Diversity and endemism of Murinae rodents in Thai limestone karsts

Alice Latinne, Surachit Waengsothorn, Prateep Rojanadilok, Kraisat Eiamampai, Kriangsak Sribuarod, & Johan R. Michaux

Laboratoire de génétique des micro-organismes, Institut de Botanique - B22, Université de Liège, Boulevard du Rectorat, 27, 4000, Liège, Belgium
Environment and Resources Technology Department, Thailand Institute of Scientific and Technological Research, Pathum Thani, Thailand
Doi Chiangdao Wildlife Research Station, Chiang Mai, Thailand
Bung Boraphet Wildlife Research Station, Nakhon Sawan, Thailand
Khlong Saeng Wildlife Research Station, Surat Thani, Thailand
Centre de Biologie et de Gestion des Populations, UMR INRA/IRD/Cirad/Montpellier SupAgro, Campus international de Baillarguet, Montferrier-sur-Lez, France

Published online: 05 Sep 2013.

To cite this article: Systematics and Biodiversity (2013): Diversity and endemism of Murinae rodents in Thai limestone karsts, Systematics and Biodiversity

To link to this article: http://dx.doi.org/10.1080/14772000.2013.818587

Please scroll down for article.

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
Research Article

Diversity and endemism of Murinae rodents in Thai limestone karsts

ALICE LATINNE1, SURACHIT WAENGSOTHORN2, PRATEEP ROJANADILOK3, KRAIRAT EIAMAMPAI4, KRIANGSAK SRIBUAROD5 & JOHAN R. MICHAUX1,6

1Laboratoire de génétique des micro-organismes, Institut de Botanique – B22, Université de Liège, Boulevard du Rectorat, 27 4000, Liège, Belgium
2Environment and Resources Technology Department, Thailand Institute of Scientific and Technological Research, Pathum Thani, Thailand
3Doi Chiangdao Wildlife Research Station, Chiang Mai, Thailand
4Bung Boraphet Wildlife Research Station, Nakhon Sawan, Thailand
5Khlong Saeng Wildlife Research Station, Surat Thani, Thailand
6Centre de Biologie et de Gestion des Populations, UMR INRA/IRD/Cirad/Montpellier SupAgro, Campus international de Baillarguet, Montferrier-sur-Lez, France

(Received 18 December 2012; revised 30 May 2013; accepted 4 June 2013)

This study aims to investigate the species diversity of rodents living in karst ecosystems of Thailand. A survey has been conducted throughout Thailand, 122 karsts sampled and 477 Murinae rodents live-trapped. Phylogenetic reconstructions were carried out using two mitochondrial markers (cytb, COI). A sequence-based species delimitation method completed by the analysis of the level of genetic divergence was then applied to define species boundaries within our dataset. The phylogenetic position of Niviventer hinpoon was also investigated and sequences obtained from the holotype specimen of this species were used to reliably identify samples of N. hinpoon. A total of 12 described Murinae species, corresponding to 17 deeply divergent genetic lineages, were encountered in limestone karsts of Thailand. Our study revealed an important genetic diversity within the traditionally recognized species Maxomys surifer (four highly divergent genetic lineages), Leopoldamys neilli (two highly divergent genetic lineages) and Berylmys bowersi (two highly divergent genetic lineages). These species could be considered as species complex and require further taxonomic work. This study also provides valuable information on the distribution of the two rodent species endemic to limestone karsts of Thailand, L. neilli and N. hinpoon. Leopoldamys neilli was the most abundant species encountered in Thai karsts during our survey. However, L. neilli specimens from western Thailand are genetically highly divergent from the remaining populations of L. neilli and could represent a separate species. Niviventer hinpoon, phylogenetically closely related to N. fulvescens, is much rarer and its distribution limited to central Thailand. Most of the other captured species are typically associated with forest ecosystems. This study suggests that limestone karsts play a key role in the preservation of the rodent species endemic to such habitat, but they would also provide refuges for the forest-dwelling Murinae rodents in deforested regions.

Key words: GMYC lineages, limestone karsts, Murinae rodents, Rattini, Southeast Asia, species delimitation, Thailand

Introduction

Southeast Asia represents only 3% of the total land area of the world but this region includes four major biodiversity hotspots that host around 18% of the total plant and animal species of the earth (Myers et al., 2000; ASEAN CFB, 2010). This region is a centre of diversity and endemism of Murinae rodents, a sub-family of the Muridae family. This highly speciose group includes 126 genera and 561 species largely distributed in the Old World (Musser & Carleton, 2005). Lecompte et al. (2008) subdivided the Murinae in several tribes, among which the Rattini tribe includes 37 genera and 170 rat species, most of them distributed in Southeast Asia. This exceptionally high diversification of Murinae in Southeast Asia is the result of several spectacular radiations during the last 10 million years (Chaimanee & Jaeger, 2001).

Southeast Asian limestone karsts were the site of several major rodent discoveries in the last few years, leading to the description of the mysterious Laonastes aenigmamus in Lao PDR (Jenkins et al., 2005), the sole remaining representative of the Diatomyidae family that occurred in Asia from 33.9 to 11.6 million years ago (Mya) (Huchon et al., 2007), and two new Murinae genera, Saxatilomys
Fig. 1. Limestone karsts in Southeast Asia (dark grey) and distribution of Southeast Asian Murinae rodents endemic to limestone karsts (red): *Leopoldamys neilli* (distribution according to Latinne et al., 2011, 2012), *Niviventer hinpoon* (Lekagul & McNeely 1988), *Saxatilomys paulinae* (Musser et al., 2005), *Tonkinomys daovantieni* (Musser et al., 2006).

and *Tonkinomys* (Musser et al., 2005, 2006). Limestone karsts are sedimentary rock outcrops consisting of calcium carbonate characterized by underground drainage systems and distinctive landforms resulting from solutional erosion and made of cone-shaped or sheer-sided hills riddled with caves and sinkholes (Gillieson, 2005; Simms, 2005). Some of the largest limestone karst regions of the world are located in Southeast Asia where they cover an area of 460 000 km² (Day & Urich, 2000) (Fig. 1). In Thailand, limestone karsts are widespread throughout most parts of the country, except for large areas on the northeastern Khorat Plateau. Karst biodiversity is characterized by an important species richness and high levels of endemic species of plants, vertebrates and invertebrates adapted to this particular environment (Vermeulen & Whitten, 1999). These high species diversity and endemism in karsts probably arise from the diverse and extreme environmental conditions of these ecosystems such as alkaline soils, rough topography, frequent drought, thin soil layers and darkness in caves (Vermeulen & Whitten, 1999; Clements et al., 2006). Limestone karsts have also a significant economic value as limestone is an essential raw material used in major industries and karsts are currently over-exploited through quarrying of cement, lime and hard core (Vermeulen & Whitten, 1999). As a result of these activities, karst ecosystems are highly threatened and some endemic species became dramatically endangered. This situation is particularly alarming as karst biodiversity suffers from a considerable lack of scientific data and remains widely unknown throughout Southeast Asia despite its high biological importance (Clements et al., 2006). Consequently, many undiscovered species could already be extinct or may become extinct within the next few years. In order to define conservation guidelines and achieve long-term preservation of karst ecosystems in Southeast Asia, it is therefore urgently necessary to increase our knowledge of this biodiversity.

Four species of Murinae rodents belonging to the Rattini tribe and endemic to limestone karsts are currently recognized in Southeast Asia, two in Thailand, *Leopoldamys neilli* (Marshall, 1976) and *Niviventer hinpoon* (Marshall, 1976), one in Lao PDR, *Saxatilomys paulinae* Musser, Smith, Robinson & Lunde, 2005, and one in Vietnam,
Waengsothorn et al., 2008 (Fig. 1). However little is currently known about these rare rodents. *Leopoldamys neilli* and *N. hinpoon* were discovered in 1973 in Saraburi and Lopburi provinces (central Thailand) (Lekagul & McNeely, 1988) and very little information about their distribution and ecological requirements were collected during more than 30 years. Due to this lack of data, the two species are currently considered Data Deficient on the IUCN Red List (Lunde & Aplin, 2008a, 2008b). A few years ago, a field survey of Waengsothorn et al. (2007) in central Thailand confirmed the presence of these two species in sparse locations of this region. More recently, two studies revealed a complex phylogeographical structure for *L. neilli* in Thailand and a wider distribution range for this species than indicated by previous records (Latinne et al., 2011, 2012) (Fig. 1).

However, the Murinae rodent diversity of Thai limestone karsts remains poorly understood and the aims of the present study were to investigate exhaustively the species composition of the rodent fauna in karst ecosystems of Thailand using DNA-based methods. As Thailand occupies a key biogeographical position at the intersection of the Indochinese and Sundaic biogeographical subregions (Hughes et al., 2003; Woodruff & Turner, 2009), we could expect to observe a faunal transition of rodent diversity from northern to southern surveyed locations.

Most Rattini genera display several external and skull diagnostic characters and can be easily identified using morphological criteria. At the intragenic level, however, species identification may be blurred by limited interspecific external morphological divergence and/or by large intraspecific morphological variations. This phenomenon combined with a lack of accurate morphological study for some species may lead to a confused taxonomy and an overabundance of species description and synonyms or, conversely, undiscovered cryptic species (two or more distinct species classified as a single species because they are at least superficially morphologically indistinguishable (Hebert et al., 2004; Bickford et al., 2007)). The genus *Rattus* and in particular the *Rattus ratus* complex (Aplin et al., 2011) represents one of the best examples of this minimal external morphological divergence among Rattini species that evolved rapidly and recently and that are differentiated only by subtle morphological variations (Rowe et al., 2011). DNA-based approaches have recently proved to be useful tools to clarify the complex taxonomy of Rattini and especially of the *Rattus* genus (Robins et al., 2007; Pages et al., 2010; Aplin et al., 2011; Bastos et al., 2011) and helped to elucidate the specific status of several *Rattus* cryptic species such as *Rattus losea* and *Rattus sakeratensis* (formerly treated as a synonym of *R. losea*) (Aplin et al., 2011) or *Rattus ratus* and *Rattus satarae* (formerly treated as a subspecies of *R. ratus*) (Pages et al., 2011). Discovery and description of cryptic species is particularly important for conservation and natural resources management as design of conservation strategies mainly relies on estimates of species richness and endemism (Bickford et al., 2007). Moreover each species within a cryptic complexes will require adapted conservation plans and their recognition is therefore essential.

In this study, we studied a large number of Murinae samples from Thai limestone karsts using DNA-based methods. Phylogenetic reconstructions were carried out using two mitochondrial markers. A sequence-based species delimitation method completed by the analysis of the level of genetic divergence was then applied to define species boundaries within our dataset. Finally, in order to name the delimited species, we used DNA sequences identified in previous studies and DNA sequences obtained from the holotype specimen of *N. hinpoon*.

### Materials and methods

#### Sample collection

In order to study the Murinae rodent biodiversity of Thai limestone karsts, a survey was conducted and 122 karst localities were sampled in 25 provinces including all karst regions of Thailand (Appendix 1 and 2, see supplemental material, which is available on the Supplementary tab of the article’s Taylor & Francis Online page at http://dx.doi.org/10.1080/14772000.2013.818587) using locally made live-capture traps (single-capture cage-traps made of wire mesh). Trapping sessions were organized so as to maximize the number and diversity of captures. At each site, traps baited with ripe banana were set in three habitat types: karst slopes, cave entrance and cave interior (provided that a cave was accessible on the site) for at least three consecutive nights. The number of traps set on karst slopes and within caves was distributed as equally as possible. All traps were checked every day in the early morning. During this survey, a total of 477 Murinae rodents were live-trapped. Captive rodents were handled using a cloth bag as described in Aplin et al. (2003). Field identifications of captured rodents were made based on geographical and morphological criteria according to Lekagul & McNeely (1988), Corbet & Hill (1992), Aplin et al. (2003) and Francis (2008) (these criteria are described in Appendix 3, see supplemental material online). Each rodent was also weighed, measured and its age/sex category determined. Finally, a small piece of ear tissue was collected and stored in 96% ethanol before the animal was released.

#### DNA extraction, amplification and sequencing

Genomic DNA was extracted from skin samples using the DNeasy Tissue Kit (Qiagen Inc.) following the manufacturer’s protocol. Two mitochondrial markers, the cytochrome b gene (cytb) and the cytochrome c oxidase sub-
unit I gene (COI), were amplified for all the samples. Universal primers L14723/H15915 (Irwin et al., 1991) were used to amplify the cyt b gene. However, due to amplification difficulties with some samples, several genus- and species-specific primers were designed to amplify a 900 base pairs (bp) fragment of the cyt b gene: Maxomys (MaxoFw: 5′-TTCTTCATGGACCTACCCYAC-3′/MaxoRv: 5′-ATGA TTGAGAAGTAGCTGTAGT-3′), Niviventer (NiviFw: 5′- TCCGAAAACTACCCCTTAC-3′/ NiviRv: 5′-GAAG TATGGGAGAATGGTACG-3′), Rattus (RatFw: 5′-CAT TGRAATCACACCC-3′/RatRv: 5′-GTGTTCTAGGGYT GKCCYYCC-3′) and Leopoldamys neilli (LneilliFw: 5′- TCCATCCACATCCTCATGTG-3′/LneilliRv: 5′-GGAG GCTAGTTGGCAAATG-3′), Primers BatL5310 (5′-CCTA CTCRGCATTTTACCTTAG-3′) and R6036R (5′-ACT TCTGGGTGTCCAAAGATCA-3′) were used to amplify COI (713 bp) (Robins et al., 2007).

PCRs were carried out in 50 μL volume containing 12.5 μL of each 2 μM primers, 1 μL of 10 mM dNTP, 10 μL of 5 × reaction buffer (Promega), 0.2 μL of 5 U/μL Promega Taq DNA Polymerase and approximately 30 ng of DNA extract. Amplifications were performed in a thermal cycler VWR Unocycler using one activation step (94°C for 4 min) followed by 40 cycles (94°C for 30 s, 48°C (COI) or 50°C (cyt b) for 60 s, and finally 72°C for 90 s) with a final extension step at 72°C for 10 min. Sequencing reactions were performed on an ABI 3730 automatic sequencer.

**Phylogenetic analyses**

Sequences were aligned in BIOEDIT 7.0.9.0 (Hall, 1999) using ClustalW algorithm. Haplotypes were identified using ARLEQUIN 3.11 (Excoffier et al., 2005). Phylogenetic reconstructions were performed on the combined dataset using the maximum likelihood (ML) and Bayesian inference (BI) approaches. Two sequences of the Eurasian harvest mouse Micromys belonging to the Rattini tribe (Lecompte et al., 2008) and available on GenBank were used as outgroups of our phylogenetic analyses. The most suitable model of DNA substitution for each locus and dataset was determined by MODELTEST 3.0 (Posada & Crandall, 1999) according to the Akaike Information Criterion (AIC). PhyML 3.0 (Guindon et al., 2010) was used to perform ML analyses. The transition/transversion ratio, the proportion of invariable sites and the gamma distribution parameter were estimated. The starting tree was determined by BioNJ analysis of the datasets. Robustness of the tree was assessed by 1000 bootstrap replicates. Bayesian analyses were performed with MrBayes 3.1.1 (Ronquist & Huelsenbeck, 2003). Metropolis-coupled Markov chain Monte Carlo (MCMC) sampling was performed with 5 chains run for 6.10^6 generations with one tree sampled every 1000 generations, using parameter settings as starting values. All trees obtained before the Markov chain reached stationary distribution (empirically determined by checking of likelihood values) were discarded as burn-in values. A 50% majority-rule consensus tree was then generated in PAUP 4.0b10 (Swofford, 1998).

**Species delimitation and identification**

To estimate putative species boundaries, we used the DNA-based approach developed by Pons et al. (2006). This method detects the transition between micro- and macroevolutionary patterns using an ultrametric tree. The null model assumes a single coalescent process for the entire tree. The alternative model, a General Mixed Yule-Coalescent (GMYC) model, combines equations that separately describe branching within populations (coalescent process) and branching between species (a Yule model that includes speciation and extinction rates). The point of highest likelihood of this mixed model estimates the switch from speciation to coalescent branching and can be interpreted as the species boundary. A more recent version of the GMYC model (multiple-threshold model) allows for multiple lineages to each have their own transition threshold within a single phylogenetic tree (Monaghan et al., 2009). GMYC analyses were performed using the SPLITs package for R (http://r-forge.r-project.org/projects/splits/). An ultrametric tree was obtained using BEAST 1.6.2 (Drummond & Rambaut, 2007) using a GTR+G model with an uncorrelated lognormal relaxed molecular clock model (Drummond et al., 2006) and a coalescent tree prior. Two runs of 25.10^6 generations, sampling every 10000th generation, were performed. Convergence of the chains to the stationary distribution was checked using TRACER 1.5 (Rambaut & Drummond, 2007). All Beast computations were performed on the freely available computational resource Bioportal at the University of Oslo (http://www.bioportal.uio.no).

To allow accurate species designation and confirm our field identification based on morphological and geographical criteria, sequences of voucher specimens available from a recent phylogenetic study of the Rattini tribe in the Indochinese region (Pages et al., 2010) were added to our dataset. Cyt b and COI sequences of a single representative of each of the 24 species delimited in this study were retrieved from GenBank. Therefore the final dataset includes 229 haplotypes from karst samples and 24 sequences from Pages et al. (2010).

According to the generic species concept, the level of cyt b net genetic distance between clusters could also be used to delimit species boundaries (Baker & Bradley, 2006). Matrices of pairwise net K2P distances for the cyt b gene between GMYC lineages within the genera Maxomys, Niviventer, Leopoldamys, Berylmys, Sundamys and Rattus were therefore calculated in MEGA 4 (Tamura et al., 2007).

**Divergence time estimates**

Divergence times among Rattini and our GMYC clusters were estimated using Bayesian inference, as implemented in BEAST 1.6.2 on the cyt b dataset. Three fossil-based
calibration points were used: (i) the split between the tribe Phloeomyini and the other tribes of Murinae at 12.3 million years (Myr) (Jacobs & Flynn, 2005) (ii) the divergence between *Apodemus mystacinus* and the species of the subgenus sylvaemus (*A. flavicollis* and *A. sylvaticus*) at 7 Myr (Aguilar & Michaux, 1996; Michaux et al., 1997); and (iii) the divergence of *Otomys/Parotomys* (*Otomynae*) from the *Arvicanthis* clade at 5.7 Myr (Winkler, 2002).

To constrain the age of these nodes in the tree, sequences of *Deomys ferrugineus* (GenBank accession number EU349745), *Gerbillus riegneri* (AJ430555), *Batomys granti* (AY324458), *Phloeomys cumingii* (DQ191484), *Apodemus mystacinus* (AF159394), *Apodemus sylvaticus* (AB033695), *Arvicanthis flavicollis* (AB032853), *Arvicanthis somalica* (EU349737), *Arvicanthis niloticus* (AF004569), *Oenomys hypoxanthus* (EU349769), *Grammomys macmilliani* (AM408345), *Otomys angoniensis* (AM408343), *Parotomys* sp. (EU349773) were added to our dataset. Two analyses with different priors and dataset were performed independently following the strategy of Aplin et al. (2011). First, to deal with the combination of inter- and intra-specific diversity in the complete dataset, we used a coalescent tree prior, a relaxed molecular clock (uncorrelated lognormal) and a GTR+G substitution model (previously estimated by ModelTest). Secondly, we used a dataset restricted to one haplotype of each GMYC clusters combined to a Yule process speciation model as tree prior, a relaxed molecular clock (uncorrelated lognormal) and a GTR+G substitution model (previously estimated by ModelTest). Both analyses were run with MCMC chain length of 8×10^7, sampling every 10^4 generations.

**Focus on *Niviventer* and *Maxomys* genera**

To confirm our field identification of *Niviventer hinpoon* and assign unequivocally a species name to this cluster, we obtained DNA sequences from a piece of skin of the holotype specimen of *N. hinpoon*. The type specimen, an adult male, is located at the Centre for Thai National Reference Collections in Bangkok (reference number: CTNRC 54-3988) and was collected by W.A. Neill at Wat Tham Phrapothisat, Kaengkhoi District, Saraburi Province in 1973 (Waengsothorn et al., 1998). Museum samples contained only small quantities of degraded DNA, we amplified a 85 bp fragment of the cytb gene that proved to be suitable for amplification of ancient DNA and to allow to discriminate among most vertebrate species (Teletchea et al., 2008) and more specifically Rattini species (Pages et al., 2010). Primers from Teletchea et al. (2008) were specifically modified for *N. hinpoon* (LNhhnp: 5′-GACAAATTCCATTTCCACC-3′/HNhhnp: 5′-TAGTTATCTGGTCTCC-3′). DNA was extracted using the QIAamp DNA Micro Kit (Qiagen) following the protocol for isolation of genomic DNA from tissues. Amplifications were carried out in 25 μL volume containing 6.25 μL of each 2 μM primers, 0.625 μL of 10 mM dNTP, 5 μL of 5× reaction buffer (Promega), 2 μL of 25 mM MgCl2, 2.5 μL of 10 mg/mL BSA, 0.5 μL of 5 μM GoTaq Hot Start Polymerase (Promega) and 2 μL of DNA extract. PCRs were performed using one activation step (95°C for 5 min) followed by 50 cycles (94°C for 30 s, 48°C for 30 s, and finally 72°C for 45 s). The precautions and criteria pertinent for ancient DNA studies (Gilbert et al., 2005) were observed during all laboratory procedures: (i) All pre-PCR work was conducted in a physically remote area in ancient DNA laboratory where Rattini samples were never processed before this study; (ii) negative extraction and PCR controls were included; (iii) several independent PCRs were performed; and (iv) nucleotide substitutions were assessed for consistency with the molecular evolutionary pattern of ancient DNA sequences. Six PCR products from three independent PCRs were sequenced and resulted in the same nucleotide sequence. Extraction and PCRs controls remained negative. For all these reasons, we are fully confident that the sequence obtained is derived from the skin sample of the holotype specimen of *N. hinpoon*.

To confirm that the fragment obtained allowed us to discriminate among *Niviventer* species, cytb sequences of *Niviventer* available in this study and on GenBank were aligned, sorted by species and reduced to the 85 bp fragment. As *Niviventer* species misidentifications are numerous on GenBank (Pages et al., 2010), sequences that seem obviously to be misidentified were removed from the dataset. The 85 bp cyt b fragment obtained for *N. hinpoon* was then compared with other *Niviventer* species to identify diagnostically informative attributes. We used the definitions of Sarkar et al. (2008) to characterize these characteristic attributes (CA s): pure attributes exist across all members of a single clade, and never in any other clade while private attributes are present across some of the members of a clade, but never in any other clade.

To compare the level of genetic diversity within the two karst endemic species, *N. hinpoon* and *L. neilli*, we calculated the haplotype diversity (h) and the nucleotide diversity (Pi) (globally as well as at synonymous and non-synonymous sites) for both species using DnaSP 5.10 (Librado & Rozas, 2009).

In order to examine the phylogenetic position of *N. hinpoon* within the *Niviventer* genus, cyt b and COI sequences of several *Niviventer* species available on GenBank were added to our dataset. As cyt b and COI sequences deriving from the same specimen were not available, the two genes were analysed separately and two analyses performed. ML and BI analyses were performed as indicated above. Sequences of *Leopoldamys* were used as outgroups.

Finally, some doubts remain on species name assignment to clusters in the *Maxomys* genus because some *Maxomys* species reported in Thailand and in the Indochinese region (i.e. *M. whiteheadi*, *M. rajah* and *M. moi*) were
absent from the dataset of Pages et al. (2010). We therefore added cytb and COI sequences of Maxomys species available on GenBank to our own Maxomys sequences. ML and BI analyses were performed as indicated above on cytb and COI datasets separately. Sequences of Leopoldamys were used as outgroups.

Results

Phylogenetic analyses

Cy tb and COI sequences were obtained, respectively, for 467 and 470 Murinae rodent samples, all of them belonging to the Rattini tribe. The final alignments included 199 cytb haplotypes, 161 COI haplotypes and 229 haplotypes for the combined dataset. All haplotype sequences have been deposited in GenBank database (GenBank Accession Numbers are listed in Appendix 4, see supplemental material online).

The ML and BI trees of the combined dataset gave congruent results at terminal nodes but present an inconsistency at a deeper node (Fig. 2). Three of the six Rattini divisions suggested by Musser & Carleton (2005) are represented in our sampling and correspond to three monophyletic groups in the Bayesian tree. The Maxomys division (Bootstrap support (BS) = 100% – Bayesian posterior probabilities (BP) = 1.0) is the first lineage to diverge, then the poorly supported Dacnomys division (BS = node not recovered – BP = 0.55) consisting of Leopoldamys and Niviventer and finally the Rattus division (BS = 81% – BP = 0.99) comprising Berylmys, Sundamys and Rattus. However, Leopoldamys and Niviventer do not form a monophyletic group in the ML tree. All terminal nodes are very well-supported in both ML and BI trees (Fig. 2).

Following our field identifications, we are able to suggest species names to the lineages belonging to Maxomys (M. surifer), Niviventer (N. hinpoon, N. fulvescens), Leopoldamys (L. sabanus, L. edwardsi, L. neilli), Berylmys (B. berdmorei, B. bowersi) and Sundamys (S. muelleri) (Fig. 2). At least four well-differentiated lineages are also observed within the Rattus genus but the field identifications of our Rattus specimens are confused and conflicting as most of these samples were identified as Rattus tanezumi or Rattus sp. in the field and therefore do not allow us to assign a species name to these lineages at this point of the analyses.

Our phylogenetic tree also indicates that some species (e.g. Maxomys surifer and Leopoldamys neilli) comprise several highly differentiated and well-supported phylogenetic lineages that could possibly belong to several distinct species.

Species delimitation and identification

In order to investigate the specific status of some lineages and estimate putative species boundaries, branch length pattern of the ultrametric tree including karsts samples and a single representative of each of the 24 Rattini species delimited in Pages et al. (2010) was analysed.

The likelihood of both single- and multiple-threshold GMYC model was significantly higher than that of the null model of uniform (coalescent) branching rates ($LR = 56.82$, $P < 0.0001$ and $LR = 64.80$, $P < 0.0001$, respectively). However, the multiple-threshold model did not fit the data better than the single-threshold model ($\chi^2 = 7.98$, d.f. = 6, $P = 0.24$), indicating the age of the switch from speciation to coalescent branching did not vary significantly among lineages.

The single-threshold GMYC model estimates 40 putative species within our dataset excluding outgroups (with a confidence interval from 36 to 48) composed of 27 distinct clusters and 13 singletons (Fig. 3). Karsts samples are distributed within 28 GMYC lineages belonging to six genera that have been labelled from Rat1 to Max6 (Fig. 3). Of the six putative species of Rattus, three could be named without ambiguity using sequences from Pages et al. (2010): Rat1 corresponds to R. tanezumi (Rattus Rattus Lineage II sensu Aplin et al., 2011), Rat5 to R. tiomanicus and Rat6 to R. andamanensis. The lineage Rat4 coincides with the lineages ‘R3’ in Pages et al. (2010) and ‘Rattus rattus Lineage IV’ in Aplin et al. (2011) to which a species name could not be assigned. Lineages Rat2 and Rat3 could not be named either but are closely associated with the lineage Rat4. The percentage of match between field and molecular identifications of our Rattus specimens is very low for R. tiomanicus and R. andamanensis (Table 1).

The clustering of remaining Pages’ sequences with some of our samples in the phylogenetic tree confirm our species designation for M. surifer, N. fulvescens, L. sabanus, L. edwardsi, L. neilli, B. berdmorei and B. bowersi (N. hinpoon and S. muelleri were not sampled in Pages et al. 2010). The percentage of match between field and molecular identifications is very high (> 90%) for all these lineages, except for L. edwardsi (68%) and N. fulvescens (57%) (Table 1). However, L. neilli, M. surifer, B. bowersi and S. muelleri seem to consist of several putative species according to the method of Pons et al. (2006) as they correspond to eight, six, two and two GMYC lineages, respectively.

The amount of genetic differentiation (cytb net K2P distance) among GMYC lineages within the genera Maxomys, Niviventer, Leopoldamys, Berylmys, Sundamys and Rattus varied greatly and ranged from 0.3 to 16%. Pairwise distances among GMYC lineages have been categorized in five main groups (0–2.5%, 2.6–5%, 5.1–7.5%, 7.6–9.9%, > 10%) for more clarity (Table 2).

Divergence time estimates

Divergence time estimates among Rattini and our GMYC clusters obtained using the complete dataset and a coalescent tree prior are given in Fig. 4. Very similar divergence time estimates were obtained with the reduced dataset and
Fig. 2. Maximum likelihood (ML) tree (GTR + G) of the two combined mitochondrial genes (cytb, COI) of rodents collected in limestone karsts. Tree topologies of ML and Bayesian trees are different and the shaded lines indicate the internodes relationships recovered by Bayesian analyses. Bootstrap support (%, 1000 replicates) and posterior probabilities of nodes are indicated above and below the branches, respectively. Node supports from within lineages were removed for clarity of presentation.
Fig. 3. Ultrametric tree obtained with BEAST including samples collected in limestone karsts (HXXX in black) and sequences from Pages et al. (2010) (green). The dashed vertical line indicates the maximum likelihood point of switch from speciation to coalescent branching estimated by a single-threshold GMYC model. The shaded zone indicated the confidence interval for this estimate (within two log-likelihood units of the maximum). Clusters corresponding to putative species are indicated in red. The 28 clusters including karst samples have been labelled from Rat1 to Max6.
Table 1. Distribution ranges and species names suggested for GMYC lineages encountered in Thai limestone karsts as well as percentages of match between field and molecular species identification.

<table>
<thead>
<tr>
<th>GMYC lineage</th>
<th>Species name</th>
<th>Distribution range in this study (provinces) (see Figs 8–13)</th>
<th>N (% of total captures)</th>
<th>Percentage of match between field and molecular species identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat1</td>
<td><em>Rattus tanezumi</em> Temminck, 1844 = <em>Rattus rattus</em> Lineage II sensu Aplin <em>et al.</em>, 2011</td>
<td>Chaiyaphum Chumphon Kanchanaburi Krabi Lampang Nakhon Ratchasima Nakhon Sawan Nakhon Si Thammarat Nan Petchabun Petchaburi Phang Nga Prachuap Khiri Khan Saraburi Surat Thani Ratchaburi</td>
<td>59 (12.4%)</td>
<td>91% (54/59)</td>
</tr>
<tr>
<td>Rat2</td>
<td><em>Rattus sp.</em> = lineages ‘R3’ sensu Pages <em>et al.</em> (2010) and ‘Rattus rattus’ Lineage IV’ sensu Aplin <em>et al.</em> (2011)</td>
<td>Chaiyaphum Kanchanaburi Loei Lopburi Nakhon Ratchasima Petchaburi Ratchaburi Saraburi</td>
<td>3 (0.6%)</td>
<td></td>
</tr>
<tr>
<td>Rat3</td>
<td><em>Rattus sp.</em> = lineages ‘R3’ sensu Pages <em>et al.</em> (2010) and ‘Rattus rattus’ Lineage IV’ sensu Aplin <em>et al.</em> (2011)</td>
<td>Chaiyaphum Kanchanaburi Loei Lopburi Nakhon Ratchasima Petchaburi Ratchaburi Saraburi</td>
<td>16 (3.6%)</td>
<td>90% (18/20) (samples identified as <em>R. tanezumi</em> were considered as correctly identified as <em>Rattus rattus</em> Lineages II and IV have similar external morphology according to Aplin <em>et al.</em> (2011) and could represent a single species (Lack <em>et al.</em>, 2012; Pages <em>et al.</em>, 2012)</td>
</tr>
<tr>
<td>Rat4</td>
<td><em>Rattus sp.</em> = lineages ‘R3’ sensu Pages <em>et al.</em> (2010) and ‘Rattus rattus’ Lineage IV’ sensu Aplin <em>et al.</em> (2011)</td>
<td>Nakorn Ratchasima</td>
<td>1 (0.2%)</td>
<td></td>
</tr>
<tr>
<td>Rat5</td>
<td><em>Rattus tiomanicus</em> (Miller, 1900)</td>
<td>Krabi</td>
<td>6 (1.2%)</td>
<td>33% (2/6)</td>
</tr>
<tr>
<td>Rat6</td>
<td><em>Rattus andamanensis</em> (Blyth, 1860)</td>
<td>Chiang Rai Lampang Nan</td>
<td>5 (1.0%)</td>
<td>40% (2/5)</td>
</tr>
<tr>
<td>Sun1</td>
<td><em>Sundamys muelleri</em> (Jentink, 1879)</td>
<td>Chumphon Krabi Nakorn Si Thammarat</td>
<td>21 (4.4%)</td>
<td>100% (36/36)</td>
</tr>
<tr>
<td>Sun2</td>
<td><em>Sundamys muelleri</em> (Jentink, 1879)</td>
<td>Krabi Nakorn Si Thammarat Phang Nga</td>
<td>15 (3.1%)</td>
<td></td>
</tr>
<tr>
<td>Ber1</td>
<td><em>Berylmys berdmorei</em> (Blyth, 1851) complex</td>
<td>Phrae</td>
<td>1 (0.2%)</td>
<td>100% (1/1)</td>
</tr>
<tr>
<td>Ber2</td>
<td><em>Berylmys bowersi</em> (Anderson, 1879) complex</td>
<td>Nakorn Si Thammarat</td>
<td>1 (0.2%)</td>
<td>100% (1/1)</td>
</tr>
<tr>
<td>Leo1</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 1</td>
<td>Kanchanaburi Uthai Thani</td>
<td>70 (14.7%)</td>
<td>100% (70/70)</td>
</tr>
<tr>
<td>Leo2</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 2</td>
<td>Lopburi Saraburi</td>
<td>22 (4.6%)</td>
<td>93% (144/155)</td>
</tr>
<tr>
<td>Leo3</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 2</td>
<td>Nakorn Ratchasima Saraburi</td>
<td>29 (6.1%)</td>
<td></td>
</tr>
<tr>
<td>Leo4</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 2</td>
<td>Khon Kaen Loei Petchabun</td>
<td>68 (14.2%)</td>
<td></td>
</tr>
</tbody>
</table>

(Continued on next page)
a Yule process speciation model as tree prior (data not shown). Most Rattini species included in our dataset diverged during Middle/Late Pliocene with the exception of the species within the *Rattus rattus* complex that diverged more recently, during Early/Middle Pleistocene.

**Niviventer genus**

The selected 85 bp fragment of the cytb gene was well amplified from the holotype specimen of *Niviventer hinpoon*. An unambiguous consensus sequence was obtained from several PCR products (GenBank Accession no. KC010174) and compared with the sequences of our own *N. hinpoon* samples collected in limestone karsts. The holotype sequence corresponds exactly to the most abundant of three distinct variants of this fragment (H1) observed among our *N. hinpoon* dataset (Fig. 5). At least four CAs were identified between *N. hinpoon* and the other *Niviventer* species distributed in Thailand (*N. confucianus, N. cremoriventer, N. fulvescens, N. langbianis, N. tenaster*) (Fig. 5) as well as the remaining *Niviventer* species for which cytb sequences were available on GenBank (data not shown). This confirms that the selected 85 bp cytb fragment is highly suitable to discriminate among *Niviventer* species without ambiguity. Therefore the lineage Niv1 could be named unequivocally as *Niviventer hinpoon*.

### Table 1. (Continued)

<table>
<thead>
<tr>
<th>GMYC lineage</th>
<th>Species name</th>
<th>Distribution range in this study (provinces) (see Figs 8–13)</th>
<th>N (% of total captures)</th>
<th>Percentage of match between field and molecular species identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leo5</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 2</td>
<td>Chaiyaphum</td>
<td>7 (1.5%)</td>
<td></td>
</tr>
<tr>
<td>Leo6</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 2</td>
<td>Chiang Rai</td>
<td>2 (0.4%)</td>
<td></td>
</tr>
<tr>
<td>Leo7</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 2</td>
<td>Nan</td>
<td>16 (3.6%)</td>
<td></td>
</tr>
<tr>
<td>Leo8</td>
<td><em>Leopoldamys neilli</em> (Marshall, 1976) complex sp. 2</td>
<td>Phrae</td>
<td>11 (2.3%)</td>
<td></td>
</tr>
<tr>
<td>Leo9</td>
<td><em>Leopoldamys edwardsi</em> (Thomas, 1882)</td>
<td>Lampang Loei Nan Petchabun Phrae</td>
<td>22 (4.6%)</td>
<td>68% (15/22)</td>
</tr>
<tr>
<td>Leo10</td>
<td><em>Leopoldamys sabanus</em> (Thomas, 1887)</td>
<td>Chumphon Krabi Nakhon Si Thammarat Prachuap Khiri Khan Surat Thani</td>
<td>23 (4.8%)</td>
<td>100% (23/23)</td>
</tr>
<tr>
<td>Niv1</td>
<td><em>Niviventer hinpoon</em> (Marshall, 1976)</td>
<td>Lopburi Nan Sawan Saraburi</td>
<td>28 (5.9%)</td>
<td>100% (28/28)</td>
</tr>
<tr>
<td>Niv2</td>
<td><em>Niviventer fulvescens</em> (Gray, 1847)</td>
<td>Loei Chaiyaphum Phrae</td>
<td>7 (1.5%)</td>
<td>57% (4/7)</td>
</tr>
<tr>
<td>Max1</td>
<td><em>Maxomys surifer</em> (Miller, 1900) complex sp. 1</td>
<td>Lampang Nan Phrae</td>
<td>1 (0.2%)</td>
<td>100% (8/8)</td>
</tr>
<tr>
<td>Max2</td>
<td><em>Maxomys surifer</em> (Miller, 1900) complex sp. 1</td>
<td>Lampang Nan Phrae</td>
<td>7 (1.5%)</td>
<td></td>
</tr>
<tr>
<td>Max3</td>
<td><em>Maxomys surifer</em> (Miller, 1900) complex sp. 2</td>
<td>Chaiyaphum Loei Saraburi Phrae Petchabun</td>
<td>12 (2.5%)</td>
<td>100% (12/12)</td>
</tr>
<tr>
<td>Max4</td>
<td><em>Maxomys surifer</em> (Miller, 1900) complex sp. 3</td>
<td>Nakhon Si Thammarat Surat Thani</td>
<td>10 (2.1%)</td>
<td>90% (9/10)</td>
</tr>
<tr>
<td>Max5</td>
<td><em>Maxomys surifer</em> (Miller, 1900) complex sp. 4</td>
<td>Chumphon</td>
<td>9 (1.9%)</td>
<td>100% (14/14)</td>
</tr>
<tr>
<td>Max6</td>
<td><em>Maxomys surifer</em> (Miller, 1900) complex sp. 4</td>
<td>Krabi Nakhon Si Thammarat Surat Thani</td>
<td>5 (1.0%)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Pairwise cyt b net K2P distance between GMYC lineages recovered by the single-threshold GMYC model. For clarity, species names of GMYC lineages with unambiguous specific status and that could be named using field identification and voucher specimens from Pages et al. (2010) were used.

<table>
<thead>
<tr>
<th>cyt b net K2P distance</th>
<th>Genus</th>
<th>Pairwise comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–2.5%</td>
<td>Leopoldamys</td>
<td>- All pairwise comparisons of Leo4, Leo5, Leo6, Leo7, Leo8</td>
</tr>
<tr>
<td></td>
<td>Maxomys</td>
<td>- Max1 vs. Max2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Max5 vs. Max6</td>
</tr>
<tr>
<td></td>
<td>Sundamys</td>
<td>- Sun1 vs. Sun2</td>
</tr>
<tr>
<td></td>
<td>Rattus</td>
<td>- All pairwise comparisons of Rat2, Rat3, Rat4</td>
</tr>
<tr>
<td></td>
<td>Niviventer</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Berylmys</td>
<td>/</td>
</tr>
<tr>
<td>2.6–5%</td>
<td>Leopoldamys</td>
<td>- Leo2 vs. Leo3, Leo4, Leo5, Leo6, Leo7, Leo8</td>
</tr>
<tr>
<td></td>
<td>Maxomys</td>
<td>- Leo3 vs. Leo4, Leo5, Leo6, Leo7, Leo8</td>
</tr>
<tr>
<td></td>
<td>Sundamys</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Rattus</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Niviventer</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Berylmys</td>
<td>/</td>
</tr>
<tr>
<td>5.1–7.5%</td>
<td>Leopoldamys</td>
<td>- Max4 vs. Max5, Max6</td>
</tr>
<tr>
<td></td>
<td>Maxomys</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Sundamys</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Rattus</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Niviventer</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Berylmys</td>
<td>/</td>
</tr>
<tr>
<td>7.6–9.9%</td>
<td>Leopoldamys</td>
<td>- Leo1 vs. Leo 2, Leo3, Leo4, Leo5, Leo6, Leo7, Leo8</td>
</tr>
<tr>
<td></td>
<td>Maxomys</td>
<td>- L. edwardsi vs. Leo2, Leo4, Leo5, Leo6, Leo7, Leo8, L. sabanus</td>
</tr>
<tr>
<td></td>
<td>Sundamys</td>
<td>- Max1 vs. Max3, Max4, Max5, Max6</td>
</tr>
<tr>
<td></td>
<td>Rattus</td>
<td>- R. tanezumi vs. Rat2, R. tiamanicus, R. sakeratensis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- R. rattus vs. R. sakeratensis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- R. norvegicus vs. R. nitidus</td>
</tr>
<tr>
<td></td>
<td>Niviventer</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Berylmys</td>
<td>/</td>
</tr>
<tr>
<td>&gt;10%</td>
<td>Leopoldamys</td>
<td>- L. edwardsi vs. Leo1, Leo3</td>
</tr>
<tr>
<td></td>
<td>Maxomys</td>
<td>- L. sabanus vs. Leo1, Leo 2, Leo3, Leo4, Leo5, Leo6, Leo7, Leo8</td>
</tr>
<tr>
<td></td>
<td>Sundamys</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Rattus</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Niviventer</td>
<td>/</td>
</tr>
<tr>
<td></td>
<td>Berylmys</td>
<td>/</td>
</tr>
</tbody>
</table>

The level of genetic diversity within N. hinpoon is much lower than within L. neilli (globally or limited to central populations), especially the nucleotide diversity which is 10 times lower within N. hinpoon than within L. neilli central lineages (Table 3). Phylogenetic analyses indicated that the N. hinpoon lineage clustered with N. fulvescens lineages and formed a well-supported group in both COI (BS = 100% – BP = 1.0) (Fig. 6) and cyt b (BS = 98% – BP = 0.99) (data not shown) trees. However, the two main COI
线状的 \textit{N. fulvescens} (泰国 + 越南 + 广西 (中国)/海南 + 贵州 (中国)) 在该群中是同系群。同系群 \textit{N. fulvescens} 在线粒体 cytb 树中也观察到。\textit{N. hinpoon} / \textit{N. fulvescens} 群是 \textit{N. cremoriventer} 线状的姐妹群，在两棵树中都与之很好支持。

\textbf{Maxomys genus}

COI (图7) 和 cytb (未显示) 的系统发生树明确认定所有 \textit{Maxomys} 样本属于支持良好的 \textit{M. surifer} 复合体 (COI: BS = 99% – BP = 1.0, cytb: BS = 100% – BP = 1.0)。我们在6个 \textit{Maxomys} GMYC 线状中没有分配到任何其他 \textit{Maxomys} 物种。四个支持良好的组对应于东南亚的不同时区：北部 + 西部泰国 (线状 Max1 和 Max2), 泰马半岛 (南部泰国 (线状 Max4, Max5 和 Max6) 和马六甲半岛), 东部的印度支那地区 (中北部和东北部泰国 (线状 Max3), 老挝, 越南和柬埔寨) 以及最后的印度尼西亚 (加里曼丹)。

\section*{Discussion}

\textbf{Species delimitation and specific status of GMYC lineages}

Fig. 5. Discrimination and identification of *N. hinpoon* among other *Niviventer* species distributed in Thailand using characteristic attributes (CAs) of a 85bp cytb fragment: *N. fulvescens* (A), *N. cremoriventer* (B), *N. confucianus* (C), *N. tenaster* (D), *N. langbianis* (E). Accession numbers of sequences retrieved from GenBank are indicated between brackets. Identical sequences are represented by just a single sequence. Identical positions as those of the reference sequence of *N. hinpoon* holotype are indicated by dots. Pure CAs and private CAs are outlined in red (solid line) and black (dashed line), respectively.

identifying clusters of organisms with similar genotypes (Bradley & Baker, 2006; Hausdorf, 2011). However, irrefutable criteria to recognize clusters representing species from others are still lacking (Sauer & Hausdorf, 2012).

The DNA-based approach of species delimitation developed by Pons et al. (2006) estimates 40 putative species belonging to seven genera within our dataset composed of karsts samples and a single representative of each of the 24
Table 3. Genetic diversity observed within N. hinpoon and L. neilli: haplotype diversity (h), nucleotide diversity (Pi), nucleotide diversity at synonymous site (Pi(s)) and nucleotide diversity at non-synonymous site (Pi(a)).

<table>
<thead>
<tr>
<th></th>
<th>cytb</th>
<th>COI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h</td>
<td>Pi</td>
</tr>
<tr>
<td>N. hinpoon</td>
<td>0.44</td>
<td>0.00141</td>
</tr>
<tr>
<td>L. neilli (all lineages)</td>
<td>0.92</td>
<td>0.04605</td>
</tr>
<tr>
<td>L. neilli (central lineages)</td>
<td>0.79</td>
<td>0.01709</td>
</tr>
</tbody>
</table>

Rattini species delimited in Pages et al. (2010). This number is much larger than the number of species described in Thailand for these genera (25: three Maxomys, six Ni- viventer, three Leopoldamys, two Berylmys, one Sundamys, two Bandicota and eight Rattus). It seems that L. neilli, M. surifer, B. bowersi and S. muelleri consist of several putative species according to this method as they correspond to eight, six, two and two GMYC lineages, respectively. However, these numerous lineages could represent real species but they could also simply represent geographically isolated populations evolving neutrally. Indeed, a limitation of this method is that a single species with a strong spatial population structure could wrongfully be split in several separate GMYC lineages (Pons et al., 2006; Fontaneto et al., 2007). Therefore the GMYC model provides an initial hypothesis of the number of species-level groups within our dataset that should be confirmed using other information (Papadopoulos et al., 2009).

The cytb net genetic distance between clusters could also be useful to delimit species boundaries. According to Bradley & Baker (2001), within rodent, cytb genetic distance values smaller than 2% indicate intraspecific variation whereas values higher than 11% denote species delimitation in great majority of cases. Intermediate values could represent either intraspecific variation or species boundaries. Our results for the Rattini tribe corroborate these findings (Table 2). The range of distances from 0 to 2.5% includes only pairwise comparisons between closely related clusters that do not represent distinct recognized species. Distances comprised between 2.6 and 7.5% correspond mostly to pairwise comparisons among the Black rat complex (the R. rattus complex sensu Aplin et al., 2011) including R. rattus (R. rattus Lineage I sensu Aplin et al., 2011), R. tanezumi (R. rattus Lineage II sensu Aplin et al., 2011), Rattus sp. R3 (R. rattus Lineage IV sensu Aplin et al., 2011), R. sakeratensis and R. tiomanicus. The range of distances between 7.6 and 9.9% contains several pairwise comparisons between species traditionally recognized as valid such as N. fulvescens, N. hinpoon, R. argentiventer, R. andamanensis, L. edwardsi and L. sabanus. Distances higher than 10% comprise pairwise comparisons between well-differentiated described species. Therefore, based on these results, a minimum threshold genetic distance of 7.5% could be indicative of species delimitation within the Rattini tribe and could be lower for the closely related species of the Rattus rattus complex.

These findings allow us to discuss below the specific status of some GMYC lineages that remains unclear within our dataset (summarized in Table 1, see Figs 8–13 for their distribution).

Leopoldamys neilli

Recent phylogeographical studies based on a combination of mitochondrial and nuclear markers indicate a strong geographical structure of the L. neilli genetic diversity (Latinne et al., 2011, 2012). The present study identified eight allopatric GMYC lineages (Leo1 to Leo8) within the species (with a confidence interval from seven to nine). Five of these clusters (Leo4, Leo5, Leo6, Leo7, Leo8) which diverged during Middle Pleistocene are separated by cytb genetic distances of less than 2.5% and the clustering may be interpreted as strong spatial population structure within a single species. The clusters Leo2 and Leo3 are differentiated from these five clusters and from each other by slightly higher genetic distance (between 2.5 and 5%). This level of differentiation is of the same order as those observed among the closely related species of the Rattus rattus complex whose taxonomic status is still confused. Therefore these values alone do not allow us to conclude that these clusters represent two separate species. Further information is required to determine the status of these lineages.

The Leo1 cluster is dated from Late Tertiary/Early Quaternary and is highly divergent from the other ones (cytb genetic distance between 7.5 and 9.9%). This level of genetic divergence is similar to the one observed among recognized Leopoldamys species in Thailand (L. edwardsi, L. sabanus and other clusters of L. neilli). Our previous phylogeographical studies also showed that this cluster formed a unique, highly supported and monophyletic clade in a phylogenetic tree based on two nuclear introns while all other L. neilli populations belong to undifferentiated clades (Latinne et al., 2012). Therefore, following these genetic arguments, the cluster Leo1 including samples from western Thailand (Kanchanaburi and Uthai Thani provinces) could represent a separate species from the seven remaining clusters of L. neilli. A careful morphological study of this highly divergent lineage is now required to determine definitively its specific status.
Fig. 6. ML tree (HKY + G) representing phylogenetic relationships within the *Niviventer* genus using COI dataset. Bootstrap support (%, 1000 replicates) and posterior probabilities of nodes are indicated above and below the branches, respectively. The dataset includes our karst samples (HNXX) and sequences retrieved from GenBank. GenBank Accession numbers are indicated.
Fig. 7. ML tree (GTR + G) representing phylogenetic relationships within the *Maxomys* genus using COI dataset. Bootstrap support (%,
1000 replicates) and posterior probabilities of nodes are indicated above and below the branches, respectively. The dataset includes our karst samples (HMXX) and sequences retrieved from GenBank. GenBank Accession numbers are indicated.
Maxomys surifer

Six GMYC lineages (Max1 to Max6) have been identified in Thailand within the species *M. surifer* (with a confidence interval from five to seven). Max1 and Max2 as well as Max5 and Max6 are separated by cytb genetic distances of less than 2.5%, values characteristic of intraspecific variation between geographically isolated populations rather than species delimitation. In contrast, four highly divergent lineages (Max1+Max2, Max3, Max4, Max5+Max6), separated by more than 7.5% of genetic distance (except...
between Max4 and Max5+Max6: between 5.1 and 7.5%), could be delimited within our dataset. Our phylogenetic analysis clearly indicates that all these clusters could not be assigned to any other *Maxomys* species currently recognized in Thailand. These findings confirm that *M. surifer* may represent a complex of several species as already suggested by Musser & Carleton (2005). All the *Maxomys* samples belonging to these four highly divergent lineages were consistent with the external description of *M. surifer* (upperparts bright orange with black spines along the back and orange collar, underparts creamy-white, bicoloured tail) and may correspond to several cryptic species. However, future morphological studies as well as the sequencing of nuclear genes are now needed to justify the species status of these lineages based on mitochondrial DNA and to determine if these species are really cryptic. A possible ecological adaptation of one of these lineages to life in limestone karsts similar to the one observed for *L. neilli* and *N. hinpoon* should also be verified.

**Sundamys muelleri**

The two GMYC lineages (Sun1 and Sun2) detected within *S. muelleri* are separated by low cytb genetic distance (less than 2.5%) and diverged recently, around 400000 years ago. Moreover, these two clusters are combined in a single lineage if the lower limit of the confidence interval of the GMYC analysis is taken into account. Therefore we consider that these two lineages do not represent separate species.

**Berylmys bowersi**

This study confirms the presence of two highly divergent genetic lineages within *B. bowersi* in Thailand as pointed out by Pages et al. (2010). The first lineage that includes northern samples has been identified as *B. bowersi* by Pages et al. (2010) and contains none of our karst samples. The second lineage (the lineage Ber2 corresponding to the lineage ‘Be2b’ in Pages et al. (2010)), that could not be named, includes southern samples from Kanchanaburi and Surat Thani provinces. However, the level of genetic divergence between these two lineages is moderate (between 5.1 and 7.5%) and lower than the distance between these two lineages and *B. berdmorei* (>10%) and is not indicative of a clear species delimitation. Several authors have already mentioned that southern populations of *B. bowersi* may represent a separate species or subspecies based on morphological criteria (Lekagul & McNeely, 1988; Francis, 2008). Further taxonomic work within *B. bowersi* is therefore required.

**Rattus sp. R3**

The clusters Rat2, Rat3, Rat4 are separated by less than 2.5% of cytb genetic distance. Moreover, Rat3 and Rat4 are combined in a single lineage if the lower limit of the confidence interval of the GMYC analysis is taken into account. Therefore, these results strongly suggest that these three lineages belong to a single species to which a name could not be assigned and identified as ‘Rattus sp. R3’ in Pages et al. (2010), ‘R. rattus Lineage IV’ in Aplin et al. (2011) and Lack et al. (2012). Recent studies revealed that Rattus sp. R3 and *R. tanezumi* could not be distinguished using nuclear markers and morphological characters and suggest possible hybridization with introgression between Rattus sp. R3 and *R. tanezumi* (Lack et al., 2012; Pages et al., 2012). Therefore, the R3 species could even be invalidated and considered as *R. tanezumi*.

**Identification of Rattini species**

Most of the Rattini species included in our dataset were identified in the field with good success using external morphological characters and geographical criteria even if misidentifications were numerous (Table 1). However, the very low percentages of match between field and molecular identifications for some *Rattus* species reflect the difficulty of identifying *Rattus* specimens using only external characters as already pointed out by Robins et al. (2007) and Pages et al. (2010).

Our study also revealed frequent misidentifications of *Leopoldamys edwardsi* as *L. sabanus* in the field. During our field survey, the *Leopoldamys* specimens from northern Thailand that clearly did not belong to *L. neilli* (because of their very large body size and tail length) were assigned either to *L. edwardsi* or *L. sabanus* following the morphological criteria based on fur colour available in the literature (Appendix 3, see supplemental material online) as both species are assumed to occur in this region. However our phylogenetic analyses revealed that all these samples, regardless of their fur colour, clearly belong to a single phylogenetic lineage (Leo9) and were genetically highly differentiated from the *L. sabanus* samples trapped in peninsular Thailand (Leo10). Based on these results and the data of Pages et al. (2010), it seems that *L. sabanus* is restricted to the southern part of Thailand with a northern range limit in the Kanchanaburi province and that this species does not occur in sympatry with *L. edwardsi* in Thailand (Fig. 12). These findings are in opposition to data available in the literature (Lekagul & McNeely, 1988; Musser & Carleton, 2005). We suspect that this contradiction is probably due to frequent misidentifications between *L. sabanus* and *L. edwardsi* in the literature as these two species are morphologically very similar and easily confused. Therefore we recommend using DNA sequences to identify unambiguously the two species in future studies.

**Rodent diversity of Thai limestone karsts**

A total of 12 described Rattini species (*R. tanezumi*, *R. tiomanicus*, *R. andamanensis*, *S. muelleri*, *B. berdmorei,*
B. bowersi, L. neilli, L. edwardsi, L. sabanus, N. hinpoon, N. fulvescens, M. surifer), corresponding to 17 highly divergent genetic lineages (see above), were encountered in limestone karsts of Thailand (Table 1). Out of these 12 species, the two karst endemic species (L. neilli and N. hinpoon) represent 53.1% of the total number of captures, L. neilli being the most abundant species and representing almost half of the total captures (47.2%). On the contrary, B. berdmorei and B. bowersi were excessively rare (less than 1% of the captures).

**Karsts as forest refuges?**

Except karst endemic species, the majority of rodent species recorded in this study are associated with forest ecosystems (R. tiomanicus, R. andamanensis, S. muelleri, L. edwardsi, L. sabanus, N. fulvescens, M. surifer). Thailand has lost about 30% of its forest cover between 1950 and 2000 and estimates of current forest loss in Thailand range from 0.7% to 1% per year (World Bank, 2004; Dupuy et al., 2012). The remaining forests that represent between 25 and 33% of the total land area of the country, depending on the calculation method (Charuppat, 1998; World Bank, 2004), are often fragmented and degraded (Dupuy et al., 2012). As karsts are mainly unsuitable for agriculture due to their topography, deforestation is generally less extensive on karst hills than in valley floors. In many regions of Thailand, limestone karsts thus appear as forested islands surrounded by agricultural fields and urban areas. This study suggests that besides hosting karst endemic species, limestone karsts could also play a key role in the preservation of the global rodent biodiversity by providing important refuges for the forest-dwelling species in deforested and highly developed regions. Moreover, karst hillside forests could help to maintain the habitat connectivity and allow movements of species among isolated forest fragments. In order to confirm these hypotheses, a precise comparison of forest-dwelling rodent occurrence between karsts and adjacent habitats using an accurate methodology will be needed in the future. The important role of forested karsts to support forest-adapted species has also been mentioned by Furey et al. (2010) in their study of bat diversity in Vietnam. These observations strengthen the necessity of protecting karst ecosystems in Southeast Asia to preserve not only the karst endemic fauna but also the forest-dwelling species.

Rattini species commonly observed in agricultural and cultivated areas of Thailand (e.g. Bandicota indica, Bandicota savilei, Rattus argentiventer, Rattus sakeratensis, Rattus nitidus) were totally absent from our inventory. Highly commensal species such as R. exulans and R. norvegicus were also lacking in Thai limestone karsts. Only one very adaptable and generalist species, R. tanezumi (or two if we consider Rattus sp. R3 as a distinct species) inhabits commonly Thai limestone karsts.

**The Indochinese–Sundaic biogeographical transition**

The transition between the Indochinese and Sundaic biogeographical subregions has long been considered to occur around the Isthmus of Kra (e.g. Kloss, 1915, 1929; Wells, 1976). However, Woodruff & Turner (2009) showed that the transition in mammal species does not coincide with a narrow zone of the Thai–Malay peninsula but instead occurs within two areas situated around 5°N (northern part of Malay peninsula) and 14°N (where the Thai peninsula meets the mainland). These authors explained this pattern of faunistic transition by numerous significant faunal compressions in the Thai–Malay peninsula due to the repeated sea-level changes that occurred during the last 5 Myr.

Our divergence time analysis allows us to date the split among some Indochinese and Sundaic Maxomys and Leopoldamys lineages. The first divergence that occurred among our Maxomys clusters did not isolate the Indochinese from Sundaic lineages but separates the northern lineages (Max1+Max2) from the northeastern (Max3) and southern ones (Max4+Max5+Max6) around 2.3 Mya. However this node is not well-supported in our phylogenetic tree. Then, the southern lineages diverged from the northeastern one around 2 Mya, within the time frame of 1 to 4 Mya predicted by Woodruff & Turner (2009) following their hypothesis. The divergence of the Sundaic L. sabanus from the Indochinese Leopoldamys species (L. edwardsi and L. neilli) was also estimated to occur within this time frame, around 3.5 Mya, corroborating the hypothesis of Woodruff & Turner (2009).

Finally, our study evidenced a high number of genetic lineages in peninsular Thailand within Maxomys surifer (3 GMYC clusters) (Fig. 13), Sundamys muelleri (2 GMYC clusters) (Fig. 10) and Leopoldamys sabanus (1 GMYC cluster subdivided in 2 lineages) that diverged during Pleistocene. The strong phylogeographical pattern of these three species in this region may be due to the numerous repetitive marine transgressions that occurred during the last 4 Myr and strongly reduced the land area (around 50%) between Krabi, Surat Thani and Nakhon Si Thammarat (Woodruff & Turner, 2009). These marine transgressions may have influenced the genetic structure of these Murinae species in the central Thai–Malay peninsula by causing repeated demographic bottlenecks.

**Karst endemic rodents of Thailand and their conservation status**

More than 20% of limestone karsts in Thailand have already been quarried for cement, lime and hard core, and many have completely disappeared from the landscape (World Bank, 2004). The cement production of Thailand has increased by about 10% every year in the early 2000s but this rate has been slightly decreasing since 2007 (ESCAP,
As a result of these activities, some endemic species are highly threatened. Due to the lack of data about their distribution range and ecology, the two karst endemic rodent species of Thailand, *L. neilli* and *N. hinpoon*, are currently listed as Data Deficient on the IUCN Red List (Lunde & Aplin, 2008a, 2008b). However, the distribution range of *L. neilli* is much more extensive than indicated by previous records. This species is present in most karstic regions of central, northern and western Thailand (Fig. 11). Accordingly, Latinne et al. (2012) recommend classifying this species as ‘Near Threatened’ on the IUCN Red List. However, if we consider that the western populations of *L. neilli* represent a separate species, this species should be listed in a higher extinction risk category and qualified for the Vulnerable category (B1ab(iii)) due to its small extent of occurrence (less than 20 000 km²), the high fragmentation of its habitat and the projected decline of the quality of this habitat in the future (threats of habitat destruction by limestone quarrying).

Contrary to *L. neilli*, our survey clearly indicated that the distribution range of *N. hinpoon* is limited to the limestone karsts of central Thailand (Saraburi, Lopburi, Nakhon Sawan provinces) (Fig. 12) where it seems to be quite common and abundant. Due to its small distribution range and the high threats that its habitat is facing in this region, *N. hinpoon* meets the same criteria as the western populations of *L. neilli* and would be qualified as Vulnerable (B1ab(iii)) on the IUCN Red List.

Our phylogenetic analyses indicate that *N. hinpoon* is closely related to *N. fulvescens*, as also suggested by morphological characters (Musser & Carleton, 2005). However, according to our phylogenetic tree combining GenBank sequences, *N. fulvescens* consists of several paraphyletic lineages, which confirms the need of a taxonomic revision of this species. At the intraspecific level, *N. hinpoon* is genetically homogeneous and, in contrast to *L. neilli* in this region, is characterized by a single mitochondrial lineage and a lower genetic diversity. *Leopoldamys neilli* and *N. hinpoon* frequently live in sympatry in the same caves of Saraburi and Lopburi provinces. Therefore competitive exclusion does not seem to occur between these two species adapted to limestone karsts which share presumably similar habitat requirements.

### Conclusions

Our study revealed an important genetic diversity within traditionally recognized species *Maxomys surifer* (four highly divergent lineages), *Leopoldamys neilli* (two highly divergent lineages) and *Berylmys boweri* (two highly divergent lineages) which could be considered as species complex. Further taxonomic work combining independent data (mitochondrial and nuclear markers, morphology) and an increased sampling is now required to confirm the specific status of these lineages. It will also be very important to determine if some of these putative cryptic species are endemic to limestone karsts as it will have very important implications for their conservation.

This study has also shown that the Murinae rodent diversity of limestone karsts in Thailand is not restricted to endemic species as karsts also host many forest-dwelling species to which they could provide forest refuge in deforested regions. Therefore limestone karsts should be protected to preserve not only endemic species, but also forest-dwelling species threatened by deforestation.

### Acknowledgements

We are grateful to B. Tontan for his help in the field throughout this survey and to S. Jittapalapong for his administrative support. We thank M. Pages for her advices about extraction and amplification of DNA from museum sample. This work was supported by a Belgian FRS-FNRS (Fonds de la Recherche Scientifique) fellowship to A. Latinne (‘Aspirant’) and to J.R. Michaux (‘Maitre de recherches’), and a financial grant from the Belgian FRS-FNRS (‘crédits pour brefs séjours à l’étranger’ to A. Latinne and J.R. Michaux and credits from the ‘Fonds de la Recherche Fondamentale Collective (FRFC)’ to J.R. Michaux), from the University of Liège (Patrimoine) and from the Communauté française de Belgique. This project was also supported by the network ‘Bibliothèque du Vivant’ funded by the CNRS, the Muséum National d’Histoire Naturelle, the INRA and the CEA (Centre National de Séquençage). This study is part of the ‘CEROPath project’ (Community Ecology of Rodents and their Pathogens in South-East Asia: effects of biodiversity changes and implications in health ecology), ANR Biodiversity ANR 07 BDIV 012, funded by the French National Agency for Research.

### References


---

2001; Shi, 2010).


