pedetidae, Gliridae and Sciuridae (in total 68 sequences of 44 species, some of which are combined into composite taxa). Information on specimens used including the list of species, collecting sites and museum catalogue numbers is given in Table 1, the GenBank Accession numbers are presented in Data S1.

DNA isolation, PCR amplification and sequencing

Total DNA was extracted from ethanol preserved tissues from liver, kidney or from dried muscles using standard protocol of proteinase K digestion, phenol-chloroform deproteinization and isopropanol precipitation (Sambrook *et al* 1989). The DNA of *Paradipus* was purified directly from the bone tissue of museum ethanol and dry specimens using silica-based spin columns of MinElute PCR Purification Kit (QIAquicky, QIAGEN) and according to recommendation of Yang *et al.* (1998).

The regions of four nuclear genes (GHR, IRBP, RAG1, BRCA1) were amplified and sequenced in 26 animals using external forward/reverse primer combinations as well as internal primers. The sequences of the original primers for amplification and sequencing are presented in the Data S1. For amplification of the exon 10 of *GHR*, in addition to the original primers, ghrEXON10 and ghrEND_R (Steppan *et al.* 2004) were used. Double-stranded polymerase chain reaction (PCR) usually entailed 30-35 thermal cycles as follows: 30 s denaturation at 94°C, 1 min annealing at 55-65 °C and 1 min extension at 72 °C. PCR products were visualized on 1 % agarose gel and then purified using DEAE Watman (UK) or NH_4EtOH. Approximately 10-40 ng of the purified PCR product was used for sequencing with each primer by autosequencing system ABI 3100-Avant using ABI PRISM®BigDye^{TM} Terminator v. 3.1 (Foster, CA, USA).

Phylogenetic analysis: molecular data

Base frequency homogeneity was tested at third codon positions separately for each locus based on values of disparity index (Kumar & Gadagkar 2001) calculated in Mega version 4.0 (Tamura *et al* 2007).

Phylogenetic reconstructions were conducted on the four loci independently and also on concatenated data. Primary analyses were conducted using information on all substitution types. To check for potential biases due to GG-content heterogeneity (see below) a second round of tree inference was performed for the combined data based on transversions only.

For phylogenetic reconstructions, maximum parsimony (MP), maximum-likelihood (ML) and Bayesian methods (BI) were used. Unweighted parsimony analysis was performed using PAUP* 4.0b10 (Swofford 2003). The following options were invoked: random addition sequence with 20 replicates, no limit for number of optimal trees, TBR branch swapping. Clade stability was assessed using 1000 bootstrap replicates using the same tree search parameters. Incongruence length difference (ILD) test (Farris *et al.* 1995) was implemented to check for significant heterogeneity among genes.

Maximum-likelihood was performed in Treefinder (Jobb 2008). The data were partitioned into gene×codon positions and separate models were used for each of the 12 partitions. This partitioning strategy was chosen based on the results of a preliminary analysis comparing five alternative partitioning schemes (see Data S1). Best models were

chosen based on BIC criterion from the set of models available in Treefinder and using the routine implemented in it. Rate heterogeneity was modeled using discrete gamma with four categories. Tree search was employed with the following options: parameter optimization simultaneous with tree search, optimized partition rates, proportional branch lengths for all partitions, maximum search depth, 10 random starting trees generated as recommended in program manual (using default settings for number of steps and tries, with the center tree obtained under the HKY model for the unpartitioned data). Bootstrap analysis (1000 pseudoreplicates) was performed using model parameters and rate values optimized for the ML tree. Transversion-based analysis was conducted with GTR2 (+G) models implemented in Treefinder. Gene tree concordance was tested in Treefinder using approximately-unbiased (AU) tests (Shimodaira 2002) contrasting the ML gene-specific topology with the tree in which relationships within Dipodoidea were modified to correspond to that of the ML tree for concatenated data.

Bayesian analyses were performed in Mrbayes 3.1.2 (Ronquist & Huelsenbeck 2003). Model complexity was assessed based on BIC criterion using Modeltest version 3.7 (Posada & Crandall 1998). Given its result, models with either two or six rate matrix parameters were chosen. Among-site rate heterogeneity was modeled using either gamma distribution or proportion of invariant sites. To correct for potential bias in branch length estimation branch length prior were adjusted following recommendations given in Brown *et al.* (2010). Transversion data were analyzed using restriction model (0/1 recoding) assuming gamma-distribution of rates. Each analysis included two independent runs of four chains (one cold plus three heated following default settings). Chain length was set at 15 million generations for the original data and at 7.5 million generations for the transversion data with sampling each 1000th generation. With these settings, the effective sample size exceeded 200 for all estimated parameters. Tracer 1.5 software (Rambaut & Drummond 2005) was used to check for convergence and determine the necessary burnin fraction which was 500 000 generations in both cases.

Levels of genetic divergence among taxa were estimated based on patristic distances calculated from the ML tree branch lengths (concatenated data, all substitution types).

Hypothesis testing was conducted in Treefinder using AU test. Eight a priori hypotheses on the relationships among lineages of Dipodoidea were tested against our best molecular tree. These hypotheses were as follows:

1. Dipodinae and Cardiocraniinae are sister taxa as inferred from morphological data (Stein 1990; Shenbrot 1992) and suggested based on fossil record (Zazhigin & Lopatin 2000b).
2. *Paradipus* is sister to Cardiocraniinae and not a part of Dipodinae (Shenbrot 1992).
3. Within Dipodinae, *Dipus* descends from the basal-most node, while *Paradipus* is sister to *Eremodipus* + *Jaculus* as proposed by Vinogradov (1937).

4. *Euchoreutes* is sister to Cardiocraniinae + Dipodinae + Allactaginae as follows from Vinogradov (1937) and Shenbrot (1992).
5. *Euchoreutes* is sister to Allactaginae as argued by Zazhigin & Lopatin (2000b).
6. *Euchoreutes* is sister to Sminthinae as inferred by Shenbrot (1992).
7. Birch-mice together with Zapodinae constitute a monophyletic group as accepted by Vinogradov (1937), Simpson (1945) etc.
8. *Allactaga* is monophyletic relative to other five-toed jerboas (*Pygeretmus* and *Allactodipus*) as routinely accepted in taxonomic accounts (e.g. Holden & Musser 2005).

Phylogenetic analysis: morphological data

The parsimony analysis of morphological data was based on a matrix, which closely follows the one used in Shenbrot (1992) with some modifications. The matrix includes 31 OTUs representing all genera and major species groups of Dipodoidea. Four morphological character partitions were included in the analysis: dentition (17 characters), auditory bulla (five characters), glans penis (seven), and accessory reproductive glands (six). The detailed description of character states is given in Data S1. As in Shenbrot (1992) the primary (ordered = O) analysis was based on a priori assumptions on polarity, character state ordering implying irreversibility of certain transformations. All information on allowed character change was coded via step-matrices (Data S1). To test the sensitivity of the phylogenetic results to the above a priori assumptions we also conducted unordered analyses (U), which impose minimum constraints on character evolution. In ordered analyses, the position of the root was inferred based on optimization in irreversible characters, while in the unordered analyses, trees were rooted using the explicitly specified hypothetical ancestor. Taking into account that the number of recognized character states was highly variable across characters (from two to nine) and across partitions (ranging from 15 in total for bulla to 58 for dentition), we also performed weighted analyses to achieve a more balanced contribution of different traits to phylogenetic reconstruction. Two methods of character weighting were used. The first method (Wc) estimated the weight of each individual character as inversely proportional to its minimum possible number of steps (equal to the number of states in a character minus one). The second one (Wp) assigned the weights on a partition-wise basis as reciprocal to the minimum possible amount of change for the partition as a whole (equal to the total number of states for all characters within a partition minus the number of characters). In both cases, weighting was expected to reduce the impact of dental traits compared to unweighted analysis. In total, six analyses employing different combinations of character ordering (O vs. U) and weighting (W0, Wc, Wp) were performed. All reconstructions were conducted in PAUP*. Starting topology was obtained via step-wise addition with random order (20 replicates). The search for optimal trees was conducted using TBR branch swapping with MULPARS option on. To sum up the output, which consists of several equally parsimonious trees, majority-rule consensus trees were reconstructed. To test clade stability, bootstrap analysis was performed with

1000 pseudoreplicates using the same options of tree search as above but with MULPARS set to off. To check for possible conflict among partitions ILD test was used as implemented in PAUP* with 1000 replicates.

To test whether morphological data are compatible with the molecular topology, all reconstructions were repeated imposing molecular-derived constraints. The trees thereby obtained were compared to the unconstrained ones using non-parametric Templeton (Wilcoxon signed-ranks) test (Templeton 1983). Following recommendations of Felsenstein (1985), the more conservative two-sided version of the test was used.

To identify which morphological characters were predominantly responsible for the conflict with molecular evidence, we assessed the level of conflict between each of the four morphological partitions and the molecular–constrained tree. Conflict was quantified as the average decrease of single-character consistency indices (number of states minus one / observed number of steps) for constraint trees versus unconstrained trees. Significance was tested using nonparametric Kruskal-Wallis test as implemented in Statistica 6.0 (StatSoft, Inc. 2001). Also, partitioned Bremer support (Baker & DeSalle 1997) was estimated for selected nodes.

The key morphological transformations including those involved in the evolution of bipedalism were mapped onto both morphological and molecular trees using MPRSets command in PAUP* (ordered data).

RESULTS

Alignment and partitioning

The complete alignment of IRBP consisted of 1107 bp (636 variable, 458 parsimony-informative) for 26 specimens of Dipodoidea and 20 outgroups. The alignment of GHR, exon 10 consisted of 927 bp (577 variable, 403 parsimony-informative) for 25 specimens of Dipodoidea and 20 outgroups. The complete alignment of BRCA, exon 11 comprised 840 bp (628 variable, 418 parsimony-informative) for 22 specimens of Dipodoidea and 15 outgroups. For RAG1 the alignment included 1143 bp (429 variable, 285 parsimony-informative) for 22 specimens of Dipodoidea and 13 outgroups. The best-fit substitution models employed for each of the 12 partitions are given in Table S1).

Base composition

Base composition at the 3rd codon position is presented in the Table S2. Individual genes demonstrated significant departure from homogeneity, which is hardly a result of multiple comparisons given highly significant ($P < 0.001$) values of disparity index. In IRBP data set, CG content was much higher in Dipodinae (except *Paradipus*) and *Salpingotus* (but not *Cardiocranius*) than in other groups; in RAG1 dataset, CG proportion was notably higher in Zapodinae and *Sicista.* There was no common trend of CG variation across genes; thus, in Dipodinae, the CG content was relatively high in IRBP but rather low in RAG1.

For data recoded into purines and pyrimidines, neither test was significant at

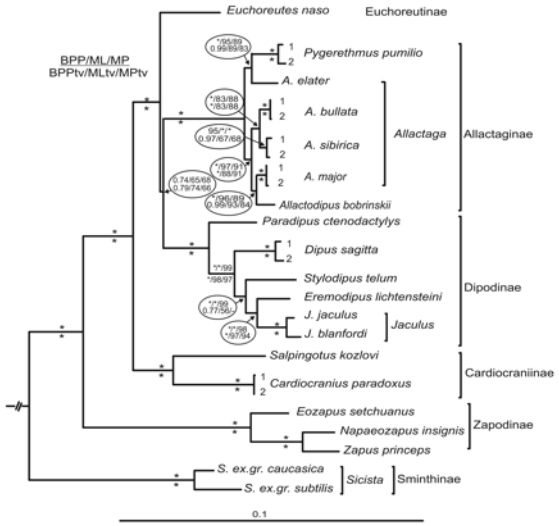


Figure 2 The maximum-likelihood (ML) phylogeny of Dipodoidea as inferred from a concatenated alignment of four nuclear genes (outgroups not shown, for the complete version see Fig. S1). Values above the branches are the Bayesian posterior probabilities (BPP) and bootstrap support (1000 pseudoreplicates) in maximum-likelihood (ML) and maximum parsimony (MP) analyses, respectively. Values below the branches refer to corresponding values of support obtained in transversion-based analyses. '*' denotes bootstrap support of 100% and BPP of 1.0; '-' indicates support values of <50%. Specimen codes are as in Table 1.

$P < 0.01$ and just five tests (four for IRBP and 1 for GHR) were significant at $P = 0.01$. The latter might be attributed to a multiple comparison artifact. Since the R/Y proportion appeared relatively stable, the RY analysis is not expected to produce biased tree topologies.

Phylogenetic results

The results of combined analysis (original data and RY recoded data) as well as those obtained based on each independent gene are presented in Figs 2 and 3. It appears that although the best topologies inferred from individual genes were not completely identical, the four alignments were combinable: in neither case the AU test rejected the H0 ($P = 0.29$, 0.27, 0.26, 0.28 for *IRBP, GHR, BRCA1, RAG1*, correspondingly) indicating that for each gene the fit of its individual ML tree was not significantly better than that of the ML tree generated from the combined data. Moreover, in neither case, a group recovered in an individual gene analysis but incompatible with the combined topology received bootstrap support higher than 61%. The results of the ILD test as well indicated the lack of conflict among the four partitions ($P = 0.155$). Therefore, we will focus on the topology inferred from the combined analysis. (For details of single-gene analyses see Data S1).

The same optimal tree topology was inferred from the concatenated alignment for Dipodoidea whatever the methods considered (MP, ML, BI) or the data coding used (original or R/Y-recoded). This tree (Fig. 2, see also Fig. S1) was well resolved with most supra-generic clades receiving high or absolute support (i.e. bootstrap support (BS) of 100% and posterior probability (PP) of 1.0). *Sicista* appears as the first group to diverge followed by the Zapodinae (BS = 100%, PP = 1). Among the Zapodinae, *Eozapus* is placed as the sister group to the American genera *Zapus* and *Napaeozapus* (BS = 100%, PP = 1). Cardiocraniinae is recovered as a monophyletic group (BS = 100%, PP = 1) branching as sister to all other jerboas (BS = 100%, PP = 1). Within the latter, the five-toed and three-toed jerboas (Allactaginae and Dipodinae) constitute a clade closely related to *Euchoreutes*. However, the position of *Euchoreutes* is just weakly supported (BS = 65-68%, PP = 0.74), and hence, based on these data the relationship between Euchoreutinae, Allactaginae and

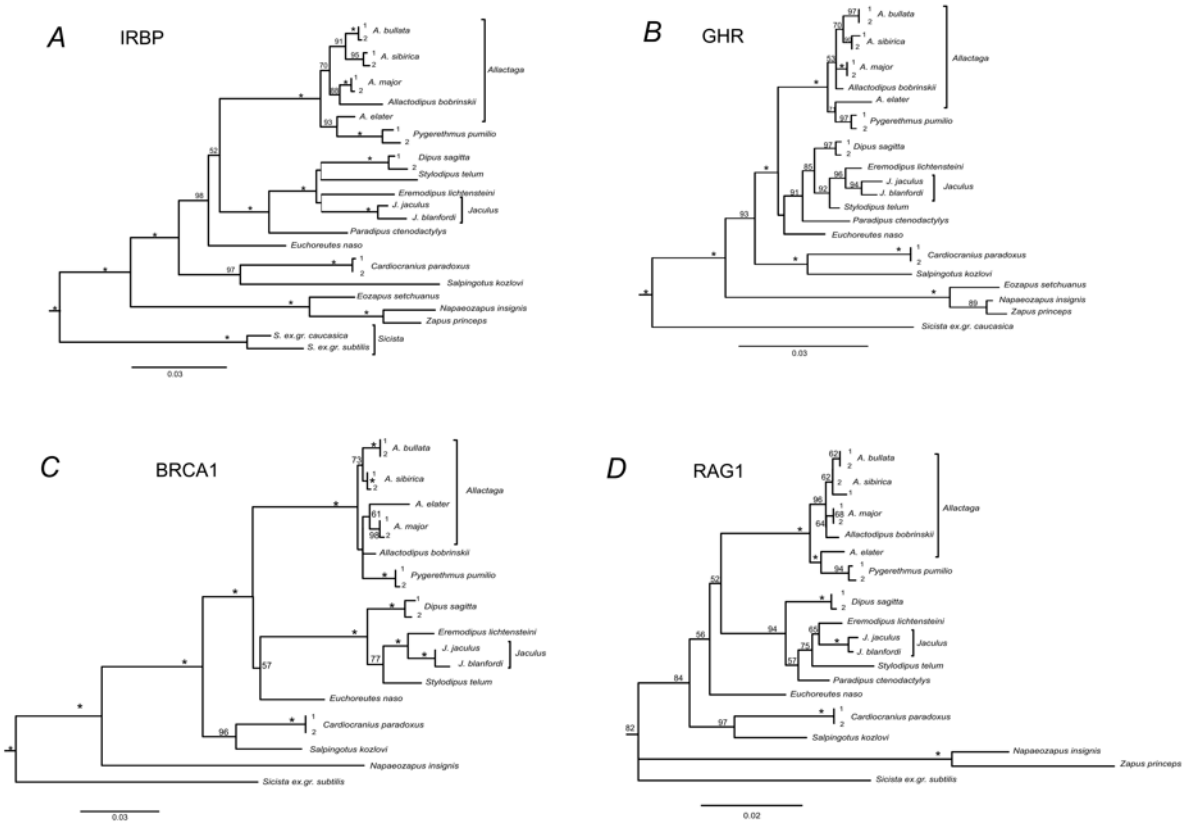


Figure 3 Maximum-likelihood (ML) gene trees for Dipodoidea based on separate analysis of –A. *IRBP*, –B. *GHR*, –C. *BRCA1,* and –D. *RAG1*. Values above the branches correspond to bootstrap support (1000 pseudoreplicates) in the ML analyses. Specimen codes are as in Table 1.

Dipodinae should be regarded rather as an unresolved trichotomy. Within the Dipodinae, the genus *Paradipus* descends from the basal-most split (BS = 97%-100%, PP = 1), *Dipus* is next to branch off, while *Stylodipus* is recovered as the sister group to a clade composed of *Eremodipus* and *Jaculus*. Within the Allactaginae, the genus *Allactaga* appears as paraphyletic in respect to both *Pygeretmus* and *Allactodipus*. Indeed, *A. elater* was most closely related *to Pygeretmus* rather than to any other *Allactaga* representatives (BS = 83%-95%, PP = 1), while *A. bobrinskii* was placed as sister to *A. major* (BS = 84%-96%, PP = 1).

The analysis based on the RY recoded alignment produced essentially the same results. Most nodes retained high support with the exception of the position of *Stylodipus*. Thus, data recoding does not result in a significant loss of phylogenetic signal. Also, it can be concluded that the observed base-compositional heterogeneity, whatever pronounced, does not introduce any bias into phylogenetic reconstructions based on the original combined dataset.

Table 2 Results of the approximately-unbiased (AU) test of a priori morphology based hypotheses on Dipodoidea phylogeny.

H0: Morphology-based hypothesis on dipodid phylogenetic relationships	**The results of the AU-test: *P* - probability that H0 is correct given molecular data**	
	All substitution types	Transversions only
Dipodinae +Cardiocraniinae is monophyletic	<0.01	<0.01
Paradipus is sister to Cardiocraniinae	<0.01	<0.01
Paradipus is sister to *Eremodipus* + *Jaculus*	<0.01	0.017
Euchoreutes is basal relative to the three other groups of jerboas	<0.01	<0.01
Euchoreutes is sister to Allactaginae	0.19	0.18
Euchoreutes is sister to Sminthinae	<0.01	<0.01
Sminthinae+Zapodinae is monophyletic	<0.01	0.029
Allactaga is monophyletic relative to *Pygeretmus* and *Allactodipus*	<0.01	<0.01

According to the results of the AU tests (Table 2) all but one a priori advanced hypothesis are rejected at least at $P < 5\%$.

A comparison of the levels of intra-group divergence as assessed from genetic distance at the basal nodes of the main jerboa clades demonstrates that the highest differentiation is observed within the Cardiocraniinae (5.7% between *Cardiocranius* and *Salpingotus*). The clade of five-toed jerboas is relatively shallow (1.7 - 2.2 % between *Pygeretmus* + *A. elater* and the rest of allactagines), whereas three-toed jerboas are significantly more divergent (3.9 - 4.6% between *Paradipus* and other dipodine genera, 3.3 - 3.8 % between *Dipus* and *Eremodipus* + *Jaculus*).

Parsimony analysis of morphological data

The data set for morphological analysis included characters of dentition, auditory bulla, glans penis and accessory reproductive glands. The ILD tests suggest significant discordance among these character partitions ($P < 0.01$ for all variants of ordering and weighting). However, this test is known to be too liberal due to its sensitivity to various confounding factors (see e.g. Barker & Lutzoni 2002); therefore, we consider the positive result as a preliminary indication of heterogeneity within morphological data. The nature of this incongruence requires further exploration. At the same time, the fact that H0 is rejected does not necessarily imply that combined approach is inadequate (Hipp *et al.* 2004); besides, some of our partitions include too few characters for their separate analysis, so here we focus on the results for the combined data.

Data re-analysis with original settings (unequal transformation costs with equal character weights and irreversible evolution postulated for many traits – O-W0) reproduced a tree very similar to the one previously obtained by Shenbrot (1992) with the use of the WISS algorithm. However, other combinations of character ordering and weighting produced topologies deviating from the former in several aspects (see Fig. 4). The following points should be noticed:

1. The position of *Euchoreutes* was sensitive to changes in character weighting – it was sister to *Sicista* with unweighted data (as in Shenbrot 1992), while down-weighting of dental traits resulted in its association with Allactaginae (O-Wp&c,

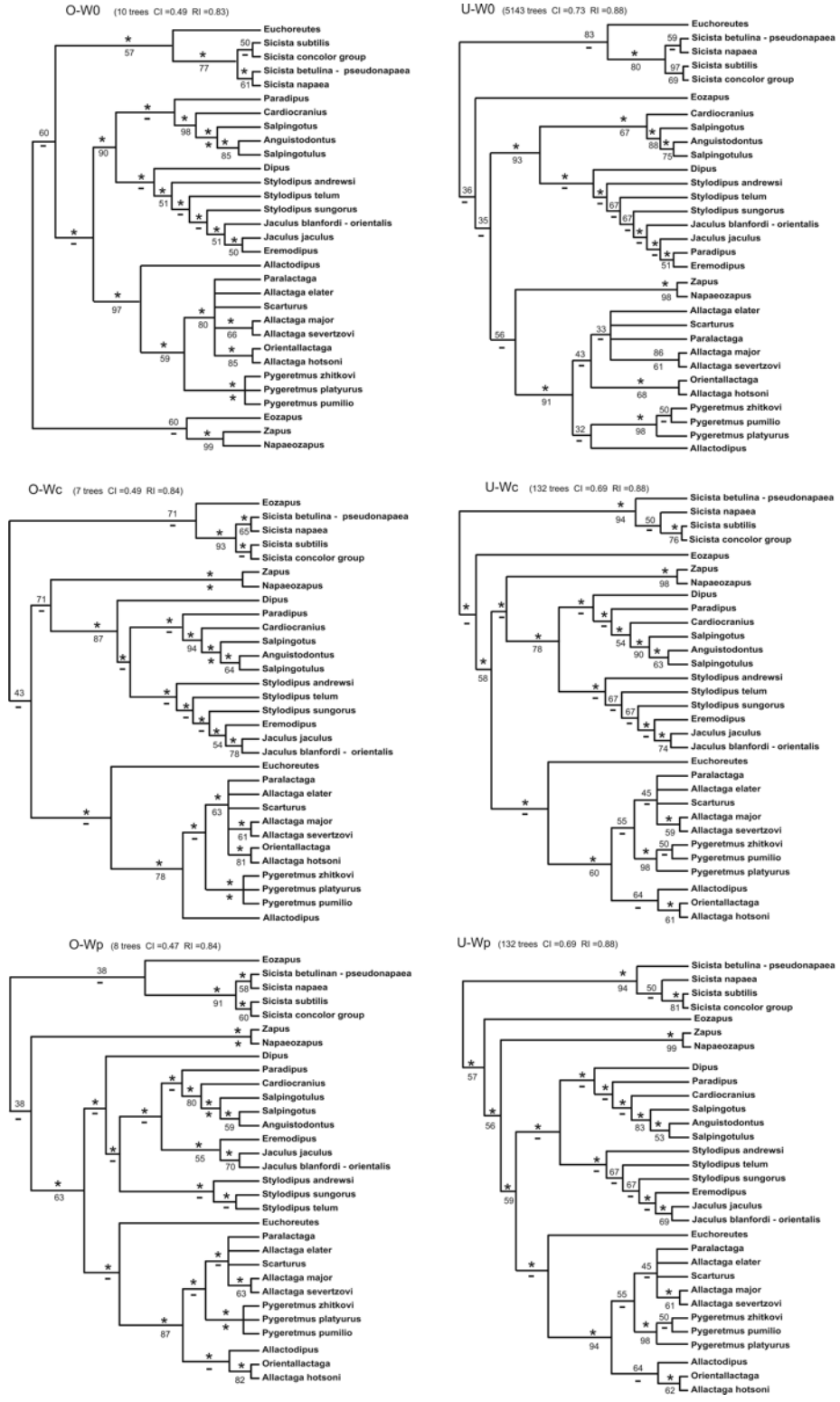


Figure 4 Phylogenetic relationships among Dipodoidea species as follows from parsimony analysis of morphological data by Shenbrot (1992, with modifications). Six analyses were performed: data were treated as either ordered (O) or unordered (U), characters were either given equal weights (W0) or weighted according to one of the two schemes (Wc or Wp, see text for details). Topologies correspond to majority-rule consensus trees inferred from all equally parsimonious trees. Compatible nodes present in <50% trees are shown as well. Numbers above branches designate percentage of most parsimonious trees supporting corresponding splits, and numbers below branches indicate bootstrap support. '*****' stand for 100%; '**-**' designate bootstrap support of <50%. Numbers of equally parsimonious trees, values of consistency index (CI) and retention index (RI) are given above each dendrogram.

U-Wp&c); no bootstrap support for any of these relationships was obtained with the exception of O-W0 analysis where *Euchoreutes* + *Sicista* clade received bootstrap support of just 57%.

2. The Cardiocraniinae and Dipodinae constitute a stable clade, which was retrieved in all analyses; it received moderate to high bootstrap support (78-93%) with the exception of analyses using Wp weighting.
3. *Paradipus* was associated with Cardiocraniinae in all variants other than U-W0 analysis in which it is placed at the very top of the dipodine clade as a sister group of *Eremodipus*, however, neither of its positions received significant bootstrap support.
4. In most cases, Allactaginae and Cardiocraniinae emerged as monophyletic groups with at least some support. In contrast, both Dipodinae sensu lato and Dipodinae sensu stricto (without *Paradipus*) never received any support above 50%. Moreover, the latter assemblage appeared to be paraphyletic in regards to Cardiocraniinae in all weighted analyses.
5. The support for the group of true jerboas (ignoring the position of *Euchoreutes*) should be considered as unstable. The clade was retrieved in half of the analyses with some bootstrap support (59-63%) in Wp only.
6. *Eozapus* showed just insignificant tendency to group with other zapodines – in all analyses other than OW0 jumping mice were paraphyletic.
7. Basal radiation appeared poorly resolved: if the basal node was not reduced to a multifurcation, basal position was occupied either by zapodines (OW0) or by *Sicista* (UWp&c).

Constrained analyses

In the constrained analyses the following well-supported features of the molecular tree were used as constraints: all subfamilies are monophyletic; *Sicista* is sister to all other dipodoids; jumping mice is the sister group to the jerboa clade; *Zapus* + *Napaeozapus* is monophyletic; Cardiocraniinae is sister to a trichotomy of Dipodinae, Allactaginae and Euchoreutinae; *Paradipus* is sister to other dipodines. If all constraints were enforced simultaneously, the trees inferred under all regimes other than OWp were significantly less parsimonious than corresponding unconstrained trees ($P < 0.05$, Templeton test, see Table S3).

The four morphological partitions differed in the level of incongruence with the molecular-based tree ($P < 0.05$, Kruskal-Wallis test; for details see Table S4). The decrease in the mean consistency index for the constrained topologies was more pronounced in dental characters as compared to other partitions. In contrast, traits of genital morphology fit nearly equally well on both molecular and morphological trees. However, insufficient character sampling precluded more rigorous testing of the hypothesis that dental partition is responsible for most of the conflict between morphological and molecular evidence.

Individual constraints demonstrated unequal contribution to the increase in tree length (see Table S5). Most constraints required few (1-2) extra steps; however,

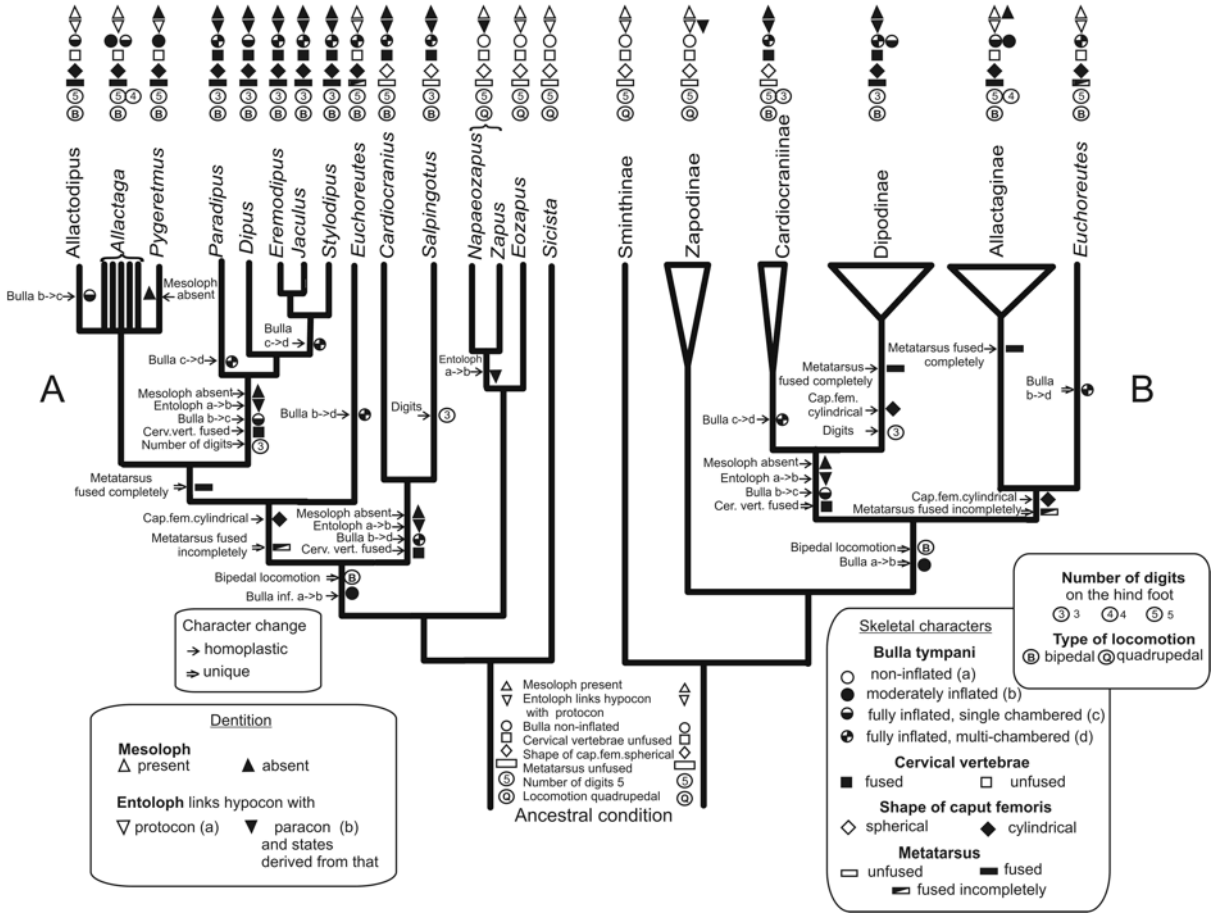


Figure 5 –A. Parsimony character mapping on molecular and –A. modified morphological trees. Dental and bulla traits are ordered (see information on characters A5, A7, B1 in the Data S1). Arrows and notes on branches indicate inferred character transformations (unique or homoplastic) and specify the derived states. Most of transformations that are mapped identically on both trees are omitted from the morphological tree. See legend for character state designations.

topologies constrained for the monophyly of Euchoreutinae + Dipodinae + Allactaginae involved 12 and eight additional steps in ordered and unordered analysis, respectively. The cost of enforcing all other constraints together was lower (five steps in both cases). This result is consistent with the fact that our morphological data robustly support the monophyly of Dipodinae + Cardiocraniinae and, hence, contradict the Euchoreutinae + Dipodinae + Allactaginae clade.

Character mapping

Character mapping was performed using both molecular and modified morphological tree (Fig. 5). The only difference between the two topologies lies in the placement of Dipodinae as sister to Cardiocraniinae in the morphological tree. This choice of morphological topology is justified by the fact that Cardiocraniinae + Dipodinae is the only morphologically well-supported grouping which is incompatible with the molecular phylogeny.

Inspection of character state changes showed that there are no apparent synapomorphies for Euchoreutinae + Dipodinae + Allactaginae in our dataset.

In contrast, Dipodinae + Cardiocraniinae was supported by the following derived character states (all dental) regardless of ordering scheme:

1. partial to complete reduction of anteroloph;
2. reduction of mesoloph, (with a parallelism in *Pygeretmus*) (Fig. 5);
3. paracone (not protocone) connected via entololoph with hypocone (or conditions derived from this), (with a parallelism in *Zapus* + *Napaeozapus*) (Fig. 5);
4. reduction of mesolophids on both m1 and m2;
5. entoconid (not hypoconid) connected via longitudinal crest with paraconid (or conditions derived from this), (with a parallelism in *Zapus* + *Napaeozapus*);
6. presence of labial crest on lower molars (or conditions derived from this).

Taken together, these changes illustrate transition to shorter and broader molar crown with opposite cusps, reduced mesostyles(-ids) and diagonally positioned longitudinal crests. Correspondingly, the molecular tree suggested that this advanced molar pattern evolved independently in three-toed and pygmy jerboas.

Cardiocraniinae and most of Dipodinae share specific bulla morphology characterized by extremely inflated mastoid; however, bullar traits do not provide unambiguous support for Dipodinae+Cardiocraniinae (see also data on partitioned Bremer support, Supplementary material Table F). Regardless of the topology used, the development of highly inflated multi-chambered bulla in Cardiocraniinae, *Euchoreutes* and Dipodinae is explained by multiple parallelisms; at the same time, this character showed less homoplasy on the morphological tree.

Traits associated with advanced bipedalism (not included in our dataset) such as cylindrical shape of caput femoris and fused metatarsus showed no homoplasy on the molecular tree and support Euchoreutinae+Dipodinae+Allactaginae; in contrast, morpological tree required parallel acquisition of derived states in at least two groups of jerboas.

DISCUSSION

Phylogenetic relationships within Dipodoidea

Our results show that internal consistency among the molecular data and its information content are high enough to produce a well-supported molecular hypothesis on phylogenetic relationships among the main dipodoid lineages. Below we address the main issues emerging from the confrontation of the molecular and the morphological data focusing on potential taxonomic implications.

Basal branching order in Dipodoidea

Our data support previous morphological (Stein 1990, Potapova 1976) and

molecular (Jansa & Weksler 2004, Montgelard *et al.* 2008; Jansa *et al.* 2009) findings confirming that the basal split is between *Sicista* and other dipodoids. The Zapodinae appears as the sister group of the monophyletic jerboas. Accordingly, this pattern contradicts the view that the nonbipedal birch-mice (Sminthinae) and jumping mice (Zapodinae) constitute a monophyletic group as proposed by Vinogradov (1937) and subsequently advocated by some morphologists (e.g. Gambaryan *et al* 1980) and paleontologists (Shevyreva 1983; Zazhigin & Lopatin 2000b).

Most paleontological scenarios agree with a divergence among major lineages in Late Eocene - Oligocene when the diversity of sminthine-like dipodoids was high (Shevyreva 1983; Zazhigin & Lopatin 2000b). However, the relationships among fossil genera as well as the ancestry of modern groups appear controversial (compare Zazhigin & Lopatin 2000b: Wang & Qiu 2000; Löpez-Antonanzas & Sen 2006). Zazhigin & Lopatin (2000b) identified two Oligocene lineages, *Heosminthus -Plesiosminthus* and *Sinosminthus -Parasminthus*, as the ancestral groups for Zapodinae-Sminthinae and Allactaginae-Dipodinae, respectively. However, our topological tests clearly rejected the monophyly of the former association, thus suggesting that fossil data should be revised to determine which of the Paleogene taxa represent stem-groups of the three – but not two – main dipodoid clades.

Monophyly of jerboas

One of the most important results stemming from our study is the monophyly of bipedal Dipodoidea including Euchoreutinae and Cardiocraniinae in congruence with many previous authors (Vinogradov 1937, Vorontzov *et al.* 1971, Stein 1990). Although this group comprises highly divergent lineages (either at a morphological point of view or a genetic one), its cohesiveness is supported by an unprecedentedly stable karyotype (2*n* = 48) shared by all genera except *Salpingotus* and *Stylodipus* (Vorontzov *et al.* 1971; Orlov & Yatsenko 1985).

Phylogenetic position of Euchoreutes

Previously, the monophyly of bipedal dipodoids was questioned, primarily due to the controversial phylogenetic position of the long-eared jerboa, *Euchoreutes naso*. It has been commonly agreed that this species is a single survivor of some ancient lineage. Vinogradov (1930, 1937) suggested that *Euchoreutes* is sister to all other bipedal taxa based on the observation that this species retains some characters intermediate between specialized jerboas (Dipodinae and Allactaginae) and sminthines (namely, generalized zygoma, less specialized pelvis and hind limb). The same pattern was afterwards recovered in the cladistic analysis by Stein (1990) but with poor bootstrap support (<60%, our re-analysis). According to an alternative view (Shenbrot 1992), *Euchoreutes* is considered as an independent derivative from ancestral dipodoid stock, which shares true synapomorphies only with *Sicista*. This implies that similarity between *Euchoreutes* and other jerboas should be treated as a result of parallel (or convergent) evolution. However, our re-analysis of morphological data indicates

that such a position of the long-eared jerboa is neither well supported nor insensitive to variations in character weights. In most of weighted analyses *Euchoreutes* demonstrates affinity to Allactaginae, but with insignificant bootstrap support. The molecular data conflict both of the above morphology-based scenarios, supporting instead the monophyly of Euchoreutinae + Dipodinae + Allactaginae. This result is in line with the opinion of Potapova (2000a) that dental traits shared by the long-eared jerboa, *Euchoreutes* and *Sicista* should be regarded not as synapomorphies but as a result of parallel adaptation to insectivory. The hypothesis that *Euchoreutes* is related to five-toed jerboas as argued by Zazhigin & Lopatin (2000b, 2005) receives no corroboration but cannot be formally rejected. Our data does not contain sufficient information to resolve the trichotomy among allactagines, dipodines and *Euchoreutes*; at the same time, one may hypothesize that the three lineages diverged in rapid succession within a relatively short time span. The earliest fossils attributed to the Euchoreutinae are found in Early Miocene (MN 3-4, 16-20 Mya) what is not significantly later than the earliest record for allactagines (Zazhigin & Lopatin 2005). Therefore, the split between the three jerboa lineages must have occurred not later than the Early Miocene.

The position of Cardiocraniinae and Dipodinae

The molecular data reject genealogical affinity between Dipodinae and Cardiocraniinae, which was supposed based on morphology (Shenbrot 1985; 1992; Stein 1990; Zazhigin & Lopatin 2000b; 2001), and suggest that the line leading to the pygmy jerboas (Cardiocraniinae) was the first to split from the common ancestor of bipedal taxa. According to the hypothetical paleontological scenario (Zazhigin & Lopatin 2000b), all extant jerboas descended from two ancestral lineages, which had separated from each other by the end of Oligocene and evolved later into Allactaginae and Lophocricetinae, respectively. Now-extinct Lophocricetinae is regarded as the immediate ancestor of the Cardiocraniinae (Shenbrot *et al.* 1995) and, indirectly, of the Dipodinae. The latter taxon is believed to have originated, in its turn, from some early Cardiocraniinae (Shenbrot 1992; Zazhigin & Lopatin 2001). Apparently, the molecular position of the Cardiocraniinae (but not of the Dipodinae) is consistent with the hypothesis of its origin from the Lophocricetinae. It remains unclear which fossil taxa might be a likely candidate for the ancestry of the Dipodinae.

Monophyly of Cardiocraniinae

A cladistic analysis of genital morphology suggested that Cardiocraniinae is not monophyletic indicating substantial differentiation between two modern branches represented by *Cardiocranius* and *Salpingotus* (Pavlinov & Shenbrot 1983). The same idea was independently proposed based on cytogenetic data (Orlov & Yatsenko 1985). The molecular evidence also demonstrates that the two genera are highly divergent, but at the same time, it consistently supports the monophyly of cardiocraniines, thus, confirming morphological results obtained with more comprehensive data sets (Shenbrot 1992). The first representatives

of Cardiocraniinae are found in the earliest Late Miocene (MN 9, 9.7-11.1 Mya – Li & Zheng 2005) or even late Middle Miocene (MN 7+8, 11.1-12.5 Mya – Zazhigin & Lopatin 2005).

Relationships among the three-toed jerboas, the Dipodinae

Considering phylogenetic relationships among dipodines, morphologists generally agree in placing *Dipus* basal to *Stylodipus*, *Jaculus* and *Eremodipus* and in treating the latter two genera as sister taxa based on their advanced molar morphology (Vinogradov 1937; Heptner 1984; Shenbrot 1992). Our molecular tree is fully congruent with this pattern. The major point of disagreement between different morphology-based phylogenies is the position of *Paradipus* – a monotypic genus, which is characterized by a unique combination of autopomorphic and shared derived features. According to Vinogradov (1937), *Paradipus* is a crown dipodine close to *Eremodipus* + *Jaculus* clade. Alternatively, it is regarded as a separate lineage just distantly related to other Dipodinae (Heptner 1984) or even as a sister group to Cardiocraniinae, thus, indicating extensive parallelism between *Paradipus* and other three-toed jerboas (Shenbrot 1992; Shenbrot *et al.* 1995; Potapova 1998). Molecular evidence rejects both the first and the last hypotheses and supports *Paradipus* as the most divergent lineage within Dipodinae. The re-analysis of morphological matrix demonstrated that the tendency for *Paradipus* to group with Cardiocraniinae is dependent on initial assumptions (character ordering and weighting).

Phylogenetic relationships within the five-toed jerboas, the Allactaginae

One of the most striking results of our molecular analysis is the lack of monophyly of the genus *Allactaga* relative to morphologically distinct *Pygeretmus* and *Allactodipus*. In our combined tree, *Pygeretmus* form a clade with *Allactaga elater,* while *Allactodipus* is grouped with *Allactaga major*. Notwithstanding the fact that this result appears highly supported (see results of the AU test) we recommend that it should be treated with caution taking into account relatively low number of informative sites available at this level of divergence what could mask potential conflict among genes. At the same time, from a morphological viewpoint, we cannot reject a hypothesis that *Allactaga* is a paraphyletic assemblage ancestral to both of the other allactagine genera. Among the morphological characters considered, only the shape of glans penis can be counted as a potential synapomorphy for *Allactaga* sensu lato. Paleontological record indicates that the first unequivocal allactagines appear in the Lower Miocene of (MN 3-4, 16-20 Mya) while the earliest fossils attributed to the genus *Allactaga* (subgenus *Paralactaga*) are found in MN11 (7.5-8.7 Mya) (Zazhigin & Lopatin 2000a). The genera *Protalactaga*, *Proalactaga* and *Paralactaga* in their interpretation by most paleontologists are rather the stages of evolutionary development of molar crown patterns (grades) than real phylogenetic clades. Therefore, the roots of different lineages within the genus *Allactaga* sensu lato could be significantly deeper than it is usually believed. It is commonly accepted that *Pygeretmus* is a rather recent (Latest Miocene) derivative of *Allactaga,* which

has evolved towards more specialized herbivore diet (Shenbrot 1984; Zazhigin & Lopatin 2000a). In contrast to that, *Allactodipus* was hypothesized to be an early offshoot of the ancestral allactagine stock, thus, representing a separate evolutionary lineage (Zazhigin & Lopatin 2000a). Our data strongly contradict this view indicating instead that *Allactodipus* lineage originated at approximately the same time as major clades within *Allactaga*. In his review of postcranial skeletal morphology in Dipodoidea, Fomin (2006) mentioned that *Allactodipus* shares some similarity in pelvis shape with *A. major* (and *A. severtzovi*). Although this finding is in line with the pattern inferred from genetic data, the phylogenetic value of this condition remains unclear.

Incongruence between molecules and morphology and evolution of bipedalism

DNA versus morphology: is the conflict real?

Our re-analysis of morphological data demonstrated that some non-trivial groupings retrieved in the previous study (i.e. *Sicista* + *Euchoreutes*, *Paradipus* + *Cardiocraniinae*) are, in fact, poorly supported and unstable. Consequently, most of the apparent discrepancies between molecular phylogeny and morphology should be attributed to topological instability in morphological reconstructions. This would be due to insufficient information content and higher homoplasy rate typical for the latter type of data (see Springer *et al.* 2007, 2008 for a discussion). However, concerning the position of Cardiocraniinae, this does not seem to be the case. Our current morphological analyses recapitulate Dipodinae + Cardiocraniinae clade with bootstrap support of up to 93%. Supposed synapomorphies for this association are presented mostly by dental features (Shenbrot 1992; our analysis). The phylogenetic signal of the dental partition may have been distorted by character non-independence or higher homoplasy rate (see Naylor & Adams 2001); however, even severe downweighting of dental characters (Wp variant) yielded the same overall topology.

Furthermore, Dipodinae + Cardiocraniinae clade was recovered with bootstrap support of 75-79% (our re-analysis) in the cladistic analysis of hind and fore limb myology by Stein (1990). The two groups also share a number of derived postcranial features such as fused cervical vertebrae (Vinogradov, 1937) and specific structure of scapulae (Fomin 2006).

Thus, different types of morphological characters (dentary, myological, osteological) support the phylogenetic position of Cardiocraninae as sister to Dipodinae what is definitely rejected by molecular data; therefore, we have to conclude that the conflict between the two kinds of data is significant.

Molecular phylogeny and locomotory evolution: apparent concordance

At the same time, some morphological features agree with molecular evidence (Fig. 5). The fact that Cardiocraniinae retain primitive condition in several traits of hind limb morphology (i.e. spherical and not cylindrical shape of caput femoris and lack of metatarsal fusion) is consistent with their early separation from the stem

of bipedal taxa (Vinogradov 1937). The same hypothesis was advocated by Fokin (1978, 1983) based on the comparative study of pelvis and hind limb myology. The latter work demonstrated that pygmy jerboas represent a mosaic of both primitive and uniquely derived traits, suggesting independent development of efficient saltatory locomotion.

Another important question is whether bipedality was acquired only once or evolved independently in several dipodoid lineages. Available paleontological data do not provide direct support for either of these hypotheses (for information on the earliest fossils of bipedal dipodoids see Zazhigin & Lopatin 2000a,b, 2005). The most parsimonious scenario, given the molecular tree, suggests that the last common ancestor of extant jerboas was bipedal. Moreover, the molecular phylogeny, while at odds with most of morphological characters, is compatible with the general trajectory of locomotory evolution in dipodoids leading from non-ricochetal *Sicista* via ricochetal but quadrupedal zapodines to primitively bipedal cardiocraniines and, finally, to highly specialized Allactaginae and Dipodinae. Thus, it appears that each stage of progressive specialization might have been attained just once.

However, although the available genetic data do not support the hypothesis of independent acquisition of key morphological adaptations to saltatory locomotion in different branches of Dipodoidea, adaptive parallelisms could have played certain role in its locomotory evolution. Thus, the finding of a Middle Miocene *Protalactaga* with incompletely fused metatarsus (Zazhigin & Lopatin 2000a) suggests that the formation of the advanced state (complete fusion) observed in modern allactagines and dipodines significantly postdates the split between them.

Possible interactions between adaptive trends in jerboa evolution

The source of the conflict between molecular and morphological partitions (as well as within the latter) remains unclear. Adaptive parallelisms are often regarded as the main reason of incongruence between morphological and molecular phylogenies (see Wiens *et al.* 2003 for a discussion). To provide a non-controversial phylogenetic history of Dipodoidea based on the molecular hypothesis, one needs to identify the factors responsible for morphological parallelisms between Cardiocraniinae and Dipodinae and/or for the lack of such parallelisms between other taxa (such as Dipodinae and Allactaginae).

The pattern of morphological diversity in Dipodoidea is a result of interplay among several adaptive trends. Trophic adaptations range from omnivory (*Sicista*, most allactagines) to specialized insectivory (*Euchoreutes*), granivory (most dipodines, Cardicraniinae) or folivory (a trend characteristic for allactagines and dipodines, with *Pygeretmus* and *Paradipus* as the most specialized forms). Even more important is the evolution of substrate adaptations. The Dipodinae and Allactaginae represent alternative trends: dipodines are predominantly psammophilous, while allactagines are better adapted to hard substrate (Shenbrot *et al*. 1995). Euchoreutinae and Cardiocraniinae are less specialized

showing more affinity for soft substrate. Allactagines demonstrate a number of unique habitat-related locomotory adaptations including specific morphology of pelvis and hind limb (Fokin 1978). In contrast to scratch-digging cardiocraniines and dipodines, allactagines are highly specialized to chisel-tooth digging, their mandible being distinctively shallow and elongate, incisors procumbent (Potapova 2000b). At the same time, Allactaginae is the only jerboa subfamily which shows no tendency to develop hyperinflated multi-chambered bulla. In other lineages, the latter condition evolved independently at least four times, namely, in Euchoreutinae, Cardiocraniinae, *Paradipus* and *Eremodipus* + *Jaculus*. Also, five-toed jerboas retain rather primitive molar morphology (narrow crown with alternating cusps, mesostyle(-id) present) contrasting advanced molar pattern (broad crown with opposite cusps), which is observed (probably, as a parallelism) in the Cardiocraniinae and Dipodinae. This differentiation between the lineages can hardly be explained exclusively by difference in diet. Instead, one may hypothesize that substrate specialization in Allactaginae (and/or other groups) may have narrowed the spectrum of adaptive options in the evolution of their skull and dentition. However, the exact nature of limitations imposed by specialized structure of jaw or bulla upon evolutionary change of the molar crown is yet unclear. Whether this effect could significantly distort the phylogenetic signal of the morphological data remains to be elucidated as well.

Taxonomic implications

According to the taxonomy proposed herein Dipodoidea is divided into three families Sminthidae, Zapodidae and Dipodidae. The same pattern was previously suggested only by Vorontsov *et al.* (1971); yet, the latter classification was based on similarity/dissimilarity between karyotypes and hence, cannot be regarded as strictly phylogenetic. In contrast to that, our system explicitly adheres to the principle of monophyly. In previous classifications the four main lineages of jerboa were assigned either familial or subfamilial rank (Fig.1). We prefer to treat them rather as subfamilies within more comprehensive Dipodidae. This solution was chosen to emphasize the monophyly of all bipedal taxa within Dipodoidea.

It should be mentioned that the system of Allactaginae remains controversial. The situation is complicated by the position of *Pygeretmus* and *Allactodipus* as sister to different branches of *Allactaga*, thus making the latter paraphyletic. Although the relationships within the five-toed jerboas appear well resolved and consistently supported by the current data, we believe that this group requires additional examination based on an extended sample of species and genes. If the phylogenetic pattern inferred here proves to be correct we will face a dilemma that we should either include both *Pygeretmus* and *Allactodipus* into *Allactaga* or elevate several subgenera within the latter (*Orientallactaga, Microallactaga*) to full genera. Morphologically, *Pygeretmus* and *Allactodipus* are well differentiated from *Allactaga*. Following this argument, we provisionally accept the second concept

(elevation of some subgenera of *Allactaga* to full generic rank). The taxonomic position and status of certain taxa for which molecular data is so far unavailable (*Scarturus* and *Paralactaga*) remain problematic. If, as might be expected based on dental similarity, it will be confirmed that they belong to the same clade as *Microallactaga* then the three lineages should be recognized at subgeneric level and retained in a single genus (under the name *Scarturus*).

The proposed system for all taxa above the species rank follows below.

Dipodoidea
Sminthidae Brandt, 1855
Sicista Gray, 1827

Zapodidae Coues, 1875
Zapodini
Zapus Coues, 1875
Napaeozapus Preble, 1899
Eozapodini Zazhigin and Lopatin, 2000
Eozapus Preble, 1899

Dipodidae Fischer de Waldheim, 1817
Cardiocraniinae Vinogradov, 1925
Cardiocraniini
Cardiocranius Satunin, 1903
Salpingotini Vinogradov, 1925
Salpingotus Vinogradov, 1922 (with subgenera *Salpingotus* s.str., *Salpingotulus* Pavlinov, 1980, *Prosalpingotus* Vorontsov and Shenbrot, 1984, *Anguistodontus* Vorontsov and Shenbrot, 1984)
Euchoreutinae Lyon, 1901
Euchoreutes Sclater, 1890
Dipodinae
Dipodini
Dipus Zimmermann, 1780
Stylodipus G. M. Allen, 1925
Eremodipus Vinogradov, 1930
Jaculus Erxleben, 1777 (with subgenera *Jaculus* s. str. and *Haltomys* Brandt, 1844)
Paradipodini Pavlinov and Schenbrot, 1983
Paradipus Vinogradov, 1930
Allactaginae Vinogradov, 1925.
Allactaga F. Cuvier, 1836 (with *A. major* and *A. severtzovi*)
Orientallactaga Shenbrot, 1984 (with *O. sibirica, O. bullata* and *O. balikunica*)
? *Scarturus* Gloger, 1841 (with *S. tetradactylus*)
? *Paralactaga* Young, 1927 (with *P. euphratica*, *P. williamsi,* ?*P. hotsoni*)
? *Microallactaga* Shenbrot, 1974 (with *M. elater* and *M. vinogradovi*)
Allactodipus Kolesnikov, 1937
Pygeretmus Gloger, 1841 (with subgenera *Pygeretmus* s.str. and *Allactagulus* Nehring, 1897)

AKNOWLEDGMENTS

DNA samples of *Eremodipus lichtenstein, Pygeretmus pumilio, Allactodipus bobrinskii and Allactaga major* were kindly provided by Dr. D. Kramerov, tissue samples of *Jaculus blanfordi* were presented by A. Abramov. We are grateful to A.A. Surov, Yu.M. Kovalskya, L.A. Khlyap and I. Tikhonov for help in collecting material. We also thank I.Ya. Pavlinov, E.G. Potapova, S.V. Fomin, K.A. Rogovin and two anonymous reviewers for useful comments. The work was supported by RFBR 11-04-00020a.

REFERENCES

Adkins, R. M., Walton, A. H. & Honeycutt, R. L. (2003). Higher-level systematics of rodents and divergence time estimates based on two congruent nuclear genes. *Molecular Phylogenetics and Evolution,* 26*,* 409–420.

Barker, F. K. & Lutzoni, F. M. (2002). The utility of the incongruence length difference test. *Systematic Biology,* 51, 625-637.

Baker, R. H. & DeSalle, R. (1997). Multiple sources of molecular characters and the phylogeny of Hawaiian drosophilids. *Systematic Biology*, 46, 654–673.

Blanga-Kanfi, S., Miranda, H., Penn, O., Pupko, T., DeBry, R. W. & Huchon, D. (2009). Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evolutionary Biology, 9*, 71 doi:10.1186/1471-2148-9-71.

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

Churakov, G., Sadasivuni, M. K., Rosenbloom, K. R., Huchon, D., Brosius, J. & Schmitz, J. (2010). Rodent evolution: back to the root. *Molecular Biology and Evolution,* 27, 1315-1326.

DeBry, R. W. (2003). Identifying conflicting signal in a multigene analysis reveals a highly resolved tree: The phylogeny of Rodentia (Mammalia). *Systematic Biology*, 52**,** 604-617.

Ellerman, J. R. (1940). *The families and genera of living rodents. Vol. 1. Rodents other than Muridae*. London: Trustees of the British Museum (Natural History).

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.

Farris, J. S., Källersjö, M., Kluge, A. G. & Bult, C. (1995). Testing significance of incongruence. *Cladistics,* 10, 315-319.

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

Fokin, I. M. (1978). The locomotion and morphology of the locomotory organs in jerboas. Leningrad: Nauka. 1 - 118 (in Russian).

Fokin, I. M. (1983). Specific features of muscular apparatus in dwarf jerboas. In: I.M. Gromov (Ed). *Materials of the 6th All-Union Conference on Rodents.* (pp. 202–204). Leningrad: Nauka. (in Russian).

Fomin, S. V. (2006). Comparative morphological analysis of the postcranial skeleton of Dipodoidea (Rodentia). PhD Thesis. Moscow, Moscow State University: Sotsvetie krasok Press (in Russian).

Gambaryan, P. P. (1983). Suprafamilial groupings in the order Rodentia. In I.M. Gromov (Ed). *Materials of the 6th All-Union Conference on Rodents.* (pp. 74–76). Leningrad: Nauka. (in Russian).

Gambaryan, P. P., Potapova, E. G. & Fokin, I. M. (1980). Morphofunctional analysis of the myology of the jerboa head. *Trudy Zoologicheskovo Instituta*, *Akademiya Nauk SSSR*, 91, 3-51 (in Russian).

Heptner, V. G. (1984). Contributions to phylogeny and classification of jerboas (Dipodidae) of the fauna of the USSR. *Archives of Zoological Museum Moscow State University*, 22, 37-60 (in Russian).

Hipp, A. L., Hall, J. C., & Sytsma, K. J. (2004). Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. Systematic Biology, 53, 81-89.

Holden, M. E. & Musser, G. G. (2005). Superfamily Dipodoidea. In: Wilson, D. E., & D. M. Reeder (Eds), Mammal Species of the World. 3nd edition, vol. 2 (pp. 871-893). Baltimore: Johns Hopkins University Press.

Huchon, D., Madsen, O., Mark, J. J., Sibbald, B., Ament, K., Stanhope, M. J., Catzeflis, F., de Jong, W. W. & Douzery, E. J. P. **(**2002). Rodent phylogeny and a timescale for the evolution of glires: evidence from an extensive taxon sampling using three nuclear genes. *Mol. Biol. Evol.,* 19, 1053–1065.

Jansa, S. & Weksler, M. (2004). Phylogeny of muroid rodents: Relationships within and among major lineages as determined by IRBP gene sequences. *Molecular Phylogenetics and Evolution,* 31, 256-276.

Jansa, S., Glarla, T. C. & Lim, B. K. (2009). The phylogenetic position of the rodent genus *Typhlomys* and the geographic origin

of Muroidea. *Journal of Mammalogy,* 90, 1083–1094.

Jobb, G. (2008). TREEFINDER version of October 2008. Munich, Germany. Distributed by the author at www.treefinder.de.

Klingener, D. (1964). The comparative myology of four dipodoid rodents (genera *Zapus*, *Napaeozapus, Sicista*, and *Jaculus*). Miscellaneous Publications, Museum of Zoology, University of Michigan, 124, 1-100.

Klingener, D. (1984). Gliroid and dipodoid rodents. In Anderson, S. & J. K. Jr., Jones (Eds), Orders and families of Recent mammals of the world (pp. 381-388). New York: John Wiley & Sons.

Kumar, S. & Gadagkar, S. R. (2001). Disparity Index: A simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. *Genetics,* 158*,* 1321-1327.

Löpez-Antonanzas, R. & Sen, S. (2006). New Saudi Arabian Miocene jumping mouse (Zapodidae): systematics and phylogeny. *Journal of Vertebrate Paleontology,* 26, 170–181.

McKenna, M. C., Bell, S. K. (1997). *Classification of mammals above the species level*. New York: Columbia University Press.

Michaux, J., & Catzeflis, F. (2000). The bushlike radiation of muroid rodents is exemplified by the molecular phylogeny of the LCAT nuclear gene. *Molecular Phylogenetics and Evolution,* 17, 280-293.

Michaux, J., Reyes, A. & Catzeflis, F. (2001). Evolutionary history of the most speciose mammals:Molecular phylogeny of muroid rodents. *Molecular Biology and Evolution,* 18, 2017-2031.

Miller, G. S. Jr., & Gidley, J. W. (1918). Synopsis of the supergeneric groups of rodents. *Journal of the Washington Academy of Sciences,* 8, 431-448.

Montgelard, C., Forty, E., Arnal, V. & Matthee, C. A. (2008). Suprafamilial relationships among Rodentia and the phylogenetic effect of removing fast-evolving nucleotides in mitochondrial, exon and intron fragments. BMC. *Evolutionary Biology,* 8, 321 doi:10.1186/1471-2148-8-321.

Musser, G. G. & Carleton, M. D. (2005). Superfamily Muroidea. In D. E. Wilson, & D. M. Reeder (Eds) *Mammal Species of the World,* 3nd edn, (2, pp. 894-1531). Baltimore: Johns Hopkins University Press.

Naylor, G. J. P. & Adams, D. C. (2001). Are the fossil data really at odds with the molecular data? Morphological evidence for Cetartiodactyla phylogeny reexamined. *Systematic Biolog*, 50, 444–453.

Ognev, S. I. (1948). *Mammals of the USSR and adjacent countries: Rodents (continued). Mammals of eastern Europe and northern Asia*. Akademiya Nauk SSSR, 6, 1-559 (in Russian).

Orlov, V. N. & V. N. Yatsenko (1985). Comparative cytogenetics and phylogeny of Dipodidae. In V. E. Sokolov V. V. Kucheruk (Eds) *Jerboas of USSR* (pp. 20-21). Moscow (in Russian)..

Posada, D., Crandall, K. A. (1998). Modeltest: Testing the model of DNA substitution. *Bioinformatics,* 14, 817–818.

Pavlinov, I. Ya. & Shenbrot, G. I. (1983). Male genital structure and supraspecific taxonomy of Dipodidae. *Trudy Zoologicheskovo Instituta, Akademiya Nauk SSSR*, Leningrad, 119, 67-88 (inRussian).

Potapova, E.G. (1976). Structure of the tongue in some Dipodoidea. *Zoologicheskii Jurnal,* 55, 1383-1389 (in Russian, English abstract).

Potapova, E. G. (1998). Evolutionary development of osseus middle ear in Rodentia, Dipodoidea. (Rodentia, Dipodoidea). *Zoologicheskii Jurnal,* 77, 80-87 (in Russian, English abstract).

Potapova, E. G. (2000a). Systematic position of *Euchoreutes* within Dipodoidea (Rodentia). In A. K. Agadzhanyan & V.N. Orlov, (Eds), *Systematics and Phylogeny of the Rodents and Lagomorphs* (p. 135-137). Moscow: KMK (in Russian).

Potapova, E. G. (2000b). Pathways of transforming mandible morphology in Dipodoidea (Rodentia) and the role of variouse adaptations in their specific features // *Zoologicheskii Jurnal*. 79, 1445-1456 (in Russian, English abstract).

Rambaut, A. & Drummond, A. (2005). Tracer version 1.5.Computer program distributed by the authors.Department of Zoology, University of Oxford,UK. Available via http://*http://evolve.zoo.ox.ac.uk/software.html*.

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

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Lab. Press.

Shenbrot, G. I. (1984). Dental morphology and phylogeny of five-toed jerboas of the subfamily Allactaginae. *Archives of Zoological Museum Moscow State University*, 22: 61-92 (in Russian).

Shenbrot, G. I. (1985). Current status and aspects of the phylogenetic study of jerboas. In V. E. Sokolov V. V. Kucheruk (Eds) *Jerboas of USSR* (pp. 21-25). Moscow (in Russian).

Shenbrot, G. I. (1992). Cladistic approach to the analysis of phylogenetic relationships among dipodoid rodents (Rodentia, Dipodoidea). *Sbornik Trudov Zoologicheskovo Muzeya MGU,* 29, 176-201 (in Russian).

Shenbrot, G. I., Sokolov, V. E., Heptner, V. G. & Koval'skaya, Yu. M. (1995). *Mammals of the fauna of Russia and contiguous countries. Dipodoid rodents*. Moscow: Nauka Publishers (in Russian); English translation (2008). *Jerboas. Mammals of Russia and Adjacent Regions.*; scientific editors: R. S. Hoffmann & D. E. Wilson. Enfield, New Hampshire: Science Publishers.

Shevyreva, N. S. (1983). Eurasian and North African Neogene rodents (Rodentia, Mammalia): The evolutionary basis for the Pleistocene and Recent Palearctic rodent fauna. In I. M. Gromov (Ed.), *History and evolution of the Recent rodent fauna (Neogene to modern times),*I (pp. 9-145). Moscow: Nauka (in Russian).

Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. *Systematic Biology,* 51, 492–508.

Simpson, G. G. (1945). The principles of classification and a classification of mammals. *Bulletin of the American Museum of Natural History,* 85, 1-350.

Springer, M. S., Burk-Herrick, A., Meredith, R., Eizirik, E., Teeling, E., O'Brien, S. J. & Murphy, W. J. (2007). The adequacy of morphology for reconstructing the early history of placental mammals. *Systematic Biology,* 56, 673—684.

Springer, M. S., Burk-Herrick, A., Meredith, R., Eizirik, E., Teeling, E., O'Brien, S. J. & Murphy, W. J. (2008). Morphology and placental mammal phylogeny, *Systematic Biology,* 57, 499—503.

StatSoft, Inc. (2001). STATISTICA (data analysis software system), Version 6. Available at: http://www.statsoft.com

Stein, B. R. (1990). Limb myology and phylogenetic relationships in the superfamily Dipodoidea (birch mice, jumping mice, and jerboas). *Zeitschrift für Zoologische Systematik und Evolutionsforschung,* 28, 299-314.

Steppan, S. J., Adkins, R. M. & Anderson, J. (2004). Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Systematic Biology,* 53, 533–553.

Swofford, D. L. (2003). PAUP* - *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sunderland, Massachusetts: Sinauer Associates.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24, 1596-1599.

Templeton, A. R. (1983). Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution,* 37, 221–244.

Vinogradov, B. S. (1925). On the structure of the external genitalia in Dipodidae and Zapodidae (Rodentia) as a classificatory character. *Proceedings of the Zoological Society of London,* 1925, 572-585.

Vinogradov, B. S. (1930). On the classification of Dipodidae (Rodentia). I. Cranial and dental characters. *Izvestiya Akademii Nauk SSSR*, 1930, 331-350 (in Russian).

Vinogradov, B. S. (1937). *Fauna of the USSR; Mammals, vol. 3, pt. 4. Jerboas*. Moscow-Leningrad: Academy Nauk SSSR Publishing House (in Russian).

Vorontsov, N. N., Malygina, N. A. & Radjabli, S. I. (1971). Chromosomes of the jerboas (Rodentia, Dipodidae). *Zoological Jurnal*, 50, 1853-1860 (in Russian).

Wang, B.-Y., & Qiu, Z.-X. (2000). Dipodidae (Rodentia, Mammalia) from the lower member of Xianshuihe Formation in Lanzhou Basin, Gansu, China. *Vertebrata*

PalAsiatica, 38, 10-35 (in English with Chinese summary).

Wiens, J. J., Chippindale, P. T. & Hillis, D. M. (2003). When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. *Systematic Biology,* 52, 501–514.

Yang, D. Y., Eng, B., Waye, J. S., Dudar, J. C. & Saunders, S. R. (1998). Technical note: improved DNA extraction from ancient bones using silica-based spin columns. *American Journal of Physical Anthropology,* 105, 539-543.

Zazhigin, V. S., & Lopatin, A. V. (2000a). The history of the Dipodoidea (Rodentia, Mammalia) in the Miocene of Asia: 3. Allactaginae. *Paleontological Journal,* 34, 553-565 (translated from *Paleontologicheskii Zhurnal,* 5, 82-94).

Zazhigin, V.S. & Lopatin, A.V. (2000b). Evolution, phylogeny, and classification of Dipodoidea. In A. K. Agadzhanyan & V. N. Orlov, (Eds), *Systematics and Phylogeny of the Rodents and Lagomorphs* (p. 50-52). Moscow: KMK (in Russian).

Zazhigin, V. S., & Lopatin, A. V. (2001). The history of the Dipodoidea (Rodentia, Mammalia) in the Miocene of Asia: 4. Dipodinae at the Miocene-Pliocene transition. *Paleontological Journal,* 35, 60-74 (translated from *Paleontologicheskii Zhurnal, 1,* 61-75).

Zazhigin, V. S. & Lopatin, A. V. (2005). The history of Dipodoidea and the aridization of Asia during the Neogen. *Bull. of Moscow Society of Naturalists,* 110, 69-71 (in Russian, English abstract).

Supporting information

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

Fig. S1 The maximum likelihood (ML) phylogeny of Dipodoidea and outgroup taxa as inferred from a concatenated alignment of four nuclear genes (designations as in Fig. 2).

Fig. S2 Relative rates for different codon positions in four nuclear genes in Dipodoidea as estimated in ML analysis.

Table S1 Models for the three codon positions of the four nuclear genes employed in maximum likelihood (ML) and Bayesian (BI) analyses (ML⁄BI).

Table S2 CG-content and proportion of purines in the 3rd codon positions (%) of the four nuclear exons in Dipodoidea. The values given in bold denote the most significant deviations from the rest of the sample.

Table S3 The results of comparisons between molecular-constrained and unconstrained trees (Templeton test).

Table S4 The level of incongruence of four morphological partitions with the molecular-constrained tree (Kruskal–Wallis test).

Table S5 The impact of molecular-derived constraints on the length of MP trees.

Table S6 Partitioned Bremer support for the two conflicting nodes Cardiocraniinae + Dipodinae and Euchoreutinae + Dipodinae + Allactaginae.

Data S1 Sminthinae⁄ Sicistinae controversy; GenBank accession numbers; Primers, Partitioning strategy.

Zoologica scripta

SUPPORTING INFORMATION
Molecular phylogeny and systematics of Dipodoidea: a test of morphology-based hypotheses

Vladimir S. Lebedev, Anna A. Bannikova, Marie Pagès, Julie Pisano, Johan R. Michaux and Georgy I. Shenbrot

Figure S1 Maximum likelihood (ML) phylogeny of Dipodoidea and outgroup taxa inferred from the concatenated alignment of four nuclear genes (designations as in Fig. 2).

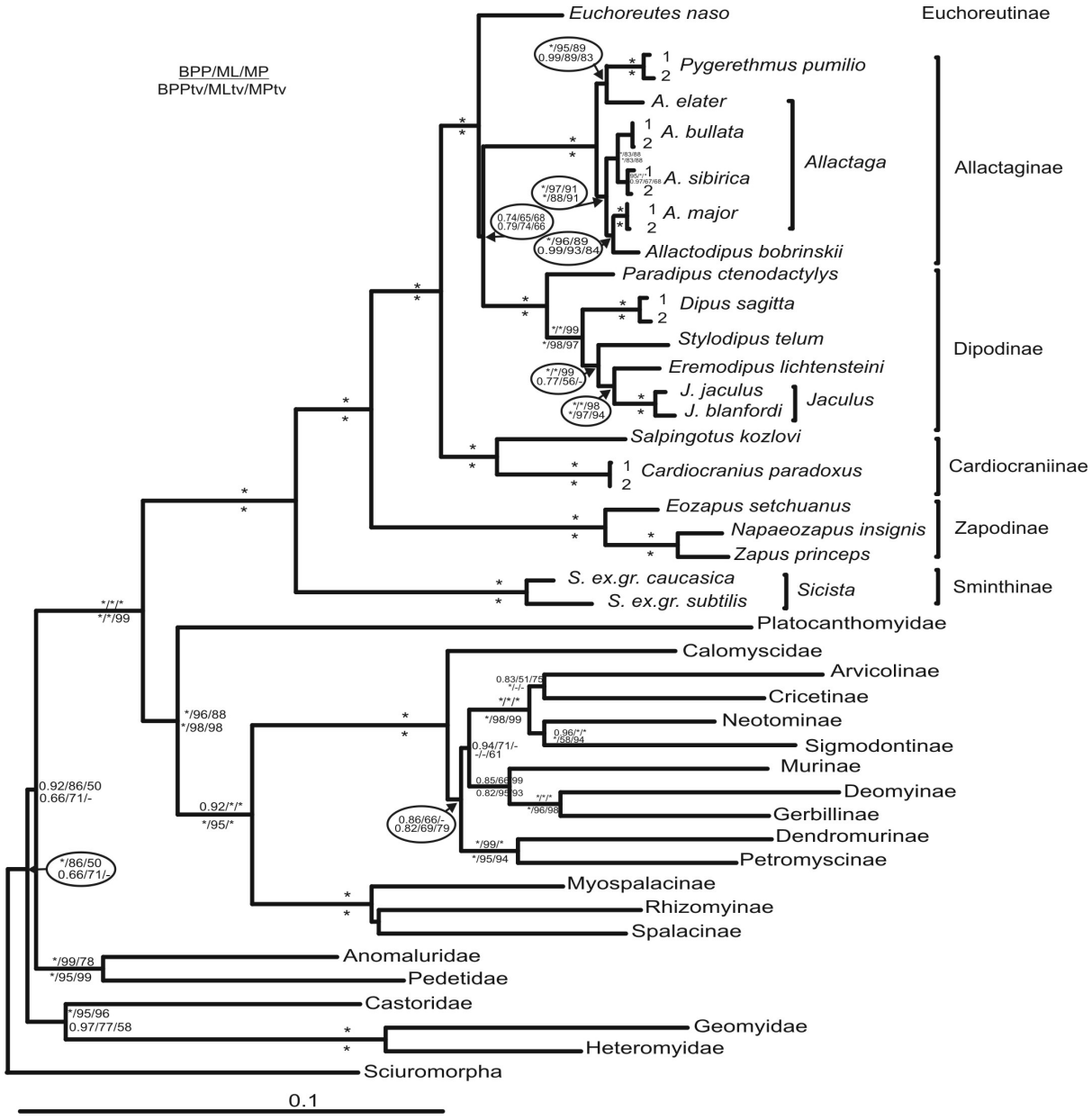


Figure S2 Relative rates of codon positions in four nuclear genes estimated in ML analysis.

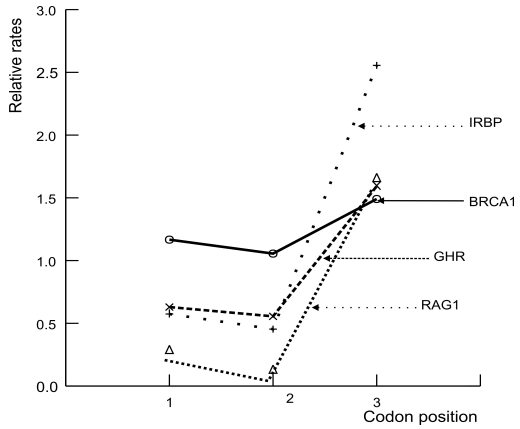


Table S1 Models of sequence evolution for the three codon positions of the four nuclear genes employed in maximum likelihood (ML) and Bayesian (BI) analyses (ML/BI). Models for ML reconstructions were determined using the routine implemented in TREEFINDER. Models for BI were selected based on the results produced by MODELTEST.

	IRBP	GHR	BRCA1	RAG1
1st	J3+G/GTR+G	HKY+G/HKY+G	HKY+G/HKY+G	J3+G/GTR+G
2nd	J1+G/GTR+G	HKY+G/HKY+G	HKY+G/HKY+G	HKY+G/HKY+I
3rd	TVM+G/GTR+G	J3+G/HKY+G	J3+G/GTR+G	HKY+G/HKY+G

Table S2 Percentage of CG-content and purines in the 3rd codon positions of the four nuclear exons in Dipodoidea. The values given in bold denote the most significant deviations from the rest of the sample.

Species	C + G				R = A + G			
	IRBP	GHR	BRCA1	RAG1	IRBP	GHR	BRCA1	RAG1
Sicista ex gr subtilis	69.9	/	37.0	**58.1**	45.9	/	51.6	50.0
Sicista ex gr caucasica	69.5	48.0	/	/	45.5	39.8	/	/
Eozapus	70.3	47.6	/	/	44.4	38.6	/	/
Zapus	70.6	49.0	/	**65.7**	44.5	39.2	/	50.7
Napaeozapus	69.2	48.5	36.8	**62.1**	45.1	39.3	47.7	48.0
Salpingotus kozlovi	**82.5**	46.5	37.8	54.3	45.1	39.4	50.2	49.7
***Cardiocranius paradoxus* 2**	70.7	48.7	36.9	52.6	44.7	39.9	51.0	49.3
***Cardiocranius paradoxus* 1**	70.9	48.5	36.7	52.5	44.8	39.8	52.2	49.0
Euchoreutes naso	70.4	46.9	37.5	53.4	44.6	39.4	49.1	51.0
***Allactaga sibirica* 1**	72.2	46.7	37.7	50.9	43.3	37.7	48.9	50.1
***Allactaga sibirica* 2**	73.0	46.9	37.1	51.6	43.0	38.7	49.3	49.7
***Allactaga bullata* 1**	72.8	46.3	38.2	51.0	43.6	38.9	50.0	49.9
***Allactaga bullata* 2**	72.5	46.1	38.4	51.3	43.4	39.1	48.9	49.9
***Allactaga major* 1**	73.0	47.0	37.1	51.7	43.2	39.3	50.0	49.7
***Allactaga major* 2**	72.7	47.2	37.4	51.7	43.4	39.4	50.0	49.4
Allactaga elater	72.9	46.1	37.1	50.9	44.7	38.4	49.8	50.4
Allactodipus bobrinskii	72.8	46.3	38.0	51.2	42.9	38.9	49.8	50.4
***Pygeretmus pumilio* 2**	74.2	45.7	34.3	52.5	44.4	39.4	49.8	49.4
***Pygeretmus pumilio* 1**	74.2	46.0	38.0	51.5	45.5	39.3	50.0	50.3
Eremodipus lichtensteini	**79.7**	49.8	41.3	50.0	44.8	38.3	49.8	50.0
***Dipus sagitta* 1**	**77.2**	46.7	41.0	50.5	46.1	39.0	51.6	49.5
***Dipus sagitta* 2**	**77.3**	46.9	40.0	50.9	45.9	39.3	50.0	49.4
Stylodipus telum	**78.4**	48.8	38.6	50.6	45.8	38.2	48.9	49.1
Jaculus jaculus	**78.7**	49.1	39.4	51.4	44.8	38.2	48.5	49.7
Jaculus blanfordi	**77.2**	47.9	38.8	51.8	44.7	37.9	48.5	49.7

Paradipus ctenodactylys	71.1	47.8		49.5	44.8	38.5		50.8
Number of pairwise disparity indexes significant at								
P<0.001	40	2	0	25	0	0	0	0
P<0.01	72	10	0	55	0	0	0	0
P<0.05	110	28	19	77	4	1	0	0

Table S3 The results of comparisons between molecular-constrained and unconstrained trees (Templeton test).

Weighting / ordering	**OW0**	**OWc**	**OWp**	**UW0**	**UWc**	**UWp**
***P* (two-sided test)**	0.016	0.007	0.37	0.015	0.013	0.004

Table S4 The level of incongruence of four morphological partitions with the molecular-constrained tree (Kruskal-Wallis test). CIunc – single-character consistency index for unconstrained tree; CIconstr – single-character consistency index for constrained tree; **Δ**CI= CIunc- CIconstr.

Partition	**Ordered analysis *Kruskal-Wallis test, *P*=0.007**			**Unordered analysis *Kruskal-Wallis test, *P*=0.026**		
	CIunc	**CIconstr**	**ΔCI***	**CIunc**	**CIconstr**	**ΔCI**
Dentition	0.679	0.567	0.112	0.765	0.630	0.135
Auditory bulla	0.504	0.441	0.063	0.588	0.592	-0.003
Glans penis	0.722	0.731	-0.009	0.885	0.872	0.012
Accessory glands	0.706	0.750	-0.044	0.708	0.750	-0.042

Table S5 The impact of molecular-derived constraints on the length of MP trees.

	Ordered analysis	**Unordered analysis**
Length of the unconstrained tree	174	139
Constrained analysis		
Constraint	**Number of extra steps (= Negative Bremer support)**	
1. Sicista is sister to other dipodoids	1	0
2. Monophyly of Zapodinae	0	1
3. Monophyly of jerboas (Cardiocraniinae + Euchoreutinae + Dipodinae + Allactaginae)	2	1
4. Monophyly of Euchoreutinae + Dipodinae + Allactaginae	**12**	**8**
5. Monophyly of Dipodinae	2	0
6. Monophyly of all Dipodinae except *Paradipus*	0	2
All constraints together	14	12
All constraints except 4	5	5

Table S6 Partitioned Bremer support for the two conflicting nodes Cardiocraniinae + Dipodinae and Euchoreutinae + Dipodinae + Allactaginae. Negative support denotes conflict.

	Cardiocraniinae +Dipodinae		**Euchoreutinae +Dipodinae +Allactaginae**	
	Ordered	**Unordered**	**Ordered**	**Unordered**
Dentition	8.87	4.02	-8.84	-10.02
Auditory bulla	-2.87	-0.90	-2.65	0.90
Glans penis	2.33	0.74	-2.52	-0.74
Accessory glands	-1.33	0.14	2.00	1.86
Total	7.00	4.00	-12.00	-8.00

Data S1

a. Sminthinae/ Sicistinae controversy.

When the genus name of a nominal family-group taxon is considered as a junior synonym of another nominal genus name, the family-group name does have not to be replaced on that account, meaning that although *Sminthus* (Nordmann, 1840) is a junior synonym of *Sicista* (Gray, 1827), Sminthinae (Brandt, 1855) should not be replaced by more recent Sicistinae (Allen, 1901). The argument in favor of Sicistinae given in MSW3 is based on Article 40.2 stating that if "… a family-group name was replaced before 1961 because of the synonymy of the type genus, the substitute name is to be maintained if it is in prevailing usage". The term "prevailing usage" refers to "the usage of the name which is adopted by at least a substantial majority of the most recent authors concerned with the relevant taxon, irrespective of how long ago their work was published" (ICZN, glossary). Although, this definition does not provide an objective criterion, the fact that Sminthinae is accepted in several important monographs on Dipodoidea (Ognev 1948, Shenbrot et al. 1995) and checklists (Pavlinov & Rossolimo 1987) argues against the opinion that Sicistinae is in prevailing usage. Since Sminthinae is an available name used for many decades, its acceptance does not threaten the stability of nomenclature; therefore, it should be considered as the valid name in accordance with the Principle of Priority.

b. GenBank accession numbers

- *Original data:*

Species	Specimen code (as in Fig. 2-3)	GenBank Accession Numbers			
		IRBP	GHR	BRCA1	RAG1
Allactaga major	1	JQ347931	JQ347911	JQ347889	JQ347867
	2	JQ347932	JQ347912	JQ347890	JQ347868
Allactaga bullata	1	JQ347929	JQ347909	JQ347887	JQ347865
	2	JQ347930	JQ347910	JQ347888	JQ347866
Allactaga sibirica	2	JQ347928	JQ347908	JQ347886	JQ347864
Allactaga elater		JQ347933	JQ347913	JQ347891	JQ347869
Allactodipus bobrinskii		JQ347934	JQ347914	JQ347892	JQ347870
Pygeretmus pumilio	1	JQ347936	JQ347916	JQ347894	JQ347872
	2	JQ347935	JQ347915	JQ347893	JQ347871
Eremodipus lichtensteini		JQ347937	JQ347917	JQ347895	JQ347873
Dipus sagitta	1	JQ347938	JQ347918	JQ347896	JQ347874
	2	JQ347939	JQ347919	JQ347897	JQ347875
Stylodipus telum		JQ347940	JQ347920	JQ347898	JQ347876
Jaculus jaculus		/	/	JQ347899	JQ347877
Jaculus blanfordi		JQ347941	/	JQ347900	/
		/	JQ347921	/	JQ347878
Paradipus ctenodactylys		JQ347942	/	/	JQ347879
		/	JQ347922	/	/
Cardiocranius paradoxus	1	JQ347926	JQ347906	JQ347884	JQ347862
	2	JQ347925	JQ347905	JQ347883	JQ347861
Salpingotus kozlovi		JQ347924	JQ347904	JQ347882	JQ347860
Euchoreutes naso		JQ347927	JQ347907	JQ347885	JQ347863
Eozapus setchuanus		/	JQ347902	/	/
Napaeozapus insignis		/	JQ347903	JQ347881	JQ347859
Sicista ex. gr. subtilis		JQ347923	/	JQ347880	JQ347858
Sicista ex. gr. caucasica		/	JQ347901	/	/

- *Sequences retrieved from GenBank*

- **DIPODOIDEA**. *Allactaga sibirica* (specimen code 1 in Fig. 2, 3): AY326076 [IRBP], AY294897 [GHR], AY294996 [BRCA1], AY241467 [RAG1]; *Jaculus jaculus*: AM407907 [IRBP], AF332040 [GHR]; *Eozapus setchuanus*: EU839445 [IRBP]; *Zapus princes*: AF297287 [IRBP], AY294935[RAG1]; *Zapus hudsonius*: AF332041 [GHR]; *Napaeozapus insignis*: AY326098 [IRBP]; *Sicista ex. gr. caucasica*: FM200058 [IRBP].
- **SCIUROMORPHA**. Gliridae – *Eliomys quercinus*: FM162056 [IRBP]; Sciuridae – *Sciurus niger*: AF332032 [GHR], AF332044 [BRCA1]; *Petaurista petaurista*: AY241473 [RAG1].
- Anomaluromorpha. **Anomaluridae** – *Anomalurus sp*.: AJ427230 [IRBP], AM407919 [GHR]; **Pedetidae** – *Pedetes surdaster*: AJ427241 [IRBP]; *Pedetes capensis*: AF332025 [GHR], DQ354448 [BRCA1], AY011882 [RAG1].
- **CASTORIDAE**. *Castor canadensis*: AJ427239 [IRBP], AF332026 [GHR], AF540622 [BRCA1], AY011880 [RAG1].
- **GEOMYIDAE**. *Thomomys talpoides*: AJ427234 [IRBP]; *Geomys bursarius*: AF332028 [GHR], AF540629 [BRCA1].
- **HETEROMYIDAE**. *Dipodomys merriami*: AJ427233 [IRBP]; *Perognathus flavus*: AF332029 [GHR], AF540638 [BRCA1].
- **PLATACANTHOMYIDAE**. *Typhlomys cinereus*: GQ272606 [IRBP], GQ272603 [GHR].
- **SPALACIDAE**. **Rhizomyinae** – *Rhizomys pruinosus*: AY326107 [IRBP], AY294899 [GHR] ; **Spalacinae** – *Spalax zemni*: U48589 [IRBP]; *Spalax* sp.: DQ354449 [BRCA1]; *Nannospalax ehrenbergi*: AY294898 [GHR], AB303250 [RAG1]; **Myospalacinae** – *Myospalax aspalax*: AY326097 [IRBP], GQ272599 [GHR].
- **CALOMYSCIDAE**. *Calomyscus baluchi*: AY163581 [IRBP], AY294901 [GHR].
- **MURIDAE**. **Deomyinae** – *Deomys ferrugineus*: AY295007 [BRCA1]; *Lophuromys flavopunctatus*: AY326091 [IRBP]; *Acomys russatus*: FM162071 [GHR]; *Uranomys ruddi*: DQ023454 [RAG1]; **Murinae** – *Rattus norvegicus*: AJ429134 [IRBP]; *Batomys granti*: AY294917 [GHR], AY295002 [BRCA1]; *Mastomys natalensis*: AY294945 [RAG1]; **Gerbillinae** – *Meriones unguiculatus*: AY326095 [IRBP], *Meriones shawi*: AF332021 [GHR], AF332048 [BRCA1]; *Taterillus emini*: DQ023453 [RAG1].
- **NESOMYIDAE**. **Petromyscinae** – *Petromyscus* sp.: DQ191517 [IRBP]; *Petromyscus monticularis*: AY294906 [GHR], AY294999 [BRCA1], AY294937 [RAG1]; **Dendromurinae**(+Cricetomyinae) – *Dendromus nyikae*: AY326083 [IRBP]; *Dendromus mesomelas*: AY294902 [GHR], AY294997 [BRCA1]; *Cricetomys gambianus*: AY294936 [RAG1];
- **CRICETIDAE**. **Sigmodontinae** – *Sigmodon alstoni*: AY163640 [IRBP]; *Oryzomys couesi*: AF332020 [GHR], AF332043 [BRCA1]; *Irenomys tarsalis*: AY294962 [RAG1]; **Neotominae** – *Peromyscus truei*: AY277413 [IRBP]; *P. leucopus*: AY294927 [GHR], AY295014 [BRCA1]; *Neotoma floridana*: AY294959 [RAG1] ; **Cricetinae** – *Phodopus sungorus*: AY163631 [IRBP]; *Mesocricetus auratus*: AF540632 [GHR], AY295013 [BRCA1]; *Cricetulus migratorius*: AY294956 [RAG1] ; **Arvicolinae** – *Eothenomys melanogaster*: AY163583 [IRBP]; *Clethrionomys gapperi*: AF540623 [GHR], AY295010 [BRCA1], AY294952 [RAG1].

c. The original primers for amplification and sequencing of *IRBP*, *GHR*, *BRCA1* and *RAG1* genes of Dipodoidea.

Primer	Sequence (5`-3`)
IRBP	
F25dip	GCGGCCATCCARCAGGTAATGAAGAG
R1175dip	GCACTGACACCTGAAACACAGAGTCCAC
F597dip	GGACATCGCCTACATCCTCAAGCAGA
R701dip	GGACACGGGCACCGTGAGGAAGAAGT
GHR	
ghr_dip_F	GCCATTCATGATAACTACAAATCTGA
ghr_dip_R	ATAGCCACAAGATGAGAGGAACTC
ghr_zap_R	ATAGCCACAAGATGCCAGGGACTC
BRCA1	
brca_All-l135F	GAGAACAGCAGTTTATTACYCACTAAAGACAG
brca_zap100F	TGTGGCACAGATGTTCRTGCCAGCTCATTACA
brca_All_982R	CTACTGGATTCACATTTTCCTCTTTCTG
brca_zap980R	CTACTGGATTCTCACTTTCCTCTTTCTGAA-
RAG1	
rag1_all-940F	GACCTGGAGAGTCCAGTGAAGTCCTTTCT
rag1_all-980F	TGAATTCCCTGATGGTRAAATGTCC
rag1_dip2117R	ACGRGTAGCATCACAAAGAGTACAAATGT
rag1-Al_613r	CTGGGTAATCATCCACAGAGGAGACAAG

d. Partitioning strategy.

Five alternative partitioning schemes were evaluated using BIC criterion. The analysis was performed in Treefinder (Jobb 2008); the tree obtained using the HKY model for the unpartitioned data was used as reference topology. Prior to BIC calculation optimal substitution model were selected for each partition from the set of models available in Treefinder. BIC imposes heavier penalty on model parameterization comparing to alternative measures such as AIC and, therefore, it is expected to favor less complex models. However, our results support the fully partitioned model, thus, highlighting significant heterogeneity of substitution patterns among genes and codon positions. At the same time, the topology of the ML tree obtained with maximum number of partitions is identical to that inferred in the unpartitioned analysis. (In the latter case, the optimum model selected by Treefinder was J3+G).

Data partitioned	Number of partitions	Total number of parameters (Except branch lengths)	Log Likelihood	BIC
By gene	4	29	-35289.04	71557.3
By codon position	3	21	-34851.34	70615.5
By gene x codon position	12	84	-34410.36	**70256.3**
As above but with 1st and 2nd positions combined	8	58	-34545.33	70310.5
Unpartitioned	1	7	-35540.90	71878.4

Chapitre 3 :

Impact des changements environnementaux, climatiques et géologiques passés sur la distribution et l'histoire évolutive des espèces :

Cas de la superfamille des Dipodoidea

Mots-clés: Biogéographie historique, phylogénie moléculaire, horloge moléculaire relâchée, modèle de dispersion-extinction-cladogenèse.

Aujourd'hui, il est communément accepté que la biodiversité est gravement menacée par les rapides changements climatiques et environnementaux survenus suite aux perturbations d'origine anthropique de ces dernières décennies (Blondel, 1995). Les patrons spatiaux sont intimement liés à l'histoire évolutive des organismes (Ricklefs, 2004). Dès lors, une des solutions pour tenter de préserver la biodiversité actuelle revient à étudier comment la biodiversité du passé a répondu aux perturbations abiotiques majeures survenues dans le passé (Cox & Moore, 2010).

De façon générale, trois types de barrières ont été identifiées comme moteurs de diversification en biogéographie : les océans, les chaînes de montagne et les déserts (Cox & Moore, 2010). A ce propos, la surrection de l'Himalaya et du Plateau Tibétain est l'un des plus grands évènements orogéniques que la planète n'ait jamais connu. Née de la collision des plaques indienne et eurasienne il y a 50-40 Ma, cette chaine de montagnes est connue pour avoir subi plusieurs grands évènements de surrection qui ont fortement impacté l'environnement et le climat en Asie (He *et al.*, 2001; Guo *et al.*, 2002; Johansson *et al.*, 2007; Bouilhol *et al.*, 2013). En particulier, la plus importante phase d'orogenèse de l'Himalaya qui a eu lieu au Miocène moyen lors de 'l'optimum climatique du Miocène moyen' est connue pour avoir déclenché de grands bouleversements climatiques et environnementaux sur tout le continent asiatique (Zachos *et al.*, 2008; Bouilhol *et al.*, 2013). Elle a en effet engendré l'aridification de la région d'Asie Centrale et par conséquent, elle a également induit d'importantes modifications environnementales (e.g. la désertification) (Guo *et al.*, 2002). Ces perturbations géologiques et leurs conséquences climatiques et environnementales sont d'ailleurs fort connues pour avoir favorisé la diversification de différents groupes d'animaux (e.g. diverses sous-familles de poisson-chat, de genres de fauvette, d'un groupe de grenouilles de la tribu des Paini (He *et al.*, 2001; Che *et al.*, 2010; Bouilhol *et al.*, 2013)) et de plantes (e.g. les espèces de fougères du genre *Lepisorus*, les espèces de plantes à fleur du genre *Caragana* (Zhang & Fritsch, 2010; Wang *et al.*, 2012)). Concrètement, l'Himalaya est donc considéré comme une « pompe à espèce » dû au fait que beaucoup d'espèces se sont formées dans l'Himalaya et ont ensuite étendu leur aire de distribution aux alentours de la chaine de montagne où d'autres espèces sont ensuite apparues (Johansson *et al.*, 2007).

Résumé de l'article

Généralement, les changements globaux climatiques, géologiques et environnementaux laissent leur empreinte sur la distribution et l'évolution des organismes vivants (e.g. Fabre *et al.* (2012b)) ; et depuis le développement d'outils génétiques (e.g. séquences de gène), leur impact peut être estimé à partir d'approches phylogénétiques. En Asie, la surrection de l'Himalaya – et les bouleversements climatiques et environnementaux que cela a entrainé – a engendré la mise en place de nombreuses barrières biogéographiques qui sont connues pour avoir favorisé divers évènements de diversification dans différents groupes (e.g. Zhang *et al.* (2006)). Comprendre quel facteur évolutif explique le mieux l'évolution du vivant peut aider à améliorer nos prédictions à propos de l'évolution des distributions spatiales des organismes vivants.

Pour mieux comprendre l'impact des changements abiotiques sur la diversification des clades de la région holarctique, nous nous sommes intéressés à l'histoire évolutive de la superfamille des Dipodoidea (Rodentia), laquelle comprend les sicistes, les souris-sauteuses et les gerboises. Essentiellement arboricoles, les sicistes (Sminthidae, 13 spp.) sont principalement distribuées dans les prairies subalpines et les forêts boréales et alpines d'Europe, de Russie et d'Asie Centrale. Quelques espèces sont exceptionnellement réparties dans des régions semi-désertiques ou steppiques (e.g. *Sicista subtilis ou S. severtzovi*). Les souris-sauteuses (Zapodidae, 5 spp.) occupent plutôt les aires boisées et riveraines des forêts de conifères d'Amérique du Nord et de Chine. Les gerboises (Dipodidae, 33 spp.) sont quant à elles principalement inféodées aux régions sèches, tels que les déserts, les semi-déserts et les steppes d'Afrique du Nord et d'Asie. Etant donné ces nombreuses espèces distribuées dans des zones arides et isolées et ces patrons de distribution disjoints répandus sur toute la région holarctique, cette superfamille constitue un groupe intéressant pour tester des scénarios biogéographiques. De plus, les Dipodoidea possèdent un large panel de fossiles. Les plus anciens fossiles connus de Dipodoidea proviennent de l'Eocène moyen d'Amérique du Nord (*Elymys* (?Zapodidae) et *Simimys* (Simimyidae)) et d'Asie (*Heosminthus* (Zapodidae ou Dipodidae) et *Sinosminthus* (Zapodidae)). Ces fossiles vont nous permettre de comparer nos résultats obtenus avec les données paléontologiques et donc, de valider l'histoire évolutive biogéographique de ce clade de rongeurs. Le but de cette étude a été d'augmenter nos connaissances sur la superfamille des Dipodoidea et de déterminer les facteurs responsables de l'évolution de ce groupe. Pour cela, nous devions obtenir la phylogénie

moléculaire la plus complète possible à partir de laquelle on pourrait ensuite reconstruire le cadre spatio-temporel au sein duquel les rongeurs Dipodoidea ont évolué.

Dans un premier temps, nous avons procédé à des analyses phylogénétiques sur la base de quatre gènes nucléaires (BRCA1, GHR, IRBP, RAG1) et d'un gène mitochondrial (cytochrome *b*). Notre échantillonnage comprenait 34 des 51 espèces appartenant à 15 des 16 genres de Dipodoidea décrits dans *Mammal Species of the World* (MSW). Il incluait également 12 espèces de rongeurs qui ont servi de groupes externes pour enraciner nos phylogénies. Ces espèces externes aux Dipodoidea ont été soigneusement sélectionnées pour reconstituer spécifiquement des nœuds de l'arbre permettant d'utiliser des calibrations fossiles lors des analyses de datation moléculaire. La congruence des gènes a été testée en reconstruisant un arbre pour chaque gène à partir d'une approche de maximum de vraisemblance (PHYML) et en comparant les cinq topologies inférées. Ensuite, les cinq gènes ont été combinés. A partir de cette supermatrice, des phylogénies moléculaires ont été reconstruites en utilisant des méthodes de maximum de vraisemblance (RAXML) et d'inférence bayésienne (MRBAYES). Enfin, des hypothèses de topologies alternatives ont été testées pour résoudre les questions restées en suspens dans Lebedev *et al.* (2012) (Chapitre 2). Certains nœuds de l'arbre ont alors été contraints de manière à tester : (i) la monophylie du genre *Allactaga* et la proximité phylogénétique des Euchoreutinae avec (ii) les Allactaginae et (iii) les Dipodinae respectivement.

L'arbre obtenu à partir de la matrice de gènes combinés a ensuite été replacé dans un cadre temporel en utilisant une méthode bayésienne d'horloge moléculaire relâchée (BEAST). Quatre fossiles de rongeurs ont été utilisés comme points de calibrations. Une procédure de « validation-croisée » a été employée pour estimer l'impact de chaque donnée paléontologique sur nos estimations de temps de divergence. Au total, 12 chronogrammes ont été reconstruits. Ensuite, les estimations des aires ancestrales ont été inférées à partir du modèle 'DEC' de 'dispersion – extinction – cladogenèse' et du programme LAGRANGE. Notre modèle géographique a été divisé en neuf aires en fonction de critères paléogéographiques et de la distribution des espèces actuelles. Le modèle 'temps – stratifié' a permis d'établir les taux de dispersion entre les aires en fonction du temps. La période étudiée a été divisée en cinq. Pour chaque intervalle de temps, les taux de dispersion appliqués étaient estimés en fonction des reconstructions paléogéographiques de la Terre et des potentielles barrières biogéographiques entre les aires. Pour tester l'impact des taux de dispersion sur nos

reconstructions biogéographiques, trois types de matrices ont été testés : la matrice nulle où tous les taux de dispersion étaient égaux à 1 et les deux autres matrices où les taux de dispersion étaient fonction de l'intervalle de temps et des barrières entre les aires. Enfin, des contraintes géographiques ont été appliquées sur le nœud correspondant à la radiation des Dipodoidea pour tester l'aire du centre d'origine de la superfamille.

Etant donné que les topologies des arbres reconstruits à partir de chacun des gènes étaient congruentes, les cinq marqueurs ont été combinés (supermatrice finale de 4974 nucléotides). La phylogénie moléculaire obtenue était fortement soutenue (i.e. toutes les probabilités postérieures étaient supérieures à 0.95 et 82% des bootstraps étaient supérieurs à 95) et correctement résolue. Les analyses de topologies alternatives ont montré que *Euchoreutes naso* (la seule et unique espèce des Euchoreutinae) se place en groupe frère du clade des 'Allactaginae – Dipodinae' et que le genre *Allactaga* est paraphylétique. Incluant 34 des 51 espèces répertoriées dans MSW, la phylogénie moléculaire la plus complète à ce jour de la superfamille des Dipodoidea a été reconstruite. Les temps de divergence estimés au cours des différentes analyses de datation ont donné des estimations identiques quelle que soit la stratégie de calibration utilisée, validant ainsi le choix des fossiles utilisés comme points de calibration et confirmant les résultats de datation obtenus. Etant donné qu'aucune des analyses n'était significativement meilleure que les autres (cf. Appendix S2), nous avons sélectionné celle qui avait le meilleur score de vraisemblance. Concernant les reconstructions d'aires ancestrales, nous avons aussi sélectionné l'analyse avec le meilleur score de vraisemblance (voir le manuscrit pour plus de détail), laquelle suggère que la région de l'Himalaya et du Plateau Tibétain constitue l'aire d'origine des Dipodoidea.

La divergence entre les Dipodoidea et son groupe frère, les Muroidea, (i.e. 'groupe du tronc', Fig. 5, p. 10) a été estimée au Paléocène supérieur (c. 57.72 Ma) en accord avec le plus vieux fossile connu de Dipodoidea, *Elymys complexus* datant du Bridgérien d'Amérique du Nord (46.2-50.3 Ma). La radiation des Dipodoidea modernes (i.e. 'groupe de la couronne', Fig. 5, p. 10) a eu lieu pendant l'Eocène moyen (c. 40.62 Ma) dans les régions de l'Himalaya, du Plateau Tibétain et d'Asie Centrale. D'ailleurs, les plus anciens fossiles connus de Dipodoidea répertoriés sur le continent asiatique – *Heosminthus* et *Sinosminthus* – ont été identifiés comme datant de l'Eocène moyen. Cette radiation serait associée aux importants évènements géologiques qui ont eu lieu à cette époque dans la région asiatique. La surrection de l'Himalaya due à la collision de l'Inde avec l'Asie est en effet déjà connue comme moteur

de diversification chez d'autres groupes d'organismes. Par ailleurs, d'autres évènements importants de diversification ayant marqué l'histoire évolutive des Dipodoidea correspondent aussi à des perturbations paléogéographiques et climatiques passées qui ont eu lieu en Asie. Notamment, la divergence entre les familles de Zapodidae et Dipodidae a eu lieu en Asie Centrale durant la transition Eocène-Oligocène, lorsque la température globale diminuait fortement et que la 'faune de l'Oligocène' dominée par des rongeurs et lagomorphes remplaçait 'la grande faune de l'Eocène' qui était dominée par des espèces périssodactyles. Au Miocène inférieur eut lieu la phase la plus importante de la surrection de la région de l'Himalaya et du Plateau Tibétain, ce qui a entrainé l'aridification et la modification des habitats de la région et des alentours. Nos reconstructions biogéographiques ont montré que les Zapodidae et les Dipodinae ont évolué suite à ces perturbations abiotiques ; leurs radiations eurent effectivement lieu en Asie Centrale pendant le Miocène inférieur. [Les scénarios biogéographiques complets des Zapodinae et des Dipodinae sont détaillés dans l'article.]

De façon générale, cette étude a permis de mettre en évidence le fait que l'histoire évolutive biogéographique de la superfamille des Dipodoidea est fortement liée aux perturbations géologiques, climatiques et environnementales survenues depuis l'Eocène moyen à la suite de la surrection de l'Himalaya et du Plateau Tibétain. Tous ces changements ont favorisé la mise en place de nouvelles niches, lesquelles ont à leur tour permis la diversification des Dipodoidea. Cette étude met donc en évidence l'importance des évènements géologiques et paléoclimatiques sur la diversification des mammifères de la Paléarctique.

Article

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

Julie PISANO[1,2], Fabien L. CONDAMINE[3], Vladimir LEBEDEV[4], Anna BANNIKOVA[5], Jean-Pierre QUERE[2], Gregory I. SHENBROT[6], Marie PAGES†[1,2] and Johan R. MICHAUX†

Journal of Biogeography (2015), 42(5): 856-870, DOI: 10.1111/jbi.12476

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

Keywords

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

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 placed 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.

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 (Bouihol *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), Allactaginae (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; Pages *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 $MgCl_2$, 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_c) implemented in PARTITIONFINDER 1.1.1 (Lanfear *et al.*, 2012). Branch lengths were estimated independently for each subset by setting 'branchlengths = 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 '*allcompat*' 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 BEAUTI (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% quantile = 35.95; 97.5% quantile = 54.05].

FC2: Following Steppan *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% quantile = 10.39; 97.5% quantile = 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.*

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 – *Douglasciurus jeffersoni*	37.9 (35.49–42.09)	37.91 (35.51–41.94)	**37.5 (32.84–43.97)**	**37.39 (33.11–44.43)**	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 – *Progonomys* (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)	**12.94 (12.12–14.34)**	**12.91 (12.11–14.29)**	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 – *Apodemus jeanteti* & *Apodemus dominans*	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)	**7.32 (6.88–8.05)**	**7.33 (6.9–8.08)**	7.29 (6.87–8.05)	7.29 (6.89–8.11)
14	FC4 – *Sicista primus* (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)	**17.7 (16.69–19.56)**	**17.72 (16.73–19.7)**
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 Cardiocraniinae	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 Cardiocraniinae	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)
3 8	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)

sylvaticus [offset = 4.376; log(mean) = 2.4345; 2.5% quantile = 7.0; 97.5% quantile = 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% quantile = 17.01; 97.5% quantile = 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, 'ucld.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 (Holden & 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 Mounts, without North Africa); C, Siberia (from Ural Mounts 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

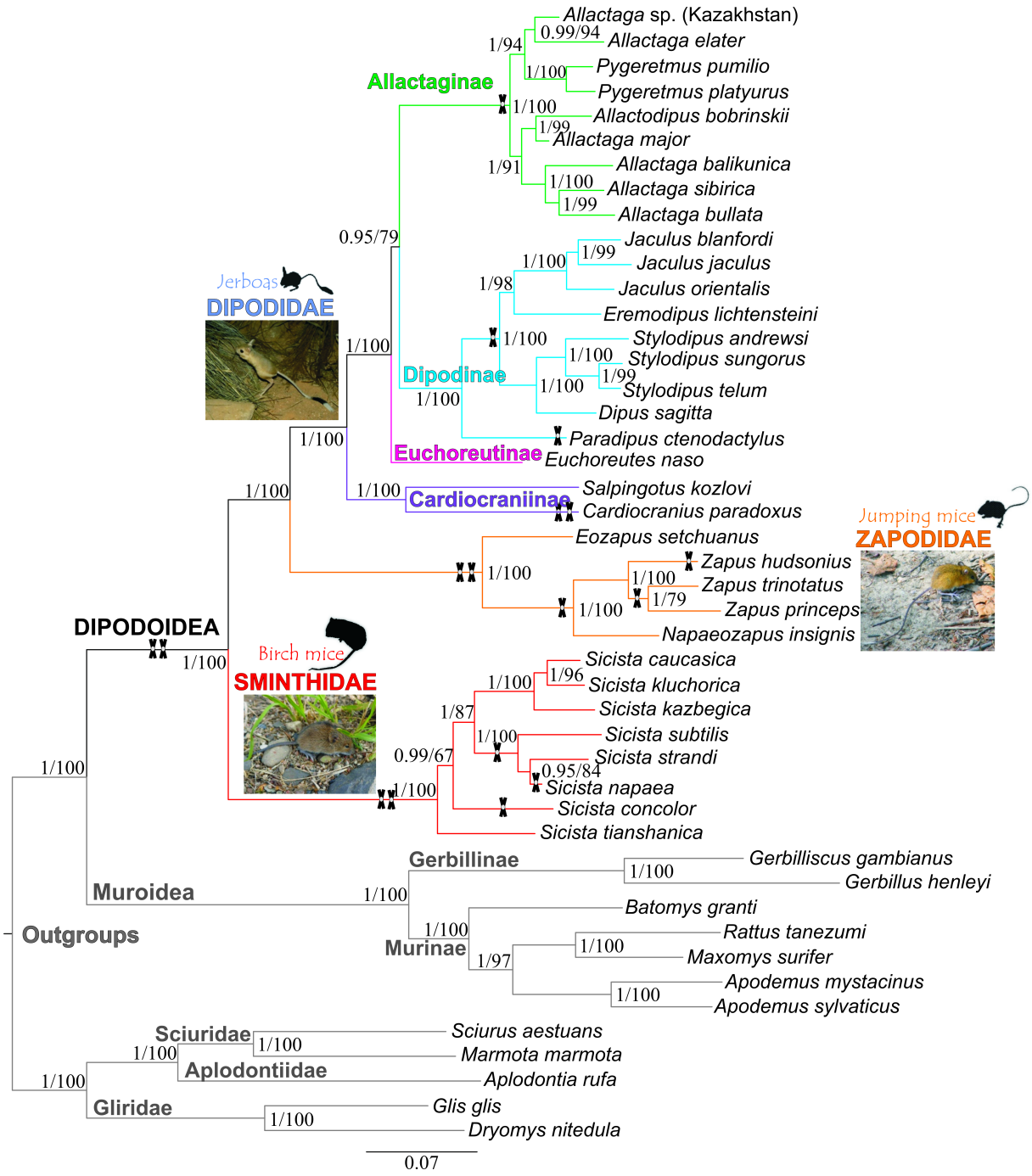


Figure 1 Phylogenetic relationships among 34 species of Dipodoidea obtained using the species-level matrix (Bayesian inference tree). Analyses were performed using the partitioned dataset of the combined *cytb*, *IRBP*, *GHR*, *BRCA1* and *RAG1* genes. Bayesian inferences and maximum likelihood analyses 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 from © Wikimedia commons.

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 ≥ 0.95,

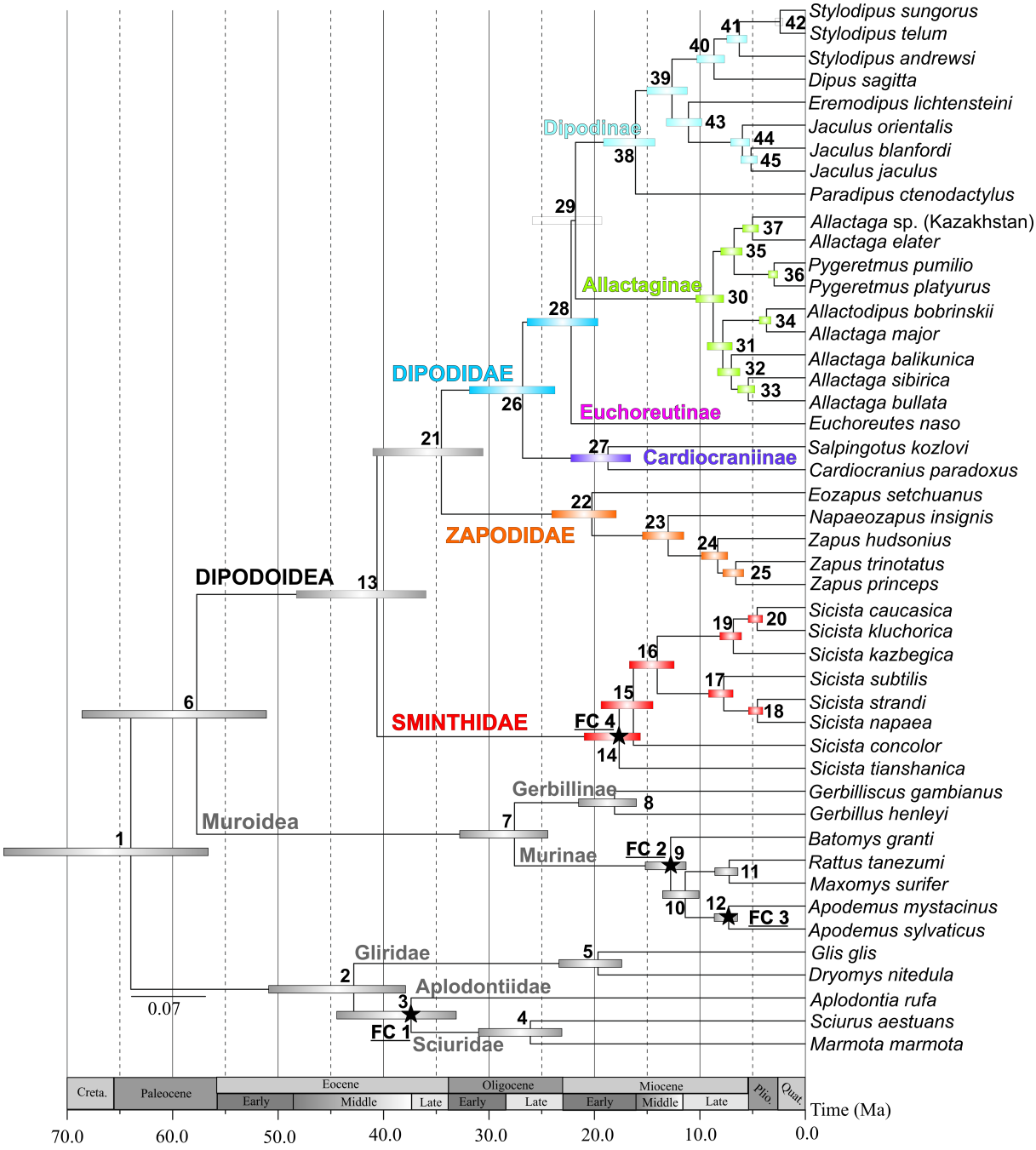


Figure 2 Dated phylogeny of Dipodoidea. The figure shows the maximum clade credibility tree with median ages from the Bayesian uncorrelated lognormal method that is discussed in further details in this study. Black stars indicate fossil calibrations of node. Numbers at nodes refer to Table 1 (see '*NoFC1*' 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.

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 ($BF_{\text{No constraint tree vs. H(1,2,3)}} > 10$) (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).

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 details 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
Root DG	**−112.8**	**−108.6**	**−107.4**
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

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.

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 ten dating strategies was significantly better than the others (see Appendix S2 for BF and likelihood scores). Consequently, we selected the analysis that produced the highest likelihood score, i.e. the chronogram obtained without the use of the fossil *Douglassciurus jeffersoni* and using the birth–death model of speciation ('No FC1' 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

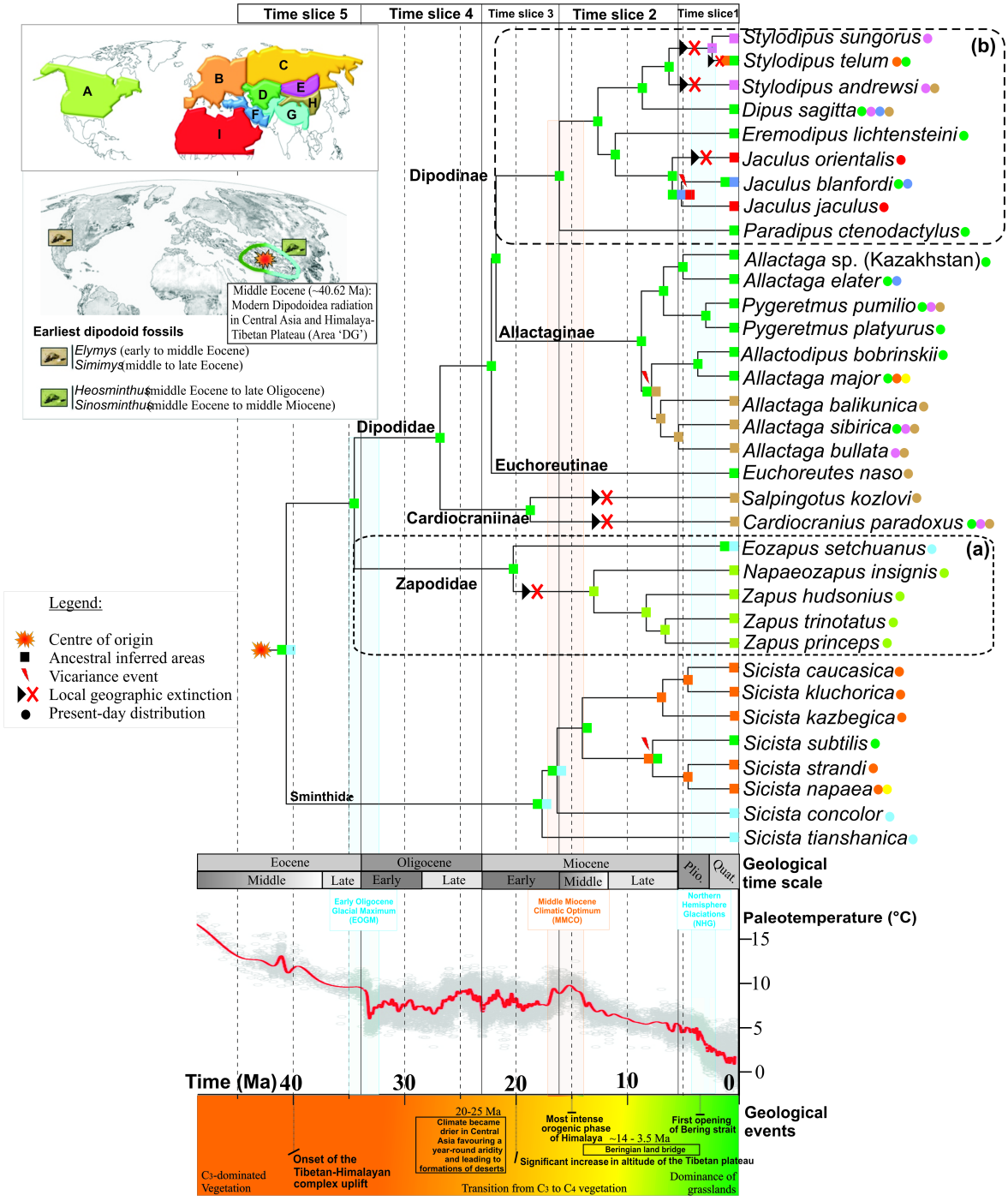


Figure 3 Temporal and geographical history of Dipodoidea based on results of the dispersal–extinction–cladogenesis (DEC) analysis for which the root of Dipodoidea was constrained with areas 'DG' and inferred using the M2 stratified model. The maximum clade credibility tree with the highest likelihood 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, for 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_3 to C_4 grasses (Cerling *et al.*, 1997).

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 ($\log L_D$ (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 ($\log L_{DG}$ (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 congruent 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 *E. 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 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 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.

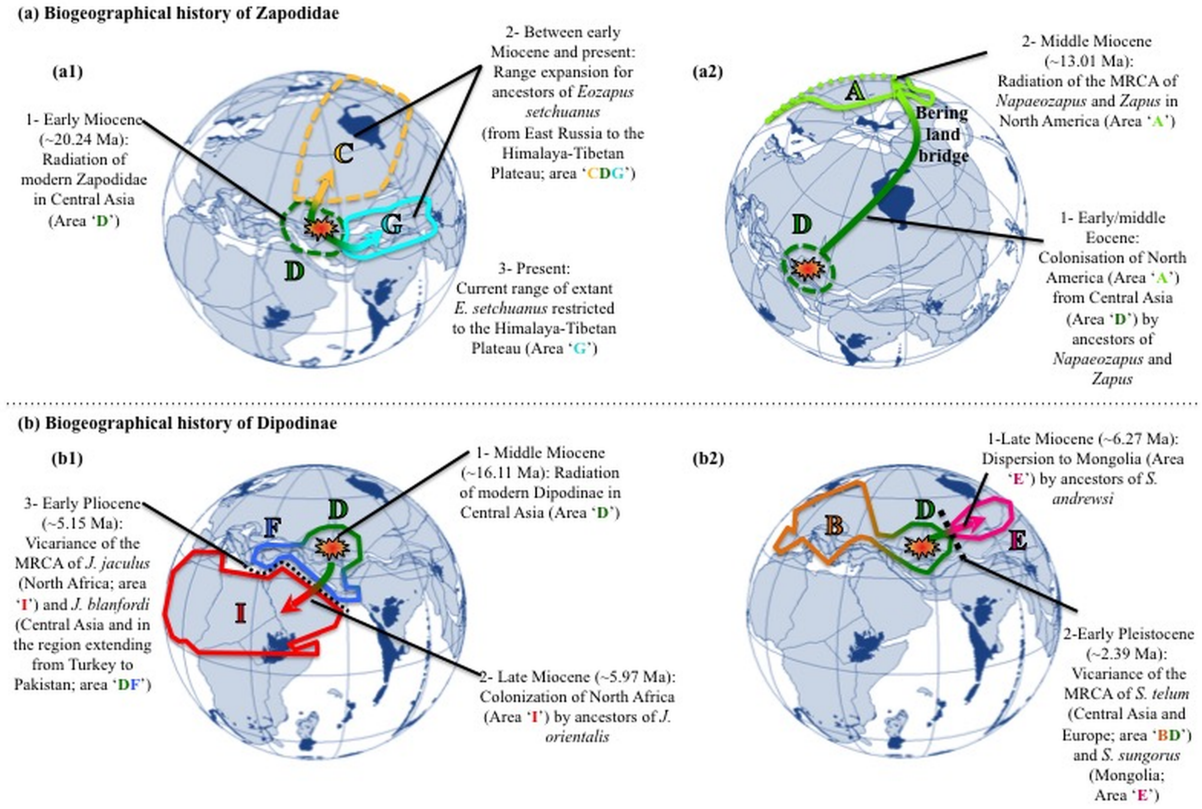


Figure 4 Biogeographical scenarios for the distribution patterns of (a) Zapodidae and (b) Dipodinae, with specific palaeogeographical maps. Concerning areas, the colours and the alphabet code are the same as those in Fig. 3. Dotted lines refer to ancestral areas. Red splashes refer to the centre of origin of clades. (a1) During early Miocene occurred the radiation of modern Zapodidae in Central Asia. Ancestors of the Asian *Eozapus setchuanus* expanded their range across East Russia and Himalaya-Tibetan Plateau. Nowadays, *E. setchuanus* is exclusively distributed in the Himalaya-Tibetan Plateau. (a2) Between 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 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).

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 during 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 & Fig. 4a).

Between *c.* 14 to 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 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* 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 sampling, 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. Catzeflis, G. Dobigny J.-M. Duplantier, W. Fuwen, L. Granjon, G. Musser, the Burke Museum (Seattle, USA), and the National Museum of Natural History (MNHN; Paris, France) for donations of dipodoid tissues. Analyses were performed at the CBGP HPC computational platform, maintained by A. Dehne-Garcia. 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 ECOEVOBIO-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. Boulder, CO.

Bouilhol, 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, Washigton, DC.

Fabre, P.H., Irestedt, M., Fjeldså, J., Bristol, R., Groombridge, J.J., Irham, M. & Jønsson, 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 geographic 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 10th 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) Erlianomys 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*, **109,** 7793–7798.

McKenna, M.C. & Bell, S.K. (1997) *Classification of mammal above the species level*. Columbia University Press, New York.

Meredith, R.W., Janecka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simao, T.L., Stadler, T., Rabosky, D.L., Honeycutt, R.L., Flynn, J.J., Ingram, C.M., Steiner, C., Williams, T.L., Robinson, T.J., Burk-Herrick, A., Westerman, M., Ayoub, N.A., Springer, M.S. & Murphy, W.J. (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.

Pages, M., Chaval, Y., Herbreteau, 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., Hohna, 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 Publishing Corporation, 06082 Enfield.

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 muroid 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 liguanqiao 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 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 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

Journal of Biogeography

SUPPORTING INFORMATION

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

Julie Pisano, Fabien L. Condamine, Vladimir Lebedev, Anna Bannikova, Jean-Pierre Quéré, Gregory I. Shenbrot, Marie Pagès and Johan R. Michaux

Appendix S1 Taxon sampling used for this study and GenBank accession numbers. *Accession numbers of sequences retrieved from GenBank. ᵒSequences obtained by V. Lebedev and A. Bannikova.

Sample information						GenBank Accession Number					
Family	Subfamily	Genera	Species	Laboratory sample number	Curator	Cyt *b*	IRBP	GHR	BRCA1	RAG1	Reference
Dipodidae	Allactaginae	*Allactaga*	*balikunica*	M2005_125	V. Lebedev	ᵒKM397180	ᵒKM397136	ᵒKM397227	ᵒKM397274	ᵒKM397313	This study
			bullata	ZMMU S179572	V. Lebedev	ᵒKM397179	JQ347929*	JQ347909*	JQ347887*	JQ347865*	Lebedev *et al.* 2012
			elater	D2007_30	L. Khlyap	ᵒKM397178	JQ347933*	JQ347913*	JQ347891*	JQ347869*	Lebedev *et al.* 2012
			major	ZMMU 01/08_32	V. Lebedev	-	JQ347931*	JQ347911*	JQ347889*	JQ347867*	Lebedev *et al.* 2012
			sp. (Kazakhstan)	T-1163	F. Catzeflis	-	KM397135	KM397226	KM397273	KM397312	This study
			sibirica	XJ0309SD03	J-P. Quéré	KM397214	KM397134	KM397225	KM397272	KM397311	This study
		Allactodipus	*bobrinskii*	t.c. Abo	D. Kramerov	ᵒKM397181	JQ347934*	JQ347914*	JQ347892*	JQ347870*	Lebedev *et al.* 2012
		Pygeretmus	*platyurus*	148231	V. Lebedev	-	ᵒKM397165	-	-	ᵒKM397314	This study
			pumilio	t.c. Apy	D. Kramerov	ᵒKM397182	JQ347936*	JQ347916*	JQ347894*	JQ347872*	Lebedev *et al.* 2012
	Cardiocraniinae	*Cardiocranius*	*paradoxus*	ZMMU S188830	V. Lebedev	ᵒKM397211	JQ347926*	JQ347906*	JQ347884*	JQ347862*	Lebedev *et al.* 2012
		Salpingotus	*kozlovi*	ZMMU S183051	V. Lebedev	ᵒKM397210	ᵒKM397159	ᵒKM397253	ᵒKM397301	ᵒKM397341	Lebedev *et al.* 2012
	Dipodinae	*Dipus*	*sagitta*	XJ0309SA14	J-P. Quéré	KM397196	KM397148	KM397240	KM397284	KM397327	This study
				XJ0309AD01	J-P. Quéré	-	KM397149	KM397241	-	KM397328	This study
		Eremodipus	*lichtensteini*	t.c. Eli	D. Kramerov	ᵒKM397200	JQ347937*	JQ347917*	JQ347895*	JQ347873*	Lebedev *et al.* 2012
		Jaculus	*blanfordi*	San-Diego, t.c. 2006	V. Lebedev	ᵒKM397195	JQ347941*	-	JQ347900*	-	Lebedev *et al.* 2012
				ZIN 96242	V. Lebedev	-	-	JQ347921*	-	JQ347878*	Lebedev *et al.* 2012

			jaculus	JJ257	A. Ben Faleh	KM397184	KM397138	KM397229	KM397276	KM397316	This study
				GAN204	A. Ben Faleh	KM397188	KM397142	KM397233	KM397280	KM397320	This study
				406	A. Ben Faleh	KM397186	KM397140	KM397231	KM397278	KM397318	This study
			deserti	N4284	A. Ben Faleh	KM397187	KM397141	KM397232	KM397279	KM397319	This study
				399	A. Ben Faleh	KM397185	KM397139	KM397230	KM397277	KM397317	This study
			orientalis	JO271	A. Ben Faleh	KM397189	KM397143	KM397234	KM397281	KM397321	This study
				JO272	A. Ben Faleh	KM397190	KM397166	KM397235	-	KM397322	This study
				423	A. Ben Faleh	KM397191	KM397144	KM397236	KM397282	KM397323	This study
				424	A. Ben Faleh	KM397192	KM397145	KM397237	KM397283	KM397324	This study
				444	A. Ben Faleh	KM397193	KM397146	KM397238	-	KM397325	This study
				445	A. Ben Faleh	KM397194	KM397147	KM397239	-	KM397326	This study
		Paradipus	*ctenodactylus*	3049	J-P. Quéré	KM397183	KM397137	KM397228	KM397275	KM397315	This study
		Stylodipus	*andrewsi*	ZMMU S189199	V. Lebedev	♂ KM397198	♂ KM397163	♂ KM397242	♂ KM397285	♂ KM397329	This study
			sungorus	ZMMU S189200	V. Lebedev	♂ KM397199	♂ KM397164	♂ KM397243	♂ KM397286	♂ KM397330	This study
			telum	K2003	V. Lebedev	♂ KM397197	JQ347940*	JQ347920*	JQ347898*	JQ347876*	Lebedev *et al* 2012
	Euchoreutinae	*Euchoreutes*	*naso*	ZMMU S179636	V. Lebedev	♂ KM397212	JQ347927*	JQ347907*	JQ347885*	JQ347863*	Lebedev *et al* 2012
Sminthidae		*Sicista*	*caucasica*	T-0764	F. Catzeflis	KM397201	KM397150	KM397244	KM397287	KM397331	This study
			concolor	MK0509BL02	J-P. Quéré	KM397207	KM397167	KM397252	KM397294	KM397339	This study
			kazbegica	T-0761	F. Catzeflis	KM397202	KM397151	KM397245	KM397288	KM397332	This study
				T-0762	F. Catzeflis	KM397203	KM397152	KM397246	KM397289	KM397333	This study
			kluchorica	T-0763	F. Catzeflis	KM397206	KM397157	KM397251	KM397293	KM397338	This study
			napaea	T-MNHN1999-451	J-P. Quéré	-	KM397155	KM397249	-	KM397336	This study
				T-MNHN1999-452	J-P. Quéré	-	KM397156	KM397250	KM397292	KM397337	This study
			strandi	ZMMU S181441	V. Lebedev	KM397209	♂ KM397158	-	-	♂ KM397340	This study
			subtilis	YuMK 1559	V. Lebedev	♂ KM397208	JQ347923*	-	JQ347880*	JQ347858*	Lebedev *et al* 2012
			tianshanica	NT0609AQ10	J-P. Quéré	KM397204	KM397153	KM397247	KM397290	KM397334	This study
				NT0609BR03	J-P. Quéré	KM397205	KM397154	KM397248	KM397291	KM397335	This study
Zapodidae		*Eozapus*	*setchuanus*	MK0509AZ01	J-P. Quéré	KM397176	KM397132	KM397223	KM397270	KM397309	This study
				RT0406AG10	J-P. Quéré	KM397177	KM397133	KM397224	KM397271	KM397310	This study
		Napaeozapus	*insignis*	3033	G. Musser	KM397168	KM397124	KM397215	KM397262	KM397302	This study
				3034	G. Musser	KM397169	KM397125	KM397216	KM397263	KM397303	This study
		Zapus	*hudsonius*	3035	G. Musser	KM397170	KM397126	KM397217	KM397264	KM397304	This study
				3036	G. Musser	KM397171	KM397127	KM397218	KM397265	KM397305	This study
			princeps	UWBM 76718	S. Birks	KM397175	KM397131	KM397222	KM397269	-	This study
				UWBM 78056	S. Birks	KM397174	KM397130	KM397221	KM397268	KM397308	This study
			trinotatus	UWBM 78818	S. Birks	KM397172	KM397128	KM397219	KM397266	KM397306	This study
				UWBM 80356	S. Birks	KM397173	KM397129	KM397220	KM397267	KM397307	This study
Aplodontiidae		*Aplodontia*	*rufa*	T-1462	F. Catzeflis	AJ389528*	AJ427238*	-	-	KM397344	Published and this study
Gliridae	Glirinae	*Glis*	*glis*	NHMC80.5.49.6	P. Lymberakis	-	-	AF332030*	AF332045*	-	Published
						-	-	-	KM397298	-	This study
						AJ225031*	AJ427235*	AM407916*	-	AB253972*	Published

	Leithiinae	*Dryomys*	*nitedula*	T-0768	F. Catzeflis	AJ225116*	AJ427236*	KM397257	KM397299	-	Published and this study
Muridae	Gerbillinae	*Gerbilliscus*	*gambianus*	T-913	F. Catzeflis	AJ430561*	-	-	-	-	Published
				N4142	D.Gauthier	-	KM397162	KM397256	KM397296	KM397343	This study
		Gerbillus	*henleyi*	T-1165	F. Catzeflis	-	KM397160	KM397254	KM397295	-	This study
	Murinae	*Apodemus*	*mystacinus*	262	J. Michaux	KM397213	-	KM397259	-	-	This study
						-	AB303229*	-	-	AB285448*	Published
		Apodemus	*sylvaticus*	102	J. Michaux	-	-	KM39725	-	-	This study
						AB033695*	AB032863*	-	-	AB164041*	Published
		Batomys	*granti*			AY324458*	DQ191496*	AY294917*	AY295002*	AY241461*	Published
		Maxomys	*surifer*	R4223	CeroPath	HM217445*	HM217682*	KM397261	-	KM397347	Published and this study
		Rattus	*tanezumi*	R4377	CeroPath	HM217452*	HM217689*	KM397260	KM397300	KM397346	Published and this study
Sciuridae	Sciurinae	*Sciurus*	*aestuans*	T-1758	F. Catzeflis	-	FM162057*	FM162078*	KM397297	KM397345	Published and this study
						AJ389530*	-	-	-	-	Published
	Xerinae	*Marmota*	*marmota*	S10_1244	D. Allainé	-	KM397161	KM397255	-	KM397342	This study
						AF100711*	-	-	-	-	Published

Journal of Biogeography

SUPPORTING INFORMATION

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

Julie Pisano, Fabien L. Condamine, Vladimir Lebedev, Anna Bannikova, Jean-Pierre Quéré, Gregory I. Shenbrot, Marie Pagès and Johan R. Michaux

Appendix S2 Supplementary information for phylogenetic and dating analyses.

Table S2 Primers and PCR cycling conditions. Five coding genes were used to reconstruct the evolutionary history of Dipodoidea. The *IRBP* and *GHR* genes were amplified into two overlapping fragments, *IRBP* – F1/F2 and *GHR* – F1/F2. On the 'designation' column, 'fw' or 'F' and 'rev' or 'R' stand for 'forward' and 'reverse', respectively.

Designation	Gene name	Primer sequence 5'→ 3'	Annealing tempera-ture	Reference
Cytb	**Cytochrome *b* (1118 bp)**			
L7-fw		ACCAATGACATGAAAAATCATCGTT	50 °C	(Irwin *et al.*, 1991)
H6-rev		TCTCCATTTCTGGTTTACAAGAC		
***IRBP* – F1**	**Interphotoreceptor retinoid-binding protein (1188 bp) IRBP fragment 1**			
I1-fw		ATGGCCAAGGTCCTCTTGGATAACTACTGCTT	60 °C	(Poux & Douzery, 2004)
J2-rev		CGCAGGTCCATGATGAGGTGCTCCGTGTCCTG		
J2-Dipo-rev		CCRCTGCCCTCCCATGTCT		This study
***IRBP* – F2**	**IRBP fragment 2**			
I2-Dipo-fw		ATCCCCTATGTCATCTCCTAYYTG	52 °C	This study
I2-fw		ATCCCCTATGTCATCTCCTACYTG		(Poux & Douzery, 2004)
J1-rev		CGCAGGTCCATGATGAGGTGCTCCGTGTCCTG		
***GHR* – F1**	**Growth hormone receptor (928 bp) GHR fragment 1**			
GHRexon 10-fw		GGRAARTTRGAGGAGGTGAACACMATCTT	58 °C	(Lecompte *et al.*, 2008)
GHR8 -rev		TTGGCATCTGACTCACAGAAGTAGG		
GHR8-Dipo-rev		TTGGCATCTGCCTCACAGAAGTAGG		This study
***GHR* – F2**	**GHR fragment 2**			
GHR7-Dipo-fw		AAGCTGATCTCTTGTGCCTTGAYCAGAA	53 °C	This study
GHR7-fw		AAGCTGATCTCTTGTGCCTTGACCAGAA		(Lecompte *et al.*, 2008)
GHR2-rev		GATTTTGTTCAGTTGGTCTGTGCTCAC		
BRCA1	**Breast cancer type 1 susceptibility gene (699 bp)**			
brca_zap 100F		TGTGGCACAGATGTTCRTGCCAGCTCATTACA	57°C (zap100F+zap980R)	(Lebedev *et al.*, 2012)
brca_All_ 982R		CTACTGGATTCACATTTTCCTCTTTCTG		
brca_zap 980R		CTACTGGATTCTCACTTTCCTCTTTCTGAA	55°C (zap100F+all_982R)	

BRCA1-Outg-fw	AGCCAACAGRGCAGATGGGC	60 °C	This study
BRCA1-Outg-rev	AAATCTRCYTTCTTGATAAAATC	60 °C	This study
RAG1	**Recombination activating gene (1056 bp)**		
rag1_all-940F	GACCTGGAGAGTCCAGTGAAGTCCTTTCT	57°C	(Lebedev *et al.*, 2012)
rag1_all-980F	TGAATTCCCTGATGGTRAAATGTCC		
rag1_dip 2117R	ACGRGTAGCATCACAAAGAGTACAAATGT		
RAG1-Outg-rev	CAAGATTTTGAGAGGCTTCC		This study

Table S3 Best substitution models and partitioning schemes used for molecular phylogenetic and dating analyses of Dipodoidea. Estimations of models and partitions were performed on the species-level dataset (i.e. outgroups and Dipodoidea represented by one single specimen per species), for which all five coding genes were concatenated. The partitioning scheme and the relative substitution model of each partition were obtained using PARTITIONFINDER v1.1.1. In the 'subset partitions' column, the code used for partitions is '*abbreviation-of-gene-names_codon-position*'. In the 'subset partitions', information refers to '*gene name_codon position*'. In the 'subset sites' column, information refers to '*1st BasePair-LastBasePair/sampledEvery3BasePair*'.

Subset	Subset partitions	Best model	Subset sites
1	BRCA1_1, BRCA1_2, BRCA1_3, GHR_3, RAG1_3	GTR + G	1-693/3, 2-693/3, 3-693/3, 1812-2736/3, 3925-4973/3
2	Cytb_1	SYM + G	694-1809/3
3	IRBP_2, RAG1_2, Cytb_2	GTR + I + G	695-1809/3, 2738-3922/3, 3924-4973/3
4	Cytb_3	GTR + G	696-1809/3
5	GHR_1, GHR_2, IRBP_1, RAG_1	GTR + G	1810-2736/3, 1811-2736/3, 2737-3922/3, 3923-4973/3
6	IRBP_3	GTR + G	2739-3922/3

Table S4 Alternative topological hypotheses tested for this study of the evolutionary history of Dipodoidea. 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 by (3) constraining the monophyly of *Allactaga*. Bayes factors on likelihood scores were calculated using TRACER v1.5. Bayes factor values > 20 were considered to be indicative of strong support for one hypothesis versus another; those > 150 indicative of very strong support. ESS, effective sample size.

	Run statistics				**Bayes factors (BF)**							
	ESS	**Likelihood**	**Harmonic mean**	**Standard error**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
A - Monophyletic "*Euchoreutes naso*–Allactaginae" - run 1	10836.59	-47067.37	-47081.552	+/- 0.239	-	-0.054	-0.529	-0.002	67.671	67.999	-941.065	-940.756
B - Monophyletic " *Euchoreutes naso*–Allactaginae" - run 2	10376.54	-47067.40	-47081.498	+/- 0.203	0.054	-	-0.476	0.052	67.725	68.052	-941.011	-940.702
C - Monophyletic " *Euchoreutes naso*–Dipodinae" - run 1	10235.01	-47067.29	-47081.023	+/- 0.218	0.529	0.476	-	0.528	68.2	68.528	-940.536	-940.226
D - Monophyletic " *Euchoreutes naso*–Dipodinae" - run 2	10099.36	-47067.41	-47081.55	+/- 0.253	0.002	-0.052	-0.528	-	67.673	68.001	-941.063	-940.754
E - Monophyletic Allactaga group - run 1	9667.51	-47135.32	-47149.223	+/- 0.217	-67.671	-67.725	-68.2	-67.673	-	0.328	-1008.736	-1008.427
F - Monophyletic Allactaga group - run 2	10578.44	-47135.36	-47149.551	+/- 0.232	-67.999	-68.052	-68.528	-68.001	-0.328	-	-1009.064	-1008.755
G - No constrain - run1	10840.23	-46126.29	-46140.487	+/- 0.247	941.065	941.011	940.536	941.063	1008.736	1009.064	-	0.309
H - No constrain - run2	11433.85	-46126.29	-46140.796	+/- 0.226	940.756	940.702	940.226	940.754	1008.427	1008.755	-0.309	-

Table S5 Rare genomic changes (RGC) highlighted in DNA sequences of Dipodoidea. 'bp' refers to 'base pairs'.

Gene	Kind of RGC	Voucher	Taxa	Sequence
BRCA1	Deletion of 3 bp	UWBM78056	*Zapus princeps*	GAGACTCCTGGGCAT---GACAGTCTTGC
		UWBM78818	*Z. trinotatus*	TGAAGTAGCTGGGAC---AGAAGTTTCAG
	Deletion of 3 bp	3049	*Paradipus ctenodactylus*	TTCCAGAAGCGGTGA---GTTAACTTTTACTGA
	Deletion of 6 bp		**Allactaginae**	
			Allactaga elater	GATGAGTTGTTAACT------GATACAAACAACATG
		T-1163	*A.* sp.	GATGAGTTGTTAACT------GATACAAACAACATG
		ZMMU 01/08_32	*A. major*	GATGAGTTGTTAACT------GATACAAACAACATG
			A. balikunica	GATGAGTTGTTAACT------GATACAAACAACATG
		ZMMU S179572	*A. bullata*	GATGAGTTGTTAACT------GATACAAACAACATG
		3041	*A. sibirica*	GATGAGTTGTTAACT------GATACAAACAACATG
		t.c. Apy	*Pygeretmus pumilio*	GATGAGTTGTTAACT------GATACAAACAACATG
		t.c. Abo	*Allactodipus bobrinskii*	GATGAGTTGTTAACT------GATACAAACAGCATG
	Deletion of 3 bp		**Zapodidae**	
		3040	*Eozapus setchuanus*	TGAAGTAGCTGGGAG---AGAAGTTTCAGATGA
		3033	*Napaeozapus insignis*	CGAAGTAGCTGGGAC---AGAAGTTTCAGATGA
		UWBM78056	*Zapus princeps*	TGAAGTAGCTGGGAC---AGAAGTTTCAGATGA
		UWBM78818	*Z. trinotatus*	TGAAGTAGCTGGGAC---AGAAGTTTCAGATGA
		3035	*Z. hudsonius*	TGAAGTAGCTGGGAC---AGAAGTTTCAGATGA
	Deletion of 3 bp		**Sminthidae**	
		T-0764	*Sicista caucasica*	GTAAATGGATATCCT---TCTTCCAAGAAAACA
		T-0763	*S. kluchorica*	GTAAATGGATATCCT---TCTTCCAAGAAAACA
		T-0761	*S. kazbegica*	GTAAATGGATATCCT---TCTTCCAAGAAAACA
		T-MNHN1999-452	*S. napaea*	GTAAATGGATATCCT---CCTTCCAAGAAAACA
		YuMK 1559	*S. subtilis*	GTAAATGGATATCCT---TCTTCCAAGAAAACA
		3045	*S. concolor*	GTAAATGGATATCCT---TCTTCCAAGAAAACA
		NT0609AQ10	*S. tianshanica*	GTAAATGGATATCCT---TCTTCCAAGAAAACA
	Deletion of 3 bp		**Sminthidae**	
		T-0764	*Sicista caucasica*	AAAACAGACTTGGTC---ACTGATCCTCATCAT
		T-0763	*S. kluchorica*	AAAACAGACTTGGTC---ACTGATCCTCATCAT
		T-0761	*S. kazbegica*	AAAACAGACTTGGTC---ACTGATCCTCATTAT
		T-MNHN1999-452	*S. napaea*	AAAACAGACTTGGTC---ACTGATCCTCAT---
		YuMK 1559	*S. subtilis*	AAAACAGACTTGGTC---ACTGATCCTCAT---
		3045	*S. concolor*	AAAACAGACTTGGTC---ACTGATCCTCATCAT
		NT0609AQ10	*S. tianshanica*	AAAACAGACTTGGTC---ACTGATCCTCATCAT
	Deletion of 3 bp	T-MNHN1999-452	*S. napaea*	---ACTGATCCTCAT---GCTTTAATATGTCAA
		YuMK 1559	*S. subtilis*	---ACTGATCCTCAT---GCTTTAATATGTCAA
	Deletion of 3 bp		**Zapodidae**	
		3040	*Eozapus setchuanus*	CGTAAAAGAAGGACT---TCATGCCTTCAGCCT
		3033	*Napaeozapus insignis*	CGCAAAAGGAGGACT---TCATGCCTTCAACCT
		UWBM78056	*Zapus princeps*	CGTAAAAGGAGGACT---TCATGCCTTCAACCT
		UWBM78818	*Z. trinotatus*	CGTAAAAGGAGGACT---TCATGCCTTCAACCT
		3035	*Z. hudsonius*	CGTAAAAGGAGGACT---TCATGCCTTCAACCT
	Deletion of 3 bp	3033 (*N. insignis*)	**All Dipodoidea**	GTCTGCTCCAAACAA---GAGAGTAATATTAAAG
GHR	Deletion of 3 bp	T-MNHN1999-452	*Sicista napaea*	ATTTCCCAGACAGTG---TGTGTGATGGTGCCT
	Insertion of 3 bp	3035	*Zapus hudsonius*	AATGTTTCCCCTGCTACTACTCAGCAGTCCAGC
	Deletion of 9 bp	7908	*Salpingotus kozlovi*	GTCTCCATGCAGTTG---------AGTTC ACTGGCCAAC
	Deletion of 6 bp	3045	*Sicista concolor*	GCCAAAAAGTGCATT------GCCCCTCACATGGAA
	Insertion of 3 bp	3033 (*N. insignis*)	**All Dipodoidea**	CCCCCATGTGACATCATGCATCCAGAAGTGATC
IRBP	Insertion of 3 bp	7908	*Salpingotus kozlovi*	TATGAGCCCAGCACCCTCCTCGAGGCTTCCCAG
	Deletion of 9 bp	3033	*Napaeozapus insignis*	ACCCACAAGGAGCTG---------TTGCAGCAGAACATC
		UWBM78056	*Zapus princeps*	ACCCACAAGGAGCTG---------TTGCAGCAGAACATC
		UWBM78818	*Z. trinotatus*	AYCCACAAGGAGCTG---------TTACAGCAGAACATC
		3035	*Z. hudsonius*	ACCCACAAGGAGCTG---------TTGCAGCAGAACATC
	Insertion of 3 bp		**All Dipodinae except *P. ctenodactylus***	
		San-Diego, t.c. 2006	*Jaculus blanfordi*	GGGCCCAGGGAAACATCGTCCCCGGGGCCTGAG
		JJ257	*J. jaculus*	GGGCCCAGGGAAACATCGTCCCCGGGGCCTGAG
		JO271	*J. orientalis*	GGGCCCAGGGAAACGGCGTCCCCGGGGCCTGAG
		t.c. Eli	*Eremodipus lichtensteini*	GGGCCCAGAGAAGCATCGTCCCCGGGGCCTGAG
		3039	*Dipus sagitta*	GGGCCCAGAGAAACGTCGTCCCCGGGGCCTGAG
			Stylodipus telum	GGGCCCAGAGAAACCTCGTCCCCGGGGCCTGAG
		94	*S. sungorus*	GGGCCCAGAGAAACCTCATCCCCGGGGCCTGAG
		53	*S. andrewsi*	GGGCCCAGAGAAGCCTCGTCCCCGGGGCCTGAG

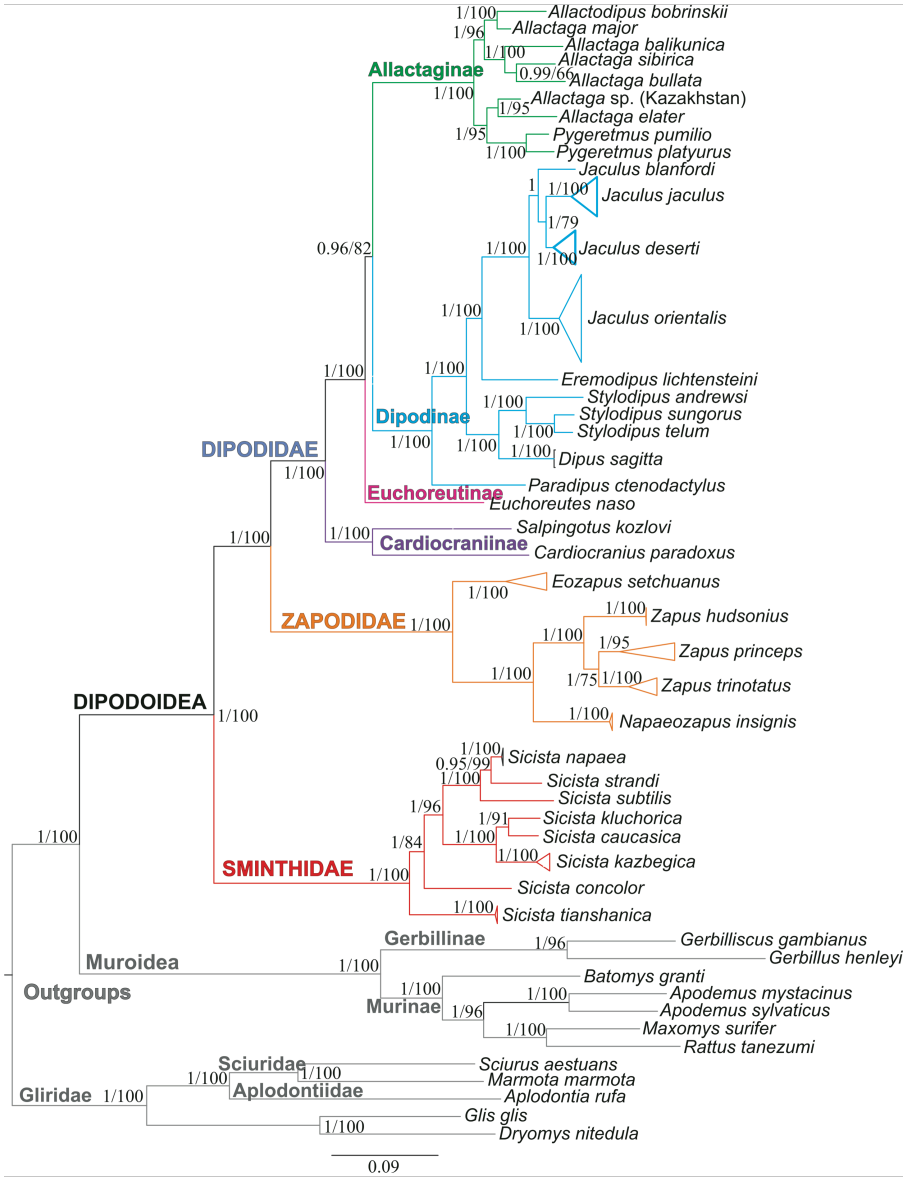
Table S6 Summary results for phylogeny, estimates of divergence times, and biogeography of Dipodoidea. Outgroups and Dipodoidea are depicted in grey and white, respectively. These analyses were performed on the species-level dataset (i.e. outgroups and Dipodoidea represented by one single specimen per species), for which all five coding genes were concatenated. The node numbers refer to those in Fig. 2 in the main text. For each node, we indicated (i) the phylogenetic support: bootstrap (BP) and posterior probabilities (PP) values, (ii) the estimated median ages with the 95% highest posterior density of the maximum clade credibility tree discussed in this study, and (iii) the ancestral range reconstruction, the likelihood score (ln *L*) and the relative probability (RP) at each node obtained using *Lagrange*. For geographical analysis, only the most likely scenario of nodes is given. '*NA*' stands for 'not available' because, for geographical range evolution analyses, outgroups were removed.

Nodes	Phylogeny	Chronogram		**Biogeographical reconstructions**								
				M0			M1			M2		
	PP/BP	Estimated Age	95% HPD	Range(s)	Ln *L*	RP	Range(s)	Ln *L*	RP	Range (s)	Ln *L*	RP
1	1	63.97	56.64-76.01	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	1	42.82	37.92-50.89	NA	NA	NA	NA	NA	NA	NA	NA	NA
3	1	37.39	33.11-44.43	NA	NA	NA	NA	NA	NA	NA	NA	NA
4	1	26.08	23.09-30.99	NA	NA	NA	NA	NA	NA	NA	NA	NA
5	1	19.66	17.41-23.37	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	1	57.72	51.11-68.59	NA	NA	NA	NA	NA	NA	NA	NA	NA
7	1	27.59	24.42-32.78	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	1	18.1	16.02-21.51	NA	NA	NA	NA	NA	NA	NA	NA	NA
9	1	12.78	11.31-15.18	NA	NA	NA	NA	NA	NA	NA	NA	NA
10	1	11.39	10.09-13.54	NA	NA	NA	NA	NA	NA	NA	NA	NA
11	1	7.22	6.4-8.58	NA	NA	NA	NA	NA	NA	NA	NA	NA
12	1	7.27	6.44-8.64	NA	NA	NA	NA	NA	NA	NA	NA	NA
13	**1**	40.62	35.97-48.27	**DG**	**113.7**	**0.40**	**DG**	**-108.6**	**0.48**	**DG**	**-108.1**	**0.47**
14	1	17.66	15.63-20.98	DG	114	0.30	DG	-109.7	0.34	DG	-108.4	0.36
15	1	16.31	14.44-19.38	BDG	-113.8	0.38	DG	-109.5	0.40	DG	-108.4	0.35
16	1	14.05	12.44-16.69	BD	-113.5	0.52	BD	-109.1	0.61	BD	-107.9	0.61
17	1	7.74	6.85-9.19	BD	-113.4	0.58	BD	-109.1	0.62	BD	-107.9	0.61
18	0.95	4.55	4.03-5.41	B	-113.3	0.63	B	-109	0.66	B	-107.8	0.66
19	1	6.83	6.05-8.12	B	-112.8	0.99	B	-108.6	0.98	B	107.4	0.98
20	1	4.57	4.05-5.43	B	-112.8	1	B	-108.6	1	B	-107.4	1
21	1	34.52	30.56-41.02	D	-113.8	0.37	D	-109.2	0.56	D	-108.1	0.49
22	1	20.24	17.92-24.06	D	-114.4	0.21	D	-110.2	0.19	D	-108.9	0.22
23	1	13.01	11.52-15.46	A	-113	0.80	A	-108.9	0.75	A	-107.7	0.74
24	1	8.31	7.36-9.88	A	-112.8	0.98	A	-108.6	0.97	A	-107.4	0.97
25	1	6.6	5.84-7.84	A	-112.8	1	A	-108.6	0.99	A	-107.4	0.99
26	1	26.82	23.74-31.87	D	-114	0.31	D	-109.8	0.32	D	-108.4	0.36
27	1	18.72	16.57-22.24	H	-113.6	0.47	H	-109.6	0.38	D	-108.5	0.32
28	1	22.21	19.67-26.39	DH	-113.8	0.37	DH	-109.6	0.36	D	-108.4	0.35
29	0.95	21.8	19.3-25.9	D	-113.4	0.56	D	-109	0.66	D	-107.7	0.70
30	1	8.75	7.75-10.4	D	-113.8	0.39	D	-109.6	0.38	D	-108.3	0.38
31	1	7.83	6.93-9.31	DH	-113.9	0.33	DH	-109.5	0.40	DH	-108.2	0.42
32	1	7.01	6.21-8.34	H	-113.3	0.60	H	-109.1	0.62	H	-107.8	0.65
33	1	5.41	4.79-6.43	H	-113.5	0.49	H	-109.3	0.51	H	-108	0.52
34	1	3.68	3.26-4.38	D	-113.4	0.53	D	-109.2	0.54	D	-108	0.53
35	1	6.76	5.98-8.03	D	-112.9	0.91	D	-108.7	0.90	D	-107.5	0.89
36	1	2.96	2.62-3.52	D	-113.3	0.62	D	-109.1	0.59	D	-107.9	0.58
37	0.99	5.01	4.44-5.95	D	-112.9	0.94	D	-108.7	0.94	D	-107.4	0.93
38	1	16.11	14.26-19.14	D	-113	0.82	D	-108.8	0.84	D	-107.5	0.83
39	1	12.65	11.2-15.03	D	-113.2	0.69	D	-109	0.68	D	-107.8	0.67
40	1	8.66	7.67-10.29	D	-114.2	0.25	D	-110.1	0.24	D	-108.8	0.23
41	1	6.27	5.55-7.45	D	-113.8	0.37	D	-109.6	0.36	D	-108.5	0.33
42	1	2.39	2.12-2.84	E	-113.4	0.54	E	-109.3	0.49	E	-108	0.51
43	1	11.09	9.82-13.18	D	-113	0.83	D	-108.8	0.80	D	-107.6	0.80
44	1	5.97	5.29-7.09	D	-112.9	0.90	D	-108.7	0.89	D	-107.5	0.88
45	1	5.15	4.56-6.12	DFI	-112.9	0.89	DFI	-108.7	0.88	DFI	-107.5	0.86

Table S7 Selection of the most likely dipodoid dating analysis using Bayes factors. Bayes factors (BFs) on likelihood scores were calculated using TRACER v1.5. Two models have been tested: the birth–death model (BD) and the Yule model of speciation (Yule). The cross-validating process involved the removal one by one of each fossil calibration used in this study. BFs showed that no analysis was better than another; all BFs were < 10 suggesting no strong support for one hypothesis versus another. Letters from A to J refer to distinct dating inference. FC(s), fossil calibration(s); ESS, effective sample size; BD, birth–death model.

	Run statistics				**Bayes factors (BF)**									
	ESS	**Likelihood**	**Harmonic mean**	**Standard error**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
A – BD - All FCs	1582.26	-43112.01	-43161.39	+/- 0.845	-	-3.058	-3.769	0.61	1.541	-1.821	0.589	1.611	-0.581	-3.373
B – BD - No *Apodemus jeanteti/A. dominans* FC	1733.55	-43110.79	-43158.332	+/- 0.803	3.058	-	-0.711	3.669	4.6	1.238	3.647	4.67	2.477	-0.315
C – BD - No *Douglasciurus jeffersoni* FC	1525.02	-43109.87	-43157.621	+/- 0.982	3.769	0.711	-	4.38	5.311	1.949	4.358	5.381	3.189	0.396
D – BD - No *Progonomys* FC	824.93	-43113.0.33	-43162	+/- 0.927	-0.61	-3.669	-4.38	-	0.931	-2.431	-0.021	1.001	-1.191	-3.984
E – BD – No *Sicista primus* FC	1417.61	-43114.78	-43162.931	+/- 0.958	-1.541	-4.6	-5.311	-0.931	-	-3.362	-0.952	0.07	-2.122	-4.915
F – Yule - All FCs	1114.86	-43111.02	-43159.569	+/- 0.935	1.821	-1.238	-1.949	2.431	3.362	-	2.41	3.432	1.24	-1.553
G – Yule - No *Apodemus jeanteti/A. dominans* FC	1116.13	-43112.42	-43161.979	+/- 0.973	-0.589	-3.647	-4.358	0.021	0.952	-2.41	-	1.022	-1.17	-3.962
H – Yule - No *Douglasciurus jeffersoni* FC	1428.66	-43112.03	-43163.001	+/- 0.91	-1.611	-4.67	-5.381	-1.001	-0.07	-3.432	-1.022	-	-2.192	-4.985
I – Yule –No *Progonomys* FC	1180.77	-43112.81	-43160.809	+/- 0.819	0.581	-2.477	-3.189	1.191	2.122	-1.24	1.17	2.192	-	-2.793
J – Yule – No *Sicista primus* FC	1548.79	-43110.67	-43158.017	+/- 0.943	3.373	0.315	-0.396	3.984	4.915	1.553	3.962	4.985	2.793	-

Figure S1 Phylogenetic relationships among 35 species of Dipodoidea obtained using the densely sampled matrix. 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 3.2.2 and RAXML: posterior probability (PP)/bootstrap (BP) values. Colours refer to clades and pertain to Fig. 1. Samples of *Jaculus* from North Africa (N4284, 399, 406, GAN204 and JJ257) are suspected to correspond to two cryptic species, *J. jaculus* and *J. deserti* (Ben Faleh *et al.*, 2012; Boratynski *et al.*, 2012). Two lineages were effectively recovered in each independent gene tree, further arguing for the recognition of these two cryptic species. The cryptic species, the North African *Jaculus deserti* was not included in biogeographical analyses. However, given that *J. deserti* is distributed in the same area as its sister species, *J. jaculus*, the ancestral area reconstruction of Dipodinae would not be affected by the addition of this species.



Journal of Biogeography

SUPPORTING INFORMATION

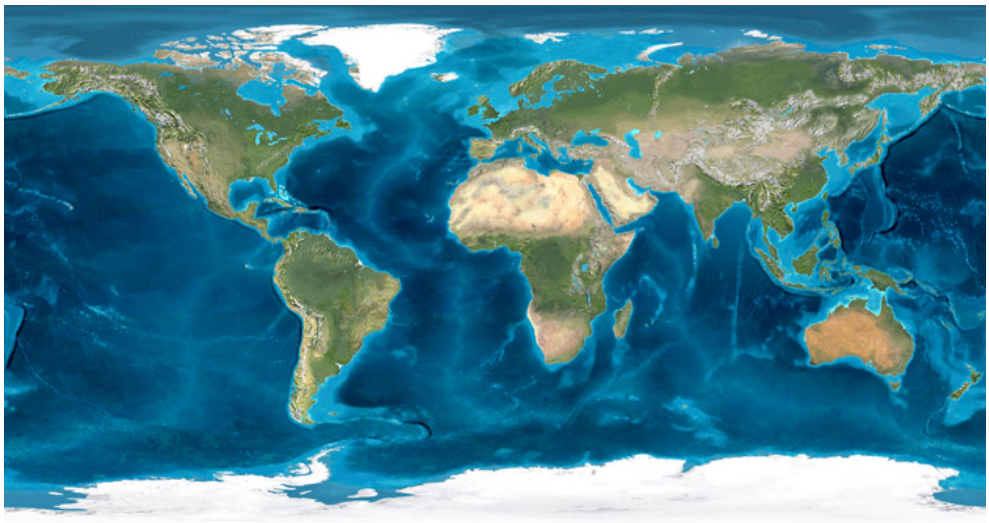
Out of Himalaya: the impact of past Asian environmental changes on the evolutionary and biogeographical history of Dipodoidea (Rodentia)

Julie Pisano, Fabien L. Condamine, Vladimir Lebedev, Anna Bannikova, Jean-Pierre Quéré, Gregory I. Shenbrot, Marie Pagès and Johan R. Michaux

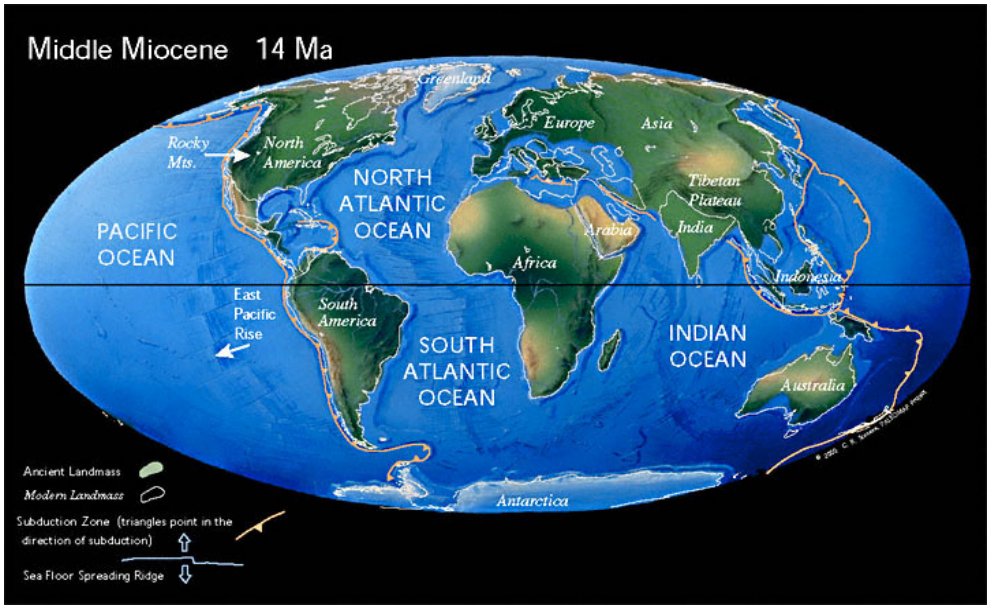
Appendix S3 Supplementary information for biogeographical analyses.

Figure S2 Paleogeographic maps reconstructed from Scotese (2001, 2004) and Blakey (2008).

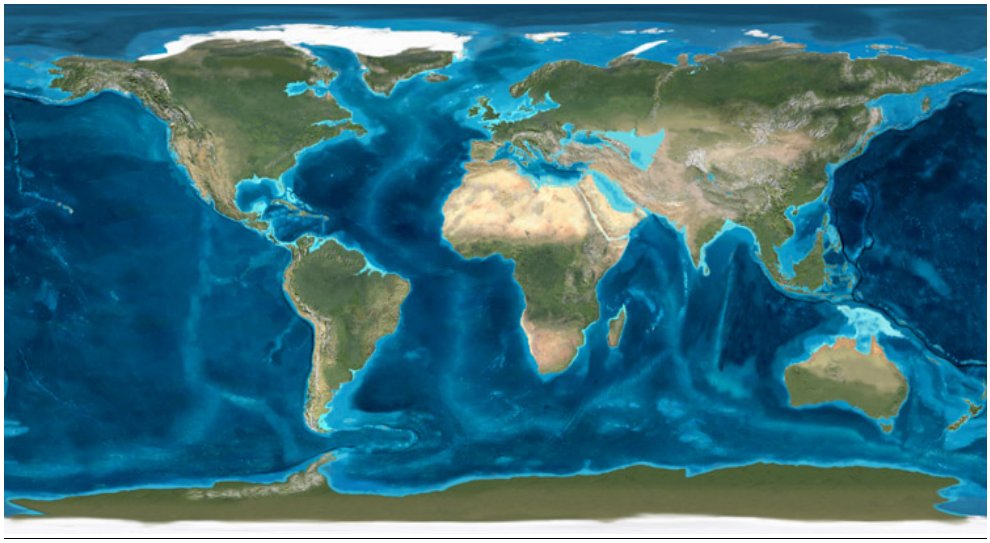
- **Pleistocene:**



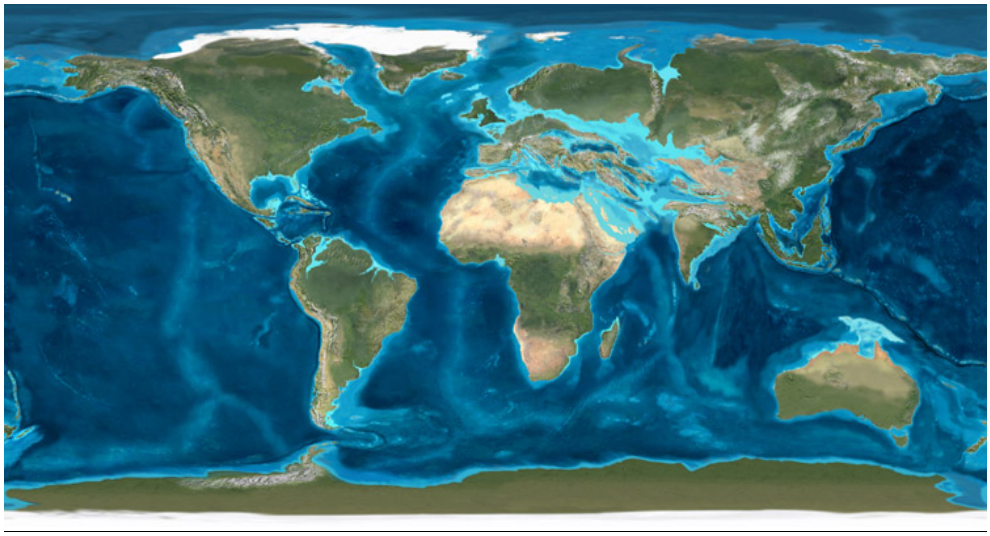
- **14 Ma:**



- **20 Ma:**



- **35 Ma:**



- **50 Ma:**

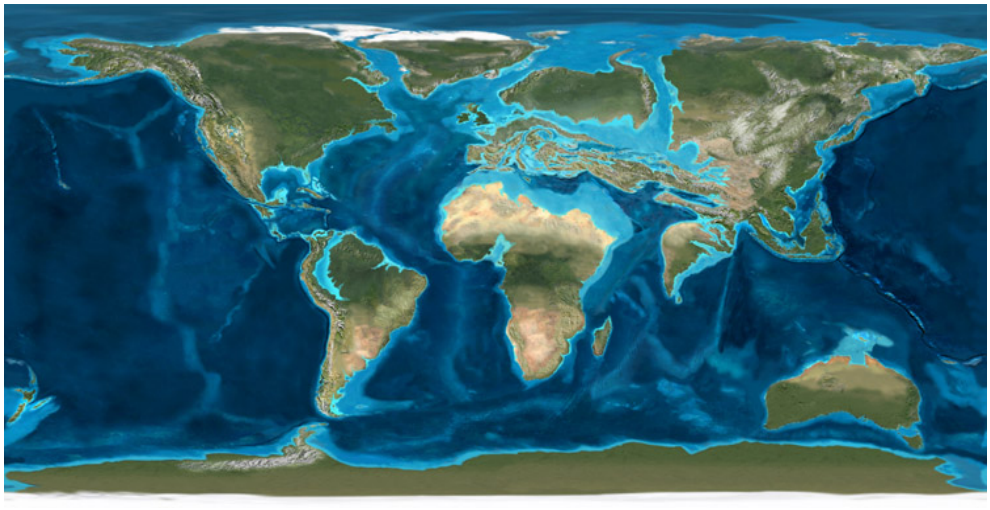
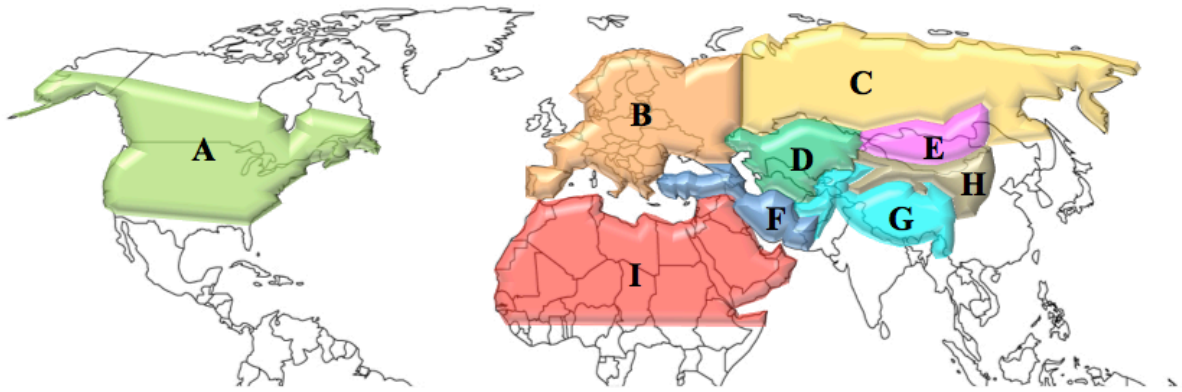


Figure S3 Geographical model.

- Nine-areas model:



The 'nine-areas model' comprises nine units: A, Nearctic (North America); B, West Palaearctic (from Western Europe to Ural Mounts, without North Africa); C, Siberia (from Ural Mounts 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).

Table S8 Species ranges.
Information was compiled from *Mammal Species of the World* (Holden & Musser, 2005) and the IUCN website (IUCN, 2012). As this information was not accurate and has to be updated, distribution ranges were refined following G. I. Shenbrot's comments.

- **Nine-areas model:**

Classification proposed in Mammal Species of the World	A Nea-rctic	B West Palearctic	C Siberia	D Central Asia*	E Altai Mt, Mongolian steppe, Yablonoi Mt	F Persian plateau, Anatolian region and caucasus, iranian plateau	G Himalaya + Tibetan plateau	H Gobi desert, Talkimakan desert	I North Africa + Arabia
Jaculus blanfordi	0	0	0	1	0	1	0	0	0
Jaculus jaculus	0	0	0	0	0	0	0	0	1
Jaculus orientalis	0	0	0	0	0	0	0	0	1
Eremodipus lichtensteini	0	0	0	1	0	0	0	0	0
Dipus sagitta	0	0	0	1	1	1	0	1	0
Stylodipus telum	0	1	0	1	0	0	0	0	0
Stylodipus sungorus	0	0	0	0	1	0	0	0	0
Stylodipus andrewsi	0	0	0	0	1	0	0	1	0
Paradipus ctenodactylus	0	0	0	1	0	0	0	0	0
Pygeretmus pumilio	0	0	0	1	1	0	0	1	0
Pygeretmus platyurus	0	0	0	1	0	0	0	0	0
Allactaga elater	0	0	0	1	0	1	0	0	0
Allactaga sp. (Kazakhstan)	0	0	0	1	0	0	0	0	0
Allactodipus bobrinskii	0	0	0	1	0	0	0	0	0
Allactaga major	0	1	1	1	0	0	0	0	0
Allactaga balikunica	0	0	0	0	0	0	0	1	0
Allactaga bullata	0	0	0	0	1	0	0	1	0
Allactaga sibirica	0	0	0	1	1	0	0	1	0
Euchoreutes naso	0	0	0	0	0	0	0	1	0
Cardiocranius paradoxus	0	0	0	1	1	0	0	1	0
Salpingotus kozlovi	0	0	0	0	0	0	0	1	0
Zapus princeps	1	0	0	0	0	0	0	0	0
Zapus trinotatus	1	0	0	0	0	0	0	0	0
Zapus hudsonius	1	0	0	0	0	0	0	0	0
Napaeozapus insignis	1	0	0	0	0	0	0	0	0
Eozapus setchuanus	0	0	0	0	0	0	1	0	0
Sicista kazbegica	0	1	0	0	0	0	0	0	0
Sicista caucasica	0	1	0	0	0	0	0	0	0
Sicista kluchorica	0	1	0	0	0	0	0	0	0
Sicista subtilis	0	0	0	1	0	0	0	0	0
Sicista napaea	0	1	1	0	0	0	0	0	0
Sicista strandi	0	1	0	0	0	0	0	0	0
Sicista concolor	0	0	0	0	0	0	1	0	0
Sicista tianshanica	0	0	0	0	0	0	1	0	0
Distributions of dipodoid species not included in this study									
Sicista armenica	0	0	0	0	0	1	0	0	0
Sicista betulina	0	1	1	0	0	0	0	0	0
Sicista caudata	0	0	1	0	0	0	0	0	0
Sicista pseudonapaea	0	0	0	0	1	0	0	0	0
Sicista severtzovi	0	1	0	0	0	0	0	0	0
Salpingotus heptneri	0	0	0	1	0	0	0	0	0
Salpingotus crassicauda	0	0	0	0	1	0	0	1	0
Salpingotus thomasi	0	0	0	0	0	1	0	0	0
Salpingotus pallidus	0	0	0	1	0	0	0	0	0
Salpingotulus michaelis	0	0	0	0	0	1	0	0	0
Pygeretmus zhitkovi	0	0	0	1	0	0	0	0	0
Allactaga euphractica	0	0	0	0	0	0	0	0	1
Allactaga hotsoni	0	0	0	0	0	1	0	0	0
Allactaga williamsi	0	0	0	0	0	1	0	0	0
Allactaga tetradactyla	0	0	0	0	0	0	0	0	1
Allactaga vinogradovi	0	0	0	1	0	0	0	0	0
Allactaga firouzi	0	0	0	0	0	1	0	0	0

*Turkmenistan, Uzbekistan, Kazakhstan

- **Taxonomic revision of Dipodoidea:**

i. Lebedev *et al.* (2012) proposed a new taxonomy of Dipodoidea, which suggests division of the superfamily into three families:

Dipodoidea
Sminthidae Brandt, 1855
Sicista Gray, 1827
Zapodidae Coues, 1875
Zapodini
Zapus Coues, 1875
Napaeozapus Preble, 1899
Eozapodini Zazhigin and Lopatin, 2000
Eozapus Preble, 1899
Dipodidae Fischer de Waldheim, 1817
Cardiocraniinae Vinogradov, 1925
Cardiocraniini
Cardiocranius Satunin, 1903
Salpingotini Vinogradov, 1925
Salpingotus Vinogradov, 1922 (with subgenera *Salpingotus* s.str., *Salpingotulus* Pavlinov, 1980, *Prosalpingotus* Vorontsov and Shenbrot, 1984; *Anguistodontus* Vorontsov and Shenbrot, 1984)
Euchoreutinae Lyon, 1901
Euchoreutes Sclater, 1890
Dipodinae
Dipodini
Dipus Zimmermann, 1780
Stylodipus G. M. Allen, 1925
Eremodipus Vinogradov, 1930
Jaculus Erxleben, 1777 (with subgenera *Jaculus* s. str. and *Haltomys* Brandt, 1844)
Paradipodini Pavlinov and Schenbrot, 1983
Paradipus Vinogradov, 1930
Allactaginae Vinogradov, 1925
Allactaga F. Cuvier, 1836 (with *A. major* and *A. severtzovi*)
Orientallactaga Shenbrot, 1984 (with *O. sibirica*, *O. bullata* and *O. balikunica*)
? *Scarturus* Gloger, 1841 (with *S. tetradactylus*)
? *Paralactaga* Young, 1927 (with *P. euphratica*, *P. williamsi*, ?*P. hotsoni*)
? *Microallactaga* Shenbrot, 1974 (with *M. elater* and *M. vinogradovi*)
Allactodipus Kolesnikov, 1937
Pygeretmus Gloger, 1841 (with subgenera *Pygeretmus* s.str. and *Alactagulus* Nehring, 1897)

ii. *Allactaga firouzi* and *A. hotsoni* are proposed as identical and conspecific (Shenbrot, 2009; Dianat *et al.*, 2013).

iii. Occurrence of two cryptic species in the North Africa: *Jaculus jaculus* and *J. deserti* (Ben Faleh *et al.*, 2012; Boratynski *et al.*, 2012).

Table S9 Matrices of dispersal rates.
For the null hypothesis (model M0), all dispersal rates were set to 1, which implies no barrier between distinct areas. Blue boxes with an "A" inside correspond to "adjacent areas".

- **M1 – Nine areas and five time slices:**

Model 1: Nine areas and five time Slices (TS)

Adjacent areas

	A	B	C	D	E	F	G	H
A	-							
B		-						
C	A	A	-					
D		A	A	-				
E			A	A	-			
F		A		A		-		
G				A		A	-	
H				A	A		A	-
I						A		

Dispersal rates: *DR=1: adjacent areas; DR=0.7: small barrier; DR=0.5: larger barrier; DR=0.3: dispersal between two areas separated by one area; DR=0.1: LDD; DR=0: dispersal inhibited.*

Time Slice 1: Present to 5,3 Ma (early Pliocene)

	A	B	C	D	E	F	G	H
A	-							
B	0.1	-						
C	0.3	1	-					
D	0.1	0.7	1	-				
E	0.1	0.3	0.7	0.7	-			
F	0	0.7	0.3	0.7	0.3	-		
G	0	0.3	0.3	0.5	0.3	0.5	-	
H	0.1	0.3	0.3	1	1	0.3	0.5	-
I	0	0.7	0.1	0.3	0.1	1	0.1	0

Time Slice 2: 5.3 to 16 Ma (middle Miocene)

	A	B	C	D	E	F	G	H
A	-							
B	0.3	-						
C	0.5	1	-					
D	0.3	0.7	1	-				
E	0.3	0.3	0.7	0.7	-			
F	0	0.5	0.3	0.7	0.3	-		
G	0	0.3	0.3	0.5	0.3	0.5	-	
H	0.1	0.3	0.3	1	1	0.3	0.5	-
I	0	0.5	0.1	0.3	0.1	0.7	0.1	0

Time Slice 3: 16 to 23 Ma (early Miocene)

	A	B	C	D	E	F	G	H
A	-							
B	0.1	-						
C	0.3	0.7	-					
D	0.1	0.5	0.7	-				
E	0.1	0.3	0.7	0.7	-			
F	0	0.3	0.3	0.5	0.3	-		
G	0	0.3	0.3	0.7	0.3	0.7	-	
H	0.1	0.3	0.3	0.7	1	0.3	0.7	-
I	0	0.5	0	0.3	0	0.7	0.3	0

Time Slice 4: 23 to 34 Ma (Oligocene)

	A	B	C	D	E	F	G	H
A	-							
B	0.3	-						
C	0.5	0.5	-					
D	0.1	0.3	0.5	-				
E	0.3	0.1	1	0.5	-			
F	0	0.3	0.1	0.3	0.1	-		
G	0	0.1	0.1	0.5	0.3	0.3	-	
H	0.1	0.1	0.3	0.5	1	0.1	0.7	-
I	0	0.1	0	0.1	0	0.3	0.1	0

Time Slice 5: 34 to 56 Ma (Eocene)

	A	B	C	D	E	F	G	H
A	-							
B	0.5	-						
C	0.5	0.1	-					
D	0.1	0.1	0.3	-				
E	0.3	0.1	1	0.3	-			
F	0	0.1	0.1	0.3	0.1	-		
G	0	0.1	0.1	0.3	0.1	0.3	-	
H	0.1	0.1	0.3	0.1	1	0.1	0.7	-
I	0	0.1	0	0	0	0.1	0	0

- **M2 – Nine areas and five time slices:**

A = adjacent areas without any significant barrier	**0.5**	**SB = Small Barrier between adjacent areas**	**0.25**	**LB = Larger Barrier between adjacent areas**	**0.125**

Model 2: Nine areas and five time slices (TS)

Adjacent areas

	A	B	C	D	E	F	G	H
A	-							
B		-						
C	A	A	-					
D		A	A	-				
E			A	A	-			
F		A		A		-		
G				A		A	-	
H				A	A		A	-
I		A				A		

Dispersal rates: *DR=0.5: adjacent areas; DR=0.25: small barrier; DR=0.125: larger barrier; DR=0.5*0.5 or 0.5*0.25 or 0.5*0.125: dispersal between two areas separated by one area; DR=0.01: LDD; DR=0: dispersal inhibited.*

Time Slice 1: Present to 5.3 Ma (early Pliocene)

	A	B	C	D	E	F	G	H
A	-							
B	**0.015625**	-						
C	**0.125**	**0.5**	-					
D	0.015625	**0.25**	**0.5**	-				
E	0.015625	0.125	**0.25**	**0.25**	-			
F	**0**	**0.125**	0.125	**0.25**	0.0625	-		
G	0	0.03125	0.0625	**0.125**	0.0625	**0.125**	-	
H	0.015625	0.125	0.125	**0.5**	**0.5**	0.015625	**0.125**	-
I	0	**0.25**	0	0.125	0.03125	**0.5**	0.0625	0

Time Slice 2: 5.3 to 16 Ma (middle Miocene)

	A	B	C	D	E	F	G	H
A	-							
B	0.03125	-						
C	**0.25**	**0.5**	-					
D	0.03125	**0.25**	**0.5**	-				
E	0.03125	0.125	**0.25**	**0.25**	-			
F	0	**0.125**	0.125	**0.25**	0.0625	-		
G	0	0.03125	0.0625	**0.125**	0.0625	**0.125**	-	
H	0.03125	0.125	0.125	**0.5**	**0.5**	0.015625	**0.125**	-
I	0	**0.25**	0	0.0625	0.015625	**0.25**	0.03125	0

Time Slice 3: 16 to 23 Ma (early Miocene)

	A	B	C	D	E	F	G	H
A	-							
B	**0.015625**	-						
C	**0.125**	**0.25**	-					
D	0.015625	**0.125**	**0.25**	-				
E	0.015625	0.0625	**0.25**	**0.25**	-			
F	**0**	**0.125**	0.03125	**0.125**	0.03125	-		
G	0	0.03125	0.0625	**0.25**	0.125	**0.25**	-	
H	0.015625	0.03125	0.125	**0.25**	**0.5**	0.0625	**0.25**	-
I	0	**0.125**	0	0.03125	0	**0.25**	0.0625	0

Time Slice 4: 23 to 34 Ma (Oligocene)

	A	B	C	D	E	F	G	H
A	-							
B	**0.03125**	-						
C	**0.25**	**0.125**	-					
D	0.03125	**0.125**	**0.125**	-				
E	0.03125	0.0625	**0.5**	**0.125**	-			
F	**0**	**0.125**	0.015625	**0.125**	0.015625	-		
G	0	0.015625	0.015625	**0.125**	0.25	**0.125**	-	
H	0.0625	0.015625	0.25	**0.125**	**0.5**	0.0625	**0.5**	-
I	0	**0.125**	0	0.015625	0	**0.125**	0.015625	0

Time Slice 5: 34 to 56 Ma (Eocene)

	A	B	C	D	E	F	G	H
A	-							
B	**0.25**	-						
C	**0.25**	0.03125	-					
D	0.03125	0.03125	**0.125**	-				
E	0.0625	0.03125	**0.5**	**0.125**	-			
F	**0**	0.0625	0.015625	**0.125**	0.015625	-		
G	0	0.015625	0.015625	**0.125**	0.03125	**0.125**	-	
H	0.03125	0.015625	0.25	0.0625	**0.5**	0.0625	**0.25**	-
I	0	0.015625	0	0	0	0.015625	0	0

Chapitre 4 :

Estimation de la différentiation nucléaire entre lignées mitochondriales et impact sur l'identification d'une zone de contact secondaire :

Cas des populations de campagnol roussâtre (*Myodes glareolus*) de Finlande

Mots-clés: Génétique des populations, discordance mito-nucléaire, clines de fréquences alléliques, microsatellites, *Myodes glareolus*

L’ hybridation interspécifique est un moteur de biodiversité considérable qui concerne ~10% des espèces de mammifères (Mallet, 2005). Elle peut conduire à l’extinction d’une espèce par assimilation ou à l’émergence d’une nouvelle espèce génétiquement intermédiaire entre les deux espèces parentales (Rhymer & Simberloff, 1996; Brelsford *et al.*, 2011). Une des conséquences les plus courantes de ces évènements d’hybridation entre espèces est l’introgression mitochondriale (Dowling *et al.*, 2008; Zink & Barrowclough, 2008; Melo-Ferreira *et al.*, 2009; Toews & Brelsford, 2012; Pages *et al.*, 2013). Dans certains cas, notamment lorsque les flux géniques entre les populations de l’espèce introgressée sont faibles, l’introgression du génome mitochondrial sera totale alors que le génome nucléaire n’aura pas, ou peu, été introgressé (Melo-Ferreira *et al.*, 2009; Hedrick, 2010; Boratynski *et al.*, 2011; Boratynski *et al.*, 2014). Des discordances mito-nucléaires seront alors observées au sein des populations introgressées.

La plupart des études phylogéographiques se sont quasi exclusivement basées sur l’étude du génome mitochondrial et ont eu pour objectif d’identifier et d’expliquer les patrons liés aux distributions spatiales des lignées mitochondriales, mettant ainsi en évidence des zones de contact entre des lignées mitochondriales (e.g. Saetre *et al.* (2008); Melo-Ferreira *et al.* (2009)). Par la suite, certaines études ont tenté de mieux comprendre les patrons de la variabilité génétique nucléaire entre ces lignées mitochondriales en utilisant des approches de phylogénie moléculaire (e.g. Boratynski *et al.* (2011); Boratynski *et al.* (2014)). Or, ces études ont utilisé des marqueurs (séquences de gènes nucléaires) dont les taux d’évolution sont trop lents pour pouvoir estimer le degré de différenciation entre des populations d’une même espèce. Une des solutions pour tenter de mieux comprendre les patrons de distribution spatiaux et de diversité génétique nucléaire au niveau populationnel est d’utiliser des approches de génétique des populations et de ‘clustering’ sur la base de marqueurs (e.g. des microsatellites) dont les taux d’évolution sont donc beaucoup plus rapides que ceux des séquences de gène nucléaires. Pour mettre en évidence des patrons de différenciation entre populations d’une même espèce, les approches de cline (i.e. des gradients spatiaux de fréquences alléliques) sont particulièrement intéressantes. Elles permettent d’identifier si la dispersion et par conséquent, le flux de gènes sont limités (présence d’un cline, populations différenciées) ou pas (fréquences alléliques homogènes, populations non-différenciées) entre les populations situées de part et d’autre d’une zone de contact.

Résumé de l'article

Nos connaissances de l'histoire évolutive des populations ont été considérablement améliorées grâce à l'arrivée des données moléculaires qui ont permis de mettre en évidence la distribution de la diversité génétique au sein et entre les populations. La plupart des études de phylogéographie reposent exclusivement sur l'étude du génome mitochondrial, lequel a longtemps été considéré comme le marqueur de choix pour estimer l'histoire évolutive des populations. Cependant, étant donné qu'il ne permet pas toujours de résoudre à lui tout seul l'histoire évolutive des organismes, les études récentes ont également inclus d'autres types de marqueurs tels que les microsatellites. Ces derniers ont l'avantage d'évoluer 100 à 1.000 fois plus rapidement que les séquences de gène et de ce fait, constituent des outils puissants pour étudier l'histoire évolutive récente des organismes. Dans les dernières décennies, de nombreux cas de signaux conflictuels entre le génome mitochondrial et nucléaire ont été mis en évidence. Ces discordances mito-nucléaires résultent majoritairement d'un remplacement du génome mitochondrial originel d'un taxon par le génome mitochondrial d'une autre espèce. Ces évènements peuvent résulter de divers processus évolutifs neutres (e.g. la dispersion sexe-spécifique, l'expansion démographique postérieur à de l'introgression en petites populations, les croisements asymétriques) ou encore de processus adaptatifs (e.g. l'introgression adaptative du génome mitochondrial permettant de s'adapter à de nouvelles conditions thermiques).

Cette question de discordance mito-nucléaire est particulièrement intéressante en ce qui concerne le campagnol roussâtre, *Myodes glareolus*. Ce rongeur Paléarctique présente plusieurs lignées mitochondriales dont une particulière qui est caractérisée par le génome mitochondrial d'une espèce voisine, le campagnol de la Taïga (*Myodes rutilus*). Cette lignée introgressée (mitotype 'RUT') s'étend du centre de la Suède, à travers le Nord de la Finlande jusqu'au centre de la Russie. Récemment, la phylogéographie de six populations finlandaises de campagnol roussâtre a été étudiée à partir d'un gène mitochondrial (le cytb) et de 6 gènes nucléaires. Ces études ont confirmé la discordance mito-nucléaire existant au sein des campagnols roussâtres en décrivant la présence de deux mitotypes en Finlande (le mitotype introgressé RUT au Nord et un mitotype non-introgressé 'GLA' au Sud) au sein desquels aucune différenciation nucléaire n'a été observée. Il a été précédemment proposé que ces deux mitotypes auraient survécu au dernier maximum glaciaire dans deux régions séparées.

Elles auraient ensuite recolonisé la Finlande par deux voies séparées. Toutefois, si cette hypothèse de contact secondaire était correcte, une différenciation nucléaire – même faible – entre les mitotypes GLA et RUT aurait du être observée mais il se peut que l'absence de différenciation entre les deux mitotypes n'aient pas pu être observé à cause du choix des marqueurs utilisés (i.e. les gènes nucléaires). Sur la base d'autres marqueurs moléculaires, il convenait donc de tester cette hypothèse de « deux évènements de recolonisation » (H1) contre une hypothèse alternative où il n'y aurait eu qu' « un seul évènement de recolonisation de la Finlande » (H2). Sous l'hypothèse H2, la zone de contact mitochondrial entre GLA et RUT résulterait d'évènements d'hybridation entre le campagnol de la Taïga (*M. rutilus*) et une population génétiquement homogène de *M. glareolus* qui aurait recolonisé la Finlande lors d'un seul événement de recolonisation postglaciaire. Cette hypothèse suppose que les fréquences alléliques entre les populations finlandaises sont homogènes.

Notre échantillonnage comprenait 442 spécimens de campagnol roussâtre de 22 populations réparties sur toute la Finlande (avec un accent particulier sur la zone de contact) et 39 spécimens de campagnol de la Taïga pour s'assurer de l'identification spécifique des individus. Nous avons séquencé une portion du gène du cytochrome *b* (cytb) pour identifier l'appartenance au mitotype GLA ou RUT et génotypé 17 marqueurs microsatellites pour estimer la diversité génétique nucléaire sur toute la Finlande. La diversité génétique (A, H_o, H_e, F_{is}) de chacune des populations a été estimée à partir d'analyses classiques de génétique des populations (FSTAT, GENETIX, GENEPOP). Leur structure génétique a été estimée en utilisant les outils classiques de « clustering » tels que l'ACP, la DAPC, STRUCTURE et GENELAND. La présence éventuelle de la barrière nucléaire aux flux géniques entre les mitotypes GLA et RUT a été testée en estimant l'isolement par la distance (GENEPOP) 'au sein' et 'entre' les deux grands groupes étudiés et surtout, en utilisant des analyses de cline (CFIT) de fréquences alléliques de marqueurs neutres qui estiment des changements de fréquences alléliques entre les populations.

Parmi les 442 spécimens, 300 appartenaient au mitotype GLA et 142 au mitotype RUT. Les 22 populations étaient significativement mais faiblement différenciées les unes des autres. Etant donné cette faible différenciation génétique entre les campagnols roussâtres de Finlande, les analyses de génétique des populations et de « clustering » n'ont pas permis de valider l'une ou l'autre des hypothèses (H1 ou H2). En effet, tandis que certains résultats (analyses de la différenciation génotypique entre paires de population, GENELAND et IBD)

soutenaient l'hypothèse de deux évènements de recolonisation (H1), d'autres (ACP, DAPC, STRUCTURE) semblaient plutôt en faveur d'un seul événement de recolonisation (H2). Nous avons pu cependant déterminer que la différenciation génétique des campagnols roussâtre était corrélée avec la distance géographique et que par conséquent, la dispersion spatiale du campagnol roussâtre était limitée dans l'espace. Les conflits observés entre les diverses approches utilisées seraient dus à un phénomène d'isolement par la distance et à une différenciation génétique nucléaire faible entre les campagnols de Finlande. Par conséquent, pour valider définitivement une des deux hypothèses, nous avons du utilisé une autre approche. L'approche des clines génétiques neutres a montré que le cytb et 16 des 17 microsatellites présentaient des changements de fréquences alléliques entre les mitotypes GLA et RUT et que, conséquemment, il y a bien eu contact secondaire entre les deux mitotypes de Finlande. La Finlande a donc connu deux évènements de recolonisation postglaciaire indépendants.

Lors de cette étude, nous avons vu que les allèles GLA et RUT du cytb étaient diagnostiques (i.e. ils étaient présents dans une population et totalement absent dans l'autre). Ce type de patron peut être observé dans le cas de barrières physiques ou suite à des pressions de sélection. L'hypothèse de la barrière physique ne semblait pas la plus probable car si c'était le cas, des différenciations nucléaires nettes auraient également dû être observées. De plus, lorsqu'on observe la position de la zone de contact entre GLA et RUT, elle ne semble pas associée à des barrières environnementales particulières (i.e. fragmentation de la forêt, présence d'une rivière). L'hypothèse adaptative semblerait plus probable. En effet, il semble que cette introgression du génome mitochondrial du campagnol de la Taïga ait facilité la dispersion de ces campagnols roussâtres dans des habitats plus froids et secs qui ne sont habituellement pas habités par le campagnol roussâtre mais qui sont plutôt caractéristiques de l'habitat du campagnol de la Taïga. Il a été montré que le génome mitochondrial peut subir des pressions de sélection induites par exemple, par un nouvel environnement thermique. Il conviendrait cependant de tester spécifiquement cette hypothèse avant de pouvoir la valider. Nous pouvons cependant affirmer que l'introgression mitochondriale de *M. rutilus* dans le mitotype RUT a certainement eu lieu récemment, peut être pendant le dernier maximum glaciaire ou lors de la recolonisation postglaciaire étant donné que l'hybridation interspécifique entre *M. glareolus* et *M. rutilus* semble rare voire inexistante de nos jours.

Article

Testing the presence of a secondary contact between two distant mitochondrial lineages of the bank vole (*Myodes glareolus*) in central Finland

Julie Pisano[1,2]*, Nathalie Charbonnel[2], Jean-François Cosson[2], Sylvain Piry[2], Maxime Galan[2], Otso Huitu[3], Heikki Henttonen[4], Raphaël Leblois[2§], Johan R. Michaux[1,5§]

Molecular Ecology (in prep.)

[1] Laboratory of conservation genetics, Institute of Botany (B22), University of Liège, 4000 Liège (Sart-Tilman), Belgium
[2] INRA, UMR 1062 CBGP (INRA/IRD/Cirad/Montpellier SupAgro), 755 avenue du campus Agropolis, 34988 Montferrier-sur-Lez cedex, France
[3] Natural Resources Institute Finland, Suonenjoki Unit, FI-77600 Suonenjoki, Finland
[4] Natural Resources Institute Finland, Vantaa Unit, FI-01301 Vantaa, Finland
[5] CIRAD, UPR AGIRs, F-34398 Montpellier, France
[§] Co-senior authors

Keywords

Mito-nuclear discordance, contact zone, isolation by distance, cline, microsatellites

ABSTRACT

The understanding of the evolutionary history of populations has been dramatically enhanced by the acquisition of molecular data that have revealed the distribution of genetic variation within and among populations. In particular, mito-nuclear discordance cases are particularly interesting to investigate. Here, we particularly focused on a contact zone between two well-differentiated mitotypes of the bank vole (*Myodes glareolus*), one of which carries the mitochondrial genome of another species, *M. rutilus.* Our objective was to test if Finland was colonised by one or two differentiated populations of bank voles. We examined patterns of differentiation and allele frequency changes for 17 microsatellite loci and one mitochondrial marker (the *cytochrome b* gene). Overall, because of isolation by distance and low level of differentiation among bank voles, differentiation and clustering analyses failed at clearly identifying the nuclear genetic structure among bank voles. Consequently, they prevented to estimate the occurrence of a secondary contact. In this case of conflicting signals among approaches, neutral genetic cline approaches may help to detect fine scale genetic discontinuities. They are particularly relevant when there is low differentiation between populations. Indeed, cline analyses enabled to highlight the spatial pattern of genetic differentiation of Finnish bank vole populations. Most importantly, they provided arguments in favour of the occurrence of a secondary contact zone between mitotypes GLA and RUT. So, we concluded that Finland was certainly colonised by two differentiated populations that certainly recolonized Finland at the end of the last glacial maximum.

INTRODUCTION

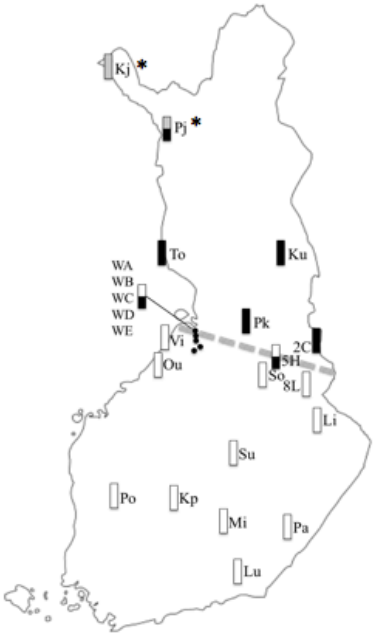
The understanding of the evolutionary history of populations has been dramatically enhanced by the acquisition of molecular data that have revealed the distribution of genetic variation within and among populations. Molecular phylogeographical studies are commonly based on the mitochondrial DNA (mtDNA), which has long been considered as the marker of choice for estimating population history because of rapid coalescence time, maternal inheritance, and high rate of variability and mutation (Zink & Barrowclough 2008). When used as a single marker, mtDNA may however provide an erroneous inference of population history because of selection, lack of recombination and possible confusing effects of pseudogenes (Dowling *et al.* 2008). Face to these problems, recent studies included a wider array of genetic markers, including nuclear genes and in particular, microsatellites. These neutral markers are co-dominant, and have a high polymorphism level compared to those of nuclear DNA (nuDNA) sequences (Zink & Barrowclough 2008). Moreover, they have the particularity to evolve 100-1000 times faster than other sequence genes, providing consequently a powerful tool for analysing recent and contemporary events. So, testing phylogeographical hypotheses is improved when combining microsatellites and mtDNA markers.

In the past decades, numerous cases of conflicting signals between mtDNA and nuDNA markers have been observed (126 cases listed by Toews and Brelsford (2012)). These patterns – well known as mito-nuclear discordance – mostly implicate the complete replacement of the original mitochondrial genome of a taxon by the mitochondrial genome of another species (Toews & Brelsford 2012). Mito-nuclear discordance can result from various neutral evolutionary processes (Rheindt & Edwards 2011) including 'asymmetrical introgression' (i.e. directional bias in gene flow and introgression of genes from one species into another) such as i) 'sex-biased dispersal' caused by the introgression of genes from the less-dispersal sex of an invaded species to the genome of the invasive species during a colonisation process, ii) 'demographic expansion following introgression in small populations' caused by genetic drift whereas allele frequencies remain unchanged, iii) 'hybridization asymmetry regarding sex' implying that females of one species are more discriminating in mate choice than males and than females from the other species, or iv) 'sex biased fitness of hybrids' (Funk & Omland 2003; Hedrick 2010; Petit & Excoffier 2009; Rheindt & Edwards 2011; Wirtz 1999). In addition, non-neutral processes, such as adaptive introgression of mtDNA, have also been reported as causes of mito-nuclear discordance. For example, adaptive processes can sometimes favour certain mtDNA variants that are better adapted to local conditions such as thermal tolerance (Ballard & Whitlock 2004; Dowling *et al.* 2008; Irwin 2012; Silva *et al.* 2014).

This question of mito-nuclear discordance is particularly interesting with regard to the bank vole, *Myodes glareolus*. This rodent has a wide range in the Palearctic, stretching from the British Isles through continental Europe and Russia and

A.

B.



C.

Locality		*N*
Kj	Kilpisjärvi	20*
Kp	Korpilahti	19
Ku	Kuusamo	19
Li	Lieksa	23
Lu	Luumäki	15
Mi	Mikkeli	18
Ou	Oulainen	14
Pa	Parikkala	20
Pj	Pallasjärvi	19
		19*
Pk	Puolanka	21
Po	Parkano	16
So	Sotkamo	17
Su	Suonenjoki	20
To	Tornio	21
Vi	Vihanti	19
WA to WE	Transect across the mtCZ	WA: 20 WB: 24 WC: 25 WD: 26 WE: 21
2C, 5H & 8L	Sites in central Finland	2C: 20 5H: 19 8L: 28

Figure 1 Sampling sites and mitochondrial affiliations. (A) Spatial distributions of *Myodes glareolus* and *M. rutilus* in Fennoscandia (IUCN 2015). (B) Sampling sites of vole populations. White rectangles indicate the presence of mitotype GLA, black ones the mitotype RUT, white and black ones populations with specimens from mitotypes GLA and RUT. Grey rectangles represent populations of *Myodes rutilus,* in grey and black sites where both species were sampled. The grey-dotted line represents the contact zone between mitotypes GLA and RUT. (C) Sampling localities. *N* indicates sample size, 'mtCZ' stands for 'mitochondrial contact zone', and the asterisks refer to sample sizes of *M. rutilus* populations.

extending in the North beyond the Arctic Circle (IUCN 2015). A peculiar mitochondrial lineage of *M. glareolus* ranges from central Sweden, through northern Finland and up to central and southern Russia; it carries the mtDNA of a northern neighbour and closely related species, the red vole, *Myodes rutilus* (Abramson *et al.* 2009; Boratynski *et al.* 2011; Deffontaine *et al.* 2005; Potapov *et al.* 2007; Tegelström 1987). Recently, the phylogeography of six Finnish populations of *M. glareolus* exploring one mitochondrial (cytochrome *b, cytb*) and six nuclear coding genes confirmed the mito-nuclear discordance among Finnish bank voles (Boratynski *et al.* 2011; Boratynski *et al.* 2014). While nuclear gene sequences clustered all bank voles together (with no relation to their mitochondrial affiliation), two mitochondrial lineages were identified: the introgressed lineage (mitotype RUT) that ranges from central to North Finland, and the non-introgressed lineage (mitotype GLA) that is distributed from central to South Finland (Boratynski *et al.* 2011; Boratynski *et al.* 2014). The Fennoscandian contact zone (CZ) between these two mitotypes is localised in central Finland (Fig. 1) (Boratynski *et al.* 2011; Boratynski *et al.* 2014; Deffontaine *et al.* 2005). Deffontaine *et al.* (2005) suggested that the mitotypes GLA and RUT survived the last glacial maximum in two separated areas. They would have consequently recolonized Finland independently and the region of central Finland should therefore be considered as a secondary CZ between mitotypes GLA and RUT. However, given that no nuclear differentiation has been evidenced (Boratynski *et al.* 2011; Boratynski *et al.* 2014), the hypothesis of secondary contact seemed unlikely.

Here, we describe the phylogeographic structure of Finnish bank vole populations to emphasize the drivers of this mito-nuclear discordance and evaluate the status of central Finland as secondary CZ. Compared to previous studies, the sampling of *M. glareolus* was greatly improved, especially in the vicinity of the CZ between mitotypes GLA and RUT (442 specimens of *M. glareolus* in 22 widely distributed sites versus 201 and 131 specimens from 6 sites in Boratynski *et al.* (2011) and Boratynski *et al.* (2014)). To properly identify species membership, 39 specimens of *M. rutilus* were also collected. We sequenced a portion of the *cyb* gene to assign bank vole specimens to mitotypes GLA or RUT and we genotyped 17 microsatellites to estimate the nuclear genetic diversity along the whole Finland. Accordingly, we tested two hypotheses about the origin and post-glacial dispersal of Finnish bank vole populations: (H1) under the '*two-colonisation events hypothesis*', mitotypes GLA and RUT colonised Finland independently and the CZ was formed by secondary contact after a period of geographic isolation that certainly occurred during the Quaternary glaciations (Abramson *et al.* 2009; Deffontaine *et al.* 2005). A disruption in nuclear allelic frequencies is expected (even slightly) at nuclear genes between mitotypes GLA and RUT. Its strength should depend on the level of differentiation of the two mitotypes that was established during the whole colonisation process, and on dispersal, time and selection. (H2) Under the '*one colonisation event hypothesis*', the mitochondrial CZ would result from a mitochondrial introgression that occurred in Finland within a genetically homogeneous nuclear population that colonised Finland during a single event. Homogeneous allelic frequencies at nuclear genes are expected under this hypothesis.

MATERIAL AND METHODS

Sample collection

ETHICAL STATEMENT According to the *Finnish Act on the Use of Animals for Experimental Purposes* (62/2006) and a further decision by the *Finnish Animal Experiment Board* (16th May, 2007), the use of traps that instantly kill the animal for vole capture is not considered an animal experiment and therefore does not require any animal ethics license from the *Finnish Animal Experiment Board*. Permits (23/5713/2001, 4/5713/2007 and 7/5713/2013) for capturing protected species were however granted by the Finnish Ministry of the Environment. The main target of this study, the bank vole is not protected in Finland (i.e. it is not included in the Red List of Finnish Species).

STUDIED AREA *M. glareolus* samples were collected at 22 sites distributed from North to South Finland with particular efforts made along the CZ, and *M. rutilus* samples were caught at two sites from North Finland (Figure 1). Trapping was performed in spring and autumn between 2008 and 2013. At each trapping site, a mean of 50 - 100 standard metal mouse snap traps were set in line, in clusters of three traps with an inter-trap distance of 1-2 m and an intercluster distance of 10-20 m. Traps were set for one to four nights and checked and re-set if necessary once a day (for more details see Henttonen (2000); Henttonen and

Wallgren (2001); Korpela *et al.* (2013); Korpela *et al.* (2014)). At each sampling site, 14 to 28 specimens were finally sampled, resulting in a total of 442 samples of *M. glareolus* and 39 samples of *M. rutilus*.

Molecular data collection

Total DNA was extracted and purified from ethanol- and dried-preserved tissues using the Qiagen® DNEASY BLOOD & TISSUE KIT according to the manufacturer's instructions (Hilden, Germany). We sequenced the mitochondrial cytochrome *b* gene (*cytb*, 1.14 kb) and genotyped nineteen unlinked microsatellite loci for all *M. glareolus* and *M. rutilus* samples.

The *cytb* gene was amplified using universal primers designed for mammal species: L7 (5-ACCAATGACATGAAAAATCATCGTT-3), H6 (5-TCTCCATTTCTGGTTTACAAGAC-3) (Irwin *et al.* 1991). When amplifications failed, specific primers designed specifically for *M. glareolus* were used: LCLE-1 (5-ATCATCAACCACKCATTTATT-3), MCLE-3 (5-ATTATGCCTGCTATTGGTATG-3) (Deffontaine *et al.* 2005). PCR 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 10X buffer, 0.5 mM of extra $MgCl_2$, 100 µM of each dNTP, and 0.2 µM of each primer. The *cytb* cycling conditions were as follows: one activation step at 94°C for 4 min followed by 40 denaturation cycles at 94°C for 30 sec, annealing at 50°C for 30 sec, elongation at 72°C for 1 min 30 sec, 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 set at 95°C for 15 min. PCR products of *cytb* were sequenced by the *National Centre of Sequencing* (CNS, Evry, France) as part of the "Bibliothèque du Vivant (BdV)" project. Sequences were corrected using CODONCODE ALIGNER (CodonCode Corporation), aligned by eye and translated in amino acids using SEAVIEW (Galtier *et al.* 1996) to ensure sequence orthology.

Nineteen microsatellites were amplified using primers and conditions described in Guivier *et al.* (2010). Genotyping was carried out using an ABI3130 automated DNA sequencer (Applied Biosystems). Alleles were scored using GENEMAPPER® (Applied Biosystems).

Identifying mitochondrial lineages: mitotypes GLA and RUT

We used the basic local alignment search tool (BLAST) to identify specimens belonging to mitotypes GLA and RUT, to examine nucleotide sequences of *cytb*, and to ensure that no human contamination occurred during pre-PCR processes (Altschul *et al.* 1990). DNASP was then applied to infer haplotypes from *cytb* sequences (Rozas 2009). The relationships among haplotypes were displayed using a Median-Joining method, as implemented in NETWORK (Bandelt *et al.* 1999).

Population genetic diversity

Null allele frequencies were estimated for each microsatellite using FreeNA (Chapuis & Estoup 2007). Loci with null allele frequency higher than 0.05 were removed from the final dataset (Chapuis & Estoup 2007). Genetic variation was estimated for each site. Allelic richness estimates (*A*) were inferred based on the smallest sample size (i.e. 14 individuals) using FSTAT (GOUDET 1995). Observed (H_o) and unbiased expected (H_e) heterozygosities and inbreeding coefficient (F_{is}) estimates were computed using GENETIX (Belkhir *et al.* 1996-2004). Heterogeneity deficit and excess were tested with GENEPOP using global tests of Hardy-Weinberg permitting to detect deficit or excess of heterozygotes (dememorisation step = 10^4, batches = 20, iterations per batch = $5x10^3$) (Rousset 2008). To identify geographic areas of high genetic diversity, spatial representations of expected heterozygosity (*He*) and allelic richness estimated on the smallest samples size (*A*) were computed by calculating, for each cell of a grid, the inverse-distance weighted mean of each population, *He* and *A* respectively (Piry's code). The distance used for weighting is the distance between cell's centroid and population. Then, we tested if there were significant gradual decrease of genetic diversity along one (H2) or two (H1) geographical axis, which could be a sign of recent colonisation events with founder effects. So, Pearson and Spearman tests were used to test correlations between H_e or *A* and the geographical distances, as implemented in R v3.1 (R Core Team 2014). Geographical distances were calculated from two sites: (*a*) the southernmost site 'Luumäki', because the mitotype GLA could have been isolated in the South (in central Europe) during the last glacial maximum (LGM), and (*b*) the easternmost site '2C', because the mitotype RUT could have been isolated in the East (in the Ural mountains) during the LGM (Deffontaine *et al.* 2005).

Population genetic structure analyses

We wanted to test if bank vole populations were geographically structured in two major groups referring to mitotypes GLA and RUT (H1) or did not exhibit any particular pattern of clustering (H2). First, methods that are free of assumptions with regards to Hardy-Weinberg equilibrium (HWE) or linkage equilibrium (LE) – i.e. principal component analysis (PCA) and discriminant analysis of principal components (DAPC) – were applied using the *R* package ADEGENET (Jombart 2008). The *K-means* clustering algorithm was used for identifying the number of homogeneous genetic groups '*K*' (i.e. clade) maximising genetic variation between groups. The Bayesian Information Criterion (BIC) was applied to compare clustering solutions. The 'a-score optimisation' estimated the optimal number of principal components to include in the DAPC. We ran the analyses with a number of *K* ranging from two to 22, to test respectively gene flow between mitotypes and between sampling sites.

Second, the Bayesian clustering method assuming HWE and LE implemented in STRUCTURE 2.3.4 was used to identify genetically homogeneous clusters (*K*) (Pritchard *et al.* 2000). Using the admixture model of ancestry, analyses were run ten times for each *K* value (*K* varying from one to ten). Each run consisted of a burn-in of 5×10^5 iterations, followed by 10^6 iterations. The thinning

was fixed to 10^2. We applied the correlated allele frequencies model first on the dataset including all specimens of *M. glareolus* and *M. rutilus* and ten microsatellites common to both species. We next restricted the analysis to *M. glareolus* only and used all microsatellite loci. Remaining parameters were set to default values. Outputs were inferred using STRUCTURE HARVESTER (Earl & vonHoldt 2012). The graph of the log probability of the data [Ln P(D)] was used to identify the optimal number of clusters (*K**) among bank voles.

Delineating spatial nuclear genetic discontinuities

The clustering method implemented in GENELAND was applied (Guillot *et al.* 2005) to test the influence on geographic spatialisation on genetic clustering. Taking into account spatial coordinates, the algorithm allocates individual genotypes of bank voles into *K* clusters under the hypothesis of HWE and LE. To infer the best *K*-value and check the consistency of our results, ten different MCMC were run (number of iterations = 10^7, thinning = 10^3, burn-in = 50%). The maximum rate of Poisson process was fixed to 300 and considered with an uncertainty attached to spatial coordinates of 5 km. The maximum number of nuclei in the Poisson-Voronoi tessellation was fixed to 700. We used both Dirichlet models of independent and correlated allele frequencies. To estimate if a reproductive barrier could be observed between mitotypes GLA and RUT (i.e. two spatially distinct populations, H1), *K* was only explored using small values. First, the uncorrelated allele frequency model was applied, and *K* was explored with values ranging from one to four with a starting value of two. Second, *K* was identically explored but the correlated allele frequency model was rather used. Finally, the correlated allele frequency model was applied, and *K* was constrained with very small values, i.e. one and two (referring to mitotypes GLA and RUT) with a starting value of one. The posterior probabilities of population membership, the number of populations, the posterior density of the model were computed for all runs.

Testing specifically the occurrence of a disruption in allelic frequencies in the CZ

ISOLATION BY DISTANCE The model of isolation by distance (IBD) was applied to test geographic-genetic correlation. It estimated the relationship between differentiation parameters (a_r) and geographical distances (Rousset 2000). The main property of IBD is that the genetic differentiation increases with geographic distances because of spatially limited dispersal. In addition, models of IBD also enabled to test CZ for barriers to nuclear gene flow, i.e. the presence of dispersal and gene flow between the two mitotypes. In this context, if there were no barrier between populations situated on both sides of the CZ, we should observe similar IBD patterns in 'within-' and 'between-mitotype' analyses (similar slope and intercept estimates for both types of analyses). If a slight barrier to nuclear gene flow between mitotype was present (e.g. because of assortative mating or reduced fitness of hybrids), the barrier should increase the genetic differentiation between both populations (higher mean differentiation value for the 'between-mitotype' analysis and the regression line

of the 'between-mitotype' analysis should be above those of the 'within-mitotype' analyses) (Rousset 1999a, b). On the contrary, if there were no gene flow because of a complete reproductive barrier, no IBD patterns should be observed under the 'between-mitotype analysis' (and the slope value would thus be null).

IBD analyses were therefore performed 'within' and 'between' mitotypes GLA and RUT using GENEPOP (Rousset 2008) and R script that modified Mantel tests to calculate rank correlation coefficients and to make permutation "within" or "between" groups as in Sagnard *et al.* (2011) and Perez *et al.* (2012) (R Core Team, 2014). Mantel tests that enabled to test the correlation between geographical and genetic distances matrices were performed using 10^4 permutations. First, we considered only pairs of specimens belonging to the same mitotype (the "within-mitotype" analysis, specimens of GLA and RUT were analysed separately). Then, we selected only pairs of individuals from distinct mitotypes (the "between-mitotype" analysis, which was performed strictly between pairs of GLA and RUT).

NEUTRAL GENETIC CLINES ANALYSES

We tested for disruption patterns in allelic frequencies using the maximum-likelihood method implemented in CFIT7. We fitted genetic clines to the *cytb* and the microsatellite data (Gay *et al.* 2008). Clines – i.e. representations of allelic frequencies in relation to geographical distance – enabled to examine patterns of genetic variation across a CZ and to evaluate the strength of a barrier to gene flow (here, between mitotypes GLA and RUT). Clines were fitted by constraining shape parameters to compare concordances (i.e. slopes) and coincidences (i.e. centres) of mitochondrial and microsatellite data. Because we studied a CZ between two intraspecific groups and subsequently, because alleles were not expected to be fixed within parental populations, we used the allelic count data with semi-diagnostic alleles (i.e. alleles that are more present in one mitotype than in the other) and thus, applied the scaled Logit cline shape as described in Rieux *et al.* (2013). For the Logit cline shape, the sigmoid is characterised by two main parameters, the centre c and the slope b, and three additional parameters, the down and up asymptotic frequencies (h_1 and h_2 for which values range between 0 and 1) and the angle. By setting the 'angle' parameter to zero (implying that the parameter should not be fitted), we could switch from *2D* (i.e. two dimensions) to *1D* analyses. The cline sigmoid should be at right angle to the CZ. Consequently, spatial coordinates of population sites were warped by a tilt angle of 20° in relation to the spatial delimitation between mitotypes GLA and RUT.

Clines are based on bi-allelic system. So, concerning the cytb dataset, we considered that GLA and RUT constitute the bi-allelic system. Because microsatellites are multi-allelic, we transformed microsatellite data into bi-allelic ones, as in Daguin *et al.* (2001); Gay *et al.* (2008) and Rieux *et al.* (2013). Accordingly, microsatellite alleles were clustered into two groups – in relation to mitotypes GLA and RUT – using a correspondence analysis (CA) performed using the R package ADEGENET (Jombart

Table 1 Comparison of the quality of fit of the various cline models of the spatial pattern microsatellite frequencies for *Myodes glareolus* in the contact zone between the two mitotypes, GLA and RUT. *c* indicates the centre of the cline, *b* the slope, *h* the allelic frequencies (up and down) and θ the cline angle. *CMsD* refers to combined microsatellite data (combined information of the 17 loci) ; *MsD* indicates microsatellite data (17 loci) ; *MtD* stands for mitochondrial data. A. Descriptions of the cline models. B. Pairwise comparisons of the models.

A.

Model	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**
Free parameters	*c, b, h1, h2, θ*	*c, b, h1, h2*	*c, b, h1, h2*	*h1, h2*	*h1, h2*	*h1, h2*
Constrained parameters	/	θ	/	c, b	c, b	c_{MsD}, b_{MsD} c_{MtD}, b_{MtD}
Data	CMsD & MtD	CMsD & MtD	MsD	MsD	MsD & MtD	MsD & MtD

B.

Data	**Model**	**Number of parameters**	**Maximum Log likelihood**	**Delta of corrected Deviance**	**AICc**	**ΔAICc**
CMsD & MtD	Null	1	-8307.138	343.330	345.330	
	Cline M1	10	-7931.329	34	54	
	Cline M2	9	-7931.875	34.449	52.449	1.55
	Saturated	44	-7890.022		88	
MsD	Null	1	-7668.533	498.051	500.051	
	Cline M3	68	-7465.389	306	442	
	Cline M4	36	-7445.389	286.763	358.763	83.24
	Saturated	374	-7141.713		748	
MsD & MtD	Null	1	-7946.030	813.941	815.941	
	Cline M5	38	-7722.538	560.123	636.123	
	Cline M6	40	-7542.803	356	436	200.12
	Saturated	396	-7229.337		792	

2008). It allowed assigning each allele to a cluster-specific compound allele (alleles *A* and *a*) according to its coordinates on the first axis of the CA. The genotypic differentiation test implemented in GENEPOP (Rousset 2008) was used to infer genotype frequencies in all sites and consequently, to estimate the frequency of the diagnostic allele for each population site.

We tested several models that varied in the number of constrained parameters (summarised in Table 1), and three datasets: (i) '*cytb* (MtD) + combined microsatellite dataset (CMsD, all frequencies of allele *A* and *a* of the 17 microsatellite loci were respectively clustered together)' to compare the cline angle of mitochondrial and microsatellite data, (ii) '17-microsatellite dataset (MsD)' to estimate if all microsatellite loci had the same slope and centre, and (iii) '*cytb* + 17-microsatellite dataset' (MtD + MsD) to evaluate if mitochondrial and microsatellite data had the same slope and centre. The most complex model (M3, applied on MsD) had 68 parameters (number of loci (= 17) x number of parameters in the cline analysis (= 4; *c, b,* h_1, h_2)) and the simplest model (M2, applied on 'CMsD + MtD') had 9 parameters (number of loci (= 2; mitochondrial and combined microsatellite dataset) x number of parameters in the

cline analysis (=4; *c, b, h_1, h_2*) + 'angle constrained for both loci' (=1)). The best model was chosen based on the AICc, i.e. the corrected Akaike Information Criterion that takes into account sample size and overdispersion of the data by correcting the deviance with the variance of inflation factor $\hat{c}$ (Lebreton *et al.* 1992).

Then, we performed cline analyses on each microsatellite locus independently. Cline shape, centre and slope significance were tested against a null model assuming homogeneous allele frequency over the studied area using the AICc. Representations of 95% confidence interval for model parameters *c* and *b* were obtained from the CFIT outputs using MATHEMATICA 8.0.1.0 (Wolfram Research, Inc., 2010).

RESULTS

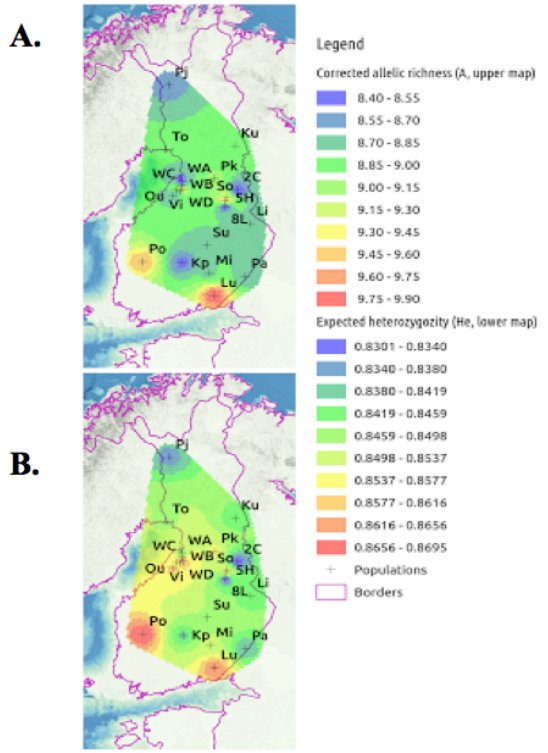
Assignment to mitotypes GLA and RUT

A segment of 519 bp of *cytb* was sequenced for 442 specimens of *M. glareolus* and 39 specimens of *M. rutilus*. The BLAST analyses showed that among all bank voles, 300 belonged to the mitotype GLA and 142 to the mitotype RUT. All *cytb* sequences from red voles belonged to the mitotype RUT. DNASP and NETWORK results confirmed these assignments to mitotypes (data not shown). The mitotype GLA was found from South to Central Finland while the mitotype RUT was distributed from Central to North Finland (Fig. 1).

Population genetic diversity

FREENA showed that two loci (Cg2F2, Cg16H5) had potential null alleles ($r \geq 0.05$). Accordingly they were excluded from the final dataset. Microsatellite genotyping revealed that seven loci could not be amplified on *M. rutilus* specimens. In total, our microsatellite dataset included genotypes obtained at 17 loci for 442 samples of *M. glareolus* and at 10 loci for 39 samples of *M. rutilus*.

Figure 2 Correlation between geographic and genetic distances of bank vole sampling site populations in relation to the unbiased expected heterozygosity (H_e) and the allelic richness (A) estimated on the smallest sample size. Colours indicate the degree of A and H_e from the lowest (blue) to the highest (red) estimates.



F_{IS} estimates suggested HWE at all sites except for Tornio, Oulainen, WB, WC and WD, which displayed significant positive F_{is} estimates ($P < 0.05$) indicating a deficit of heterozygotes (see Table S1 for A, H_e, H_o and F_{is} estimates per population site). Projections of A and H_e estimates across Finland showed that the higher genetic diversity levels were found at the southern sites, Luumäki and Parkano, and that genetic diversity decreased northward (Fig. 2). However, Spearman and Pearson tests rejected any significant correlation between these genetic indices and

Table 2 Genetic differentiation estimates (F_{ST}) between Finnish bank vole populations performed using GENEPOP. ($^{\#}$) Non-significant *P*-values ($P > 0.01$). (*) Significant *P*-values ($0.001 < P < 0.01$). (In bold) Very highly significant *P*-values ($P < 10^{-7}$). Others are all highly significant ($P < 0.001$).

	Pj	To	Ku	Pk	WA	WB	WC	WD	WE	2C	Vi	5H	Ou	So	8L	Li	Su	Po	Kp	Mi	Pa
Pallasjärvi(Pj)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Tornio (To)	0.013	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Kuusamo (Ku)	0.032	0.025	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Puolanka (Pk)	0.025	0.021	0.018	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WA site (WA)	**0.039**	**0.032**	0.038	0.031	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WB site (WB)	**0.037**	**0.033**	**0.028**	0.023	0.017	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WC site (WC)	**0.042**	**0.026**	**0.030**	0.021	0.021	0.009	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WD site (WD)	0.027	0.023	0.027	0.016	0.016	0.015	0.008	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WE site (WE)	**0.036**	0.022	0.021	0.010	0.020	0.011	0.011	-0.002$^{\#}$	-	-	-	-	-	-	-	-	-	-	-	-	-
"2C" site (2C)	**0.047**	**0.034**	**0.031**	0.024	**0.040**	**0.037**	0.024	**0.025**	0.028	-	-	-	-	-	-	-	-	-	-	-	-
Vihanti (Vi)	**0.037**	0.029	0.028	0.023	0.012	0.010	0.012	0.002$^{\#}$	0.007	0.036	-	-	-	-	-	-	-	-	-	-	-
"5H" site (5H)	**0.032**	0.029	0.034	0.013	0.036	0.029	0.025	0.016	0.012	0.032	0.020	-	-	-	-	-	-	-	-	-	-
Oulainen (Ou)	0.031	0.025	0.027	0.016	0.014	0.012	0.014	0.006$^{\#}$	0.006$^{\#}$	0.031	0.001$^{\#}$	0.020	-	-	-	-	-	-	-	-	-
Sotkamo (So)	0.033	**0.027**	0.029	0.015	0.024	0.019	0.014	0.011*	0.011*	0.037	0.011	0.021	0.017*	-	-	-	-	-	-	-	-
"8L" site (8L)	**0.043**	**0.036**	**0.033**	0.017	**0.034**	0.024	0.024	0.017	0.015	0.026	**0.025**	0.023	0.020	0.021	-	-	-	-	-	-	-
Lieksa (Li)	**0.039**	**0.040**	**0.032**	0.019	**0.027**	0.020	0.028	0.014	0.009	0.040	0.013	0.021	0.023	0.019	0.015	-	-	-	-	-	-
Suonenjoki (Su)	0.039	**0.032**	0.031	**0.018**	0.031	0.030	0.021	0.020	0.013	**0.038**	0.020	0.026	0.020	0.014	0.021	**0.019**	-	-	-	-	-
Parkano (Po)	0.028	0.022	0.027	0.017	0.022	**0.025**	0.021	0.012	0.009	**0.039**	0.018	0.021	0.010*	0.021	0.021	0.025	0.014	-	-	-	-
Korpilahti (Kp)	**0.054**	**0.044**	0.038	0.027	**0.046**	**0.039**	**0.036**	0.031	0.022	**0.064**	0.028	**0.031**	0.028	0.028	**0.042**	0.031	0.023	0.018	-	-	-
Mikkeli (Mi)	**0.040**	**0.039**	0.040	0.029	0.020	**0.029**	0.027	0.025	0.022	0.044	0.021	0.029	0.022	0.010*	0.036	0.029	0.021	0.022	0.030	-	-
Parikkala (Pa)	**0.042**	**0.041**	0.033	0.022	0.027	**0.026**	0.023	0.017	0.015	0.039	0.015	0.028	0.015	0.023	**0.028**	0.021	0.021	0.016	**0.042**	0.028	-
Luumaki (Lu)	0.033	**0.037**	0.023	0.016	0.027	0.023	0.019	0.016	0.011	0.040	0.019	0.018	0.013*	0.012*	0.027	0.019	0.016	0.011*	0.027	0.021	0.020

geographical distances from sites '2C' or Luumäki ($P > 0.05$).

Population genetic structure

All populations were significantly differentiated from each other ($P < 0.01$, Table 2) except some pairwise comparisons between sites that were geographically close from each other: WD & WE, WD & Vihanti, WD & Oulainen and WE & Oulainen ($P > 0.01$). We calculated the proportion of very highly significant ($P < 10^{-7}$) and non-significant ($P > 0.01$) F_{ST} estimates within and between mitotypes GLA and RUT based on the results of the genotypic differentiation for each population pair (excluding population sites from central Finland that exhibited both mitotypes). We found four times more very highly significant F_{ST} estimates between mitotypes (17 highly significant F_{ST} estimates between GLA and RUT versus 5 and 3 respectively within GLA and RUT). Only two non-significant F_{ST} were observed and both were identified within RUT.

PCA did not reveal any clear genetic differentiation between mitotypes (Fig. S1). DAPC analyses estimated an optimal number of clusters of $K = 3$, but clustering solutions for $K = 2$, 4 or 5 had nearly equivalent BIC (Fig. S1). Whatever the K considered (2, 3, 4 or 5), no clear discrimination between clusters was observed. When testing specifically the mitochondrial ($K = 2$, Fig. S1) or the sampling ($K = 22$, Fig. S1) structure, no clear distinct clusters were detected. We could only hypothesise about a slight geographical differentiation in relation to the latitudinal axis (Fig. S1).

In STRUCTURE, when inferring clusters using specimens from both species and the 10-microsatellite dataset, two distinct clusters corresponding to *M. glareolus* and *M. rutilus* were consistently observed (Fig. S2 and S3). However, when inferring homogeneous genetic clusters using bank vole specimens only and the 17-microsatellite dataset, we faced convergence difficulties. We were unable to estimate K^* (e.g. see variances of values of Ln P(D) observed for $K = 2$, 9 and 10 in Fig. S2) and no clear cluster could be detected on barplots whatever the K value considered. We only observed a very slight signal of differentiation, consistent among runs, distinguishing a portion of the mitotype RUT (the northernmost sites) from the others (Fig. S3). Because of these convergence issues, the method of Evanno *et al.* (2005) for estimating K^* was not applied.

WE also encountered difficulties in detecting clusters among Finnish bank vole populations when using Geneland. In the case of the uncorrelated and correlated allele frequency models with K varying from one to four, posterior distributions of K respectively displayed a mode at $K = 1$ and $K = 4$ in all replicates (Table S2). Investigating results for $K = 4$ under the correlated allele frequency model, we noted that the spatial delimitation between the different clusters were very different in all runs (Fig. 3). When forcing GENELAND to explore small K values ($K = 1$, 2), posterior distributions displayed a peak at $K = 2$ for all runs (Fig. 3, Table S2). Although the spatial delimitation between the two clusters greatly varied between runs, we noted in most runs with $K = 2$ that the genetic discontinuity in central

Figure 3 Maps of population membership inferred from GENELAND posterior probabilities. Population memberships were inferred using the model of correlated allele frequencies for (A) $K = 4$ and (B) $K = 2$. Black dots represent specimens' geographical positions. Colours refer to distinct clusters. The dotted line indicates the contact zone between mitotypes GLA and RUT.

Finland roughly corresponded to the CZ between mitotypes GLA and RUT.

Testing specifically the occurrence of a disruption in allelic frequencies in the CZ

IBD ANALYSES IBD patterns were highly significant on the whole bank vole dataset and when considering each mitotype separately: (i) within all 442-Finnish bank voles (intercept = 0.0357, slope = 0.0025 [0.00024; 0.0048], $P_{(slope)} < 10^{-2}$), (ii) within all 142-RUT specimens (intercept = 0.0348, slope = 0.0021 [-0.0035; 0.0078], $P_{(slope)} < 10^{-2}$), and (iii) within all 300-GLA specimens (intercept = 0.0361, slope = 0.0010 [-0.0018; 0.0045], $P_{(slope)} < 10^{-2}$). The 'between-group' pattern was also significant, with a slightly higher slope estimate than those of the 'within-group' analyses (intercept = 0.0312, slope = 0.0032 [0.00052; 0.0062], $P_{(slope)} = 0.03$). Intercept estimates were mostly similar.

CLINE ANALYSES All polymorphic loci (except Cg6G11) and the *cytb* displayed a cline pattern (ΔAICc > 2 between null and cline models) (Table S3). Analyses in *2D* did not differ significantly from those performed in *1D* (ΔAICc < 2 between M1 and M2, Table 1). Consequently, we proceeded with the cline analyses in *1D*. Cline analyses performed on each

microsatellite locus independently yielded very few information given that each locus demonstrated very large 95% confidence interval of *c* and *b* parameter estimates. Cline model comparisons between M3 and M4 showed that all microsatellite loci had distinct slopes and centres and those between M5 and M6 suggested that microsatellites and *cytb* had also different slopes and centres (ΔAICc > 2, Table 1).

DISCUSSION

Population genetic structure of Finnish bank voles

Because of low intraspecific differentiation among bank voles, the bulk of our population genetic differentiation and clustering analyses did not permit to clearly identify if there is or not a secondary contact zone between mitotypes GLA and RUT. We thus encountered difficulties to estimate which hypothesis best explains the colonisation of Finland. Indeed, while some results (Genotypic differentiation for each population pair, GENELAND and IBD) provided supports for two colonisation events (H1), others (the spatial representations of *A* and H_e estimates, PCA and DAPC, and STRUCTURE) tipped the scale in favour of one single colonisation (H2). Overall, all these results did not show either clear signals of differentiation or clear signals of absence of differentiation between mitotypes GLA and RUT. We were only confident about the fact that the genetic differentiation among Finnish bank voles was correlated with the geographic distances, and consequently, that the dispersal of bank voles was spatially restricted. In this context, to detect fine scale genetic discontinuities when there is low differentiation between populations, we tested the approach of neutral genetic cline (Lenormand *et al.* 1998). Given that the *cytb* and all except one microsatellite loci displayed breaks in allelic frequencies between mitotypes, we were finally able to demonstrate that the Finnish contact zone between mitotypes GLA and RUT still show signs of a past contact between two differentiated lineages (Barton & Hewitt 1985). The nuclear barrier is very weak compared to the abrupt changes in mitochondrial lineages but its occurrence confirmed that Finland was colonised by two populations, the mitotypes GLA and RUT.

Conflicting signals among approaches

Inferring population structure when differentiation between populations is weak is a difficult analytical problem (Putman & Carbone 2014). This is especially true when dispersal generates continuous differentiation pattern such as IBD (Leblois, pers. comm.). Under conditions of low population differentiation (F_{st} < 0.03), the Bayesian clustering method implemented in STRUCTURE may fail to properly identify the number of clusters (Latch *et al.* 2006). Indeed, we showed that investigating the genetic structure using only *M. glareolus specimens* (and the 17-microsatellite dataset), STRUCTURE encountered convergence difficulties to identify the number of *K*; while when incorporating red vole specimens (using the 10-microsatellite dataset), no such issues were observed and two clusters referring to *M. glareolus* and *M. rutilus* were clearly detected (Fig. S2). Accordingly, following Latch *et al.* (2006), it seems that the reason undermining STRUCTURE to estimate *K* within Finnish bank voles was that: even if

F_{st} indices among Finnish bank vole populations were almost all significant (Table 2), ~30 % of F_{st} values were below '0.01' and ~70 % were below '0.03', levels of differentiation were too low for applying STRUCTURE on our dataset. Moreover, Guillot & Santos (2009) showed that irregular spatial localisation of population borders, fluctuations in the number of identified clusters and variations in the individual population memberships could be signs of IBD patterns. Using GENELAND, we actually observed one, two or four spatial genetic clusters for which the position in space differed in each run (Fig. 3) and we identified IBD among Finnish bank voles. Thus, even if microsatellites are often cited as being very useful fur studying recent evolutionary events, biologists must be aware that each clustering method present varying degrees of weaknesses when based on microsatellites and that consequently, some limitations could prevent their use in some cases, as observed in this study.

The barrier to mitochondrial gene flow between GLA and RUT in Finland

Traditionally, it has been assumed that genetic variation within the mitochondrial genome was selectively neutral (Ballard & Kreitman 1995). However, there are more and more supports towards the assumption of positive selection in maintaining polymorphism in the mitochondrial genome (Review in Dowling *et al.* (2008)). In particular, it has been suggested that the mitochondrial genome might evolve adaptively to selection imposed from the prevailing local environment and, more precisely, from a novel thermal environment (Ballard & Whitlock 2004).

In this study, we found that the *cytb* locus demonstrated a steep and sharp cline and that mitochondrial alleles GLA and RUT were diagnostic (i.e. alleles are representatives of one population and are absent in the other). Actually, abrupt delimitations between distinct mitochondrial lineages could arise from physical barriers or from selective pressure (Barton & Gale 1993; Kingston *et al.* 2012; Kruuk *et al.* 1999). First, if the hypothesis of a physical barrier to dispersal was the most likely, we would also found a clear differentiation on the nuclear genome between mitotypes GLA and RUT. Moreover, when we looked at the position of the CZ, it does not seem to be associated with any peculiar environmental barrier (e.g. fragmentation of forested areas or rivers). Second, the cline analyses enabled to detect a strong pattern delimitating the northern mitotype RUT and the southern mitotype GLA. The mitotype RUT is distributed in northern colder and drier habitats than animals carrying their own mtDNA (GLA) and interestingly, the species (*M. rutilus*) that shares his mtDNA with the mitotype RUT is well-known as a northern Holarctic species well adapted to cold conditions of the tundra and taiga forests (IUCN 2015). Accordingly, it seems that this mtDNA introgression from *M. rutilus* to mitotype RUT would have facilitated the dispersion into more severe habitats than usually inhabited by the bank vole as suggested by Boratynski *et al.* (2014). Given this context and our results, we would rather be in favour of the adaptive hypothesis. However, no studies found supports for selection (Boratynski *et al.* 2011; Boratynski *et al.* 2014; Filipi *et al.* 2015). Nevertheless, the absence of selection signal could not be considered as an

evidence of absence of selection. Adaptive hypotheses of mitochondrial variation to different climate are generally difficult to sustain (Dowling *et al.* 2008) and in addition, past selective events are, in general, challenging to detect (Toews & Brelsford 2012). So, we believe that one fruitful avenue in solving the above adaptive hypothesis will be to explore specifically whether the mitochondrial introgression of *M. rutilus* into *M. glareolus* respond adaptively to thermal selection.

The secondary contact zone between GLA and RUT

We support the hypothesis that Finland was colonised by two differentiated populations, which colonisation event certainly occurred during the last post-glacial maximum (Abramson *et al.* 2009; Deffontaine *et al.* 2005). The question whether the mitochondrial introgression between *M. glareolus* and *M. rutilus* occurred during the last glacial maximum (or during the post-glacial recolonisation) or after the secondary contact is difficult to answer given that our sampling and analyses did not allowed testing this issue. However, we know that a population of *M. rutilus* is found in central Finland at the northern Finnish range of the mitotype GLA and that interspecific hybridisation appears to be totally absent in this site (pers. comm. H. Henttonen). By the way, generally speaking, there is no evidence for hybridisation of these two species in nature so far (Osipova & Soktin 2008). In addition, even if hybridisation between *M. glareolus* and *M. rutilus* is forced (because hybridisations *in situ* have already been tested), it has been showed that male hybrid offspring are sterile (Osipova & Soktin 2008). Thus, we are of the opinion that the capture of the mtDNA may have certainly occurred before Finland colonisation and that selection has certainly provided an advantage to disperse in other habitats that are usually not inhabited by the bank vole.

Acknowledgements

This work have been partly supported by the European programs GOCE-CT-2003-010284 EDEN and FP7-261504 EDENext. This article is registered with the EDENext Steering Committee as EDENext413 (http://www.edenext.eu/). The sequencing was supported by the "Bibliothèque du Vivant (BdV)" network, which was funded by the National Centre for Scientific research (CNRS), the National Museum of Natural History (MNHN, Paris, France) and the National Sequencing Centre (CEA, Evry, France). We thank the GenSeq platform of the labex CeMEB (Montpellier, France) for the genotyping. We are particularly grateful to Thomas Lenormand (CEFE, CNRS, Montpellier, France) for its help in performing neutral cline analyses.

AUTHOR CONTRIBUTIONS

J.P., R.L., N.C. and J.R.M. conceived the ideas; J.P., H.H., O.H., and J.R.M. collected the data; J.P. and M.G. achieved the molecular work; J.P. and R.L. analysed the data; and J.P. led the writing with revisions of all co-authors.

REFERENCES

Abramson NI, Rodchenkova EN, Kostygov AY (2009) Genetic variation and phylogeography of the bank vole (*Clethrionomys glareolus*, Arvicolinae, Rodentia) in Russia with special reference to the introgression of the mtDNA of a closely related species, red-backed vole (*Cl. rutilus*). *Russian Journal of Genetics* **45**, 533-545.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Ballard JWO, Kreitman M (1995) Is mitochondrial DNA a strictly neutral marker? *Trends in Ecology and Evolution* **10**, 485-488.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology* **13**, 729-744.

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular biology and evolution* **16**, 37-48.

Barton N, Hewitt G (1985) Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics* **16**, 113-148.

Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations., Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Boratynski Z, Alves PC, Berto S*, et al.* (2011) Introgression of mitochondrial DNA among Myodes voles: consequences for energetics? *BMC Evolutionary Biology* **11**, 355.

Boratynski Z, Melo-Ferreira J, Alves PC*, et al.* (2014) Molecular and ecological signs of mitochondrial adaptation: consequences for introgression? *Heredity* **113**, 277-286.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**, 621-631.

Daguin C, Bonhomme F, Borsa P (2001) The zone of sympatry and hybridization of *Mytilus edulis* and *M. galloprovincialis*, as described by intron length polymorphism at locus *mac-1*. *Heredity* **86**, 342-354.

Deffontaine V, Libois R, Kotlik P*, et al.* (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (Clethrionomys glareolus). *Molecular Ecolology* **14**, 1727-1739.

Dowling DK, Friberg U, Lindell J (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in Ecology & Evolution* **23**, 546-554.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359-361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Filipi K, Markova S, Searle JB, Kotlik P (2015) Mitogenomic phylogenetics of the bank vole Clethrionomys glareolus, a model system for studying end-glacial colonization of Europe. *Mol Phylogenet Evol* **82 Pt A**, 245-257.

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34**, 397-423.

Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Cabios* **12**, 543-548.

Gay L, Crochet PA, Bell DA, Lenormand T (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution* **62**, 2789-2806.

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Guillot G, Estoup A, Mortier F, Cosson JF (2005) A spatial statistical model for landscape genetics. *Genetics* **170**, 1261-1280.

Guillot G, Santos F (2009) A computer program to simulate multilocus genotype data with spatially autocorrelated allele

frequencies. *Mol Ecol Resour* **9**, 1112-1120.

Hedrick PW (2010) Cattle ancestry in bison: explanations for higher mtDNA than autosomal ancestry. *Molecular Ecology* **19**, 3328-3335.

Henttonen H (2000) Long-term dynamics of the bank vole Clethrionomys glareolus at Pallasjärvi, northern Finnish taiga. In: *Bank vole biology: Recent advances in the population biology of a model species* (ed. Bujalska GaH, L.). Polish Journal of Ecology.

Henttonen H, Wallgren H (2001) Small rodent dynamics and communities in the birch forest zone of northern Fennoscandia. In: *Nordic mountain birch forest ecosystem* (ed. Wielgolaski FE), pp. 261-278. CRC Press.

Irwin DE (2012) Local adaptation along smooth ecological gradients causes phylogeographic breaks and phenotypic clustering. *The American Naturalist* **180**, 35-49.

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution* **32**, 128-144.

IUCN (2015) *The IUCN Red List of Threatened Species*. http://www.iucnredlist.org

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405.

Kingston SE, Jernigan RW, Fagan WF, Braun D, Braun MJ (2012) Genomic variation in cline shape across a hybrid zone. *Ecol Evol* **2**, 2737-2748.

Korpela K, Delgado M, Henttonen H, *et al.* (2013) Nonlinear effects of climate on boreal rodent dynamics: mild winters do not negate high-amplitude cycles. *Global Change Biology* **19**, 697-710.

Korpela K, Helle P, Henttonen H, *et al.* (2014) Predator-vole interactions in northern Europe: the role of small mustelids revised. *Proceedings of the Royal Society of Biological Sciences* **281**.

Kruuk LEB, Baird SJE, Gale KS, Barton NH (1999) A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics* **153**, 1959-1971.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics* **7**, 295-302.

Lebreton JD, Burnham KP, Clobert J, Anderson DR (1992) Modeling survival and testing biological hypotheses using marked animals - a unified approach with case-studies. *Ecological Monographs* **62**, 67-118.

Lenormand T, Guillemaud T, Bourguet D, Raymond M (1998) Evaluating gene flow using selected markers: a case study. *Genetics* **149**, 1383-1392.

Osipova OV, Soktin AA (2008) Experimental simulation of ancient hybridization between bank and red voles. *Doklady Biological Sciences* **420**, 169-171.

Perez M, Leblois R, Livoreil B, *et al.* (2012) Effects of landscape features and demographic history on the genetic structure of Testudo marginata populations in the southern Peloponnese and Sardinia. *Biological Journal of the Linnean Society* **105**, 591-606.

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution* **24**, 386-393.

Potapov SG, Illarionova NA, Andreeva TA, *et al.* (2007) Transfer of mitochondrial genome of the northern redbacked vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*) in northwestern Europe. *Doklady Biological Sciences* **417**, 435-438.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Putman AI, Carbone I (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol Evol* **4**, 4399-4428.

Rheindt FE, Edwards SV (2011) Genetic introgression: an integral but neglected component of speciation in birds. *The Auk* **128**, 620-632.

Rieux A, Lenormand T, Carlier J, de Lapeyre de Bellaire L, Ravigne V (2013) Using neutral cline decay to estimate contemporary dispersal: a generic tool and its application to a major crop pathogen. *Ecol Lett* **16**, 721-730.

Rousset F (1999a) Genetic differentiation in populations with different classes of individuals. *Theoretical Population Biology* **55**, 297-308.

Rousset F (1999b) Genetic differentiation within and between two habitats. *Genetics* **151**, 397–407.

Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58-62.

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Rozas J (2009) DNA sequence polymorphism analysis using DnaSP. *Methods in Molecular Biology* **537**, 337-350.

Sagnard F, Deu M, Dembele D*, et al.* (2011) Genetic diversity, structure, gene flow and evolutionary relationships within the Sorghum bicolor wild-weedy-crop complex in a western African region. *Theorical and Applied Genetics* **123**, 1231-1246.

Silva G, Lima FP, Martel P, Castilho R (2014) Thermal adaptation and clinal mitochondrial DNA variation of European anchovy. *Proceedings of the Royal Society B* **281**.

Team RC (2014) *R: a language and environment for statistical computing* R Foundation for Statistical Computing, Vienna, Austria.

Team RDC (2007) *A language and environment for statistical computing*, Vienna, Austria.

Tegelström H (1987) Transfer of mitochondrial DNA from the Northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*Clethrionomys glareolus*). *Journal of Molecular Evolution* **24**, 218-227.

Toews DP, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21**, 3907-3930.

Wirtz P (1999) Mother species-father species: unidirectional hybridization in animals with female choice. *Animal Behaviour* **58**, 1-12.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* **17**, 2107-2121.

Molecular Ecology

SUPPORTING INFORMATION

Testing the presence of a secondary contact between two distant mitochondrial lineages of the bank vole (*Myodes glareolus*) in central Finland

Julie Pisano, Nathalie Charbonnel, Jean-François Cosson, Sylvain Piry, Maxime Galan, Otso Huitu, Heikki Henttonen, Raphaël Leblois, Johan R. Michaux

Figure S1 Scatterplots from PCA and DAPC for 442 bank vole specimens from the 22 sampling sites. (A) PCA of the first two components. Dots represent specimens and colours indicate sampling sites, which were structured in a North-South axis. (B) Graph of BIC values in relation to the number of clusters and complemented by the *a*-score optimization graph that allows identifying the optimal number of principal components to use in the DAPC. (C) DAPC inferred from the first two components for $K = 2$. (D) DAPC inferred from the first two components for $K = 22$. Dots represent specimens and colours indicate sampling sites, which were structured in a North-South axis.

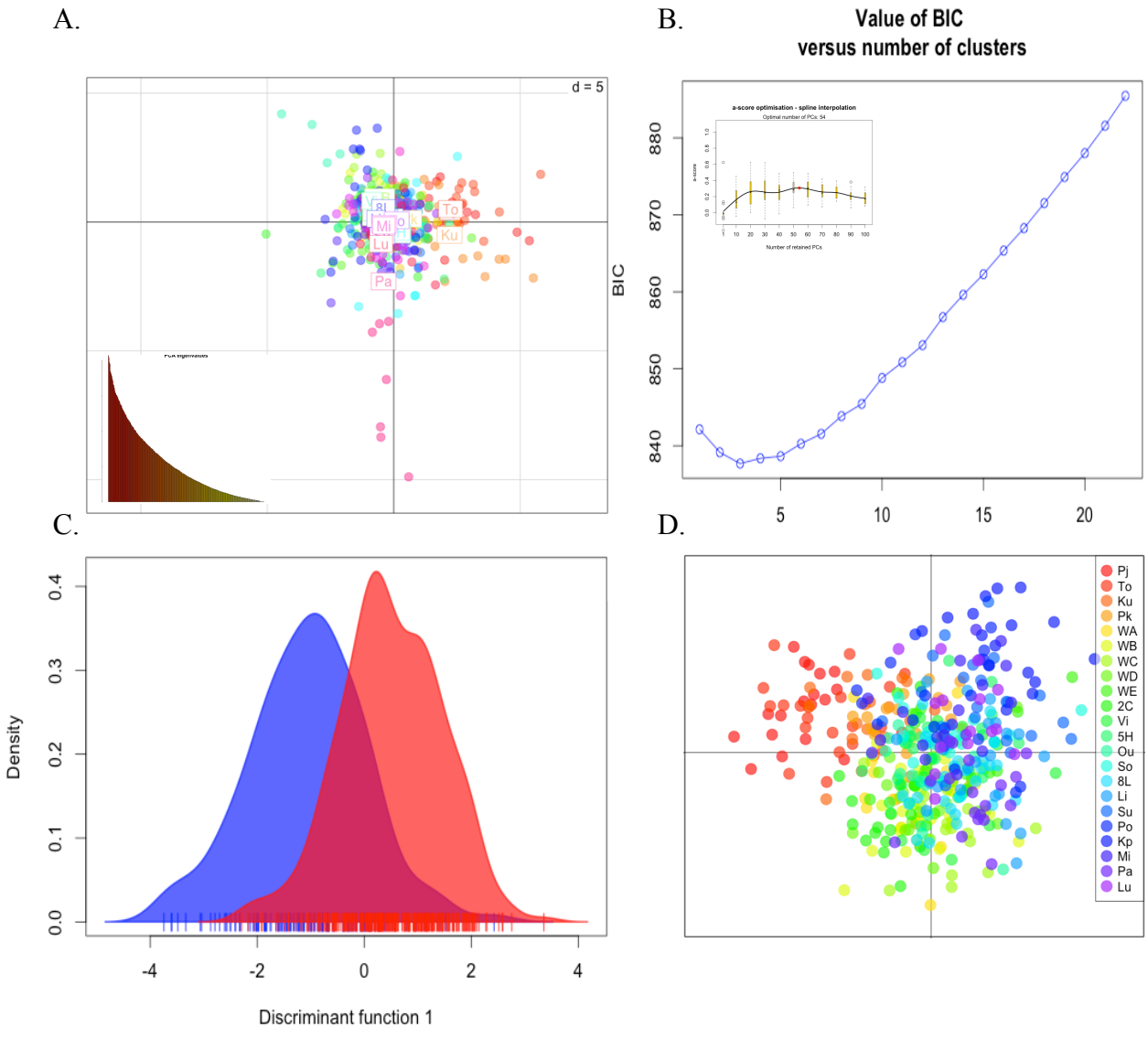
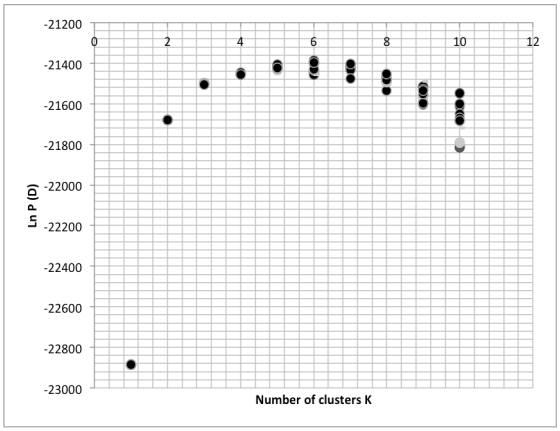


Figure S2 Mean posterior probability values per cluster. Structure analyses were run ten times for each *K*-value (*K* was estimated from one to ten) under the model with admixture and correlated allele frequencies. Analyses were performed on the (A) 10-microsatellite dataset including all specimens of *M. glareolus* and *M. rutilus*, and (B) 17-microsatellite dataset including only *M. glareolus* specimens.

A.



B.

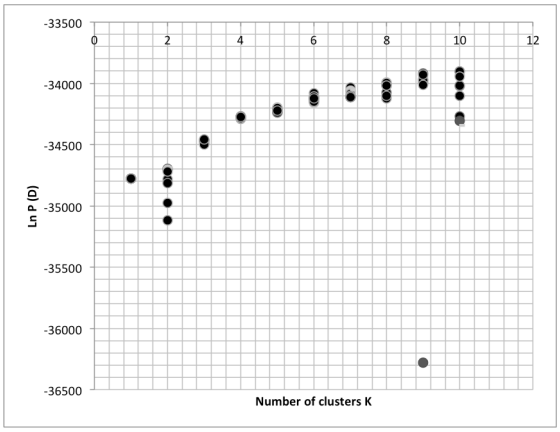
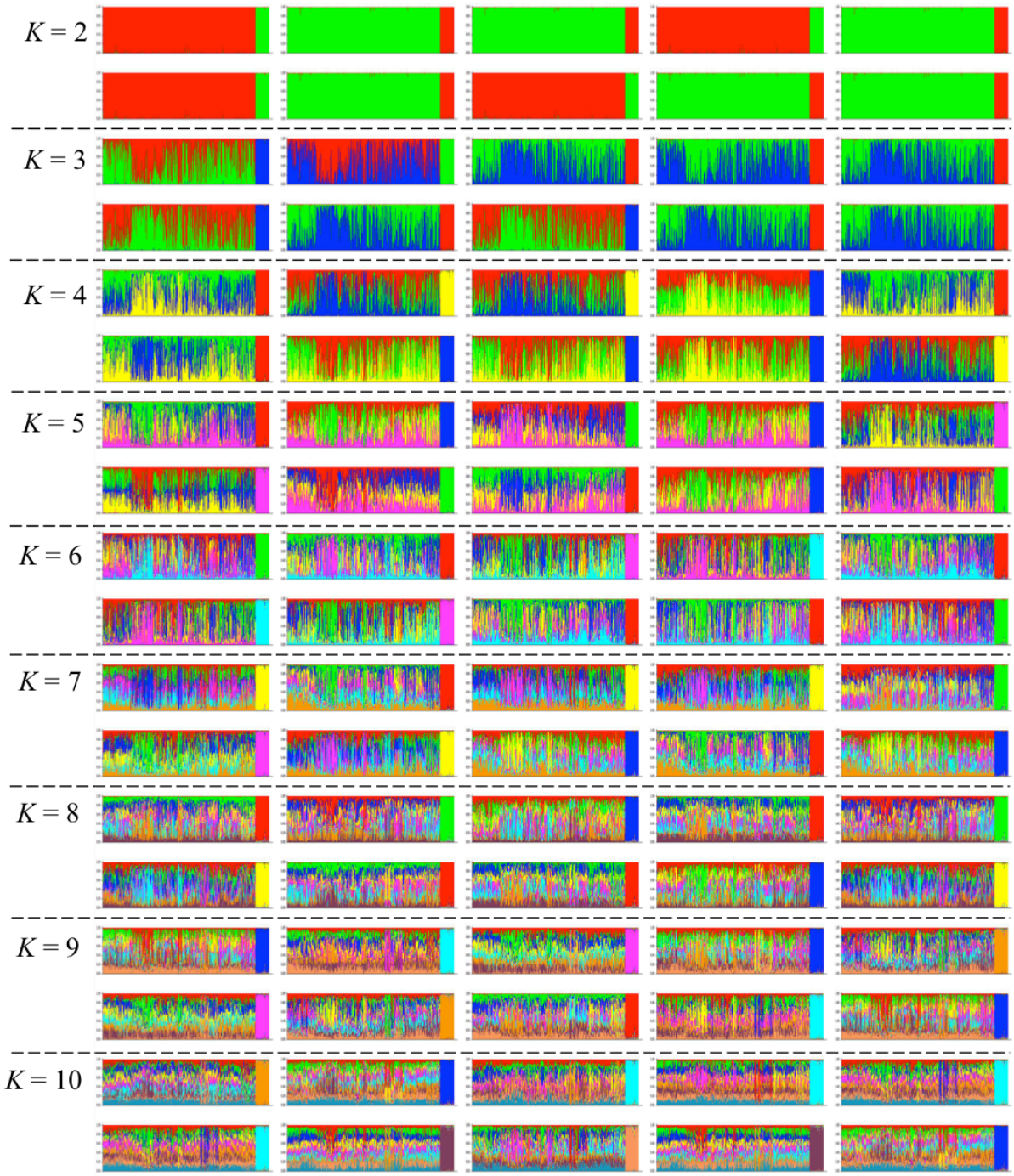


Figure S3 Estimated population structure from *STRUCTURE* analyses for $K = 2$ to $K = 10$. Analyses were run ten times for each K-value (K was estimated from one to ten) under the model with admixture and correlated allele frequencies. Each individual is represented by a thin vertical line divided in K coloured segments representing the estimated membership fractions in the distinct K clusters. Specimens of *M. glareolus* are arranged based on a North-South gradient. Analyses were performed on the (A) 10-microsatellite dataset including all specimens of *M. glareolus* and *M. rutilus* (*M. rutilus* specimens are situated on the right of barplots), and (B) 17-microsatellite dataset including only *M. glareolus* specimens.

A.



B.

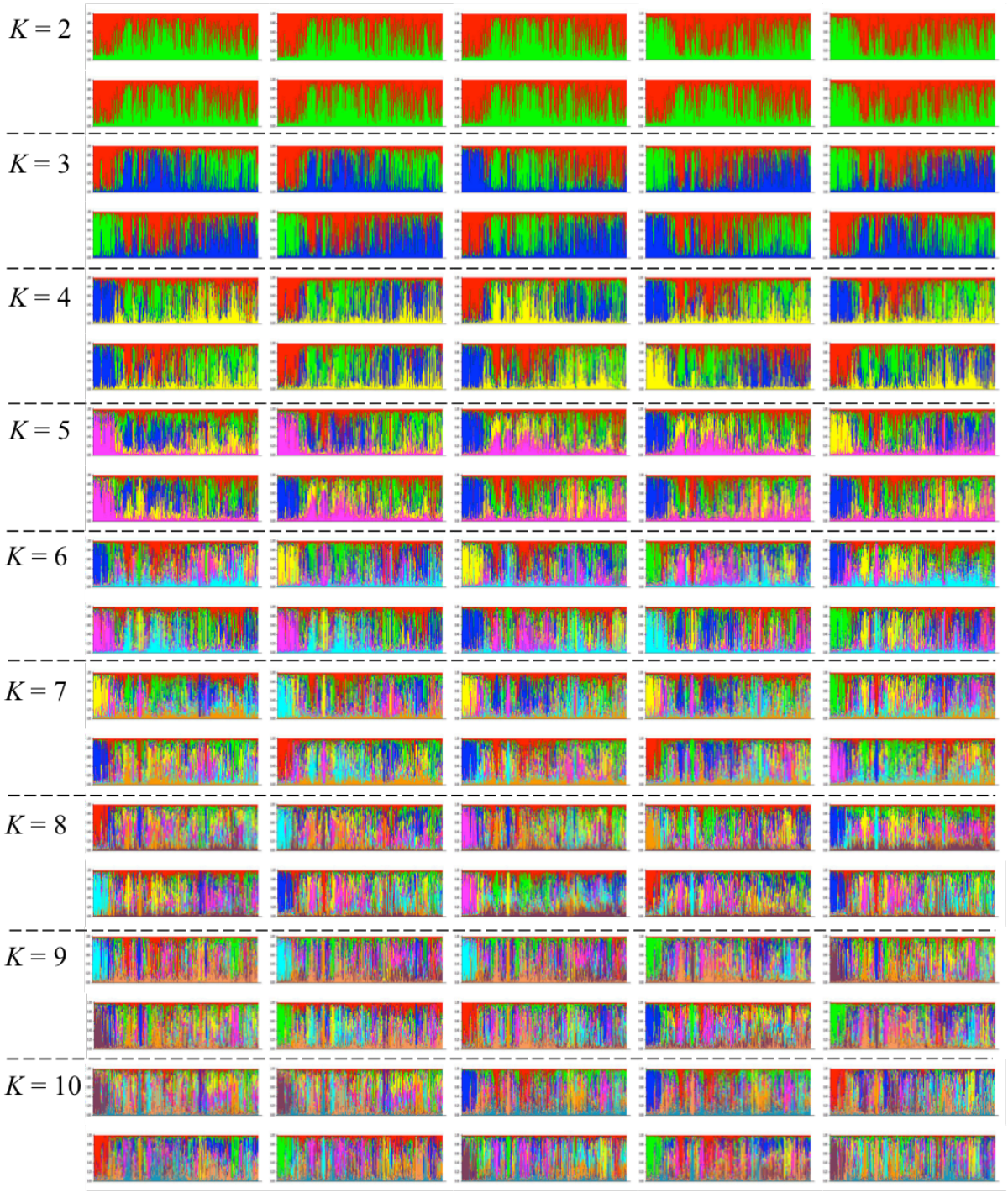


Table S1 Genetic variation at the 17-microsatellite loci for sampling site populations of *M. glareolus* and at 10-microsatellite loci for sampling site populations of *M. rutilus*. For each site, the sample size (N), the allelic richness estimated for a sample size of 14 individuals (*A*), the observed (H_o) and unbiased expected (H_e) heterozygosity and the inbreeding coefficient of Weir & Cockerham (F_{is}) are given. Significant F_{is} values ($P \leq 0.01$) are indicated with an asterisk '*'. Hashtags '#' indicate *M. rutilus* sampling sites.

Sampling sites	N	*A*	H_e	H_o	F_{is}
Pallasjärvi (Pj)	19	8.6	0.84	0.83	0.013
Tornio (To)	21	8.9	0.85	0.79	0.068*
Kuusamo (Ku)	19	9.0	0.84	0.85	-0.010
Puolanka (Pk)	21	9.1	0.85	0.85	0.011
Site WA	20	8.4	0.84	0.83	0.024
Site WB	24	8.6	0.85	0.82	0.040
Site WC	25	9.2	0.87	0.82	0.050*
Site WD	26	9.4	0.86	0.81	0.055*
Site WE	21	8.8	0.86	0.85	0.002
Site 2C	20	8.4	0.84	0.81	0.023
Site 5H	19	9.2	0.85	0.85	0.006
Site 8L	28	8.4	0.83	0.83	0.009
Vihanti (Vi)	19	8.9	0.86	0.85	0.004
Oulainen (Ou)	14	8.7	0.86	0.81	0.063*
Sotkamo (So)	17	9.6	0.86	0.84	0.024
Lieksa (Li)	23	8.7	0.84	0.84	0.006
Suonenjoki (Su)	20	8.7	0.85	0.85	-0.004
Parkano (Po)	16	9.6	0.87	0.87	0.005
Korpilahti (Kp)	19	8.5	0.84	0.86	-0.027
Mikkeli (Mi)	18	8.8	0.85	0.85	0.005
Parikkala (Pa)	20	8.7	0.84	0.82	0.018
Luumäki (Lu)	15	9.9	0.86	0.84	0.027
# Kilpisjärvi (Kj)	20	6.4	0.71	0.66	0.067
# Pallasjärvi (Pj)	19	8.1	0.79	0.76	0.042

Table S2 Membership probability and the associated number of population inferred using Geneland. (A) The allele frequency model was set to "uncorrelated", and *K* was explored with values ranging from one to four with a starting value of two. (B) *K* was explored in the same way but the allele frequency model was set to "correlated". (C) The correlated allele frequency model was applied, and *K* was constrained with values ranging from one to two with a starting value of one.

Analysis	Run	Number of population	Support for the number of population	Likelihood
A	7	1	69 %	-44974.00
	8	1	68 %	-44988.84
	2	1	70 %	-45021.90
	4	1	70 %	-45051.42
	10	1	71 %	-45102.89
	5	1	71 %	-45109.75
	6	1	72 %	-45137.64
	3	1	72 %	-45151.34
	1	1	72 %	-45192.13
	9	1	70 %	-45289.63
B	4	4	100 %	-37007.98
	5	4	100 %	-37504.40
	6	4	100 %	-37518.25
	7	4	100 %	-37961.03
	3	4	100 %	-38458.52
	9	4	100 %	-38474.17
	10	4	100 %	-38475.53
	1	4	100 %	-38550.05
	8	4	100 %	-38557.14
	2	4	100 %	-38599.64
C	5	2	100 %	-42005.12
	1	2	100 %	-42147.01
	8	2	100 %	-42154.10
	9	2	100 %	-42157.42
	3	2	100 %	-42158.73
	10	2	100 %	-42174.25
	7	2	100 %	-42254.37
	6	2	100 %	-42309.30
	2	2	100 %	-42374.04
	4	2	100 %	-43017.63

Table S3 Testing cline pattern in nuclear and mitochondrial markers. Clines were fitted using scaled logit function (equation in Appendix S6 Rieux *et al.* (2013)). For each Locus, the cline shape, the centre and the slope were tested using the corrected Akaike Criterion against a null model assuming homogeneous allele frequency over the study area. '/' indicates non-significant deviance between the null and the cline models (consequently, *c and b* was not inferred).

Marker	Model	Number of parameters	Maximum Log likelihood	% of Deviance	AICc	ΔAICc	Centre *c*	Slope *b*
Cg10A11	Null	1	-434.386		442.064			
	Cline	4	-412.103	35.536	425.489	16.574	499.800	0.005
	Saturated	22	-394.335					
Cg15F7	Null	1	-472.685		517.698			
	Cline	4	-462.509	32.997	512.596	5.102	356.203	0.038
	Saturated	22	-446.010					
Cg1F11	Null	1	-412.831		643.275			
	Cline	4	-391.867	23.176	616.711	26.564	345.992	2.128
	Saturated	22	-380.279					
Cg13G2	Null	1	-468.405		525.157			
	Cline	4	-454.069	32.232	515.146	10.011	412.239	0.673
	Saturated	22	-437.953					
Cg16E4	Null	1	-435.354		582.684			
	Cline	4	-412.853	26.990	558.672	24.012	390.873	0.013
	Saturated	22	-399.358					
Cg5F11	Null	1	-461.838		651.798			
	Cline	4	-440.212	25.587	627.371	24.427	414.075	0.618
	Saturated	22	-427.419					
Cg13F9	Null	1	-442.524		668.645			
	Cline	4	-436.671	23.897	665.828	2.817	826.601	0.004
	Saturated	22	-424.722					
Cg3A8	Null	1	-464.918		511.034			
	Cline	4	-455.986	32.880	507.254	3.780	413.039	1.135
	Saturated	22	-439.546					
Cg16A3	Null	1	-459.945		477.910			
	Cline	4	-441.549	34.792	464.875	13.035	460.036	0.011
	Saturated	22	-424.153					
Cg13C12	Null	1	-462.878		653.906			
	Cline	4	-446.299	25.561	636.557	17.349	184.142	0.003
	Saturated	22	-433.518					
Cg17C9	Null	1	-441.933		680.118			
	Cline	4	-421.168	23.461	654.256	25.863	411.704	0.075
	Saturated	22	-409.437					
Cg4F9	Null	1	-455.262		1175.758			
	Cline	4	-449.392	13.963	1166.625	9.133	339.500	6.729
	Saturated	22	-442.410					
Cg3F12	Null	1	-448.973		556.549			
	Cline	4	-435.999	29.146	546.525	10.025	556.356	0.011
	Saturated	22	-421.426					
Cg14E1	Null	1	-415.359		319.514			
	Cline	4	-389.400	47.094	305.670	13.844	367.258	0.034
	Saturated	22	-365.853					
Cg6G11	Null	1	-466.146		486.231			
	Cline	4	-458.449	34.655	484.236	1.995	/	/
	Saturated	22	-441.121					
Cg2C5	Null	1	-457.954	34.518	479.612	3.790	360.502	0.011
	Cline	4	-448.567		475.822			

	Saturated	22	-431.308					
Cg12A7	Null	1	-467.144		460.625	19.473		
	Cline	4	-441.198	36.669	441.152		415.504	0.048
	Saturated	22	-422.864					
cytb	Null	1	-277.497		29.790			
	Cline	4	-97.760	359.474	17.790	12	368.034	-0.081
	Saturated	22	-87.624					

Chapitre 5:

Discussion Générale & Perspectives

La biodiversité sur la Terre n'est pas répartie de façon homogène dans le temps, à travers les régions et entre les groupes biologiques de l'arbre de vie. De nos jours, la diversité du vivant est unanimement reconnue comme étant en péril. L'activité humaine et les changements environnementaux associés nous amènent aujourd'hui à parler de la 6ème grande extinction (Raymond & Godelle, 2010; Barnosky *et al.*, 2011). Dans ce contexte, afin de pouvoir mettre en place des stratégies de conservation de la diversité biologique, il est important de bien comprendre quels sont les facteurs qui influencent l'histoire évolutive des organismes (Stephens & Wiens, 2003; Ricklefs, 2004; Mittelbach *et al.*, 2007; Rabosky, 2009; Wiens, 2011; Rabosky, 2013; Title & Burns, 2015). Lors de cette thèse, me suis intéressée à l'évolution de la biodiversité à deux échelles : macroévolutive (les Dipodoidea) et microévolutive (les populations de campagnol roussâtre de Finlande). Je me suis posée la question de savoir ce qui a influencé l'histoire évolutive biogéographique de la superfamille des Dipodoidea et de ce qui régit la distribution spatiale des populations du campagnol roussâtre (*M. glareolus*) en Finlande.

5.1 Mieux comprendre la biodiversité des Dipodoidea

5.1.1 Quelle est la cause des controverses taxonomiques issues du siècle dernier ?

Avant cette thèse, la phylogénie des Dipodoidea n'avait jamais été étudiée qu'à partir de marqueurs morphologiques, et ceux-ci avaient conduit à de nombreuses controverses à propos de la taxonomie et de la systématique de cette superfamille (e.g. Vinogradov (1930); Shenbrot (1992)). Une hypothèse permettant d'expliquer ces controverses est que les marqueurs morphologiques étudiés lors du siècle dernier étaient sensibles aux homoplasies, mais cela n'a jamais pu être vérifié car les homoplasies ne sont pas directement détectables dans des phylogénies morphologiques. Ainsi, une manière de les détecter est de comparer des phylogénies morphologiques avec des phylogénies moléculaires robustes (Wake *et al.*, 2011).

Comme les études précédentes, nos phylogénies reconstruites à partir de caractères morphologiques étaient discordantes entre elles (voir chapitre 2 ; Lebedev *et al.* (2012)). Par conséquent, pour tester l'hypothèse de l'homoplasie, notre phylogénie moléculaire fortement soutenue et résolue a été comparée avec celle reconstruite à partir des marqueurs

morphologiques et il n'a pas été étonnant de remarquer que les phylogénies moléculaire et morphologique étaient discordantes. Dans Lebedev *et al.* (2012), nous avons montré que la raison de ces discordances observées entre les phylogénies morphologiques était due à des états de caractères semblables chez des espèces non-directement apparentées (i.e. à du parallélisme de convergence), lesquels ont donc brouillé le signal phylogénétique. Par conséquent, étant donné que l'homoplasie des caractères morphologiques amène à des discordances de topologies, il semble que les marqueurs morphologiques ne soient pas les plus adaptés pour reconstruire la phylogénie des Dipodoidea. Donc, cette thèse a démontré l'importance d'utiliser des marqueurs moléculaires (e.g. des séquences de gènes) pour étudier la phylogénie au sein des Dipodoidea.

5.1.2 Vers une nouvelle systématique et taxonomie des Dipodoidea

La taxonomie décrite dans *Mammal Species of the World* suggérant que les 51 espèces des 16 genres de Dipodoidea se regroupent en la seule famille des Dipodidae a longtemps servi de référence. Hors, cette ancienne taxonomie a été rejetée par Lebedev *et al.* (2012), qui ont proposé une nouvelle taxonomie de la superfamille des Dipodoidea au sein de laquelle trois familles et 19 genres sont désormais reconnus. L'étude biogéographique des Dipodoidea développée dans le cadre de cette thèse incluant 14 espèces supplémentaires (Pisano *et al.*, 2015) a permis de confirmer la taxonomie de Lebedev *et al.* (2012). Néanmoins, étant donné que cette phylogénie, la plus exhaustive à ce jour, n'incluait que 34 des 51 espèces et 15 des 19 genres (Pisano *et al.*, 2015), tous les liens de parenté n'ont pas pu être résolus.

Selon l'ancienne taxonomie (Holden & Musser, 2005), il nous manquait le genre *Salpingotulus* dans nos études sur les Dipodoidea. Ce dernier était précédemment décrit comme faisant partie intégrante de la sous-famille des Cardiocraniinae au même titre que le genre *Salpingotus* et *Cardiocranius*. Cependant, selon la nouvelle taxonomie, *Salpingotulus* fait désormais partie du genre *Salpingotus*. Ainsi, la sous-famille des Cardiocraniinae ne comprend plus que deux genres. Aussi, selon l'ancienne taxonomie (Holden & Musser, 2005), la sous-famille des Allactaginae ne comprenait que les genres *Allactaga, Pygeretmus* et *Allactodipus*. Or, suite à la découverte de la paraphylie du genre *Allactaga* lors de cette thèse, de nouveaux genres ont été proposés dans la nouvelle taxonomie : *Orientallactaga, Scarturus, Paralactaga et Microallactaga.* Par conséquent, la sous-famille des Allactaginae ne

comprendrait plus uniquement trois mais sept genres. Cependant, pour justifier l'utilisation de chacun de ces noms de genres, il faudrait donc reconstruire une phylogénie – de préférence moléculaire de manière à éviter les homoplasies liées aux caractères morphologiques (cf. 5.1.1) – la plus complète possible des Allactaginae pour avoir une meilleure connaissance des relations phylogénétiques entre les espèces de cette sous-famille et déterminer si ces nouveaux noms de genre respectent bien la phylogénie.

De nouvelles espèces de Dipodoidea ont récemment été proposées. Une nouvelle espèce de gerboise, *Allactaga toussi* a été suggérée sur base de caractères morpho-anatomiques (Darvish *et al.*, 2008). Or, nous avons vu que les caractères morphologiques ont de fortes probabilités d'être soumis à de l'homoplasie. D'ailleurs, ces homoplasies nous ont amené à considérer le genre *Allactaga* comme paraphylétique. Par conséquent, avant de la valider en tant que nouvelle espèce de gerboise, il serait intéressant d'inclure des spécimens de *A. toussi* dans une phylogénie moléculaire des Allactaginae et de procéder par exemple à des analyses complémentaires aux approches de phylogénie moléculaire classique en appliquant par exemple la méthode développée par Pons *et al.* (2006). Cette méthode permet d'estimer sans *a priori* la délimitation entre les espèces sur la base de séquences de gènes en établissant la limite 'intraspécifique – interspécifique' dans la phylogénie. Cette approche pourrait d'ailleurs être utilisée pour établir le statut taxonomique d'autres espèces de gerboises. Notamment, Ben Faleh *et al.* (2010) ont suggéré l'existence d'une espèce cryptique de *Jaculus* en Tunisie*, J. deserti* sur base de l'étude du cytochrome *b*, de caractères morphométriques et cytogénétiques, et nos résultats (Chapitre 3, Annexe 2, Pisano *et al.* (2015)) semblent d'ailleurs appuyer cette hypothèse. Cependant, afin de confirmer l'existence de cette espèce cryptique, il conviendrait d'estimer la délimitation entre les espèces de *Jaculus* de Tunisie sur la base de données moléculaires en appliquant par exemple, la méthode de Pons *et al.* (2006).

En conclusion, grâce à notre échantillonnage exhaustif des Zapodidae et des Dipodinae, les relations phylogénétiques entre chacune des espèces de souris-sauteuses et de gerboises dipodines ont pu être reconstruites, validant ainsi la nouvelle taxonomie de ces deux groupes. Il reste donc à récolter l'échantillonnage le plus complet possible des Sminthidae, des Cardiocraniinae, et des Allactaginae de manière à inférer la phylogénie moléculaire de ces familles ou sous-familles de Dipodoidea et à terme, valider la taxonomie de ces groupes.

5.1.3 Perspectives : Pourquoi les Dipodoidea sont-ils si peu diversifiés par rapport à leur groupe frère, les Muroidea ?

Nous l'avons vu lors de cette thèse : la biodiversité n'est pas répartie de façon homogène à travers le temps et l'espace, et de nombreux facteurs biotiques et abiotiques peuvent influencer ces patrons spatiaux et temporels de diversité. Cependant, ce que nous n'avons pas eu le temps d'aborder lors de cette thèse est de comprendre: pourquoi la biodiversité n'est pas répartie de façon homogène à travers les groupes taxonomiques ? Pourquoi certains clades contemporains présentent une richesse spécifique incroyable alors que d'autres sont pauvres en espèces ? Aussi, il serait intéressant de savoir ce qui cause ces disparités de richesse spécifique entre les clades et quels sont les facteurs biotiques et abiotiques qui ont influencé leur spéciation et/ou leur extinction.

L'un des exemples les plus frappants d'inégalité de richesse spécifique entre des groupes frères de Mammifères concerne les deux superfamilles constitutives du sous-ordre des Myomorpha, les Dipodoidea (51 espèces) et les Muroidea (~1517 espèces, plus de 50% de la diversité de l'ordre des Rongeurs) (Musser & Carleton, 2005; Fabre *et al.*, 2012; Pisano *et al.*, 2015). Actuellement, les Muroidea sont présents dans presque tous les continents sauf l'Antarctique et occupent presque tous les habitats disponibles sur Terre, s'étendant au Nord dans les forêts de Toundra à travers les forêts de zones tempérées et tropicales, et même jusqu'aux régions sèches des déserts et des savanes (Musser & Carleton, 2005; Schenk *et al.*, 2013; Myers *et al.*, 2015). Comme décrit précédemment, les Dipodoidea – ~2% de la diversité des Muroidea – occupent au contraire majoritairement la région Eurasienne et de façon plus anecdotique, l'Afrique du Nord et l'Amérique du Nord (Holden & Musser, 2005; Pisano *et al.*, 2015). La majorité des espèces de Dipodoidea (particulièrement les gerboises qui représentent 60% de la diversité des Dipodoidea) sont en réalité adaptées à des environnements secs et chauds tels les déserts, les semi-déserts et les steppes (Holden & Musser, 2005; Pisano *et al.*, 2015). Nous pouvons donc nous demander : pourquoi les Dipodoidea possèdent si peu d'espèces en comparaison avec leur groupe frère, les Muroidea ? Des tests de diversification peuvent permettre de répondre à cette question. Ces approches sont basées sur la phylogénie moléculaire pour estimer la trajectoire de la diversité au cours du temps (Morlon *et al.*, 2011; Stadler, 2011). Elles proposent de tester plusieurs hypothèses pour tenter de mieux comprendre ces différences de richesse spécifique entre clades (Alfaro *et al.*, 2009; Benton, 2009; Rabosky *et al.*, 2012; Mahler *et al.*, 2013).

Comparer deux groupes frères dont la richesse spécifique contemporaine est significativement différente représente une belle opportunité pour mieux comprendre comment et pourquoi les diversités des espèces évoluent différemment à travers le temps. Les groupes-frères sont issus du même ancêtre commun et ont eu le même temps pour évoluer. Par conséquent, l'hypothèse de la « spéciation en fonction du temps » qui suppose que le groupe le plus riche est celui qui a bénéficié de plus de temps pour se diversifier peut directement être rejetée (Stephens & Wiens, 2003; Benton, 2009). Par contre, il se peut, par exemple, que les Muroidea soient plus riches en nombre d'espèces grâce à des taux de diversification plus élevés, qu'ils aient connu des habitats qui ont favorisé leur dynamique de diversité…

Plusieurs hypothèses ont été proposées pour tenter de comprendre ces disparités de richesse spécifique entre les clades : (1) L'hypothèse de « diversité-dépendance » suggère que le nombre d'espèces dans un écosystème donné soit limité dans l'espace à cause d'une capacité d'accueil limitée, de ressources limitées ou encore de compétition entre les espèces (e.g. les Canidae d'Amérique du Nord ont été remplacés par les Felidae qui sont de meilleurs compétiteurs en milieu de savane) (Rabosky, 2009; Rabosky, 2013; Silvestro *et al.*, 2015). Phillimore and Price (2008) a d'ailleurs montré que les taux de spéciation pouvaient augmenter en relation avec les disponibilités de niches écologiques vacantes et diminuer au fur et à mesure où les niches se remplissaient, et que les taux de spéciation d'un clade pouvaient augmenter lorsque les taux de diversification du clade le plus riche en espèce diminuaient (i.e. que la spéciation diminuait et que l'extinction augmentait). (2) L'hypothèse des « hotspots de diversification » propose qu'un clade ou une région soit plus diversifiée simplement grâce à des taux de spéciation plus élevés et/ou d'extinction plus faibles (Mittelbach *et al.*, 2007). Par exemple, le gradient latitudinal de diversité montre que la richesse spécifique de groupes de mammifères (e.g. Primates, Lagomorphes, Rongeurs) et d'insectes (Papilionidae) augmente des pôles vers les tropiques grâce à des taux de spéciation élevés et d'extinction faibles en région tropicale (Condamine *et al.*, 2012; Rolland *et al.*, 2014). (3) L'hypothèse de « l'extinction massive » suggère que certains clades ou régions ont connu des évènements d'extinction importants qui ont réduit leur richesse spécifique de façon drastique (Morlon *et al.*, 2011). Par exemple, en congruence avec ce que suggère le registre fossile des cétacés, Morlon *et al.* (2011) ont montré que la diversité des cétacés a soudainement fortement augmenté vers 30 Ma et qu'il y a 10 Ma, leur diversité a soudainement diminué (Fig. 2 - Introduction). Les périodes d'extinction pourraient donc aussi

expliquer pourquoi certains groupes possèdent actuellement si peu d'espèces (Condamine & Hines, 2015).

Pour comprendre cette disparité entre les Muroidea et les Dipodoidea, plusieurs étapes sont nécessaires. La première est de collecter un jeu de données moléculaires composé du plus grand nombre d'espèces possible de Muroidea de la région Holarctique (i.e. qui sont distribuées sur l'aire de distribution des Dipodoidea) et de Dipodoidea. J'ai réalisé ce travail pendant ma thèse : nous avons pour cela sélectionné un gène mitochondrial (le cytochrome *b*) et quatre gènes nucléaires (BRCA1, GHR, IRBP, RAG1) qui ont déjà prouvé leur utilité pour résoudre de façon robuste la phylogénie de divers groupes de rongeurs (Fabre *et al.*, 2012; Pisano *et al.*, 2015). La matrice des gènes concaténés fait ~5000 nucléotides. Plus de 80% de la diversité spécifique des Dipodoidea (i.e. 41 des 51 espèces) et ~20% de la diversité totale des Muroidea (i.e. 276 des 1547 espèces) ont été collectées (Table S2). La seconde étape sera de reconstruire une phylogénie moléculaire robuste en utilisant des méthodes d'inférence bayésienne. Cette phylogénie sera ensuite placée dans un cadre temporel grâce à des approches de datation moléculaire calibrée à l'aide de données fossiles. La troisième étape sera de tester les hypothèses évoquées plus haut.

5.2 Mieux comprendre la biodiversité des campagnols roussâtres (*Myodes glareolus*) de Fennoscandie

Le patron phylogéographique du campagnol roussâtre a souvent été étudié sur base d'un marqueur mitochondrial (le gène du cytochrome *b*). Plusieurs lignées mitochondriales ont ainsi été mises en évidence en Europe et notamment, dans la péninsule de Fennoscandie (Deffontaine *et al.*, 2005; Kotlik *et al.*, 2006; Deffontaine *et al.*, 2009; Colangelo *et al.*, 2012). Dans cette région nordique, il existe deux fortes barrières limitant le flux génique mitochondrial : l'une en Finlande, et l'autre en Suède (Deffontaine *et al.*, 2005).

Lors de cette thèse, nous nous sommes intéressés à la barrière limitant le flux génique mitochondrial située en Finlande, i.e. la zone de contact entre les mitotypes GLA et RUT (cf. Chap. 4, lesquels sont aussi respectivement nommés 'lignée Est-européenne' et 'lignée Oural' dans Deffontaine *et al.* (2005)). Nous avons vu que cette zone de contact entre les mitotypes

GLA et RUT, correspondait bien à une zone de contact secondaire. Ce patron a été difficile à mettre en évidence du fait que les populations étaient caractérisées par de faibles taux de différenciation nucléaire. De façon générale, le mitotype GLA présente une faible diversité génétique sur l'ensemble de son aire de distribution, si bien que l'étude du cytochrome *b* des populations de GLA qui avait pour but d'identifier la zone refuge de GLA n'a pas permis de mettre en évidence un gradient de diversité nucléotidique même sur un large échantillonnage qui recouvrait la quasi-totalité de son aire de répartition (Deffontaine, 2008). Or, selon Hewitt (1999), pour identifier une zone refuge, il faut rechercher des populations qui présentent une diversité génétique plus élevée que les autres. Dès lors, cela pourrait expliquer pourquoi Deffontaine *et al.* (2005, 2008) n'ont pas pu clairement définir où le mitotype GLA aurait survécu aux dernières glaciations du Quaternaire. Si on émet l'hypothèse que la recolonisation postglaciaire de ce mitotype GLA s'est faite rapidement depuis la zone refuge à partir d'un petit effectif, cela pourrait expliquer le fait que l'on observe aujourd'hui une faible diversité génétique au sein de cette lignée (Hewitt, 1996).

Nous avons également observé lors de cette thèse que le mitotype RUT présente une très faible différenciation génétique nucléaire avec le mitotype GLA. Nous pensons dès lors que l'origine du mitotype RUT que l'on retrouve du centre de la Russie à travers la Finlande, jusqu'au centre de la Suède (Fig. 10) proviendrait d'un phénomène d'hybridation entre le mitotype GLA et *Myodes rutilus*. Le mitotype RUT (i.e. la lignée Oural) a une large zone de sympatrie avec *M. rutilus* en Russie et au Nord de la Finlande ; tandis que le mitotype GLA possède uniquement une petite aire de sympatrie avec *M. rutilus* en Finlande près de la zone de contact. Deux hypothèses avaient précédemment été proposées pour expliquer l'introgression mitochondriale de *M. rutilus* vers le mitotype RUT : une hybridation interspécifique (a) contemporaine ou (b) ancienne (i.e. datant du dernier âge glaciaire ou de la recolonisation postglaciaire) (Abramson *et al.*, 2009; Melnikova *et al.*, 2012). En fait, il n'y a pas de preuves de l'hybridation entre ces deux espèces dans la nature (Osipova & Soktin, 2008), hormis un seul individu sur deux marqueurs microsatellites (Melnikova *et al.*, 2012). D'ailleurs, même si l'hybridation est forcée en milieu expérimental, les femelles hybrides F1 sont certes fertiles et le rétrocroisement possible (ce qui justifie d'ailleurs l'introgression mitochondriale) mais les mâles F1 hybrides sont stériles (Osipova & Sotkin, 2006). Aussi, étant donné la confirmation de la zone de contact secondaire entre les mitotypes GLA et RUT, notre étude semble plutôt en faveur d'une hybridation ancienne entre *M. glareolus* – peut-être plus particulièrement, le mitotype GLA – et *M. rutilus*. D'ailleurs, la présence d'une

population de *M. rutilus* au sein du mitotype GLA en Finlande où aucun événement d'hybridation interspécifique n'a jamais été détecté (comm. pers. H. Henttonen) soutient cette hypothèse. Dès lors, l'hypothèse d'une hybridation interspécifique contemporaine semble peu probable, comme le suggère d'ailleurs Deffontaine (2008). Ainsi, l'introgression du génome mitochondrial de *M. rutilus* vers le mitotype GLA résulterait d'événements d'hybridation anciens.

Une autre barrière aux flux de gènes mitochondriaux entre des populations du campagnol roussâtre a été identifiée en Fennoscandie (Deffontaine *et al.*, 2005). Elle se situe au centre de la Suède. Elle différencie au Nord, la lignée introgressée (mitotype RUT) et au Sud, la lignée Ouest-européenne (Fig. 10) (Deffontaine *et al.*, 2005). Deffontaine *et al.* (2005) ont suggéré que cette lignée de campagnols aurait trouvé refuge dans les Carpates lors de la dernière glaciation. Par conséquent, la zone de contact suédoise devrait également être une zone de contact secondaire. Cette zone de contact suédoise est déjà connue comme une zone de suture (Jaarola *et al.*, 1999), car plusieurs zones de contact intraspécifiques de vertébrés (e.g. le campagnol agreste (*Microtus agrestis*), la musaraigne commune (*Sorex araneus*), l'ours brun (*Ursus arctos*)) se rassemblent dans cette région (Jaarola & Tegelström, 1995; Taberlet *et al.*, 1995; Fredga, 1996). Dès lors, il serait intéressant de tester cette hypothèse de contact secondaire en utilisant les mêmes marqueurs et le même protocole d'analyses que celui développé lors de cette thèse pour d'abord confirmer le contact secondaire entre ces deux lignées et ensuite, comparer la diversité génétique de ces deux zones de contact. Cela permettrait par exemple de déterminer si les campagnols de la lignée Ouest-européenne sont plus différenciés par rapport aux individus du mitotype RUT que les spécimens du mitotype GLA. Etant donné que la lignée Ouest-européenne aurait trouvé refuge dans les Carpathes et le mitotype RUT dans l'Oural, on attendrait un degré de différenciation élevé au sein de cette zone de contact suédoise entre les lignées de campagnol roussâtre situées de part et d'autres de la zone de contact.

Vu que les deux zones de contact fenno-scandinaves du campagnol roussâtre ont en commun le mitotype RUT et qu'elles diffèrent par les lignées qui entrent en contact, les processus évolutifs impliqués dans chacune des zones de contact pourraient donc être comparés, ce qui permettraient au final de mieux comprendre l'histoire évolutive récente du campagnol roussâtre.

Conclusion

Etudier la biodiversité en utilisant des approches macroévolutives et microévolutives est très intéressant car cela permet d'avoir un regard large sur la manière avec laquelle les espèces, les populations et leurs génomes évoluent. Au terme de cette thèse, je comprends d'autant mieux que se préoccuper de la préservation de la biodiversité actuelle n'est pas chose facile car de nombreux facteurs et diverses forces peuvent impacter l'évolution de la biodiversité. Sauver cette biodiversité en péril constitue donc un vrai challenge.

GLOSSAIRE

Abiotiques, facteurs Facteurs physico-chimiques du milieu (e.g. le climat, les perturbations géologiques) pouvant influencer la diversité du vivant

Aire ancestrale Zone géographique précédemment occupée par les ancêtres des taxa étudiés

Allopatrie Isolement géographique de deux populations d'une même espèce

Biodiversité Terme décrivant la variabilité parmi les êtres vivants et exprimant la diversité à plusieurs niveaux (génétique, morphologique, écologique, habitat etc.). La biodiversité englobe l'ensemble des êtres vivants sur la Terre.

Biogéographie Discipline qui étudie la diversité du vivant dans le temps et l'espace, et qui cherche à expliquer les raisons des répartitions géographiques observées

Biogéographie historique Science à l'interface entre l'écologie, la géologie, paléogéographie et la systématique (1) qui vise à reconstruire l'origine et l'histoire évolutive de divers taxa dans le temps et l'espace et (2) qui tente de mettre en évidence les facteurs abiotiques (e.g. mise en place d'une barrière géographique) ayant participé à la répartition actuelle des taxa. Elle permet de mieux comprendre 'où, quand et comment' les taxa sont nés et se sont diversifiés à la surface de la Terre. Elle se base donc principalement sur deux grands principes : l'évolution géologique – i.e. l'évolution des aires paléogéographiques découlant de la dérive des plaques tectoniques, des changements géologiques qui en résultent, et des changements climatiques passés – et l'évolution biologique – i.e. l'évolution des taxa impliquant la survie, l'adaptation, la dispersion, l'extinction, la diversification.

Biologie évolutive Domaine de la biologie qui vise à comprendre les scénarios et les mécanismes de l'évolution des organismes vivants

Biotiques, facteurs Ensemble des interactions du vivant sur le vivant (e.g. interactions entre espèces)

Caractère Tout attribut observable d'un organisme

Centre d'origine Zone géographique dans laquelle les taxa sont apparus

Clade Groupe monophylétique

Convergence Ressemblance apparue indépendamment dans divers taxa non-apparentés

Dispersion (en biogéographie) Processus de spéciation biogéographique où certaines populations périphériques iront établir des populations fondatrices dans de nouveaux espaces *< s'oppose à la vicariance >*

Evolution Transformation des organismes vivants via des changements de caractères génétiques ou morphologiques au cours des générations

Extinction géographique locale Extinction sur l'aire géographique précédemment occupée par les ancêtres des taxa, mettant en évidence des évènements de dispersion

Homoplasie Similarité (morphologique, moléculaire) chez plusieurs espèces ne provenant pas de l'héritage d'un ancêtre commun (e.g. convergence, parallélisme, réversion)

In situ Du latin, qui signifie 'sur place'

Introgression Incorporation de gènes d'une espèce dans le génome d'une autre espèce par hybridation

Monophylie Ensemble de taxa incluant un ancêtre hypothétique et tous ses descendants

Parallélisme de convergence Ressemblance apparue chez divers taxa proches mais non-héritée de l'ancêtre commun

Paraphylie Un groupe incluant un ancêtre hypothétique et une partie de ses descendants

Phylogénie Discipline qui vise à reconstruire les relations de parenté entre taxa. Elle répond à la question « qui est plus proche parent de qui ? »

Phylogéographie Discipline visant à établir les liens de parenté entre des populations géographiquement éloignées de la même espèce afin, par exemple, d'établir un scénario de colonisation de l'aire géographique occupée par certaines populations

Population Ensemble d'individus d'une même espèce coexistant dans un même habitat

Recombinaison Processus par lequel l'information génétique contenues dans deux chromosomes homologues est réorganisée pour aboutir à la formation de nouveaux chromosomes constituant un mélange des chromosomes parentaux

Synapomorphie Caractère dérivé partagé entre deux ou plusieurs taxa hérité de l'ancêtre commun

Systématique Science des classifications biologiques et des relations évolutives entre les organismes

Taxon Ensemble des organismes reconnus et définis dans la classification hiérarchisée

Taxonomie Science des lois de la classification visant à décrire, identifier et classer les organismes vivants

Vicariance Processus de spéciation biogéographique induit par une scission de l'aire de distribution d'une espèce *< s'oppose à la dispersion >*

Zone de contact secondaire Région où des lignées ayant évolué en allopatrie durant une certaine période de temps (e.g. durant le dernier âge glaciaire) reviennent en contact (e.g. lors de recolonisation postglaciaire).

REFERENCES BIBLIOGRAPHIQUES

Abramson, N.I., Rodchenkova, E.N. & Kostygov, A.Y. (2009) Genetic variation and phylogeography of the bank vole (*Clethrionomys glareolus*, Arvicolinae, Rodentia) in Russia with special reference to the introgression of the mtDNA of a closely related species, red-backed vole (*Cl. rutilus*). *Russian Journal of Genetics*, **45**, 533-545.

Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G. & Harmon, L.J. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci U S A*, **106**, 13410-4.

Almeida, E.A.B., Pie, M.R., Brady, S.G. & Danforth, B.N. (2012) Biogeography and diversification of colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. *Journal of Biogeography*, **39**, 526-544.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Rebb, C.A. & Saunders, N.C. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Baer, C.F., Miyamoto, M.M. & Denver, D.R. (2007) Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nature Review Genetics*, **8**, 619-631.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B. & Ferrer, E.A. (2011) Has the Earth's sixth mass extinction already arrived? *Nature*, **471**, 51-7.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. & Courchamp, F. (2012) Impacts of climate change on the future of biodiversity. *Ecol Lett*, **15**, 365-77.

Ben Faleh, A., Cosson, J.F., Tatard, C., Ben Othmen, A., Said, K. & Granjon, L. (2010) Are there two cryptic species of the lesser jerboa Jaculus jaculus (Rodentia: Dipodidae) in Tunisia? Evidence from molecular, morphometric, and cytogenetic data. *Biol J Linn Soc* **99**, 673-686.

Benton, M.J. (2009) The red queen and the court jester: species diversity and the role of biotic and abiotic factors through time. *Science*, **323**, 728-732.

Bergsten, J. (2005) A review of long-branch attraction. *Cladistics*, **21**, 163-193.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol*, **22**, 148-55.

Blaxter, M. (2003) Counting angels with DNA. *Nature*, **421**, 122-124.

Blondel, J. (1986) *Biogéographie évolutive*, Masson edn, Paris.

Blondel, J. (1995) *Biogéographie. Approche écologique et évolutive*.

Boratynski, Z., Alves, P.C., Berto, S., Koskela, E., Mappes, T. & Melo-Ferreira, J. (2011) Introgression of mitochondrial DNA among Myodes voles: consequences for energetics? *BMC Evolutionary Biology*, **11**, 355.

Boratynski, Z., Melo-Ferreira, J., Alves, P.C., Berto, S., Koskela, E., Pentikainen, O.T., Tarroso, P., Ylilauri, M. & Mappes, T. (2014) Molecular and ecological signs of mitochondrial adaptation: consequences for introgression? *Heredity*, **113**, 277-286.

Bouilhol, 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.

Brelsford, A., Mila, B. & Irwin, D.E. (2011) Hybrid origin of Audubon's warbler. *Molecular Ecology*, **20**, 2380-2389.

Bromham, L. & Penny, D. (2003) The modern molecular clock. *Nature Reviews Genetics*, **4**

Carlson, E.A. (2004) *Mendel's legacy - the origin of classical genetics*. Cold Spring Harbor Laboratory Press New-York.

Chapuis, M.P. & Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621-631.

Che, J., Zhou, W.-W., Hu, J.-S., Yan, F., Papenfuss, T.J., Wake, D.B. & Zhang, Y.-P. (2010) Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. *PNAS*, **107**, 13765-13770.

Chen, S., Xing, Y., Su, T., Zhou, Z., Dilcher, E.D.L. & Soltis, D.E. (2012) Phylogeographic analysis reveals significant spatial genetic structure of *Incarvillea sinensis* as a product of mountain building. *BMC Plant Biology*, **12**

Colangelo, P., Aloise, G., Franchini, P., Annesi, F. & Amori, G. (2012) Mitochondrial DNA reveals hidden diversity and an ancestral lineage of the bank vole in the Italian peninsula. *Journal of Zoology*, **287**, 41-52.

Condamine, F.L. & Hines, H.M. (2015) Historical species losses in bumblebee evolution. *Biology letters*, **11**, 2014.1049.

Condamine, F.L., Rolland, J. & Morlon, H. (2013a) Macroevolutionary perspctives to environmental change. *Ecology letters*, **16**, 72-85.

Condamine, F.L., Sperling, F.A., Wahlberg, N., Rasplus, J.Y. & Kergoat, G.J. (2012) What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol Lett*, **15**, 267-77.

Condamine, F.L., Toussaint, E.F.A., Cotton, A.M., Genson, G.S., Sperling, F.A.H. & Kergoat, J.K. (2013b) Fine-scale biogeographical and temporal diversification processes of peacock swallowtails (Papilio subgenus Achillides) in the Indo-Australian Archipelago. *Cladistics*, **29**, 88-111.

Condamine, F.L., Toussaint, E.F.A., Clamens, A.-L., Genson, G., Sperling, F.A.H. & Kergoat, G.K. (2015) Deciphering the evolution of birdwing butterflies 150 years after Alfred Russel Wallace. *Scientific Reports*, **5**, 11860.

Cornell, H.V. (2013) Is regional species diversity bounded or unbounded? *Biol Rev Camb Philos Soc*, **88**, 140-65.

Cox, C.B. & Moore, P.D. (2010) *Biogeography: an ecological and evolutionary approach*, eighth edn.

Cutter, A.D. (2013) Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Mol Phylogenet Evol*, **69**, 1172-85.

Darlington, P.J. (1957) *Zoogeography; the geographical distribution of animals*, Wiley edn, New York & London.

Darlu, P. & Tassy, P. (1993) *La reconstruction phylogénétique: concepts et méthodes*, Masson edn, Paris.

Darvish, J., Hajjar, T., Moghadam Matin, M., Haddad, F. & Akbarirad, S. (2008) New species of five-toed jerboa (Rodentia: Dipodidae, Allactaginae) from North-East Iran. *Journal of Sciences, Islamic Republic of Iran*, **19**, 103-109.

Daxner-Höck, G. (2001) New zapodids (Rodentia) from Oligocene-Miocene deposits in Mongolia. Part 1. *Senckenbergiana lethaea*, **81**, 359-389.

Daxner-Höck, G. & Wu, W. (2003) Plesiosminthus (Zapodidae, Mammalia) from China and Mongolia: migrations to Europe. In: *Distribution and migration of Tertiary mammals in Eurasia* eds. J.W.F. Reumer and W. Wessels), pp. 127-151. Backhuys, Rotterdam.

DeBiasse, M.B. & Hellberg, M.E. (2015) Discordance between morphological and molecular species boundaries among Caribbean species of the reef sponge Callyspongia. *Ecol Evol*, **5**, 663-75.

Deffontaine, V., Ledevin, R., Fontaine, M.C., Quéré, J.P., Renaud, S., Libois, R. & Michaux, J.R. (2009) A relict bank vole lineage highlights the biogeographic history of the Pyrenean region in Europe. *Molecular Ecology*, **18**, 2489-2502.

Deffontaine, V., Libois, R., Kotlik, P., Sommer, R., Nieberding, C., Paradis, E., Searle, J.B. & Michaux, J.R. (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (Clethrionomys glareolus). *Molecular Ecolology*, **14**, 1727-1739.

Delsuc, F. & Douzery, E.J. (2004a) Les méthodes probabilistes en phylogénie moléculaire : (2) L'approche bayésienne. *Biosystema*, **22**, 75-86.

Delsuc, F. & Douzery, E.J. (2004b) Les méthodes probabilistes en phylogénie moléculaire: (1) Les modèles d'évolution des séquences et le maximum de vraisemblance. *Biosystema*, **22**, 59-74.

Donoghue, P.C. & Benton, M.J. (2007) Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution*, **22**

Douzery, E.J., Blanquart, S., Criscuolo, A., Delsuc, F., Douady, C., Lartillot, N., Philippe, H. & Ranwez, V. (2010) Phylogénie moléculaire. *Biologie évolutive*, pp. 183-243. De Boeck, Paris.

Dowling, D.K., Friberg, U. & Lindell, J. (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in Ecology & Evolution*, **23**, 546-54.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PloS Biology*, **4**, e88.

Ellerman, J.R. (1940) Superfamily Dipodoidea. *The families and genera of living rodents*, pp. 560-598. Oxford University Press, London.

Emry, R.J. & Korth, W.W. (1989) *Rodents of the Bridgerian (Middle Eocene) Elderberry Canyon local fauna of Eastern Nevada*. Smithsonian Institution Press, Washigton, D.C.

Fabre, P.-H., Hautier, L., Dimitrov, D. & Douzery, E.J.P. (2012a) A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evolutionary Biology*, **12**, 88.

Fabre, P.-H., Pagès, M., Musser, G.G., Fitriana, Y.S., Fjeldså, J., Jennings, A., Jønsson, K.A., Kennedy, J., Michaux, J., Semiadi, G., Supriatna, N. & Helgen, K.M. (2013) A new genus of rodent from Wallacea (Rodentia: Muridae: Murinae: Rattini), and its implication for biogeography and Indo-Pacific Rattini systematics. *Zoological Journal of the Linnean Society*, **169**, 408-447.

Fabre, P.H., Irestedt, M., Fjeldså, J., Bristol, R., Groombridge, J.J., Irham, M. & Jønsson, K.A. (2012b) Dynamic colonization exchanges between continents and islands drive diversification in paradise-flycatchers (Terpsiphone, Monarchidae). *Journal of Biogeography*, **39**, 1900-1918.

Fahlbusch, V. (1992) The Neogene mammalian faunas of Ertemte and Harr Obo in Inner Mongolia (Nei Mongol), China. - 10. Eozapus (Rodentia). *Senckenbergiana lethaea*, **72**, 199-217.

Felsenstein, J. (1985) Confidence limits on phyloge- nies : an approach using the bootstrap. *Evolution*, **39**

Fokin, I.M. (1978) *The locomotion and morphology af the locomotory organs in jerboas*. Nauka.

Fortelius, M. (2013) New and Old Worlds Database of Fossil Mammals (NOW). In, University of Helsinki.

Grant, B.S., Cook, A.D., Clarke, C.A. & Owen, D.F. (1998) Geographic and temporal variation in the incidence of melanism in peppered moth populations in America and Britain. *Journal of Heredity*, **89**, 465-471.

Green, M. (1977) Neogene Zapodidae (Mammalia: Rodentia) from South Dakota. *Journal of Paleontology*, **51**, 996-1015.
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.
Hall, E.R. (1930) Rodents and lagomorphs from the later Tertiary of Fish Lake Valley, Nevada. *University of California Press* **18**, 227-231.
Harry, M. (2001) *Génétique moléculaire et évolutive*. Maloine, Paris.
He, S., Cao, W. & Chen, Y. (2001) The uplift of qinghai-xizang (tibet) plateau and the vicariance speciation of glyptosternoid fishes (silurifoumes: sisoridae). *Science in China*, **44**, 650-651.
Hedrick, P.W. (2010) Cattle ancestry in bison: explanations for higher mtDNA than autosomal ancestry. *Molecular Ecology*, **19**, 3328-3335.
Henry, J.-P. & Gouyon, P.-H. (1998) *Précis de génétique des populations*, Masson edn.
Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F. & Yoder, A.D. (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Mol Phylogenet Evol*, **54**, 291-301.
Hillis, D.M. (1987) Molecular versus morphological approaches to systématics. *Annual Review of Ecology and Systematics*, **18**, 23-42.
Ho, S.Y.W. & Phillips, M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367-380.
Holden, M.E. & Musser, G.G. (2005) Superfamily Dipodoidea. *Mammal Species of the World. A Taxonomic and Geographic Reference (3rd ed)*. Johns Hopkins University Press.
Huang, X. (1992) Zapodidae (Rodentia, Mammalia) from the middle Oligocene of Ulantatal, Nei Mongol. *Vertebrata PalAsiatica*, **30**, 249-303.
Hudson, R.R. & Coyne, J.A. (2002) Mathematical consequences of the genealogical species concept. *Evolution*, **56**
Huelsenbeck, J.P., Ronquist, F., Nielsen, R. & Bollback, J.P. (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, **294**, 2310-2314.
Hunt, T., Bergsten, J., Levkanicova, Z., Papadopoulou, A., St. John, O., Wild, R., Hammond, P.M., Ahrens, D., Balke, M., Bocakova, M., Bocak, L. & Vogler, A.P. (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, **318**, 1913-1916.
IUCN (2015) *The IUCN Red List of Threatened Species*. Available at: http://www.iucnredlist.org (accessed
Jenkins, P., Kilpatrick, C., Robinson, M. & Timmins, R. (2005) Morphological and molecular investigations of a new family, genus and species of rodent (Mammalia: Rodentia: Hystricognatha) from Laos. *Systematics and Biodiversity*, **2**, 419-454.
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.
Jombart, T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-5.
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. (2010) New material of dipodid rodents (Dipodidae, Rodentia) from the early Miocene of Gashunyinadege, Nei Mongol, China. *Journal of Vertebrate Paleontology*, **30**, 1860-1873.

Kimura, Y. (2011) The earliest record of birch mice from the Early Miocene Nei Mongol, China. *Naturwissenschaften*, **98**, 87-95.
Korth, W.W. (1993) Miosicista angulus, a new sicistine rodent (Zapodidae, Rodentia) from the Barstovian (Miocene) of Nebraska. *Transactions of the Nebraska Academy of Sciences*, **20**, 97-101.
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.
Lecointre, G. & Le Guyader, H. (2001) *Classification phylogénétique du vivant*, Belin edn, Paris.
Lenormand, T., Guillemaud, T., Bourguet, D. & Raymond, M. (1998) Evaluating gene flow using selected markers: a case study. *Genetics*, **149**, 1383-1392.
LI, H. & Chi, H. (1981) Two new rodents from Neogene of Chilong Basin, Tibet. *Vertebrata PalAsiatica*, **19**, 245-255.
Li, Q. & Zheng, S.H. (2005) Note on four species of dipodids (Dipodidae, Rodentia) from the late Miocene Bahe Formation, Lantian, Shaanxi. *Vertebrata PalAsiatica*, **43**, 283-296.
Lomolino, M.V., Riddle, B.R., Whittaker, R.J. & Brown, J.H. (2010) *Biogeography*, 4th edn. Sinauer Associates, Inc., Sunderland.
Lopez Antonanzas, R. & Sen, S. (2006) New Saudi Arabian miocene jumping mouse (Zapodidae): systematics and phylogeny. *Journal of Vertebrate Paleontology*, **26**, 170-181.
Lyon, M. (1901) A comparison of the osteology of the jerboas and jumping mice. *Proc. U. S. Natl. Mus.*, **23**, 659-668.
Mahler, D.L., Ingram, T., Revell, L.J. & Losos, J.B. (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, **341**, 292-295.
Mallet, J. (2005) Hybridization as an invasion of the genome. *Trends Ecol Evol*, **20**, 229-37.
Martin, R.A. (1989) Early Pleistocene zapodid rodents from the Java local fauna of north-central South Dakota. *Journal of Vertebrate Paleontology*, **9**, 101-109.
Martin, R.A. (1994) A preliminary review of dental evolution and paleogeography in the Zapodid rodents, with emphasis on Pliocene and Pleistocene taxa. *Rodent and Lagomorph families of Asian origins and diversification* (ed. by Y. Tomida, C.K. Li and T. Setoguchi), pp. 99-113. National Science Monographs, Tokyo.
May, R.M. (1988) How many species are there on Earth? *Science*, **241**, 1441-1449.
Mayr, E. (1963) *Animal species and evolution*, Harvard University Press edn, Cambridge.
McKenna, M.C. & Bell, S.K. (1997) *Classification of mammal above the species level*. Columbia University Press, New-York.
Melo-Ferreira, J., Alves, P.C., Freitas, H., Ferrand, N. & Boursot, P. (2009) The genomic legacy from the extinct Lepus timidus to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Mol Ecol*, **18**, 2643-58.
Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., McCain, C.M., McCune, A.R., McDade, L.A., McPeek, M.A., Near, T.J., Price, T.D., Ricklefs, R.E., Roy, K., Sax, D.F., Schluter, D., Sobel, J.M. & Turelli, M. (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol Lett*, **10**, 315-31.
Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B. & Worm, B. (2011) How many species are there on Earth and in the ocean? *PLoS Biology*, **9**, e1001127.
Morlon, H. (2014) Phylogenetic approaches for studying diversification. *Ecology letters*, **17**, 508-525.
Morlon, H., Potts, M.D. & Plotkin, J.B. (2010) Inferring the dynamics of diversification: a coalescent approach. *PloS Biol*, **8**, e1000493.

Morlon, H., Parsons, T.L. & Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record. *Proc Natl Acad Sci U S A*, **108**, 16327-16332.
Musser, G.G. & Carleton, M.D. (2005) Superfamily Muroidea. *Mammal Species of the World. A Taxonomic and Geographic Reference (3rd ed)*. Johns Hopkins University Press.
Myers, P., Espinosa, R., Parr, C.S., Jones, T., Hammond, G.S. & Dewey, T.A. (2015) *The Animal Diversity Web*. Available at: http://animaldiversity.org (accessed
Pages, M., Bazin, E., Galan, M., Chaval, Y., Claude, J., Herbreteau, V., Michaux, J., Piry, S., Morand, S. & Cosson, J.F. (2013) Cytonuclear discordance among Southeast Asian black rats (Rattus rattus complex). *Mol Ecol*, **22**, 1019-34.
PaleobiologyDatabase (2015) *Paleobiology Database*. Available at: https://paleobiodb.org/ (accessed
Parmesan, C. (2006) Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637-669.
Pennisi, E. (2005) What determines species diversity? *Science*, **309**, 1.
Petit, R.J. & Excoffier, L. (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, **24**, 386-393.
Philippe, H., Zhou, Y., Brinkmann, H., Rodrigue, N. & Delsuc, F. (2005) Heterotachy and long-branch attraction in phylogenetics. *BMC Evol Biol*, **5**, 50.
Phillimore, A.B. & Price, T.D. (2008) Density-dependent cladogenesis in birds. *PloS Biol*, **6**, e71.
Pisano, J., Condamine, F.L., Lebedev, V., Bannikova, A., Quéré, J.-P., Shenbrot, G.I., Pagès, M., Michaux, J.R. & Riddle, B. (2015) Out of Himalaya: the impact of past Asian environmental changes on the evolutionary and biogeographical history of Dipodoidea (Rodentia). *Journal of Biogeography*, **42**, 856-870.
Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., Kamoun, S., Sumlin, W. & Vogler, A. (2006) Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology*, **55**, 595-609.
Potapov, S.G., Illarionova, N.A., Andreeva, T.A., Baskevich, M.I., Okulova, N.M., Lavrenchenko, L.A. & Orlov, V.N. (2007) Transfer of mitochondrial genome of the northern redbacked vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*) in northwestern Europe. *Doklady Biological Sciences*, **417**, 435-438.
Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.
Purvis, A. (2008) Phylogenetic Approaches to the Study of Extinction. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 301-319.
Qiu, Z. (2003) The Neogene mammalian faunas of Ertemte and Harr Obo in Inner Mongolia (Nei Mongol), China. *Senckenbergiana lethaea*, **83**, 135-147.
Qiu, Z. & Storch, G. (2000) The early Pliocene Micromammalian Fauna of Bilike, Inner Mongolia, China (Mammalia: Lipotyphla, Chiroptera, Rodentia, Lagomorpha). *Senckenbergiana lethaea*, **80**, 173-229.
Quental, T.B. & Marshall, C.R. (2010) Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol Evol*, **25**, 434-41.
Quéré, J.-P. & Le Louarn, H. (2011) *Les rongeurs de France, faunistique et biologie (Edition revue et augmentée)*. Quae.
Rabosky, D.L. (2009) Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol Lett*, **12**, 735-43.
Rabosky, D.L. (2013) Diversity-Dependence, Ecological Speciation, and the Role of Competition in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 481-502.

Rabosky, D.L. & Lovette, I.J. (2008) Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution*, **62**, 1866-1875.
Rabosky, D.L., Slater, G.J. & Alfaro, M.E. (2012) Clade age and species richness are decoupled across the eukaryotic tree of life. *PloS Biol*, **10**, e1001381.
Rabosky, D.L., Title, P.O. & Huang, H. (2015) Minimal effects of latitude on present-day speciation rates in New World birds. *Proceedings B*, **282**
Raymond, M. & Godelle, B. (2010) Applications de la biologie évolutive. *Biologie évolutive* (ed. by D. Boeck), Paris.
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.
Rhymer, J.M. & Simberloff, D. (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83-109.
Ricklefs, R.E. (2004) A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, **7**, 1-15.
Rieux, A., Lenormand, T., Carlier, J., de Lapeyre de Bellaire, L. & Ravigne, V. (2013) Using neutral cline decay to estimate contemporary dispersal: a generic tool and its application to a major crop pathogen. *Ecol Lett*, **16**, 721-730.
Rofes, J., Garcia-Ibaibarriaga, N., Murelaga, X., Arrizabalaga, A., Iriarte, M.-J., Cuenca-Bescos, G. & Villaluenga, A. (2012) The southwesternmost record of Sicista (Mammalia: Dipodidae) in Eurasia, with a review f the palaeogeography and palaeoecology of the genus in Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **348-349**, 67-73.
Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014) Faster speciation and reduced extinction in the tropics contribute to the Mammalian latitudinal diversity gradient. *PLoS Biol*, **12**, e1001775.
Roughgarden, J. (1979) *Theory of population genetics and evolutionary ecology*. MacMillan Pub, New-York.
Rousset, F. (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
Saetre, G.-P., Borge, T., Lindell, J., Moum, T., Primmer, C.R., Sheldon, B.C., Haavie, J., Johnsen, A. & Ellegren, H. (2008) Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. *Molecular Ecology*, **10**, 737-749.
Sanderson, M.J. & Hufford, L. (1996) *Homoplasy: the recurrence of similarity in evolution*. Academic Press.
Schemske, D.W., Mittelbach, G.G., Cornell, H.V., Sobel, J.M. & Roy, K. (2009) Is There a Latitudinal Gradient in the Importance of Biotic Interactions? *Annual Review of Ecology, Evolution, and Systematics*, **40**, 245-269.
Schenk, J.J., Rowe, K.C. & Steppan, S.J. (2013) Ecological opportunity and incumbency in the diversification of repeated continental colonizations by muroid rodents. *Systematic Biology*, **62**, 837-864.
Schulze, E.D. & Mooney, H.A. (1993) *Biodiversity and ecosystem function*. Springer-Verlag, New-York.
Shahin, A.A.B. (2005) Growth and maturation of metatarsals and their taxonomic significance in the jerboas Allactaga and Jaculus (Rodentia: Dipodidae). *Acta Zoologica*, **86**, 81-90.
Shenbrot, G.I. (1992) Cladistic approach to the analysis of phylogenetic relationships among dipodoid rodents. *Arch. Zool. Mosc. State Univ.*, **29**, 176-201.
Shenbrot, G.I., Sokolov, V.E., Heptner, V.G. & Koval'skaya, Y.M. (2008) *Jerboas: Mammals of Russia and adjacent regions*. Science Publishing Corporation.

Silvestro, D., Antonelli, A., Salamin, N. & Quentale, T.B. (2015) The role of clade competition in the diversification of North American canids. *PNAS*, **112**, 8684–8689.
Smith, A.B. & Peterson, K.J. (2002) Dating the time of origin of major clades: molecular clocks and the fossil record. *Annual Review of Earth and Planetary Sciences*, **30**, 65-88.
Springer, M.S., Stanhope, M.J., Madsen, O. & de Jong, W.W. (2004) Molecules consolidate the placental mammal tree. *Trends in Ecology & Evolution*, **19**, 430-8.
Stadler, T. (2011) Mammalian phylogeny reveals recent diversification rate shifts. *Proc Natl Acad Sci U S A*, **108**, 6187-6192.
Stein, B.R. (1990) Limb myology and phylogenetic relationships in the superfamily Dipodoidea (birch mice, jumping mice, and jerboas). *Z. Zool. Syst. Evol.*, **28**, 299-314.
Stephens, P.R. & Wiens, J.J. (2003) Explaining species richness from continents to communities: the time-for-speciation effect in Emydid turtles. *Am. Nat.*, **161**, 112-128.
Steppan, S.J., Adkins, R.M. & Anderson, J. (2004) Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Systematic Biology*, **53**, 533-553.
Stucky, R.K. & McKenna, M.C. (1993) Mammalia. *Tho fossil record 2* (ed. by M.J. Benton), pp. 43-52. Chapman & Hall, London.
Sulimski, A. (1962) Two new rodents from Weze 1 (Poland). *Acta Palaeontologica Polonica*, **7**, 503-511.
Tegelström, H. (1987) Transfer of mitochondrial DNA from the Northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*Clethrionomys glareolus*). *Journal of Molecular Evolution*, **24**, 218-227.
Tesakov, A.S. (2001) New data on Plio-Pleistocene dipodine jerboas of the genus Plioscirtopoda from Eastern Europe (Mammalia: Dipodidae). *Zoosystematica Rossica*, **9**, 237-244.
Thomas, O. (1896) On the genera of rodents: an attemp to bring up to date the current arrangement of the order. *Proceedings of the Zoological Society of London*, 1012-1028.
Tilman, D. (2000) Causes, consequences, and ethics of biodiversity. *Nature*, **405**
Title, P.O. & Burns, K.J. (2015) Rates of climatic niche evolution are correlated with species richness in a large and ecologically diverse radiation of songbirds. *Ecol Lett*, **18**, 433-40.
Toews, D.P. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907-30.
Tong, Y. (1997) *Middle Eocene small mammals from liguanqiao basin of Henan province and Yuanqu basin of Shanxi province*.
Vinogradov, B. (1930) On the Classification of Dipodidae (Rodentia). *Bulletin de l'Académie des Sciences de l'URSS*, **4**, 331-470.
Vinogradov, B.S. (1937) Fauna of the USSR. *Mammals*. Academy Nauk SSSR Moscow.
Wake, D.B., Wake, M.H. & Specht, C.D. (2011) Homoplasy: from detecting pattern to determinign process and mechanism of evolution. *Science*, **331**, 1032-1035.
Wan, Q.H., Wu, H., Fujihara, T. & Fang, S.G. (2004) Review: Which genetic marker for which conservation genetics issue? *Electrophoresis*, **25**, 2165-2176.
Wang, B. (1985) Zapodidae (Rodentia, Mammalia) from the Lower Oligocene of Qujing, Yunnan, China. *Mainzer Geowissenschaftliche Mitteilungen*, **14**, 345-367.
Wang, B. & Qiu, Z. (2000) Dipodidae (Rodentia, Mammalia) from the lower member of Xianshuihe formation in Lanzhou Basin, Gansu, China. *Vertebrata PalAsiatica*, **38**, 10-35.

Wang, L., Schneider, H., Zhang, X.-C. & Xiang, Q.P. (2012) The rise of the Himalaya enforced the diversification of SE Asian ferns by altering the monsoon regimes. *BMC Plant Biology*, **12**

Wang, X.-M., Qiu, Z.-D., Li, Q., Yukimitsu, T., Kimura, Y., Tseng, Z.J. & Wang, H.-J. (2009) A new early to late Miocene fossiliferous region in central Nei Mongol: Lithostratigraphy and biogeography in Aoerban strata. *Vertebrata PalAsiatica*, **47**, 111-134.

Wiens, J.J. (2011) The causes of species richness patterns across space, time, and clades and the role of "ecological limits". *The Quarterly Review of Biology*, **86**, 75-96.

Wilson, D. & Reeder, D. (2005) *Mammal species of the world*, Third edn. Johns Hopkins University Press, Baltimore.

Wilson, R.W. (1936) A Pliocene rodent fauna from Smiths Valley, Nevada. *Carnegie Institution publications*, **473**, 15-34.

Winge, H. (1887) Jordfundne og nulevende Gnavere (Rodentia) fra Lagoa Santa, Minas Geraes, Brasilien. *Museo Lundi (Copenhagen)*, **1**, 1-178.

Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686-693.

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.

Zazhigin, V.S. & Lopatin, A.V. (2000) The history of Dipodoidea (Rodentia, Mammalia) in the Miocene of Asia: 3. Allactaginae. *Paleont. Zh.*, **34**, 553-565.

Zhang, M.L. & Fritsch, P.W. (2010) Evolutionary response of Caragana (Fabaceae) to Qinghai–Tibetan Plateau uplift and Asian interior aridification. *Plant Systematic & Evolution*, **288**, 191-199.

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*, **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.

Zink, R.M. & Barrowclough, G.F. (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107-21.

Zuckerkandl, E. & Pauling, L. (1965) Molecules as documents of evolutionary history. *Journal of theoretical biology*, **8**

ANNEXES

Table S1 Registre fossile des Dipodoidea. Les données paléontologiques ont été collectées dans la littérature et sur des bases de données paléontologiques en ligne : '*Paleobiology Database'* (PBDB) et '*New and Old World Database'* (NOW) (Fortelius, 2013; PaleobiologyDatabase, 2015). Au total, les informations paléontologiques de 170 espèces de Dipodoidea ont été collectées. L'interval de temps durant lequel le genre était présent a été estimé sur base des espèces les plus récentes et les plus anciennes qui lui étaient apparentées.

Genre	Nombre d'espèces éteintes répertoriées	Espèce (Les espèces contemporaines sont indiquées en gras)	Classification taxonomique proposée …			Aire de distribution	Intervalle de temps durant lequel le genre était présent	Bases de données en ligne (où des informations paléontologiques ont été collectées)	Références
			Sur le web et/ou dans la littérature	Dans McKenna & Bell (1997)	Dans 'Mikko's Phylogeny Archive'				
Allactaga	8	*A. sp, A. ucrainica, A. jaculus, A. saltator,* ***A. euphratica, A. elater (1,95-), A. sibirica (1,95-),*** *A. anderssoni,* ***A. major (0,427-),*** *A. minor, A. nogaiskiensis, A. praejaculus*	Dipodidae - Allactaginae	Dipodoidea - Dipodidae - Allactaginae	Dipodidae - Allactaginae	Russia, Romania, Mongolia, Turkey, Kazakhstan, Ukraine, Azerbaijan	5.332 -	PBDB & NOW	
Allosminthus	2	*A. ernos, A. majusculus*	Zapodidae (Sicistinae; Martin, 1994)	Dipodoidea - Dipodidae - Sicistinae	Zapodidae - Sicistinae	China	Late Eocene	NOW	Daxner-Höck (2001)
Arabosminthus	2	*A. isabellae, A. quadratus*	Zapodidae (Sicistinae; Martin, 1994)	Dipodoidea - Dipodidae - Sicistinae	Zapodidae - Sicistinae	Saudi Arabia	20 - 15.2	NOW	Lopez Antonanzas and Sen (2006)
Banyuesminthus	2	*B. disconjugatus, B. uniconjugatus*	Zapodidae		Zapodidae	China	Mid-Late Eocene	NOW	Tong (1997); Daxner-Höck (2001)
Bohlinosminthus	1	*B. parvulus*	Dipodidae			China	26 - 23.8	NOW	
Brachyscirtetes	2	*B. robustus, B. wimani*	Dipodidae - Allactaginae	Dipodoidea - Dipodidae - Allactaginae	Dipodidae - Allactaginae	China, Kazakhstan	7.1 - 3.4	PBDB & NOW	Qiu and Storch (2000); Qiu (2003); Wang *et al.* (2009)
Cardiocranius	1	*C. pusillus*	Dipodidae - Cardiocraniinae	Dipodoidea - Dipodidae - Dipodinae - Cardiocraniini	Dipodidae - Dipodinae - Cardiocraniini	China	11.2 -	NOW	Li and Zheng (2005)

Dipus	3	*D. fraudator, D. sp., D. qiu*	Dipodidae	Dipodoidea - Dipodidae - Dipodinae - Dipodini	Dipodidae - Dipodinae - Dipodini	China	7.1 -	NOW	Wang *et al.* (2009)
Elymys	1	*E. complexus*	? Zapodidae	Dipodoidea - Dipodidae	Zapodidae	Nevada	50.3 - 46.2	PBDB	Emry and Korth (1989)
Eozapus	4	*E. similis, E. intermedius, E. sp., E. janossyi*	Dipodoidea - Zapodidae (Zapodinae; Martin, 1994)	Dipodoidea - Dipodidae - Zapodinae	Zapodidae - Zapodinae	Spain, China, Austria, Germany, Turkey, Poland, France, Hungary	11.608 -	PBDB & NOW	Fahlbusch (1992)
Euchoreutes	1	*E. sp.*	Dipodidae	Dipodoidea - Dipodidae - Euchoreutinae		Mongolia	4.667 -	NOW	
Gobiosminthus	1	*G. qiui*	Zapodidae		Zapodidae - Sicistinae - Sicistini	China	Middle Oligocene	/	Huang (1992)
Heosminthus	4	*H. verus, H. minutus, H. primiverus, H. sp.*	Dipodidae (PBDB), or Zapodidae (NOW; (Sicistinae; Martin, 1994))		Zapodidae - Sicistinae (Phylogenetically untreated taxa)	China, Turkey, Mongolia	43.6 - 23.03	PBDB & NOW	Wang (1985); Daxner-Höck (2001)
Heterosminthus	14	*H. orientalis, H. mongoliensis, H. sp., H. gansus, H. simplicidens, H. firmus, H. jucundus, H. gabuniai, H. intermedius, H. nanus, H. erbajeva, H. lanzhouensis, H. honestus, H. mugodzharicus*	Dipodidae - Lophocricetinae (PBDB; Kimura, 2010) or Zapodidae (NOW; Daxner-Höck, 2001; Wang et al, 2009; (Sicistinae; Martin, 1994))	Dipodoidea - Dipodidae - Sicistinae - Sicistini	Zapodidae - Sicistinae - Sicistini (all but *H. gansus* that is classified in Sicistinae only)	China, Libya, Mongolia, Kazakhstan, Moldova, Turkey	24.6 - 5.332	PBDB, NOW & Mikko's phylogeny archive	Daxner-Höck (2001); Kimura (2010)
Himalayactaga	1	*H. liui*	Dipodidae	Dipodoidea - Dipodidae - Allactaginae	Dipodidae - Allactaginae	China	8.2 - 7.1	NOW	LI and Chi (1981)
Jaculus	1	*J. sp,* ***J. orientalis (3,56-)***	Dipodidae - Dipodinae	Dipodoidea - Dipodidae - Dipodinae - Dipodini	Dipodidae - Dipodinae - Dipodini	Ethiopia, Tanzania, Kenya	2.588 -	PBDB & NOW	

Javazapus	1	*J. weeksi*	Dipodoidea - Zapodidae (Zapodinae; Martin, 1994)	Dipodoidea - Dipodidae - Zapodinae	Zapodidae - Zapodinae	South Dakota	1.8 - 0.3	PBDB	Martin (1989)
Litodonomys	4	*L. huangheensis, L. sp., L. xishuiensis, L. minimus*	Zapodidae (NOW; Wang *et al*., 2009), or Dipodidae - Lophocricetinae (Kimura, 2010)		Zapodidae	China	26 - 14.059	NOW	Wang and Qiu (2000); Kimura (2011)
Lophocricetus	10	*L. grabaui, L. pusillus, L. sp., L. minusculus, L. orientalis, L. sarmaticus, L. sibiricus, L. vinogradovi, L. xianensis, L. gansus*	Dipodidae (PBDB), or Zapodidae (NOW; Wang *et al.*, 2009; (Lophocricetinae; Martin, 1994))	Dipodoidea - Dipodidae - Sicistinae - Lophocricetini	Zapodidae - Sicistinae - Lophocricetini (Phylogenetically untreated taxa)	China, Mongolia, Kazakhstan, Ukraine, Russia	11.2 - 3.6	PBDB & NOW	Qiu and Storch (2000)
Macrognathomys	3	*M. gemmacolis, M. nanus, M. sp.*	Dipodoidea - Zapodidae (Sicistinae; Martin, 1994)	Dipodoidea - Dipodidae - Sicistinae - Sicistini	Zapodidae - Sicistinae - Sicistini (Phylogenetically untreated taxa)	Oregon, Nebraska, South Dakota, Nevada, Kansas, Oklahoma	15.97 - 4.9	PBDB	Hall (1930)
Megasminthus	1	*M. sp.*	Dipodoidea - Zapodidae (Sicistinae; Martin, 1994)	Dipodoidea - Dipodidae - Zapodinae	Zapodidae - Zapodinae	Kazakhstan	17 - 11.2	NOW	
	3	*M. gladiofex, M. tiheni, M. sp.*				South Dakota, Nebraska, Canada	23.03 - 5.332	PBDB	Green (1977)
Miosicista	1	*M. angulus*	Dipodoidea - Zapodidae - Sicistinae	Dipodoidea - Dipodidae - Sicistinae - Sicistini	Zapodidae - Sicistinae - Sicistini (Phylogenetically untreated taxa)	Nebraska	15.97 - 13.6	PBDB	Korth (1993)
Napaeozapus	1	*N. sp.*, ***N. insignis***	Dipodoidea - Zapodidae (Zapodinae; Martin, 1994)	Dipodoidea - Dipodidae - Zapodinae	Zapodidae - Zapodinae	Pennsylvania, Virginia, West Virginia, Maryland	2.588 -	PBDB	
Omoiosicista	1	*O. fui*	Dipodidae - Lophocricetinae			China	Early Miocene	/	Kimura (2010)
Paracricetulus	1	*P. Schaubi*	Zapodidae			China	12.5 - 11.2	NOW	
Paralactaga	6	*P. anderssoni, P. grauberi, P. sp., P. minor, P. suni, P. varians*	Dipodidae			China, Ukraine, Kazakhstan	23.03 - 1.4	PBDB & NOW	Qiu and Storch (2000); Qiu (2003);

									Wang *et al.* (2009)
Paralophocricetus	2	*P. Progressus, P. saraicus*	Dipodidae			Russia	7.1 - 4.35	NOW	
Parasminthus	10	*P. asiaecentralis, P. tangingoli, P. sp., P. parvulus, P. quartus, P. xiningensis, P. lajeensis, P. huangshuiensis, P. debruijni, P. cubitalus*	Dipodidae (PBDB), or Zapodidae (NOW; (Sicistinae; Martin, 1994))		Zapodidae - Sicistinae - Sicistini	Czech Republic, China, Kazakhstan	34 - 15.2	PBDB, NOW& Mikko's phylogeny archive	Huang (1992)
Plesiosminthus	13	*P. barsboldi, P. geringensis, P. schaubi, P. promyarion, P. winistoerferi, P. clivosus, P. myarion, P. xiningensis, P. huangshuiensis, P. lajeensis, P. sp., P. vegrandis, P. asiaticus*	Dipodoidea - Zapodidae (Sicistinae; Martin, 1994)	Dipodoidea - Dipodidae - Sicistinae - Sicistini	Zapodidae - Sicistinae - Sicistini (Phylogenetically untreated taxa)	Nebraska, China, Spain, Germany, Canada, Switzerland, South Dakota, Colorado, Texas, Russia, France, Mongolia	30.8 - 11.608	PBDB & NOW	Daxner-Höck and Wu (2003); Wang *et al.* (2009); Kimura (2010)
Pliopygerethmus	2	*P. Brachydens, P. sp.*	Dipodidae			Ukraine, Russia	1.95 - 0.621	NOW	
Plioscirtopoda	3	*P. sp., P. stepanovi, P. novorossica*	Dipodidae - Dipodinae	Dipodoidea - Dipodidae - Dipodinae - Dipodini		Kazakhstan, Russia, Ukraine	3.158 - 0,76	NOW	Tesakov (2001)
Pliozapus	1	*P. solus*	Dipodoidea - Zapodidae (Zapodinae; Martin, 1994)	Dipodoidea - Dipodidae - Zapodinae	Zapodidae - Zapodinae	Nevada, Oregon, Nebraska	13.6 - 4.9	PBDB	Wilson (1936)
Primisminthus	3	*P. jinus, P. shangenus, P. yenus*	Zapodidae		Zapodidae	China	Mid-Late Eocene	NOW	Tong (1997); Daxner-Höck (2001)
Proalactaga	1	*P. sp.*	Dipodidae	Dipodoidea - Dipodidae - Allactaginae	Dipodidae - Allactaginae	Asia	Late Miocene	PBDB	McKenna and Bell (1997)
Protalactaga	8	*P. lantianensis, P. grabaui, P. shevyrevae, P. sefriouii, P. tunggurensis, P. sp., P. maghrebiensis, P. major*	Dipodidae	Dipodoidea - Dipodidae - Allactaginae	Dipodidae - Allactaginae	Saudi Arabia, China, Turkey, Mongolia, Morocco, Russia, Kazakhstan	23.03 - 5.332	PBDB & NOW	Li and Zheng (2005); Wang *et al.* (2009)
Protozapus	2	*P. sp., P. intermedius*	Zapodidae		Zapodidae	China	3.4 - 2.6	NOW &	

			(Zapodinae; Martin, 1994)					Mikko's phylogeny archive	
Pygeretmus	2	*P. sp., P. pygmaeus*	Dipodidae	Dipodoidea - Dipodidae - Allactaginae	Dipodidae - Allactaginae	Mongolia, Kazakhstan	3.158 - 2.192	NOW	
Salpingotus	1	*S. primitivus*	Dipodidae - Cardiocraniinae	Dipodoidea - Dipodidae - Dipodinae - Cardiocraniini	Dipodidae - Dipodinae - Cardiocraniini	China	11.2 -	-	Li and Zheng (2005)
Schaubeumys	5	*S. aralensis, S. woodi, S. grangeri, S. sp., S. cartomylos*	Dipodoidea - Zapodidae			Canada, South Dakota, Kazakhstan, Nebraska	28.4 - 13.6	PBDB	
Scirtodipus	3	*S. sp., S. kalbica, S. kazakhstanica*	Dipodidae	Dipodoidea - Dipodidae - Dipodinae - Dipodini	Dipodidae - Dipodinae - Dipodini	Mongolia, Kazakhstan	7.1 - 4.033	NOW	
Scirtopoda	1	*S. sp.*	Dipodidae			Mongolia	2.917 - 2.192	NOW	
Shamosminthus	3	*S. tongi, S. sodovis, S. sp.*	Zapodidae		Zapodidae - Sicistinae (- Sicistini for *S. tongi*)	China, Mongolia	31.5 - 28.4	NOW	Huang (1992); Daxner-Höck (2001)
Sicista	7	*S. primus, S. wangi, S. sp., S. pliocaenica, S. praeloriger, S. bagajevi, S. vinogradovi,* ***S. betulina (0.568-), S. subtilis (1.95-)***	Dipodoidea - Zapodidae (Sicistinae; Martin, 1994)	Dipodoidea - Dipodidae - Sicistinae - Sicistini	Zapodidae - Sicistinae - Sicistini (Phylogenetically untreated taxa)	China, Russia, Hungary, Romania, Greece, Germany, France, Austria, Czech Republic, Kazakhstan, Ukraine, Mongolia, Slovakia, Turkey, Bulgaria, Moldova, Poland, Netherlands, UK, Belgium, Denmark, Italy,	17 -	PBDB & NOW	Qiu and Storch (2000); Wang *et al.* (2009); Kimura (2011); Rofes *et al.* (2012)

						Serbia, Switzerland			
Simiacritomys	1	*S. whistleri*	? Zapodidae	Dipodoidea - Dipodidae	Dipodidae	California	40.4 - 37.2	PBDB	Kelly (1992)
Simimys	3	*S. simplex, S. sp., S. landeri*	Dipodoidea - Simimyidae			California, Texas, Nebraska	46.2 - 33.9	PBDB	Kelly (1992)
Sinodonomys	1	*S. simplex*	Dipodidae - Lophocricetinae			China	Early Miocene	/	Kimura (2010)
Sinosminthus	2	*S. sp., S. inapertus*	Zapodidae (Sicistinae; Martin, 1994)		Zapodidae - Sicistinae (Phylogenetically untreated taxa)	Kazakhstan	43.6 - 15.2	NOW	Tong (1997); Daxner-Höck (2001)
Sinozapus	2	*S. volkeri, S. sp.*	Dipodidae (PBDB), or Zapodidae (NOX database), or Zapodine dipodid (Qiu and Storch, 2000)			China	11.2 - 3.4	PBDB & NOW	Qiu and Storch (2000)
Sminthoides	2	*S. fraudator, S. sp.*	Dipodidae (PBDB), or Zapodidae (NOX database)	Dipodoidea - Dipodidae - Dipodinae - Dipodini	Dipodidae - Dipodinae - Dipodini	China	7.1 - 2.588	PBDB & NOW	Qiu and Storch (2000); Qiu (2003)
Sminthozapus	3	*S. janossyi, S. betfianus, S. sp.*	Dipodoidea - Zapodidae - Zapodinae	Dipodoidea - Dipodidae - Zapodinae	Zapodidae - Zapodinae	Germany, Poland, Hungary, Romania	5.332 - 0.781	PBDB	Sulimski (1962); Martin (1994)
Tatalsminthus	1	*T. khandae*	Dipodoidea - Zapodidae		Zapodidae	Mongolia	33.9 - 31,5	NOW	Daxner-Höck (2001)
Tyrannomys	1	*T. harkseni*	Dipodoidea - Zapodidae (Sicistinae; Martin, 1994)	Dipodoidea - Dipodidae - Sicistinae - Sicistini	Zapodidae - Sicistinae - Sicistini (Phylogenetically untreated taxa)	South Dakota	1.8 - 0.3	PBDB	Martin (1989)

Zapus	6	*Z. burti, Z. rinkeri, Z. sandersi, Z. sykesae, Z. sp.,* ***Z. hudsonius****, Z. adamsi*	Dipodoidea - Zapodidae (Zapodinae; Martin, 1994)	Dipodoidea - Dipodidae - Zapodinae	Zapodidae - Zapodinae	Kansas, Nebraska, West Virginia, Georgia, South Dakota, Oklahoma, Pennsylvania, Florida, Utah, Michigan, Missouri, Virginia	4.9 -	PBDB	Martin (1989)
Zindapiria	1	*Z. sp.*	Dipodidae			Pakistan	24.6 - 23.8	NOW	

REFERENCES BIBLIOGRAPHIQUES DE LA TABLE S1

Abramson, N.I., Rodchenkova, E.N. & Kostygov, A.Y. (2009a) Genetic variation and phylogeography of the bank vole (*Clethrionomys glareolus*, Arvicolinae, Rodentia) in Russia with special reference to the introgression of the mtDNA of a closely related species, red-backed vole (*Cl. rutilus*). *Russian Journal of Genetics*, **45**, 533-545.

Abramson, N.I., Rodchenkova, E.N., Fokin, M.V., Rakitin, S.B. & Gileva, E.A. (2009b) Recent and ancient introgression of mitochondrial DNA between the red (*Clethrionomys rutilus*) and bank (*Clethrionomys glareolus*) voles (Rodentia, Cricetidae). *Doklady Biological Sciences*, **425**, 147-150.

Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G. & Harmon, L.J. (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc Natl Acad Sci U S A*, **106**, 13410-4.

Almeida, E.A.B., Pie, M.R., Brady, S.G. & Danforth, B.N. (2012) Biogeography and diversification of colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. *Journal of Biogeography*, **39**, 526-544.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Rebb, C.A. & Saunders, N.C. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Baer, C.F., Miyamoto, M.M. & Denver, D.R. (2007) Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nature Review Genetics*, **8**, 619-631.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B. & Ferrer, E.A. (2011) Has the Earth's sixth mass extinction already arrived? *Nature*, **471**, 51-7.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. & Courchamp, F. (2012) Impacts of climate change on the future of biodiversity. *Ecol Lett*, **15**, 365-77.

Ben Faleh, A., Cosson, J.F., Tatard, C., Ben Othmen, A., Said, K. & Granjon, L. (2010) Are there two cryptic species of the lesser jerboa Jaculus jaculus (Rodentia: Dipodidae) in Tunisia? Evidence from molecular, morphometric, and cytogenetic data. *Biol J Linn Soc* **99**, 673-686.

Benton, M.J. (2009) The red queen and the court jester: species diversity and the role of biotic and abiotic factors throuh time. *Science*, **323**, 728-732.

Bergsten, J. (2005) A review of long-branch attraction. *Cladistics*, **21**, 163-193.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol*, **22**, 148-55.

Blaxter, M. (2003) Counting angels with DNA. *Nature*, **421**, 122-124.

Blondel, J. (1986) *Biogéographie évolutive*, Masson edn, Paris.

Blondel, J. (1995) *Biogéographie. Approche écologique et évolutive*.

Boratynski, Z., Alves, P.C., Berto, S., Koskela, E., Mappes, T. & Melo-Ferreira, J. (2011) Introgression of mitochondrial DNA among Myodes voles: consequences for energetics? *BMC Evolutionary Biology*, **11**, 355.

Boratynski, Z., Melo-Ferreira, J., Alves, P.C., Berto, S., Koskela, E., Pentikainen, O.T., Tarroso, P., Ylilauri, M. & Mappes, T. (2014) Molecular and ecological signs of mitochondrial adaptation: consequences for introgression? *Heredity*, **113**, 277-286.

Bouilhol, 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.

Brelsford, A., Mila, B. & Irwin, D.E. (2011) Hybrid origin of Audubon's warbler. *Molecular Ecology*, **20**, 2380-2389.

Bromham, L. & Penny, D. (2003) The modern molecular clock. *Nature Reviews Genetics*, **4**

Carlson, E.A. (2004) *Mendel's legacy - the origin of classical genetics*. Cold Spring Harbor Laboratory Press New-York.

Che, J., Zhou, W.-W., Hu, J.-S., Yan, F., Papenfuss, T.J., Wake, D.B. & Zhang, Y.-P. (2010) Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. *PNAS*, **107**, 13765-13770.

Chen, S., Xing, Y., Su, T., Zhou, Z., Dilcher, E.D.L. & Soltis, D.E. (2012) Phylogeographic analysis reveals significant spatial genetic structure of *Incarvillea sinensis* as a product of mountain building. *BMC Plant Biology*, **12**

Colangelo, P., Aloise, G., Franchini, P., Annesi, F. & Amori, G. (2012) Mitochondrial DNA reveals hidden diversity and an ancestral lineage of the bank vole in the Italian peninsula. *Journal of Zoology*, **287**, 41-52.

Condamine, F.L. & Hines, H.M. (2015) Historical species losses in bumblebee evolution. *Biology letters*, **11**, 2014.1049.

Condamine, F.L., Rolland, J. & Morlon, H. (2013a) Macroevolutionary perspctives to environmental change. *Ecology letters*, **16**, 72-85.

Condamine, F.L., Sperling, F.A., Wahlberg, N., Rasplus, J.Y. & Kergoat, G.J. (2012) What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol Lett*, **15**, 267-77.

Condamine, F.L., Toussaint, E.F.A., Cotton, A.M., Genson, G.S., Sperling, F.A.H. & Kergoat, J.K. (2013b) Fine-scale biogeographical and temporal diversification processes of peacock swallowtails (Papilio subgenus Achillides) in the Indo-Australian Archipelago. *Cladistics*, **29**, 88-111.

Condamine, F.L., Toussaint, E.F.A., Clamens, A.-L., Genson, G., Sperling, F.A.H. & Kergoat, G.K. (2015) Deciphering the evolution of birdwing butterflies 150 years after Alfred Russel Wallace. *Scientific Reports*, **5**, 11860.

Cornell, H.V. (2013) Is regional species diversity bounded or unbounded? *Biol Rev Camb Philos Soc*, **88**, 140-65.

Cox, C.B. & Moore, P.D. (2010) *Biogeography: an ecological and evolutionary approach*, eighth edn.

Cutter, A.D. (2013) Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Mol Phylogenet Evol*, **69**, 1172-85.

Darlington, P.J. (1957) *Zoogeography; the geographical distribution of animals*, Wiley edn, New York & London.

Darlu, P. & Tassy, P. (1993) *La reconstruction phylogénétique: concepts et méthodes*, Masson edn, Paris.

Darvish, J., Hajjar, T., Moghadam Matin, M., Haddad, F. & Akbarirad, S. (2008) New species of five-toed jerboa (Rodentia: Dipodidae, Allactaginae) from North-East Iran. *Journal of Sciences, Islamic Republic of Iran*, **19**, 103-109.

Daxner-Höck, G. (2001) New zapodids (Rodentia) from Oligocene-Miocene deposits in Mongolia. Part 1. *Senckenbergiana lethaea*, **81**, 359-389.

Daxner-Höck, G. & Wu, W. (2003) Plesiosminthus (Zapodidae, Mammalia) from China and Mongolia: migrations to Europe. In: *Distribution and migration of Tertiary mammals in Eurasia* eds. J.W.F. Reumer and W. Wessels), pp. 127-151. Backhuys, Rotterdam.

DeBiasse, M.B. & Hellberg, M.E. (2015) Discordance between morphological and molecular species boundaries among Caribbean species of the reef sponge Callyspongia. *Ecol Evol*, **5**, 663-75.

Deffontaine, V. (2008) *Histoire évolutive du campagnol roussâtre (Myodes (Clethrionomys) glareolus) en Eurasie*. Liège, Liège.

Deffontaine, V., Ledevin, R., Fontaine, M.C., Quéré, J.P., Renaud, S., Libois, R. & Michaux, J.R. (2009) A relict bank vole lineage highlights the biogeographic history of the Pyrenean region in Europe. *Molecular Ecology*, **18**, 2489-2502.

Deffontaine, V., Libois, R., Kotlik, P., Sommer, R., Nieberding, C., Paradis, E., Searle, J.B. & Michaux, J.R. (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (Clethrionomys glareolus). *Molecular Ecolology*, **14**, 1727-1739.

Delsuc, F. & Douzery, E.J. (2004a) Les méthodes probabilistes en phylogénie moléculaire: (1) Les modèles d'évolution des séquences et le maximum de vraisemblance. *Biosystema*, **22**, 59-74.

Delsuc, F. & Douzery, E.J. (2004b) Les méthodes probabilistes en phylogénie moléculaire : (2) L'approche bayésienne. *Biosystema*, **22**, 75-86.

Donoghue, P.C. & Benton, M.J. (2007) Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution*, **22**

Douzery, E.J., Blanquart, S., Criscuolo, A., Delsuc, F., Douady, C., Lartillot, N., Philippe, H. & Ranwez, V. (2010) Phylogénie moléculaire. *Biologie évolutive*, pp. 183-243. De Boeck, Paris.

Dowling, D.K., Friberg, U. & Lindell, J. (2008) Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in Ecology & Evolution*, **23**, 546-54.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PloS Biology*, **4**, e88.

Ellerman, J.R. (1940) Superfamily Dipodoidea. *The families and genera of living rodents*, pp. 560-598. Oxford University Press, London.

Emry, R.J. & Korth, W.W. (1989) *Rodents of the Bridgerian (Middle Eocene) Elderberry Canyon local fauna of Eastern Nevada*. Smithsonian Institution Press, Washigton, D.C.

Fabre, P.-H., Hautier, L., Dimitrov, D. & Douzery, E.J.P. (2012a) A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evolutionary Biology*, **12**, 88.

Fabre, P.-H., Pagès, M., Musser, G.G., Fitriana, Y.S., Fjeldså, J., Jennings, A., Jønsson, K.A., Kennedy, J., Michaux, J., Semiadi, G., Supriatna, N. & Helgen, K.M. (2013) A new genus of rodent from Wallacea (Rodentia: Muridae: Murinae: Rattini), and its implication for biogeography and Indo-Pacific Rattini systematics. *Zoological Journal of the Linnean Society*, **169**, 408-447.

Fabre, P.H., Irestedt, M., Fjeldså, J., Bristol, R., Groombridge, J.J., Irham, M. & Jønsson, K.A. (2012b) Dynamic colonization exchanges between continents and islands drive diversification in paradise-flycatchers (Terpsiphone, Monarchidae). *Journal of Biogeography*, **39**, 1900-1918.

Fahlbusch, V. (1992) The Neogene mammalian faunas of Ertemte and Harr Obo in Inner Mongolia (Nei Mongol), China. - 10. Eozapus (Rodentia). *Senckenbergiana lethaea*, **72**, 199-217.

Felsenstein, J. (1985) Confidence limits on phyloge- nies : an approach using the bootstrap. *Evolution*, **39**

Fokin, I.M. (1978) *The locomotion and morphology af the locomotory organs in jerboas*. Nauka.

Fortelius, M. (2013) New and Old Worlds Database of Fossil Mammals (NOW). In, University of Helsinki.

Fredga, K. (1996) The chromosome races of *Sorex araneus* in Scandinavia. *Hereditas*, **125**, 123-135.

Grant, B.S., Cook, A.D., Clarke, C.A. & Owen, D.F. (1998) Geographic and temporal variation in the incidence of melanism in peppered moth populations in America and Britain. *Journal of Heredity*, **89**, 465-471.

Green, M. (1977) Neogene Zapodidae (Mammalia: Rodentia) from South Dakota. *Journal of Paleontology*, **51**, 996-1015.

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.

Hall, E.R. (1930) Rodents and lagomorphs from the later Tertiary of Fish Lake Valley, Nevada. *University of California Press* **18**, 227-231.

Harry, M. (2001) *Génétique moléculaire et évolutive*. Maloine, Paris.

He, S., Cao, W. & Chen, Y. (2001) The uplift of qinghai-xizang (tibet) plateau and the vicariance speciation of glyptosternoid fishes (silurifoumes: sisoridae). *Science in China*, **44**, 650-651.

Hedrick, P.W. (2010) Cattle ancestry in bison: explanations for higher mtDNA than autosomal ancestry. *Molecular Ecology*, **19**, 3328-3335.
Henry, J.-P. & Gouyon, P.-H. (1998) *Précis de génétique des populations*, Masson edn.
Hewitt, G.M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.
Hewitt, G.M. (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87-112.
Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H., Johnson, J.B., Rissler, L., Victoriano, P.F. & Yoder, A.D. (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Mol Phylogenet Evol*, **54**, 291-301.
Hillis, D.M. (1987) Molecular versus morphological approaches to systématics. *Annual Review of Ecology and Systematics*, **18**, 23-42.
Ho, S.Y.W. & Phillips, M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367-380.
Holden, M.E. & Musser, G.G. (2005) Superfamily Dipodoidea. *Mammal Species of the World. A Taxonomic and Geographic Reference (3rd ed)*. Johns Hopkins University Press.
Huang, X. (1992) Zapodidae (Rodentia, Mammalia) from the middle Oligocene of Ulantatal, Nei Mongol. *Vertebrata PalAsiatica*, **30**, 249-303.
Hudson, R.R. & Coyne, J.A. (2002) Mathematical consequences of the genealogical species concept. *Evolution*, **56**
Huelsenbeck, J.P., Ronquist, F., Nielsen, R. & Bollback, J.P. (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, **294**, 2310-2314.
Hunt, T., Bergsten, J., Levkanicova, Z., Papadopoulou, A., St. John, O., Wild, R., Hammond, P.M., Ahrens, D., Balke, M., Bocakova, M., Bocak, L. & Vogler, A.P. (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, **318**, 1913-1916.
IUCN (2015) *The IUCN Red List of Threatened Species*. Available at: http://www.iucnredlist.org (accessed
Jaarola, M. & Tegelström, H. (1995) Colonization history of north European field voles (*Microtus agrestis*) revealed by mitochondrial DNA. *Molecular Ecology*, **4**, 299-310.
Jaarola, M., Tegelström, H. & Fredga, K. (1999) Colonization history in Fennoscandian rodents. *Biological Journal of the Linnean Society*, **68**, 113-127.
Jenkins, P., Kilpatrick, C., Robinson, M. & Timmins, R. (2005) Morphological and molecular investigations of a new family, genus and species of rodent (Mammalia: Rodentia: Hystricognatha) from Laos. *Systematics and Biodiversity*, **2**, 419-454.
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.
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. (2010) New material of dipodid rodents (Dipodidae, Rodentia) from the early Miocene of Gashunyinadege, Nei Mongol, China. *Journal of Vertebrate Paleontology*, **30**, 1860-1873.
Kimura, Y. (2011) The earliest record of birch mice from the Early Miocene Nei Mongol, China. *Naturwissenschaften*, **98**, 87-95.
Korth, W.W. (1993) Miosicista angulus, a new sicistine rodent (Zapodidae, Rodentia) from the Barstovian (Miocene) of Nebraska. *Transactions of the Nebraska Academy of Sciences*, **20**, 97-101.
Kotlik, P., Deffontaine, V., Mascheretti, S., Zima, J., Michaux, J.R. & Searle, J.B. (2006) A northern glacial refugium for bank voles (*Clethrionomys glareolus*). *PNAS*, **103**, 14860-14864.
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.
Lecointre, G. & Le Guyader, H. (2001) *Classification phylogénétique du vivant*, Belin edn, Paris.
LI, H. & Chi, H. (1981) Two new rodents from Neogene of Chilong Basin, Tibet. *Vertebrata PalAsiatica*, **19**, 245-255.
Li, Q. & Zheng, S.H. (2005) Note on four species of dipodids (Dipodidae, Rodentia) from the late Miocene Bahe Formation, Lantian, Shaanxi. *Vertebrata PalAsiatica*, **43**, 283-296.
Lomolino, M.V., Riddle, B.R., Whittaker, R.J. & Brown, J.H. (2010) *Biogeography*, 4th edn. Sinauer Associates, Inc., Sunderland.
Lopez Antonanzas, R. & Sen, S. (2006) New Saudi Arabian miocene jumping mouse (Zapodidae): systematics and phylogeny. *Journal of Vertebrate Paleontology*, **26**, 170-181.
Lyon, M. (1901) A comparison of the osteology of the jerboas and jumping mice. *Proc. U. S. Natl. Mus.*, **23**, 659-668.
Mahler, D.L., Ingram, T., Revell, L.J. & Losos, J.B. (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science*, **341**, 292-295.
Mallet, J. (2005) Hybridization as an invasion of the genome. *Trends Ecol Evol*, **20**, 229-37.
Martin, R.A. (1989) Early Pleistocene zapodid rodents from the Java local fauna of north-central South Dakota. *Journal of Vertebrate Paleontology*, **9**, 101-109.
Martin, R.A. (1994) A preliminary review of dental evolution and paleogeography in the Zapodid rodents, with emphasis on Pliocene and Pleistocene taxa. *Rodent and Lagomorph families of Asian origins and diversification* (ed. by Y. Tomida, C.K. Li and T. Setoguchi), pp. 99-113. National Science Monographs, Tokyo.
May, R.M. (1988) How many species are there on Earth? *Science*, **241**, 1441-1449.
Mayr, E. (1963) *Animal species and evolution*, Harvard University Press edn, Cambridge.
McKenna, M.C. & Bell, S.K. (1997) *Classification of mammal above the species level*. Columbia University Press, New-York.
Melnikova, E.N., Kshnyasev, I.A., Bodrov, S.Y., Mukhacheva, S.V., Davydova, Y.A. & Abramson, N.I. (2012) Sympatric area of *Myodes glareolus* and *M. rutilus* (Rodentia, Cricetidae): historic and recent hybridization. *Proceedings of the Zoological Institute RAS*, **316**, 307-323.

Melo-Ferreira, J., Alves, P.C., Freitas, H., Ferrand, N. & Boursot, P. (2009) The genomic legacy from the extinct Lepus timidus to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Mol Ecol*, **18**, 2643-58.

Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., McCain, C.M., McCune, A.R., McDade, L.A., McPeek, M.A., Near, T.J., Price, T.D., Ricklefs, R.E., Roy, K., Sax, D.F., Schluter, D., Sobel, J.M. & Turelli, M. (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol Lett*, **10**, 315-31.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B. & Worm, B. (2011) How many species are there on Earth and in the ocean? *PLoS Biology*, **9**, e1001127.

Morlon, H. (2014) Phylogenetic approaches for studying diversification. *Ecology letters*, **17**, 508-525.

Morlon, H., Potts, M.D. & Plotkin, J.B. (2010) Inferring the dynamics of diversification: a coalescent approach. *PloS Biol*, **8**, e1000493.

Morlon, H., Parsons, T.L. & Plotkin, J.B. (2011) Reconciling molecular phylogenies with the fossil record. *Proc Natl Acad Sci U S A*, **108**, 16327-16332.

Musser, G.G. & Carleton, M.D. (2005) Superfamily Muroidea. *Mammal Species of the World. A Taxonomic and Geographic Reference (3rd ed)*. Johns Hopkins University Press.

Myers, P., Espinosa, R., Parr, C.S., Jones, T., Hammond, G.S. & Dewey, T.A. (2015) *The Animal Diversity Web*. Available at: http://animaldiversity.org (accessed

Oliveri, I. (2010) Génétique et évolution des populations et des métapopulations. *Biologie évolutive* (ed. by D. Boeck), pp. 49-78, Bruxelles.

Osipova, O. & Sotkin, A.A. (2006) Bank and red vole hybridisation under experimental conditions. *Doklady Biological Sciences*, **410**, 381-383.

Osipova, O.V. & Soktin, A.A. (2008) Experimental simulation of ancient hybridization between bank and red voles. *Doklady Biological Sciences*, **420**, 169-171.

Pages, M., Bazin, E., Galan, M., Chaval, Y., Claude, J., Herbreteau, V., Michaux, J., Piry, S., Morand, S. & Cosson, J.F. (2013) Cytonuclear discordance among Southeast Asian black rats (Rattus rattus complex). *Mol Ecol*, **22**, 1019-34.

PaleobiologyDatabase (2015) *Paleobiology Database*. Available at: https://paleobiodb.org/ (accessed

Parmesan, C. (2006) Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637-669.

Pennisi, E. (2005) What determines species diversity? *Science*, **309**, 1.

Petit, R.J. & Excoffier, L. (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, **24**, 386-393.

Philippe, H., Zhou, Y., Brinkmann, H., Rodrigue, N. & Delsuc, F. (2005) Heterotachy and long-branch attraction in phylogenetics. *BMC Evol Biol*, **5**, 50.

Phillimore, A.B. & Price, T.D. (2008) Density-dependent cladogenesis in birds. *PloS Biol*, **6**, e71.

Pisano, J., Condamine, F.L., Lebedev, V., Bannikova, A., Quéré, J.-P., Shenbrot, G.I., Pagès, M., Michaux, J.R. & Riddle, B. (2015) Out of Himalaya: the impact of past Asian environmental changes on the evolutionary and biogeographical history of Dipodoidea (Rodentia). *Journal of Biogeography*, **42**, 856-870.

Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., Kamoun, S., Sumlin, W. & Vogler, A. (2006) Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology*, **55**, 595-609.

Potapov, S.G., Illarionova, N.A., Andreeva, T.A., Baskevich, M.I., Okulova, N.M., Lavrenchenko, L.A. & Orlov, V.N. (2007) Transfer of mitochondrial genome of the northern redbacked vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*) in northwestern Europe. *Doklady Biological Sciences*, **417**, 435-438.

Purvis, A. (2008) Phylogenetic Approaches to the Study of Extinction. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 301-319.

Qiu, Z. (2003) The Neogene mammalian faunas of Ertemte and Harr Obo in Inner Mongolia (Nei Mongol), China. *Senckenbergiana lethaea*, **83**, 135-147.

Qiu, Z. & Storch, G. (2000) The early Pliocene Micromammalian Fauna of Bilike, Inner Mongolia, China (Mammalia: Lipotyphla, Chiroptera, Rodentia, Lagomorpha). *Senckenbergiana lethaea*, **80**, 173-229.

Quental, T.B. & Marshall, C.R. (2010) Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol Evol*, **25**, 434-41.

Quéré, J.-P. & Le Louarn, H. (2011) *Les rongeurs de France, faunistique et biologie (Edition revue et augmentée)*. Quae.

Rabosky, D.L. (2009) Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol Lett*, **12**, 735-43.

Rabosky, D.L. (2013) Diversity-Dependence, Ecological Speciation, and the Role of Competition in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 481-502.

Rabosky, D.L. & Lovette, I.J. (2008) Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution*, **62**, 1866-1875.

Rabosky, D.L., Slater, G.J. & Alfaro, M.E. (2012) Clade age and species richness are decoupled across the eukaryotic tree of life. *PloS Biol*, **10**, e1001381.

Rabosky, D.L., Title, P.O. & Huang, H. (2015) Minimal effects of latitude on present-day speciation rates in New World birds. *Proceedings B*, **282**

Raymond, M. & Godelle, B. (2010) Applications de la biologie évolutive. *Biologie évolutive* (ed. by D. Boeck), Paris.

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.

Rhymer, J.M. & Simberloff, D. (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83-109.

Ricklefs, R.E. (2004) A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, **7**, 1-15.

Rofes, J., Garcia-Ibaibarriaga, N., Murelaga, X., Arrizabalaga, A., Iriarte, M.-J., Cuenca-Bescos, G. & Villaluenga, A. (2012) The southwesternmost record of Sicista (Mammalia: Dipodidae) in Eurasia, with a review f the palaeogeography and palaeoecology of the genus in Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **348-349**, 67-73.

Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014) Faster speciation and reduced extinction in the tropics contribute to the Mammalian latitudinal diversity gradient. *PLoS Biol*, **12**, e1001775.

Roughgarden, J. (1979) *Theory of population genetics and evolutionary ecology*. MacMillan Pub, New-York.

Rousset, F. (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Saetre, G.-P., Borge, T., Lindell, J., Moum, T., Primmer, C.R., Sheldon, B.C., Haavie, J., Johnsen, A. & Ellegren, H. (2008) Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. *Molecular Ecology*, **10**, 737-749.

Sanderson, M.J. & Hufford, L. (1996) *Homoplasy: the recurrence of similarity in evolution*. Academic Press.

Schemske, D.W., Mittelbach, G.G., Cornell, H.V., Sobel, J.M. & Roy, K. (2009) Is There a Latitudinal Gradient in the Importance of Biotic Interactions? *Annual Review of Ecology, Evolution, and Systematics*, **40**, 245-269.

Schenk, J.J., Rowe, K.C. & Steppan, S.J. (2013) Ecological opportunity and incumbency in the diversification of repeated continental colonizations by muroid rodents. *Systematic Biology*, **62**, 837-864.

Schulze, E.D. & Mooney, H.A. (1993) *Biodiversity and ecosystem function*. Springer-Verlag, New-York.

Shahin, A.A.B. (2005) Growth and maturation of metatarsals and their taxonomic significance in the jerboas Allactaga and Jaculus (Rodentia: Dipodidae). *Acta Zoologica*, **86**, 81-90.

Shenbrot, G.I. (1992) Cladistic approach to the analysis of phylogenetic relationships among dipodoid rodents. *Arch. Zool. Mosc. State Univ.*, **29**, 176-201.

Shenbrot, G.I., Sokolov, V.E., Heptner, V.G. & Koval'skaya, Y.M. (2008) *Jerboas: Mammals of Russia and adjacent regions*. Science Publishing Corporation.

Silvestro, D., Antonelli, A., Salamin, N. & Quentale, T.B. (2015) The role of clade competition in the diversification of North American canids. *PNAS*, **112**, 8684–8689.

Smith, A.B. & Peterson, K.J. (2002) Dating the time of origin of major clades: molecular clocks and the fossil record. *Annual Review of Earth and Planetary Sciences*, **30**, 65-88.

Springer, M.S., Stanhope, M.J., Madsen, O. & de Jong, W.W. (2004) Molecules consolidate the placental mammal tree. *Trends in Ecology & Evolution*, **19**, 430-8.

Stadler, T. (2011) Mammalian phylogeny reveals recent diversification rate shifts. *Proc Natl Acad Sci U S A*, **108**, 6187-6192.

Stein, B.R. (1990) Limb myology and phylogenetic relationships in the superfamily Dipodoidea (birch mice, jumping mice, and jerboas). *Z. Zool. Syst. Evol.*, **28**, 299-314.

Stephens, P.R. & Wiens, J.J. (2003) Explaining species richness from continents to communities: the time-for-speciation effect in Emydid turtles. *Am. Nat.*, **161**, 112-128.
Steppan, S.J., Adkins, R.M. & Anderson, J. (2004) Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Systematic Biology*, **53**, 533-553.
Stucky, R.K. & McKenna, M.C. (1993) Mammalia. *Tho fossil record 2* (ed. by M.J. Benton), pp. 43-52. Chapman & Hall, London.
Sulimski, A. (1962) Two new rodents from Weze 1 (Poland). *Acta Palaeontologica Polonica*, **7**, 503-511.
Taberlet, P., Swenson, J.E., Sandegren, F. & Bjarvall, A. (1995) Localization of a contact zone between highly divergent mitochondrial DNA lineages of the brown bear *Ursus arctos* in Scandinavia. *Conservation Biology*, **9**, 1255-1261.
Tegelström, H. (1987) Transfer of mitochondrial DNA from the Northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*Clethrionomys glareolus*). *Journal of Molecular Evolution*, **24**, 218-227.
Tesakov, A.S. (2001) New data on Plio-Pleistocene dipodine jerboas of the genus Plioscirtopoda from Eastern Europe (Mammalia: Dipodidae). *Zoosystematica Rossica*, **9**, 237-244.
Thomas, O. (1896) On the genera of rodents: an attemp to bring up to date the current arrangement of the order. *Proceedings of the Zoological Society of London*, 1012-1028.
Tilman, D. (2000) Causes, consequences, and ethics of biodiversity. *Nature*, **405**
Title, P.O. & Burns, K.J. (2015) Rates of climatic niche evolution are correlated with species richness in a large and ecologically diverse radiation of songbirds. *Ecol Lett*, **18**, 433-40.
Toews, D.P. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907-30.
Tong, Y. (1997) *Middle Eocene small mammals from liguanqiao basin of Henan province and Yuanqu basin of Shanxi province*.
Vinogradov, B. (1930) On the Classification of Dipodidae (Rodentia). *Bulletin de l'Académie des Sciences de l'URSS*, **4**, 331-470.
Vinogradov, B.S. (1937) Fauna of the USSR. *Mammals*. Academy Nauk SSSR Moscow.
Wake, D.B., Wake, M.H. & Specht, C.D. (2011) Homoplasy: from detecting pattern to determinign process and mechanism of evolution. *Science*, **331**, 1032-1035.
Wan, Q.H., Wu, H., Fujihara, T. & Fang, S.G. (2004) Review: Which genetic marker for which conservation genetics issue? *Electrophoresis*, **25**, 2165-2176.
Wang, B. (1985) Zapodidae (Rodentia, Mammalia) from the Lower Oligocene of Qujing, Yunnan, China. *Mainzer Geowissenschaftliche Mitteilungen*, **14**, 345-367.
Wang, B. & Qiu, Z. (2000) Dipodidae (Rodentia, Mammalia) from the lower member of Xianshuihe formation in Lanzhou Basin, Gansu, China. *Vertebrata PalAsiatica*, **38**, 10-35.
Wang, L., Schneider, H., Zhang, X.-C. & Xiang, Q.P. (2012) The rise of the Himalaya enforced the diversification of SE Asian ferns by altering the monsoon regimes. *BMC Plant Biology*, **12**

Wang, X.-M., Qiu, Z.-D., Li, Q., Yukimitsu, T., Kimura, Y., Tseng, Z.J. & Wang, H.-J. (2009) A new early to late Miocene fossiliferous region in central Nei Mongol: Lithostratigraphy and biogeography in Aoerban strata. *Vertebrata PalAsiatica*, **47**, 111-134.
Wiens, J.J. (2011) The causes of species richness patterns across space, time, and clades and the role of "ecological limits". *The Quarterly Review of Biology*, **86**, 75-96.
Wilson, D. & Reeder, D. (2005) *Mammal species of the world*, Third edn. Johns Hopkins University Press, Baltimore.
Wilson, R.W. (1936) A Pliocene rodent fauna from Smiths Valley, Nevada. *Carnegie Institution publications*, **473**, 15-34.
Winge, H. (1887) Jordfundne og nulevende Gnavere (Rodentia) fra Lagoa Santa, Minas Geraes, Brasilien. *Museo Lundi (Copenhagen)*, **1**, 1-178.
Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686-693.
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.
Zazhigin, V.S. & Lopatin, A.V. (2000) The history of Dipodoidea (Rodentia, Mammalia) in the Miocene of Asia: 3. Allactaginae. *Paleont. Zh.*, **34**, 553-565.
Zhang, M.L. & Fritsch, P.W. (2010) Evolutionary response of Caragana (Fabaceae) to Qinghai–Tibetan Plateau uplift and Asian interior aridification. *Plant Systematic & Evolution*, **288**, 191-199.
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*, **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.
Zink, R.M. & Barrowclough, G.F. (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107-21.
Zuckerkandl, E. & Pauling, L. (1965) Molecules as documents of evolutionary history. *Journal of theoretical biology*, **8**

Table S2 Echantillonnage taxonomique et numéros d'accession GenBank des séquences de gènes de Muroidea, de Dipodoidea et des groupes externes qui seront utilisées lors de l'étude de la dynamique de diversification des Dipodoidea et des Muroidea.

Muroidea's Family	Subfamily	Species		RAG1	IRBP	GHR	Cyt *b*	BRCA1
Spalacidae	**Myospalacinae**	*Eospalax fontanierii*	E_fontanierii	-	-	-	AF326266	-
		Myospalax aspalax	M_aspalax	-	AY326097	KC953281	AF326272	KC953192
	Spalacinae	*Spalax ehrenbergi*	S_ehrenbergi	JN414978	JN414825	AY294898	AF155871	JN414208
		Spalax nehringi	S_nehringi	-	-	-	JX451862	-
		Spalax zemni	S_zemni	-	SZU48589	-	KF021261	-
Calomyscidae		*Calomyscus bailwardi*	C_bailwardi	-	-	-	AF160604	-
		Calomyscus baluchi	C_baluchi	-	AY163581	GQ405372	AY288509	-
Cricetidae	**Cricetinae**	*Cricetulus barabensis (syn. griseus)*	C_barabensis	AY011885	AB033705	JN414756	AB033693	JN414195
		Cricetulus longicaudatus	C_longicaudatus	-	AY326082	-	AJ973386	-
		Cricetulus migratorius	C_migratorius	AY294956	KC953367	AY294926	AY288508	-
		Cricetus cricetus	C_cricetus	KC953488	AY277410	KC953253	AJ973392	KC953168
		Mesocricetus auratus	M_auratus	AY294955	AY163591	AF540632	AM904614	AY295013
		Mesocricetus brandti	M_brandti	-	-	-	AJ973380	-
		Mesocricetus newtoni	M_newtoni	-	-	-	AJ973381	-
		Mesocricetus raddei	M_raddei	-	-	-	AJ973382	-
		Phodopus campbelli	P_campbelli	-	-	-	AJ973389	-
		Phodopus roborovskii	P_roborovskii	-	-	-	GU797469	-
		Phodopus sungorus	P_sungorus	AY294954	KC953439	AF540640	AJ973390	AY295012
		Tscherskia triton	T_triton	-	-	-	AJ973388	-
	Arvicolinae = voles, Muskrat	*Alticola albicaudus*	A_albicaudus		KJ556779		KJ556729	
		Alticola argentatus	A_argentatus	-	KJ556777	-	DQ845186	-
		Alticola barakshin	A_barakshin	-	KJ556753	-	DQ845194	-
		Alticola macrotis	A_macrotis	-	-	-	AY309411	-
		Alticola semicanus	A_semicanus	-	KJ556765	-	DQ845193	JN244733
		Alticola strelzowi	A_strelzowi	-	KJ556755	-	DQ845190	-

		Alticola tuvinicus	A_tuvinicus	-	-	JN244720	-	JN244732
		Arborimus albipes	A_albipes	-	-	-	AY338815	-
		Arborimus longicaudus	A_longicaudus	-	-	-	AY338812	-
		*Arvicola amphibius (*syn. *Terrestris)*	A_amphibius	-	AY277407	AM392380	HQ728478	JX440345
		Arvicola sapidus	A_sapidus	-	JX457667	-	FJ539346	-
		Arvicola scherman	A_scherman	-	JX457669	-	JX457750	-
		Blanfordimys afghanus	B_afghanus	-	-	-	EF599109	-
		Blanfordimys bucharensis	B_bucharensis	-	-	AM392392	AM392369	-
		Chionomys gud	C_gud	-	-	JN244714	GQ352463	JN244726
		Chionomys nivalis	C_nivalis	-	AM919424	AM392378	AM392367	JX440344
		Chionomys roberti	C_roberti	-	-	JN244717	GQ352462	JN244729
		Dicrostonyx groenlandicus	D_groenlandicus	-	KJ556766	-	AJ238431	-
		Dicrostonyx torquatus	D_torquatus	-	-	AM392381	AJ238425	-
		Ellobius fuscocapillus	E_fuscocapillus	-	-	-	AF126430	-
		Ellobius talpinus	E_talpinus	-	-	GQ142005	-	-
		Ellobius tancrei	E_tancrei	-	-	GQ142006	AF119270	-
		Eolagurus luteus	E_luteus	-	-	GQ142004	-	-
		Eothenomys chinensis	E_chinensis	-	-	GQ374497	HM165437	-
		Eothenomys melanogaster	E_melanogaster	-	AY163583	AM392399	AB017254	JX440348
		Eothenomys miletus	E_miletus	-	KJ556773	-	AY426686	-
		Eothenomys proditor	E_proditor	-	KJ556772	-	AY426691	-
		Lagurus lagurus	L_lagurus	-	-	GQ142003	AF429818	-
		Lasiopodomys brandtii (syn. Microtus brandtii)	L_brandtii	-	JF906134	GQ374498	HQ123598	-
		Lasiopodomys fuscus	L_fuscus	-	JF906131	GQ374495	HQ123609	-
		Lasiopodomys mandarinus	L_mandarinus	-	AM919413	AM392396	FJ986322	-
		Lemmus amurensis	L_amurensis	-	-	-	FJ025977	-
		Lemmus lemmus	L_lemmus	-	-	-	FJ025984	-
		Lemmus sibiricus	L_sibiricus	-	AM919402	AM392398	FJ025980	-
		Lemmus trimucronatus	L_trimucronatus	-	-	-	FJ025982	-

		Microtus agrestis	M_agrestis	JX284363	JX457685	AM910792	JF318979	JX284316
		Microtus anatolicus	M_anatolicus	-	-	-	FJ767742	-
		Microtus arvalis	M_arvalis	JX284364	AM919416	AM392386	GU187386	JX284317
		Microtus cabrerae	M_cabrerae	JX284365	JX457698	-	JX457760	JX284318
		Microtus californicus	Mi_californicus	KC953523	KC953401	KC953277	EF506194	-
		Microtus canicaudus	M_canicaudus	-	-	-	AF163892	-
		Microtus chrotorrhinus	M_chrotorrhinus	-	AM919403	AM392383	AF163893	-
		Microtus daghestanicus	M_daghestanicus	-	-	GQ142009	AY513790	-
		Microtus dogramacii	M_dogramacii	-	-	-	AY513795	-
		Microtus fortis	M_fortis	-	JF906129	GQ374494	FJ986308	-
		Microtus gregalis	M_gregalis	-	-	GQ142007	GQ352466	-
		Microtus guentheri	M_guentheri	-	AM919420	AM392397	FJ767752	-
		Microtus irani	M_irani	-	-	-	FJ767750	-
		Microtus levis	M_levis	-	-	-	FJ641161	-
		Microtus limnophilus	M_limnophilus	-	AM919426	-	FJ986323	-
		Microtus longicaudus	M_longicaudus	-	AM919414	AM392379	AF205318	-
		Microtus majori	M_majori	-	AM919409	AM910796	AY513814	-
		Microtus miurus	M_miurus	-	-	-	GU809171	-
		Microtus mongolicus	M_mongolicus	-	-	-	FJ986310	-
		Microtus montanus	M_montanus	KC953524	KC953402	KC953278	DQ663650	-
		Microtus montebelli	M_montebelli	-	AM919421	AM910793	AF163900	-
		Microtus ochrogaster	M_ochrogaster	-	AM919423	AM392389	AF163901	-
		Microtus oeconomus	M_oeconomus	-	AM919418	AM392388	FJ986326	JX440347
		Microtus oregoni	M_oregoni	-	-	-	AF163903	-
		Microtus pennsylvanicus	M_pennsylvanicus	AY241463	AM919415	AF540633	AF119279	AY295009
		Microtus pinetorum	M_pinetorum	-	-	-	AF163904	-
		Microtus richardsoni	M_richardsoni	-	AM919404	AM392387	DQ663655	-
		Microtus schelkovnikovi	M_schelkovnikovi	-	AM919408	AM910794	AM910619	-
		Microtus socialis	M_socialis	-	FM162055	FM162073	GQ352468	-
		Microtus subterraneus	M_subterraneus	-	-	-	AJ717745	-

		Microtus thomasi	M_thomasi	-	AM919422	AM910797	EF666490	-
		Microtus townsendii	M_townsendii	-	-	-	AF163906	-
		Microtus transcaspicus	M_transcaspicus	-	AM919405	AM910795	-	-
		Microtus xanthognathus	M_xanthognathus	-	-	-	AF163907	-
		Myodes andersoni (syn. Phaulomys andersoni)	M_andersoni	-	KJ556763	AM392391	KJ556708	-
		Myodes californicus	M_californicus	-	KJ556775	-	AY309423	-
		Myodes centralis	M_centralis	-	KJ556738	-	DQ845185	-
		Myodes gapperi	M_gapperi	AY294952	AY326080	AF540623	FJ234423	AY295010
		Myodes glareolus	M_glareolus	-	JX457709	JF930127	HQ288418	JX440346
		Myodes regulus	M_regulus	-	-	-	AB017237	-
		Myodes rex	My_rex	-	-	-	AB017239	-
		Myodes rufocanus	M_rufocanus	-	KJ556761	GQ142002	AB676836	JF930117
		Myodes rutilus	M_rutilus	-	KJ556776	JF930130	JQ182076	JF930113
		Myopus schisticolor	M_schisticolor	-	-	GQ142001	EU165268	-
		Neodon irene (syn. Microtus irene)	N_irene	AY241464	JF906136	GQ374493	GU908338	-
		Neodon sikimensis (syn. Microtus sikimensis)	N_sikimensis	-	AY163593	GQ374496	JF906124	-
		Neofiber alleni	N_alleni	-	AM919425	AM910791	AM910618	-
		Ondatra zibethicus	O_zibethicus	AY294953	KC953427	AY294925	KC563206	AY295011
		Phaiomys leucurus (syn. Microtus leucurus)	P_leucurus	-	AM919400	AM392394	AM392371	-
		Phenacomys intermedius	P_intermedius	-	KC953438	AM392377	AF119260	-
		Proedromys bedfordi	P_bedfordi	-	JF906130	GQ374487	JF906118	-
		Proedromys liangshanensis	P_liangshanensis		JF906133	-	JF906119	-
		Prometheomys schaposchnikowi	P_schaposchnikowi	-	AM919406	AM392395	AM392372	-
		Synaptomys borealis	S_borealis	-	-	-	AF119259	-
		Synaptomys cooperi	S_cooperi	KC953586	KC953459	KC953335	DQ323957	KC953221
		Volemys musseri	V_musseri	-	JF906128	GQ374484	JF906121	-

		Volemys millicens	V_millicens	-	JF906135	GQ374481	JF906123	-
	Neotominae = woodrat, mouse	*Neotoma cinerea*	N_cinerea	KC953533	KC953409	-	JQ241243	-
		Neotoma devia	N_devia	KC953534	KC953410	-	DQ781302	-
		Neotoma floridana	N_floridana	AY294959	KC953411	-	AF294333	KC953195
		Neotoma fuscipes	N_fuscipes	-	-	-	AF337767	-
		Neotoma lepida	N_lepida	-	AY163599	-	DQ781256	-
		Neotoma magister	N_magister	-	-	-	DQ179856	-
		Ochrotomys nuttalli	O_nuttalli	KC953543	KC953422	KC953297	AY195798	KC953202
		Onychomys leucogaster	O_leucogaster	KC953550	EF989860	KC953303	-	-
		Peromyscus crinitus	P_crinitus	KC953558	KC953436	KC953310	EF028168	-
		Peromyscus keeni	P_keeni	-	-	-	DQ385716	-
		Peromyscus leucopus	P_leucopus	AY294957	EF989880	AY294927	DQ973104	AY295014
		Peromyscus maniculatus	P_maniculatus	HM072277	AY163630	EF989783	JF489123	-
		Peromyscus truei	P_truei	-	AY277413	-	HQ269735	-
		Reithrodontomys megalotis	R_megalotis	KC953572	AY277414	KC953323	AF176249	-
	Sigmondontinae = mouse, rat	*Oryzomys palustris*	O_palustris	KC953551	AY163623	KC953304	EU074640	KC953205
		Sigmodon arizonae	S_arizonae	KC953581	EU635700	KC953331	AF155423	KC953218
		Sigmodon hispidus	S_hispidus	AY241465	AY277479	AF540641	HQ326714	AY295016
	Lophiomyinae	*Lophiomys imhausi*	L_imhausi	KC953514	KC953389	-	-	-
Muridae	**Deomyinae**	*Acomys cahirinus*	A_cahirinus	-	FN984743	FN984742	Z96053	-
		Acomys chudeaui	A_chudeaui	-	-	-	FJ415533	-
		Acomys dimidiatus	A_dimidiatus	-	-	-	Z96062	-
		Acomys ignitus	A_ignitus	AY294951	KC953348	AY294923	JN247674	AY295008
		Acomys russatus	A_russatus	-	FM162053	FM162071	Z96066	-
		Deomys ferrugineus	D_ferrugineus	AY241460	AY326084	AY294922	FJ415478	AY295007
		Lophuromys flavopunctatus	L_flavopunctatus	AY294950	AY326091	AY294921	AY828236	AY295006
	Gerbillinae	*Brachiones przewalskii*	B_przewalskii	-	-	-	AB381903	-
		Desmodilliscus braueri	D_braueri	-	FN357289	-	AJ851273	-
		Desmodillus auricularis	D_auricularis	KC953494	AM910940	DQ019048	AJ851272	-

		Dipodillus campestris	D_campestris	-	-	-	JN652801	-
		Dipodillus dasyurus (syn. Gerbillus dasyurus)	D_dasyurus	-	FM162054	FM162072	-	-
		Dipodillus rupicola	D_rupicola	-	-	-	KF496284	-
		Dipodillus simoni	D_simoni	-	-	-	GU356579	-
		Gerbilliscus afra (syn. Tatera afra)	G_afra	-	-	-	AM409388	-
		Gerbilliscus brantsii	G_brantsii	-	-	-	AM409392	JF716029
		Gerbilliscus guineae(syn. Tatera guineae)	G_guineae	-	AM408334	-	AM409379	-
		Gerbilliscus (Taterona) kempi	G_kempi	-	-	-	AJ430561	JF716043
		Gerbilliscus leucogaster (syn. Tatera leucogaster)	G_leucogaster	-	-	-	AM409389	JF716051
		Gerbilliscus nigricaudus (syn. Tatera nigricauda)	G_nigricauda	-	-	-	AJ875247	-
		*Gerbilliscus robustus (*syn. *Tatera robusta)*	G_robustus	KC953587	AY326113	AY294920	AM409372	AY295005
		Gerbilliscus validus (syn. Tatera valida)	G_validus	-	-	-	AJ875299	-
		Gerbillurus paeba	G_paeba	KC953500	AM910941	KC953261	AJ430557	JF716027
		Gerbillurus tytonis	G_tytonis	EU349886	EU349845	-	AJ430559	-
		Gerbillurus setzeri	G_setzeri	-	-	-	AJ430558	-
		Gerbillurus vallinus	G_vallinus	AY294948	KC953377	AF332022	-	-
		Gerbillus andersoni	G_andersoni	-	-	-	KF496286	-
		Gerbillus gerbillus	G_gerbillus	DQ023452	EU349846	DQ019049	AJ851269	-
		Gerbillus henleyi	G_henleyi	-	-	-	JQ753050	JF716023
		Gerbillus hesperinus	G_hesperinus	-	-	-	JN652803	-
		Gerbillus hoogstraali	G_hoogstraali	-	-	-	JN021413	-
		Gerbillus latastei	G_latastei	-	-	-	GU356562	-
		Gerbillus nancillus	G_nancillus	-	-	-	KF496237	-
		Gerbillus nanus	G_nanus	KC953501	KC953378	KC953262	JQ753063	-
		Gerbillus nigeriae	G_nigeriae	-	AM408333	-	AF141226	-
		Gerbillus occiduus	G_occiduus	-	-	-	JN652805	-
		Gerbillus perpallidus	G_perpallidus	-	-	-	JN652806	-

		Gerbillus poecilops	G_poecilops	-	-	-	JQ753064	-
		Gerbillus pyramidum	G_pyramidum	-	-	-	JN652813	-
		Gerbillus tarabuli	G_tarabuli	-	-	-	JN652832	-
		Meriones chengi	M_chengi	-	-	-	AB381900	-
		Meriones crassus	M_crassus	-	-	-	AJ851267	-
		Meriones libycus	M_libycus	-	-	-	JQ927412	-
		Meriones meridianus	M_meridianus	-	JQ713292	-	AB381899	-
		Meriones rex	M_rex	-	-	-	AJ851265	-
		Meriones shawi	M_shawi	AY294947	KC953400	AF332021	-	AF332048
		Meriones tamariscinus	M_tamariscinus	-	-	-	JN604782	-
		Meriones unguiculatus	M_unguiculatus	-	AY326095	AF247184	AF159405	-
		Pachyuromys duprasi	P_duprasi	-	-	-	JF704120	JF716024
		Psammomys obesus	P_obesus	-	FN357290	-	AJ851275	-
		Psammomys vexillaris	P_vexillaris	-	-	-	AY934541	-
		Rhombomys opimus	R_opimus	-	-	-	AJ430556	-
		Sekeetamys calurus	S_calurus	-	-	-	AJ851276	-
		Tatera indica	T_indica	-	-	-	AJ430563	-
		Taterillus arenarius	T_arenarius	-	FN357288	-	AJ851261	-
		Taterillus emini	T_emini	DQ023453	KC953461	DQ019050	-	KC953224
		Taterillus gracilis	T_gracilis	-	-	-	AJ851263	JF716026
		Taterillus pygargus	T_pygargus	-	-	-	AJ851262	-
	Murinae	*Micaelamys namaquensis (syn. Aethomys namaquensis)*	A_namaquensis	AY294941	154090312	AY294914	AF141215	EU349649
		Aethomys chrysophilus	A_chrysophilus	-	AY326075	JQ694059	AJ604515	-
		Apodemus agrarius	A_agrarius	DQ023472	AB096842	DQ019054	AB096817	EU349658
		Apodemus alpicola	A_alpicola	AB285442	AB032861	-	AF159391	-
		Apodemus argenteus	A_argenteus	AB285443	AB032855	-	AB032848	-
		Apodemus chevrieri	A_chevrieri	AB285444	AB096848	-	AB096819	-
		Apodemus draco	A_draco	AB285445	AB096850	-	AB096825	-
		Apodemus flavicollis	A_flavicollis	AB285446	AB032860	AM910943	AB032853	-

		Apodemus gurkha	A_gurkha	AB285447	AB032859	-	AB032852	-
		Apodemus latronum	A_latronum	AB303239	AB096852	GU908448	AB096836	-
		Apodemus mystacinus	A_mystacinus	AB285448	AB303229	KM397259	KM397213	KC953157
		Apodemus pallipes	A_pallipes	-	-	-	AF160603	-
		Apodemus peninsulae	A_peninsulae	AB285449	AB032857	-	AB032850	-
		Apodemus ponticus	A_ponticus	AB303242	-	-	FN433634	-
		Apodemus semotus	A_semotus	AB164040	AB032862	DQ019055	AB033694	-
		Apodemus speciosus	A_speciosus	AB285450	AB032856	AB491493	AB032849	-
		Apodemus sylvaticus	A_sylvaticus	AB303243	AB032863	KM39725	AB033695	-
		Apodemus uralensis	A_uralensis	AB285451	AB096853	-	AB096837	-
		Apomys hylocetes	A_hylocetes	DQ023465	KC953357	AY294915	AY324467	AY295000
		Archboldomys luzonensis	A_luzonensis	DQ023466	DQ191495	EU349794	AY324460	EU349675
		Arvicanthis niloticus	A_niloticus	-	DQ022386	KC953243	AF004569	-
		Bandicota bengalensis	B_bengalensis	-	AM408331	AM910945	AM408336	-
		Batomys granti	B_granti	AY241461	DQ191496	AY294917	AY324458	AY295002
		Berylmys bowersi	B_bowersi	DQ023457	AM407896	AM910946	AM408337	KC953160
		Bunomys chrysocomus	B_chrysocomus	-	AM910937	AM910947	AM910934	-
		Chrotomys gonzalesi	C_gonzalesi	AY294943	DQ191503	GQ405375	AY324461	-
		Colomys goslingi	C_goslingi	-	DQ022395	AM910948	AF518372	-
		Conilurus penicillatus	C_penicillatus	DQ023467	AM910938	AM910949	AM910935	-
		Cremnomys cutchicus	C_cutchicus	-	DQ022384	-	DQ022381	-
		Dasymys incomtus	D_incomtus	-	EU292143	AM910950	AF141217	-
		Desmomys harringtoni	D_harringtoni	-	EU292144	-	AF141206	-
		Diplothrix legata	D_legata	EU349885	AB033706	EU349799	AB033696	EU349670
		Golunda ellioti	G_ellioti	-	AM408332	AM910951	AM408338	-
		Grammomys macmillani	G_macmillani	-	AY326086	AM910980	AM408345	KC953175
		Heimyscus fumosus	H_fumosus	-	DQ022397	AM910953	AF518333	-
		Hybomys univittatus	H_univittatus	DQ023462	DQ022388	DQ019059	AF141219	KC953181
		Hydromys chrysogaster	H_chrysogaster	EU349890	AM408319	AM910954	AM408339	EU349699
		Hylomyscus parvus	H_parvus	DQ023479	DQ022399	DQ019060	AF518330	-

		Hylomyscus stella	H_stella	-	AM408320	AM910955	AF518331	-
		Lemniscomys barbarus	L_barbarus	DQ023461	KC953387	DQ019062	-	KC953184
		Lemniscomys striatus	L_striatus	-	AM408321	AM910956	AF141210	-
		Leopoldamys edwarsi	L_edwarsi	-	AJ698897	-	AJ698881	-
		Malacomys edwarsi	M_edwarsi	-	DQ022392	AM910958	DQ022379	-
		Malacomys longipes	M_longipes	DQ023474	DQ022393	AM910957	AM408341	EU349656
		Mastomys erythroleucus	M_erythroleucus	KC953519	AM408335	AM910959	AF518338	KC953189
		Mastomys kollmannspergeri	M_kollmannspergeri	-	-	AM910961		-
		Mastomys natalensis	M_natalensis	EU349899	AY326093	EU349813	AF518342	EU349660
		Mastomys pernanus	M_pernanus	-	DQ022403	AM910960	AF518343	-
		Maxomys whiteheadi	M_whiteheadi	-	AY326094	KC878193	EU292150	-
		Micromys minutus	M_minutus	EU349904	EU349862	EU349818	AB033697	EU349664
		Millardia kathleenae	M_kathleenae	-	EU292145	AM910963	EU292148	-
		Millardia meltada	M_meltada	-	AM408322	AM910962	AF141221	-
		Mus crociduroides	M_crociduroides	-	AJ698894	AM910964	AJ698878	-
		Mus macedonicus	M_macedonicus	AB125829	AB125805	-	AB125770	-
		Mus minutoides	M_minutoides	-	AJ875086	-	AY057816	-
		Mus musculus	M_musculus	AY241462	JX457617	AY271378	V00711	EU349657
		Mus platythrix	M_platythrix	AB125845	AJ698895	-	AJ698880	-
		Mus spicilegus	M_spicilegus	AB125835	AJ698882	-	AB125775	-
		Mus spretus	M_spretus	AB125836	JX457619	-	AB033700	-
		Mylomys dybowskii	M_dybowskii	-	EU292146	AM910965	AF141212	-
		Myomyscus brockmani	M_brockmani	-	-	AM910966	JX126608	-
		Myomyscus verreauxii (syn. Myomys verreauxii)	M_verreauxii	-	DQ022408	AM910967	AF518355	-
		Myomyscus yemeni	M_yemeni	-	DQ022409	AM910968	AF518357	-
		Niviventer niviventer	N_niviventer	-	AM408323	AM910969	AM408344	-
		Oenomys hypoxanthus	O_hypoxanthus	DQ023464	AM408324	AM910970	AM408342	EU349654
		Pelomys fallax	P_fallax	-	DQ022391	JQ694062	DQ022382	-
		Phloeomys cumingi	P_cumingi	-	AY326103	-	DQ191484	-

		Praomys daltoni (syn. Myomys daltoni)	P_daltoni	-	DQ022406	AM910972	AF518349	-
		Praomys degraaffi	P_degraaffi	KC953562	DQ022410	KC953315	AF518359	-
		Praomys jacksoni	P_jacksoni	DQ023477	AM408326	AM910973	AF518361	EU349663
		Praomys misonnei	P_misonnei	KC953563	DQ022412	JF284233	AF518364	-
		Praomys tullbergi	P_tullbergi	DQ023478	AM408327	AM910974	AF518365	EU349662
		Praomys verschureni (syn. Malacomys verschureni)	P_verschureni	-	DQ022394	-	AF518373	-
		Pseudomys australis	P_australis	DQ023469	AM910939	AM910975	AM910936	EU349688
		Rattus norvegicus	R_norvegicus	AY294938	AB033709	-	AB033713	EU349671
		Rattus pyctoris	R_pyctoris	-	-	-	JN675511	-
		Rattus exulans	R_exulans	DQ023455	AY326105	DQ019074	DQ191486	-
		Rattus rattus	R_rattus	-	HM217606	AM910976	AB033702	-
		Rattus tanezumi	R_tanezumi	-	AB096856	GQ405393	AB096841	-
		Rhabdomys pumilio	R_pumilio	EU349916	AY326106	AY294913	AF141214	-
		Rhynchomys isarogensis	R_isarogensis	AY294944	AY326108	DQ019075	AY324462	-
		Stenocephalemys albipes (syn. Myomys albipes)	S_albipes	-	DQ022404	AM910977	AF518347	-
		Stenocephalemys albocaudata	S_albocaudata	-	DQ022414	AM910978	AF518370	-
		Stochomys longicaudatus	S_longicaudus	-	EU292147	DQ019076	EU292149	-
		Sundamys muelleri	S_muelleri	-	AY326111	AM910979	AM408340	-
		Tokudaia muenninki	T_muenninki	AB548695	AB548697	-	AB548692	-
		Tokudaia osimensis	T_osimensis	EU349918	AB033712	AM910981	AB029429	-
		Zelotomys hildegardeae	Z_hildegardeae	DQ023476	DQ022396	DQ019080	AF518375	EU349661
	Otomyinae	*Otomys angoniensis*	O_angoniensis	-	AM408325	AM910971	AM408343	-

Dipodoidea's Family	Subfamily	Species		Cyt *b*	BRCA1	GHR	IRBP	RAG1
Dipodidae	Allactaginae	*Allactaga balikunica*	*A_balikunica*	KM397180	KM397274	KM397227	KM397136	KM397313
		Allactaga bullata	*A_bullata*	KM397179	JQ347887	JQ347909	JQ347929	JQ347865
		Allactaga elater	*A_elater*	KM397178	JQ347891	JQ347913	JQ347933	JQ347869
		Allactaga euphratica	*A_euphratica*	JQ954953	-	-	-	-
		Allactaga firouzi	*A_firouzi*	JQ954939	-	-	-	-
		Allactaga hotsoni	*A_hotsoni*	JQ954961	JF938780	JF938857	JF938883	-
		Allactaga major	*A_major*	-	JQ347889	JQ347911	JQ347931	JQ347867
		Allactaga sp. (Kazakhstan)	*A_sp*	-	KM397273	KM397226	KM397135	KM397312
		Allactaga sibirica	*A_sibirica*	KM397214	KM397272	KM397225	KM397134	KM397311
		Allactaga toussi	*A_toussi*	JQ954956	-	-	-	-
		Allactaga williamsi	*A_williamsi*	JQ954945	-	-	-	-
		Allactodipus bobrinskii	*A_bobrinskii*	KM397181	JQ347892	JQ347914	JQ347934	JQ347870
		Pygeretmus platyurus	*P_platyurus*	-	-	-	KM397165	KM397314
		Pygeretmus pumilio	*P_pumilio*	KM397182	JQ347894	JQ347916	JQ347936	JQ347872
	Cardiocraniinae	*Cardiocranius paradoxus*	*C_paradoxus*	KM397211	JQ347884	JQ347906	JQ347926	JQ347862
		Salpingotus kozlovi	*S_kozlovi*	KM397210	KM397301	KM397253	KM397159	KM397341
		Salpingotus crassicauda	*S_crassicauda*	Unpublished sequences from V. Lebedev.				
	Dipodinae	*Dipus sagitta*	*D_sagitta*	KM397196	KM397284	KM397240	KM397148	KM397327
		Eremodipus lichtensteini	*E_lichtensteini*	KM397200	JQ347895	JQ347917	JQ347937	JQ347873
		Jaculus blanfordi	*J_blanfordi*	KM397195	JQ347900	JQ347921	JQ347941	JQ347878
		Jaculus jaculus	*J_jaculus*	KM397184	KM397276	KM397229	KM397138	KM397316
		Jaculus favonicus	*J_favonicus*	KM397187	KM397279	KM397232	KM397141	KM397319
		Jaculus orientalis	*J_orientalis*	KM397189	KM397281	KM397234	KM397143	KM397321

		Paradipus ctenodactylus	*P_ctenodactylus*	KM397183	KM397275	KM397228	KM397137	KM397315
		Stylodipus andrewsi	*S_andrewsi*	KM397198	KM397285	KM397242	KM397163	KM397329
		Stylodipus sungorus	*S_sungorus*	KM397199	KM397286	KM397243	KM397164	KM397330
		Stylodipus telum	*S_telum*	KM397197	JQ347898	JQ347920	JQ347940	JQ347876
	Euchoreutinae	*Euchoreutes naso*	*E_naso*	KM397212	JQ347885	JQ347907	JQ347927	JQ347863
Sminthidae		*Sicista caucasica*	*S_caucasica*	KM397201	KM397287	KM397244	KM397150	KM397331
		Sicista concolor	*S_concolor*	KM397207	KM397294	KM397252	KM397167	KM397339
		Sicista kazbegica	*S_kazbegica*	KM397202	KM397288	KM397245	KM397151	KM397332
		Sicista kluchorica	*S_kluchorica*	KM397206	KM397293	KM397251	KM397157	KM397338
		Sicista napaea	*S_napaea*	-	KM397292	KM397250	KM397156	KM397337
		Sicista strandi	*S_strandi*	KM397209	-	-	KM397158	KM397340
		Sicista subtilis	*S_subtilis*	KM397208	JQ347880	-	JQ347923	JQ347858
		Sicista tianshanica	*S_tianshanica*	KM397204	KM397290	KM397247	KM397153	KM397334
Zapodidae		*Eozapus setchuanus*	*E_setchuanus*	KM397176	KM397270	KM397223	KM397132	KM397309
		Napaeozapus insignis	*N_insignis*	KM397168	KM397262	KM397215	KM397124	KM397302
		Zapus hudsonius	*Z_hudsonius*	KM397170	KM397264	KM397217	KM397126	KM397304
		Zapus princeps	*Z_princeps*	KM397174	KM397268	KM397221	KM397130	KM397308
		Zapus trinotatus	*Z_trinotatus*	KM397172	KM397266	KM397219	KM397128	KM397306

OUTGROUP's Family	Subfamily	Species		BRCA1	GHR	IRBP	RAG1	Cyt *b*
Aplodontiidae		*Aplodontia rufa*	A_rufa	AF332045	AF332030	AJ427238	AY241468	AJ389528
Gliridae	Glirinae	*Glis glis*	G_glis	KM397298	AM407916	AJ427235	AB253972	AJ225031
	Leithiinae	*Dryomys nitedula*	D_nitedula	KM397299	KM397257	AJ427236	-	AJ225116
Sciuridae	Sciurinae	*Sciurus aestuans*	S_aestuans	KM397297	FM162078	FM162057	KM397345	AJ389530
	Xerinae	*Marmota marmota*	M_marmota	-	KM397255	KM397161	KM397342	AF100711