Author's personal copy

Conserv Genet (2011) 12:1495–1511
DOI 10.1007/s10592-011-0248-3

RESEARCH ARTICLE

Evidence of complex phylogeographic structure for the threatened rodent *Leopoldamys neilli*, in Southeast Asia

Alice Latinne · Surachit Waengsothorn · Vincent Herbreteau · Johan R. Michaux

Received: 2 November 2010 / Accepted: 10 July 2011 / Published online: 22 July 2011
© Springer Science+Business Media B.V. 2011

Abstract *Leopoldamys neilli* is a threatened murine rodent species endemic to limestone karsts of Thailand. We have studied the phylogeography of *L. neilli* using two mitochondrial markers (cytb, COI) and one nuclear fragment (bfibr), in order to assess the influence of its endemicity to karst habitat. One hundred fifteen individuals of *L. neilli* were collected in 20 localities throughout the geographic range of this species in Thailand. Our study revealed strong geographic structure of the mtDNA genetic diversity: six highly differentiated, allopatric genetic lineages were observed in our dataset. They exhibit a very high degree of genetic divergence, low gene flow among lineages and low levels of haplotype and nucleotide diversities within lineages. Our results suggest that *L. neilli*'s populations are highly fragmented due to the scattered distribution of its karst habitat. The most divergent lineage includes the populations from western Thailand, which have been separated from the other genetic lineages since at least the Early Pleistocene. The other lineages are more closely related and have diverged since the Middle Pleistocene. This study revealed an unexpected high level of genetic differentiation within *L. neilli* and highlighted the high endemicity of this species to limestone karsts. Our results enhance the importance of protecting limestone habitats to preserve not only the species but also intraspecific diversity.

Keywords Southeast Asia · *Leopoldamys neilli* · Limestone karsts · Conservation · Phylogeography · Intraspecific diversity

A. Latinne (✉) · J. R. Michaux
Institut de Botanique, B 22, Université de Liège, 4000 Liège (Sart Tilman), Belgium
e-mail: alice.latinne@ulg.ac.be

S. Waengsothorn
Environment and Resources Technology Department, Thailand Institute of Scientific and Technological Research, 35 Mu 3 Tambon Khlong Ha, Amphoe Khlong Luang, Changwat Pathum Thani 12120, Thailand
e-mail: surachit@tistr.or.th

V. Herbreteau
CIRAD, UR AGIRs (Animal et Gestion Intégrée des Risques), Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez Cedex, France

J. R. Michaux
CBGP (Centre de Biologie et de Gestion des Populations), UMR INRA/IRD/Cirad/Montpellier SupAgro, Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez Cedex, France

Introduction

Phylogeography is a field of science that has evolved significantly during the last 20 years. However some regions of the world remain poorly studied and need more attention from phylogeographers. This is particularly the case for Southeast Asia. Only 3.3% of the phylogeographic studies published between 1987 and 2006 concern taxa from this region (Beheregaray 2008). The lack of data about phylogeographic patterns and biodiversity in Southeast Asia contrasts with its high scientific interest. This region overlaps with four biodiversity hotspots and is facing huge habitat loss (Myers et al. 2000). The relative rate of deforestation is higher in Southeast Asia than in any other tropical region (Achard et al. 2002), millions of hectares of tropical forest are destroyed annually (Hansen et al. 2008). The biodiversity of this region is also threatened by industrial-scale agriculture, climate changes, biomass fires (Taylor 2010), invasive species (Peh 2010), poaching and illegal wildlife trade

Author's personal copy

(Nijman 2010), among others. Increased knowledge of Southeast Asian biodiversity is urgently needed to better preserve it (Sodhi et al. 2004; Koh and Sodhi 2010), since the most threatened species of the world's land mammals are concentrated in this region (Schipper et al. 2008; Sodhi et al. 2010). Moreover, the distribution of research effort in Southeast Asia is habitat-biased and numerous habitats, such as limestone karsts, remain poorly studied (Sodhi et al. 2004; Clements et al. 2006).

Limestone karsts are sedimentary rock outcrops consisting of calcium carbonate, created millions of years ago by calcium-secreting marine organisms and subsequently lifted above sea level by tectonic movement (Clements et al. 2006). Within Southeast Asia, karsts cover an area of 460,000 km^2—about 10% of the total land area of this region (Day and Urich 2000). Karsts are recognized as biodiversity hotspots with high levels of endemic species of plants, vertebrates and invertebrates (Vermeulen and Whitten 1999; Schilthuizen et al. 2005; Clements et al. 2006, 2008). These karsts biotopes are highly threatened by limestone quarrying, hunting, deforestation and urbanization (Vermeulen and Whitten 1999; Clements et al. 2006). Due to the scattered distribution of karsts, endemic taxa are isolated on "karst islands" with reduced gene flow between them (Clements et al. 2008). Therefore, limestone karsts constitute an interesting model for testing hypotheses of the island theory in a continental environment. At the intraspecific level, oceanic island populations diverge genetically from each other due to restricted migration and genetic drift and they are characterized by increased inbreeding and a loss of genetic diversity (Frankham 1997; Frankham et al. 2002; Eldridge et al. 1999; Nieberding et al. 2006). Consequently, it might be expected that populations of endemic taxa isolated on "karsts islands" share some of these features of oceanic island populations. A better understanding of these biogeographic processes that remain largely unexplored is essential for creating more effective biodiversity conservation policies for limestone karsts of Southeast Asia.

We have studied the phylogeography of Neill's Rat *Leopoldamys neilli* (Marshall, 1976), a murine rodent species endemic to limestone karsts of Thailand, using two mitochondrial and one nuclear markers. *L. neilli* is a large and long-tailed rat with greyish-brown fur and a white belly, and was discovered in 1973 in Saraburi province, Central Thailand (Lekagul and McNeely 1988). The species has also been recorded in a few locations in north and west Thailand (Lekagul and McNeely 1988; Waengsothorn et al. 2007). However, the limits of its geographical range are not clearly resolved and *L. neilli* may also occur in neighboring countries, though these populations may represent a separate species (Musser and Carleton 2005; Waengsothorn et al. 2009). Two additional *Leopoldamys* species, *L. edwardsi* and *L. sabanus*, also occur in Thailand. *L. neilli*, whose phylogenetic relationship among the *Leopoldamys* genus requires resolution (Musser and Carleton 2005), appears as the sister taxa to *L. edwardsi* in a recent molecular phylogenetic study (Pagès et al. 2010). *L. neilli* was classified as "Endangered" on the IUCN Red List, but is now listed as "Data deficient" since very little information is available about its biology and ecology (Lunde and Aplin 2008). Acquisition of information about its geographic range, ecological requirements and the genetic structure of its populations are thus needed to assess the conservation status of *L. neilli*.

This study was conducted to answer the following questions: (1) is there geographic structure of the genetic diversity of *L. neilli*? (2) Are populations of *L. neilli* isolated from each other on "karst islands" and what are the levels of gene flow between them? (3) What are the divergence times between the main lineages of *L. neilli*? (4) From a conservation point of view, do the populations of *L. neilli* belong to the same Management Unit (MU) or Evolutionarily Significant Unit (ESU) or should they be defined as separate MUs or ESUs and therefore managed separately?

Materials and methods

Sample collection

A total of 115 *L. neilli* were live-trapped on limestone karsts from 20 localities in seven provinces (Loei, Nan, Phrae, Saraburi, Nakhon Ratchasima, Lopburi, and Kanchanaburi) (Fig. 1) throughout the geographical range of this species in Thailand. Animals were released after taking a tissue biopsy from the ear. The skin samples were stored in ethanol. Field identifications were made based on morphological criteria according to Lekagul and McNeely (1988) and Corbet and Hill (1992).

For the phylogenetic analyses, we added to our dataset four specimens (vouchers R4485, R4486, R4517, R4527) from the CERoPath tissue collection housed at the Center of Biology for the Management of Populations of Montpellier (France) identified without ambiguity as *L. neilli* by Pagès et al. (2010) using the genuine sequence of the holotype specimen of *L. neilli* (collected by W.A. Neill in 1973 in the Saraburi Province, Kaengkhoi District, Thailand, 14°35′N × 101°8′E). Moreover, 11 voucher specimens from the CERoPath tissue collection used in a recent molecular phylogenetic study (Pagès et al. 2010) were chosen as outgroups for our phylogenetic analyses. These voucher specimens belong to three genera of the Rattini tribe and were sampled from several regions of Thailand: *Leopoldamys sabanus* (vouchers R3033, R3111), *Leopoldamys edwardsi* (vouchers R4098, R4222, R4276, R4296, R4370), *Niviventer fulvescens* (vouchers R4497,

Author's personal copy

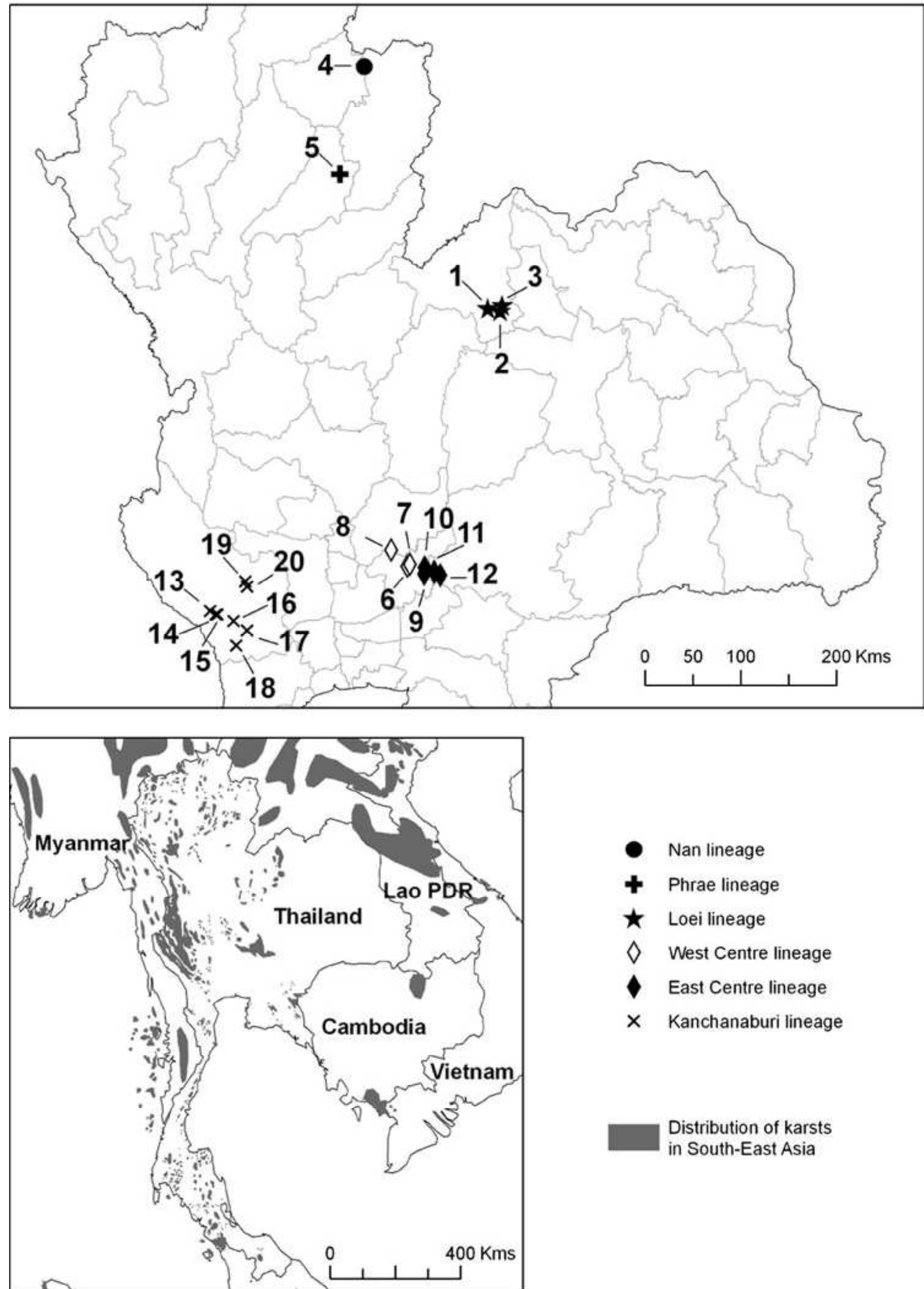


Fig. 1 Sampling localities of the 115 *L. neilli* analyzed in this study. *Symbols* correspond to mtDNA genetic clades (see Fig. 2a). *Numbers* correspond to the locality ID used in Fig. 2 and Appendix 1—Table 5

R4525), *Maxomys surifer* (voucher R4223) and *Maxomys* sp. (voucher R3116) (see Appendix 2—Table 7).

DNA extraction, PCR, and sequencing

Genomic DNA was extracted from skin samples using the DNeasy Tissue Kit (Qiagen Inc., Valencia, CA) following the manufacturer's instructions.

Two mitochondrial markers, the cytochrome *b* gene (cytb) and the cytochrome c oxidase subunit I gene (COI), were amplified for all the samples. Specific primers were specially designed to amplify a large fragment of the cytb gene (900 bp): LneilliFw (5′-TCCATCCAACATCTCAT-CATG-3′) and LneilliRv (5′-GGAGGCTAGTTGGCCA ATG-3′). Primers BatL5310 (5′-CCTACTCRGCCATTT-TACCTATG-3′) and R6036R (5′-ACTTCTGGGTGTC-CAAAGAATCA-3′) were used to amplify COI (713 bp) (Robins et al. 2007).

We also amplified a nuclear fragment, the β-fibrinogen intron 7 (bfibr) (745 bp), in a subset of 65 samples

Author's personal copy

representing the main sampling localities, using primers BFIBR1 (5′-ATTCACAACGGCATGTTCTTCAG-3′) and BFIBR2 (5′-AANGKCCACCCCAGTAGTATCTG-3′) (Seddon et al. 2001). This nuclear marker was chosen as it showed a good phylogeographic resolution in several studies of mammals (e.g. Seddon et al. 2001; Patou et al. 2010).

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 thermal cycler VWR Unocycler using one activation step (94°C/4 min) followed by 40 cycles (94°C/30 s, 50–50–58.7°C/60 s for cytb, COI and bfibr respectively, and finally 72°C/90 s) with a final extension step at 72°C for 10 min.

Sequencing reactions were performed by Macrogen Inc. (Seoul, Korea) using sequencing on ABI 3730 automatic sequencer and the same primers as for amplification.

For the nuclear bfibr gene, heterozygous states were identified as strong double peaks of similar height in both forward and reverse strands, or when the particular base corresponding to the dominant peak alternated on the two chromatograms (Hare and Palumbi 1999). The identity of the two bfibr alleles within a heterozygote was inferred through "haplotype subtraction" (Clark 1990).

Phylogenetic analyses

The cytb and COI sequences of the voucher specimens were retrieved from GenBank (see Appendix 2—Table 7 for their GenBank Accession numbers). Since the bfibr sequences of these vouchers were not available on GenBank, we amplified the β-fibrinogen intron 7 (bfibr) (745 bp) for the vouchers R4485, R4527, R3033, R3111, R4098, R4222, R4276, R4296, R4370, R4497 and R4223 using the protocol described in the previous section.

Sequences were aligned in BIOEDIT v7.0.9.0 (Hall 1999) using ClustalW algorithm. Phylogenetic reconstructions were performed using the maximum likelihood (ML) and Bayesian inference (BI) approaches. Analyses were first run independently on each locus and then on two combined dataset (cytb/COI and cytb/COI/bfibr). The two or three partitions were combined in a same dataset using the DAMBE software (Xia and Xie 2001). The most suitable model of DNA substitution for each locus and dataset was determined using MODELTEST v3.0 (Posada and Crandall 1998) according to the Akaike Information Criterion (AIC). PhyML v3.0 (Guindon and Gascuel 2003) 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 1,000 bootstrap replicates.

Bayesian analyses were performed with MRBAYES v3.1.1 (Ronquist and Huelsenbeck 2003). Metropolis-coupled Markov chain Monte Carlo (MCMC) sampling was performed with five chains run for 10^6 generations with one tree sampled every 1,000 generations, using default parameters 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 generated in PAUP v4.0b10 (Swofford 1998).

Partitioned Bremer support (PBS)

In order to assess the contribution of each data partition in the combined analysis (cytb/COI/bfibr), PBS (Baker and DeSalle 1997) of the main internal nodes within the *L. neilli* lineage were calculated. First, the data were analysed using the maximum parsimony (MP) criterion in PAUP v4.0b10 using a heuristic search strategy with 1,000 search replicates of random addition taxa sampling and tree-bisection reconnection (TBR) branch-swapping. Then, the program TREEROT v3 (Sorenson and Franzosa 2007) was used to determine the PBS of the main internal nodes within the *L. neilli* lineage.

Phylogeographic analyses and population differentiation

A median joining network was performed with NETWORK v4.5.1.6 (Bandelt et al. 1999) to explore relationships among haplotypes for the mitochondrial dataset (cytb/COI).

Haplotype (h) and nucleotide (π) diversities of the main lineages were estimated for the three loci independently using ARLEQUIN v3.11 (Excoffier et al. 2005). An Analysis of Molecular Variance (AMOVA) performed on the bfibr dataset and the two combined datasets (cytb/COI–cytb/COI/bfibr) in Arlequin v3.11 was used to assess the distribution of genetic variation among populations. The defined groups correspond to the main karst regions. The populations correspond to sampling localities within regions. Genetic distance within and between lineages was computed in MEGA v4.1 (Tamura et al. 2007) under the Kimura two parameters (K2P model) for the cytb dataset (to allow comparison with the study of Bradley and Baker (2001) concerning the genetic species concept).

Divergence times estimates

Divergence time of *L. neilli* and *L. edwardsi*, and of the main lineages of *L. neilli* (approximation of the time to the most recent common ancestor—TMRCA, Rosenberg and Feldman 2002), were estimated using BI, as implemented in the

Author's personal copy

program BEAST v1.5.3 (Drummond and Rambaut 2007) on the cytb dataset. Two fossil-based calibration points were used: (1) the *Mus/Rattus* divergence (10–12 Myr; Jacobs and Pilbeam 1980; Jaeger et al. 1986; Jacobs and Downs 1994; Jacobs and Flynn 2005) and (2) the divergence between *Apodemus mystacinus* and all the species of the subgenus *sylvaemus* (*A. flavicollis* and *A. sylvaticus*) (7 Myr; Aguilar and Michaux 1996; Michaux et al. 1997). Sequences of *Mus musculus* (GenBank accession No: AY057807), *Mus spretus* (AY057810), *Rattus rattus* (AB033702), *Rattus tanezumi* (AB096841), *Apodemus mystacinus* (AF159394, AJ311 147), *Apodemus flavicollis* (AB032853, AF159392) and *Apodemus sylvaticus* (AB033695, AF159395) were added to our cytb dataset. Analyses were performed under the TN93+G substitution model parameter (previously estimated by MODELTEST), a relaxed molecular clock (Drummond et al. 2006), and a Bayesian Skyline coalescent model (Drummond et al. 2005). Three independent runs with MCMC chain length of 6.5×10^7 were performed, sampling every 1,000th generation. Results were visualized using TRACER v1.5.

Divergence times between lineages were also estimated on the basis of the mean net K2P distance (taking into account the correction for ancestral mtDNA polymorphism) between lineages. A mutation rate for the cytb gene of 2.6%/Myr was used in this analysis (Michaux et al. 2003).

Demographic history

A Bayesian Skyline reconstruction performed in BEAST v1.5.3 and TRACER v1.5 using the cytb dataset was used to study the historical demography of the main lineages including more than 15 samples (Loei, East Centre, Kanchanaburi). The demographic histories of these three lineages were also examined using mismatch distribution of pairwise nucleotide differences for the three loci estimated in DNASP v5 (Librado and Rozas 2009). A unimodal distribution is consistent with sudden expansion of population and multimodal pattern indicates demographic stability of population. Fu's Fs (calculated in Arlequin v3.11), Fu and Li's F* & D* statistics (calculated in DNASP), and R2 (Ramos-Onsins and Rozas 2002) (calculated in DNASP) were also used to test for population growth under assumptions of neutrality and were calculated for the three loci and the three lineages independently.

Analysis of selection influence

The McDonald–Kreitman test (McDonald and Kreitman 1991) was performed in DNASP v5 to test whether our cytb and COI datasets departed significantly from neutral expectations. A Fisher's exact test was used to determine whether the ratio of synonymous and nonsynonymous substitutions differs between two categories: polymorphisms that are variable within *L. neilli* and *L. edwardsi* and polymorphisms that distinguish these two species (i.e., fixed differences).

Results

Sequence variation

A total of 22 and 16 haplotypes were identified for the cytb and COI datasets, respectively (GenBank Accession numbers: HM219573–HM219612). For bfibr, 16 different alleles were identified among our dataset (GenBank Accession numbers: HM219555–HM219570). The mitochondrial (cytb and COI combined) and the three combined genes datasets showed 29 and 36 haplotypes, respectively (see Appendix 1—Table 5 for mtDNA haplotypes/bfibr alleles distribution among sampling localities and Appendices 2 and 3 for GenBank Accession numbers and haplotype definitions).

The cytb gene (900 bp) contained 145 (293 in the dataset including outgroup sequences) variable sites and among them 123 (257 in the dataset including outgroup sequences) were informative for parsimony, whereas the COI gene (713 bp) contained 71 (179) variables sites, 64 (167) of which being parsimoniously informative. The bfibr (745 bp) contained 29 (85) variables sites and 17 (59) were informative for parsimony.

No stop codons or indels were observed in the coding genes (cytb, COI).

Phylogenetic and phylogeographic analysis

The two mitochondrial genes analysed separately yielded low-supported phylogenies, but gave globally the same topologies. Therefore, they were combined into one matrix, yielding a better-resolved phylogeny (Fig. 2a). The ML and Bayesian analyses gave congruent results although minor inconsistencies between these methods were observed and were characterized by low bootstrap support in the ML tree (node D in Fig. 2a). The Rattini divisions proposed by Musser and Carleton (2005) were observed in our phylogenetic tree. The *Maxomys* division is the first to diverge, followed by the *Dacnomys* division, here represented by *Leopoldamys* and *Niviventer* genera. The phylogeny clearly indicated that samples of *L. neilli* belong to a monophyletic group. The sister species of *L. neilli* is represented by *L. edwardsi* sequences. The haplotypes of *L. neilli* clustered into six geographically well-structured lineages: Loei lineage (samples from Loei province, localities 1–3; vouchers R4517 and R4527 were trapped in locality 3, Loei province), Phrae lineage (samples from Phrae province, locality 5; vouchers R4485 and R4486 were trapped in locality 5, Phrae

Author's personal copy

province), Nan lineage (samples from Nan province, locality 4), West Centre lineage (samples from Lopburi province and from the western part of Saraburi province, localities 6–8), East Centre lineage (samples from Nakhon Ratchasima province and from the eastern part of Saraburi province, localities 9–12) and Kanchanaburi lineage (samples from Kanchanaburi province, localities 13–20) (Fig. 1). The monophyly of these lineages is very well supported with BS: 100% - BP: 1.0. The Kanchanaburi lineage is separated from all the other regions, which are nested in a single well supported group (BS: 88% - BP: 1.0). Within this group, the East and West Centre lineages form a subgroup (BS: 85% - BP: 0.97) sister to another weakly supported subgroup including the Nan, Phrae and Loei lineages (BS: 47%- BP: -). The bfibr phylogeny also supported the monophyly of the three *Leopoldamys* species, but intraspecific relationships in *L. neilli* were not resolved (Fig. 2b). The ML and Bayesian consensus trees combining nuclear and mitochondrial datasets recovered globally the same topology as the mitochondrial tree (Fig. 2c). They confirmed the monophyly of all the lineages, with a strong support (BS: 100% - BP: 1.0) and the separation of the Kanchanaburi lineage from all the other regions. The weakly supported group including the Nan, Phrae and Loei lineages in the ML analysis of the mitochondrial dataset (node D in Fig. 2a) is not observed in the combined dataset (Fig. 2c).

The MP consensus tree (length = 444) of the combined dataset (cytb/COI/bfibr) recovered similar topology than ML and BI trees. PBS analysis showed conflict between mitochondrial and nuclear partitions for most of the lineages within *L. neilli* (negative value of PBS for the bfibr) (Table 1). All the nodes are well supported by the two mitochondrial genes (with the exception of the Phrae + - Loei lineage (node D in Fig. 2c) for COI) but the bfibr provides support for only two nodes: the Phrae + Loei lineage (node D) and the Kanchanaburi lineage (node K).

The median joining network for the mitochondrial dataset (Fig. 3) generated six geographically well-structured haplogroups that are clearly isolated and separated by a very high number of mutational steps. The structure of network for the Kanchanaburi and Loei lineages appeared heterogeneous, which would signal a stable population.

Genetic diversity and population differentiation

Haplotype and nucleotide diversities were calculated for the six main lineages for the three markers independently. The results are presented in Table 2 and indicate that all the lineages show a very low genetic diversity for all the markers (between 0 and 0.81 for haplotype diversity and between 0 and 0.58% for nucleotide diversity). The genetic divergence (K2P distance) within each lineage is also very low (between 0 and 0.4% for all markers). No haplotypes/alleles are shared among the six main geographic regions, with the exception of one bfibr allele shared between the East and West Centre lineages (see Appendix 1—Table 5 for mtDNA haplotypes/bfibr alleles distribution among sampling localities).

Fig. 2 **a** Maximum likelihood tree summarizing the phylogenetic relationships among the studied populations based on the mitochondrial dataset (GTR + G). Bootstrap support (1,000 replicates)/posterior probabilities of nodes are indicated above the branches. Node support values from within lineages were removed for clarity of presentation. The geographic distribution (locality ID) of each haplotype has been included in *brackets* (see Fig. 1 and Appendix 1—Table 5 for numbers and localities correspondences). See Table 4 for the divergence time estimates of nodes marked by a *letter* in Fig. 2a. **b** ML tree summarizing the phylogenetic relationships among the studied populations based on the nuclear marker (bfibr) (GTR + G). Bootstrap support (1,000 replicates)/posterior probabilities of nodes are indicated above the branches. Only nodes with BS > 750 and/or BP > 0.95 were indicated. The geographic distribution (locality ID) of each haplotype has been included in *brackets* (see Fig. 1 and Appendix 1—Table 5 for numbers and localities correspondences) **c** ML tree summarizing the phylogenetic relationships among the studied populations based on the three combined gene dataset (GTR + G). Bootstrap support (1,000 replicates)/posterior probabilities of nodes are indicated above the branches. Node support values from within lineages were removed for clarity of presentation. The geographic distribution (locality ID) of each haplotype has been included in *brackets* (see Fig. 1 and Appendix 1—Table 5 for numbers and localities correspondences). See Table 1 for the PBS of nodes marked by a *letter* in Fig. 2c ▶

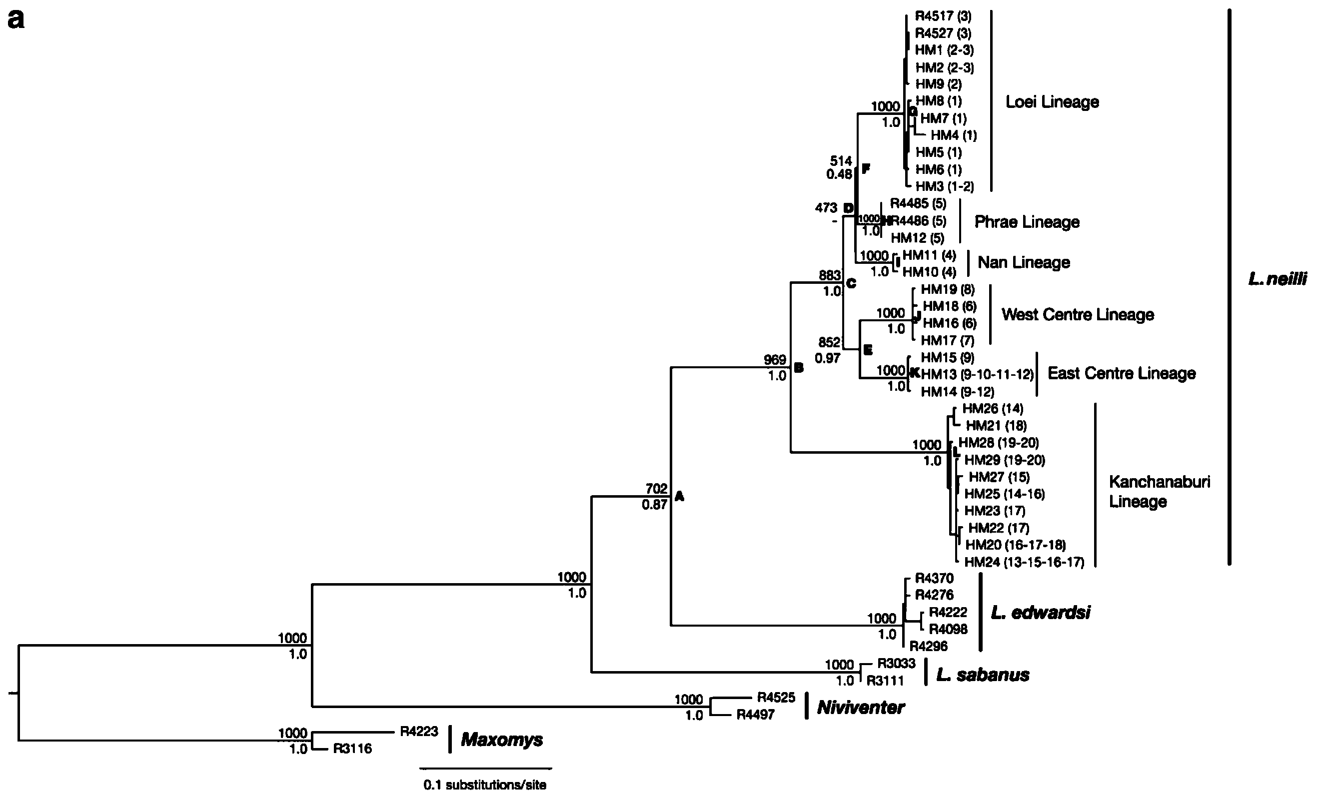
The AMOVA performed on the bfibr, mitochondrial (cytb/COI) and on the three combined genes (cytb/COI/bfibr) datasets showed that 54, 96, and 94% respectively of variation is explained by differences among geographical groups, whereas 2, 2, 2% of this variation are explained by differences among populations within groups and 44, 2, 4% of this variation are explained by differences within populations, respectively. These results are confirmed by the high genetic divergence (K2P distance, calculated only on the cytb dataset) observed between the lineages (Table 3).

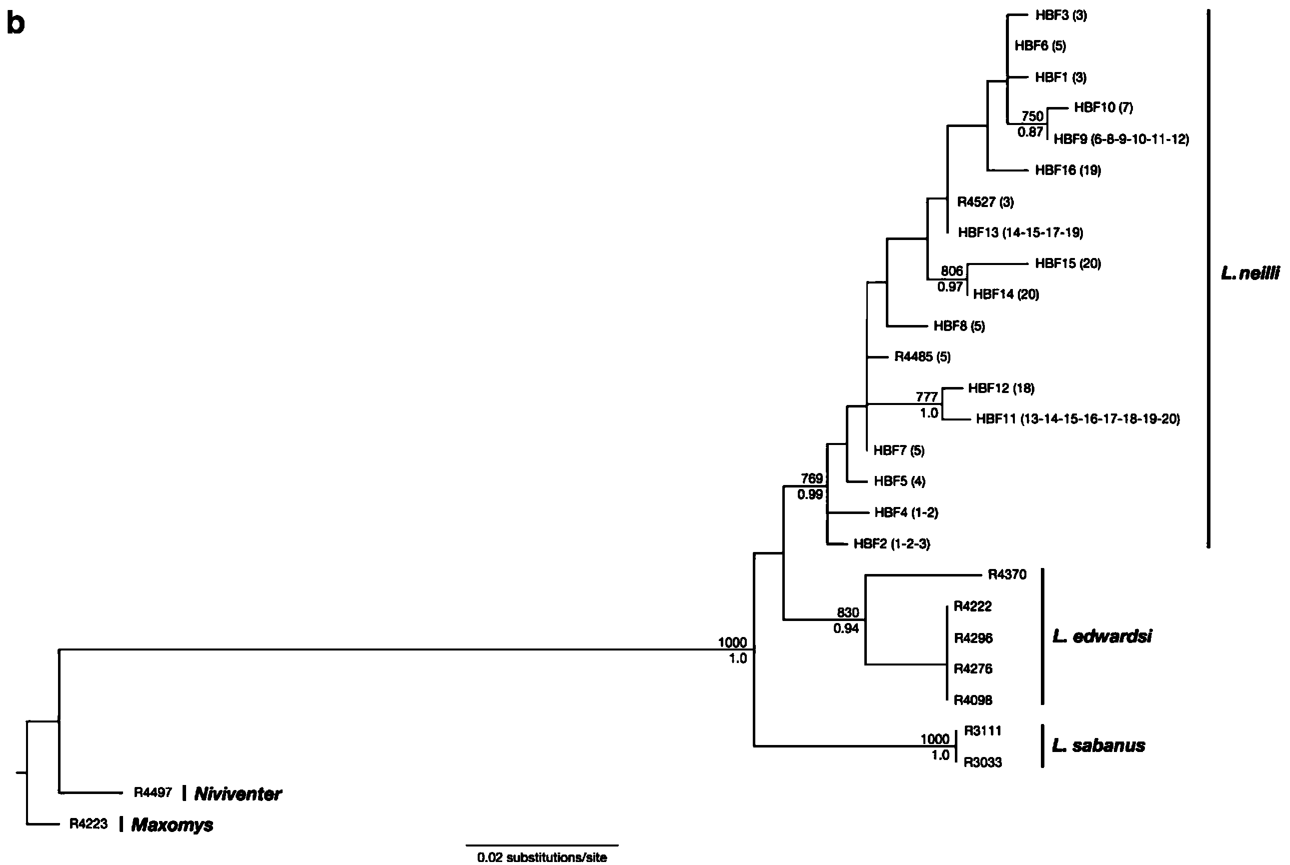
Divergence time analysis

The most recent common ancestor between *L. neilli* and *L. edwardsi* was estimated to occur during the Late Pliocene, around 3.2 Myr (Table 4). The most recent common ancestor for the main lineages of *L. neilli* was estimated to live around 2.2 Myr ago (Table 4). Therefore, our results suggest that the separation between the Kanchanaburi lineage from the other lineages of *L. neilli* appeared during the Early Pleistocene. The divergence estimations obtained for the remaining lineages (Loei, Nan, Phrae, East, and West Centre), suggested that they appeared during the Middle Pleistocene (Table 4).

Divergence time estimates based on the mean net K2P distance between lineages were similar to those obtained with the BI. Using a mutation rate of 2.6%/Myr for the cytb gene,

Author's personal copy





Author's personal copy

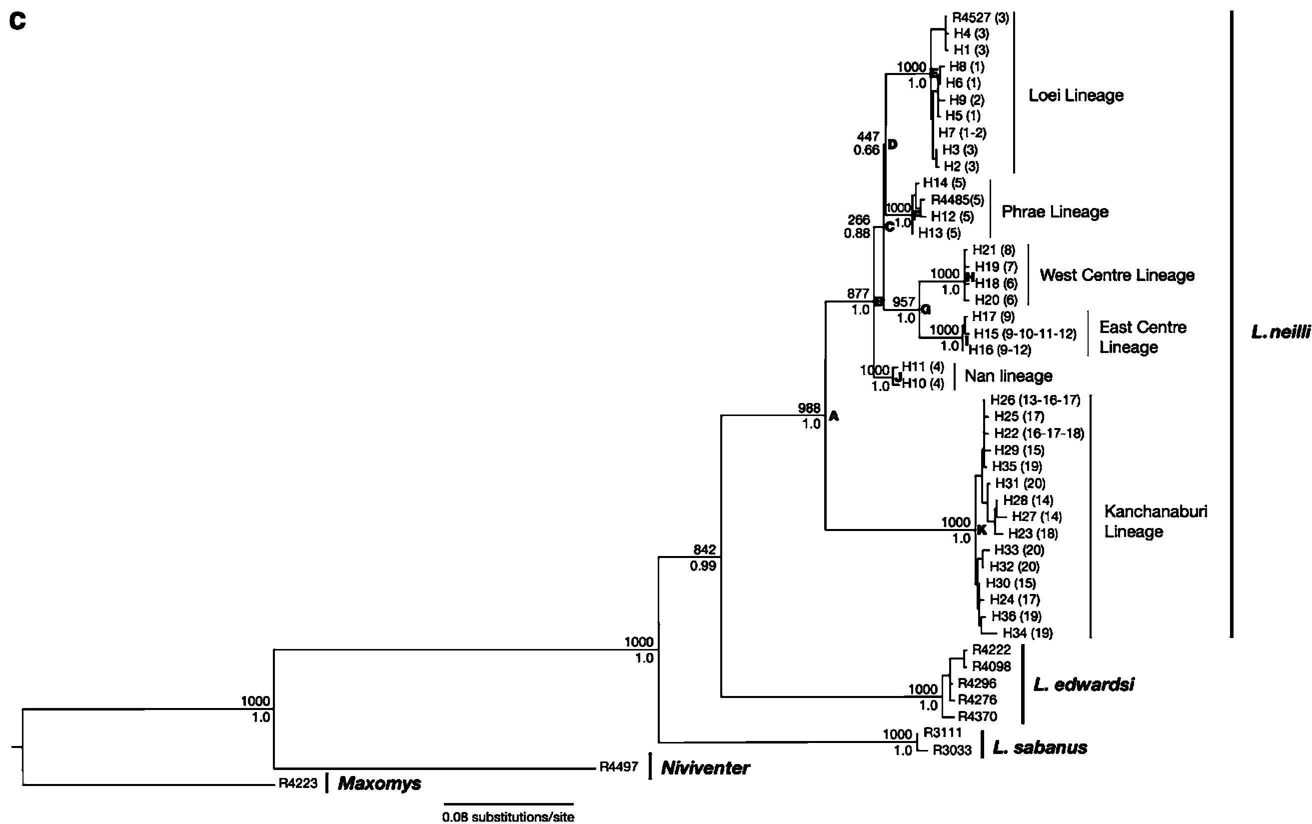


Fig. 2 continued

Table 1 Results from PBS analysis. Positive values indicate support for lineages, while negative values indicate a lack of support

Lineages	Corresponding node in the phylogenetic tree (Fig. 2c)	cytb	COI	bfibr
Leopoldamys neilli	A	46.28	44.11	2.61
Loei + Phrae + E–W Centre + Nan	B	9.85	7.79	−0.64
Loei + Phrae + E–W Centre	C	1.13	0.27	−0.4
Loei + Phrae	D	1.05	−0.15	0.1
Loei	E	7.58	4.81	−0.39
Phrae	F	7.05	2.64	−0.69
East + West Centre	G	3.02	2.37	−3.39
West Centre	H	11.28	8.78	−1.06
East Centre	I	11.19	4.37	−3.56
Nan	J	1.66	5.91	−1.58
Kanchanaburi	K	36.25	14.54	0.21

and taking into account the correction for ancestral mtDNA polymorphism, the following molecular dates were obtained: 2.7 Myr for the separation between *L. neilli* and *L. edwardsi* (7% of K2P distance), 2.15 Myr for the separation between the Kanchanaburi lineage and other lineages (5.6% of K2P distance), and 0.81 Myr for the separation between the Centre lineages (East and West Centre) and the Northern lineages (Loei, Phrae, and Nan) (2.1% of K2P distance).

Demographic history

Past demographic trends of the Kanchanaburi, Loei and East Centre lineages were inferred using Bayesian Skyline Plot reconstruction. It appears that these lineages have experienced a long period of constant population size (Kanchanaburi and Loei lineages) or slight population expansion (East Centre lineage) but no real population trend was evident once the highest probability density (HPD) limits were taken into account (data not shown). A tendency for stable populations for all three lineages was confirmed with mismatch distributions, which revealed a multimodal pattern for each locus dataset (data not shown).

Fu and Li's F* and D* were not significant for all lineages; Fu's Fs values were not significant and had values around 0 or lightly positive for the three lineages (cytb: between 0.87 and 1.44–COI: between −0.84 and −0.28–bfibr: between 1.63 and 2.84). These results also suggest stable populations or small declines. The R2 index was not significant for the three loci, thus indicating stable populations.

Author's personal copy

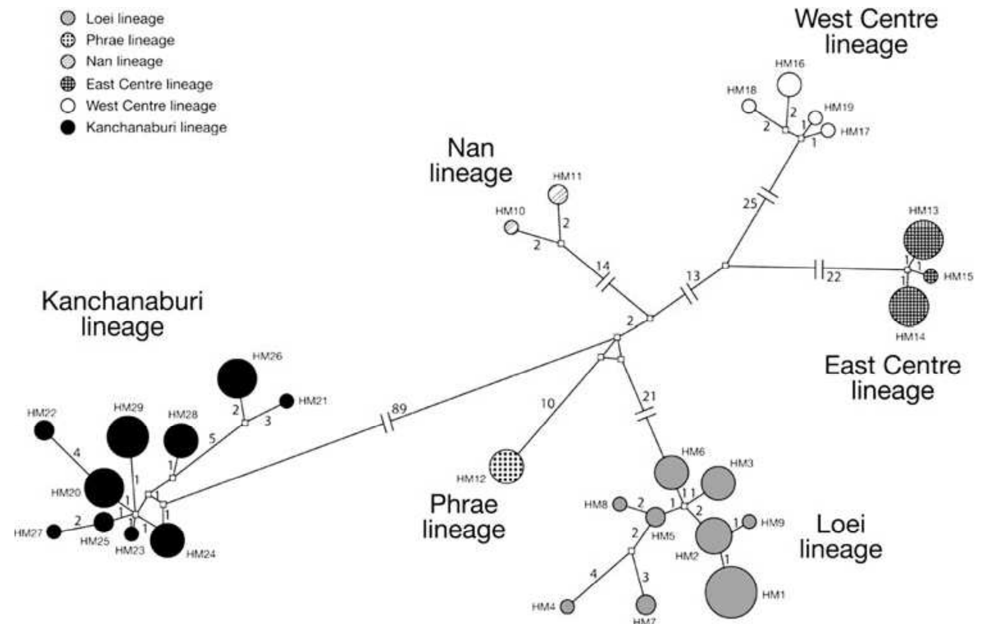


Fig. 3 Median-joining network summarizing the relationships among the studied populations based on the mitochondrial dataset. *Circles* represent haplotypes obtained in this study and the size of *circle* is proportional to the number of individuals sharing a haplotype. *Squares* represent median vectors, i.e. presumed unsampled or missing intermediate haplotypes. *Numbers* on branches represent the number of mutational steps between haplotypes

Table 2 Genetic diversity observed within the main genetic lineages of *Leopoldamys neilli* for the three loci

Locus	Lineages	Sample size	Number of haplotypes	Haplotype diversity (h ± SD)	Nucleotide diversity (π ± SD)	Genetic divergence within each lineage (K2P distance)
cytb	Loei	40	6	0.79 ± 0.04	0.0024 ± 0.001	0.001
	Nan	3	2	0.67 ± 0.31	0.0022 ± 0.002	0.003
	Phrae	6	1	0.00 ± 0.00	0.0000 ± 0.000	0.000
	East Centre	17	3	0.59 ± 0.06	0.0013 ± 0.001	0.000
	West Centre	6	3	0.60 ± 0.21	0.0011 ± 0.001	0.001
	Kanchanaburi	43	7	0.81 ± 0.03	0.0032 ± 0.002	0.003
COI	Loei	40	4	0.19 ± 0.08	0.0007 ± 0.0007	0.001
	Nan	3	2	0.67 ± 0.31	0.0009 ± 0.0012	0.001
	Phrae	6	1	0.00 ± 0.00	0.0000 ± 0.0000	0.000
	East Centre	17	1	0.00 ± 0.00	0.0000 ± 0.0000	0.000
	West Centre	6	3	0.73 ± 0.15	0.0021 ± 0.0017	0.002
	Kanchanaburi	43	5	0.61 ± 0.08	0.0019 ± 0.0013	0.002
bfibr	Loei	12	4	0.68 ± 0.10	0.0054 ± 0.0032	0.004
	Nan	3	1	0.00 ± 0.00	0.0000 ± 0.0000	0.000
	Phrae	5	3	0.70 ± 0.22	0.0058 ± 0.0040	0.003
	Centre East	11	1	0.00 ± 0.00	0.0000 ± 0.0000	0.000
	Centre West	6	2	0.33 ± 0.21	0.0002 ± 0.0000	0.000
	Kanchanaburi	28	5	0.64 ± 0.09	0.0057 ± 0.0033	0.002

Table 3 Genetic divergence between lineages [K2P distance (below diagonal) and net K2P distance (above diagonal)] for the cytb dataset

	Loei	Nan	Phrae	East Centre	West Centre	Kanchanaburi
Loei		0.023	0.015	0.031	0.032	0.065
Nan	0.025		0.015	0.033	0.038	0.057
Phrae	0.015	0.017		0.023	0.028	0.061
East Centre	0.031	0.034	0.023		0.027	0.065
West Centre	0.033	0.040	0.029	0.027		0.071
Kanchanaburi	0.067	0.059	0.062	0.067	0.073	

Author's personal copy

Table 4 Divergence time estimates [in 1,000 years (kyr)] of clades observed in mitochondrial phylogenetic tree (Fig. 2a). Mean values and 95% HPD for estimates of time to the most recent common ancestor are presented

	Corresponding node in the phylogenetic tree (Fig. 2a)	Mean (kyr)	95% HPD (kyr)
TMRCA (*L. neilli* + *L. edwardsi*)	A	3232.6	1852.3–4693.4
TMRCA (*L. neilli*)	B	2235.2	1223.3–3301.3
TMRCA (L + P + N + E–W C)	C	1148.1	658.7–1736.6
TMRCA (Loei + Phrae + Nan)	D	793.8	408.3–1242.5
TMRCA (East–West Centre)	E	786.3	359.1–1247.4
TMRCA (Loei + Phrae)	F	682	328–1078.6
TMRCA (Loei)	G	237.3	98.9–402
TMRCA (Phrae)	H	65.6	6.8–150.6
TMRCA (Nan)	I	76.3	6–172.8
TMRCA (East Centre)	J	110.9	30.3–211.9
TMRCA (West Centre)	K	48	3–112.1
TMRCA (Kanchanaburi)	L	173.5	67.8–302.9

Analysis of selection influence

The McDonald–Kreitman test showed no significant evidence of selection on the cytb and COI dataset.

Discussion

Phylogeographic structure of *Leopoldamys neilli*

Our study revealed a strong geographic structure of the mtDNA genetic diversity for *L. neilli*: six highly differentiated genetic lineages that corresponded to particular regions of Thailand (Loei, Nan, Phrae, Lopburi + Western Saraburi, Nakhon Ratchasima + Eastern Saraburi and Kanchanaburi provinces) (Fig. 1). This structure is highly supported by the AMOVA analysis that showed that at least 96% of the mtDNA variation is distributed among the six main geographic regions. These six mtDNA lineages exhibit a very high degree of genetic divergence (Table 2) and gene flows among them are extremely low for mtDNA.

In contrast, the bfibr dataset does not show a clear phylogeographic structure and does not support lineages corresponding to the six mtDNA lineages with the exception of the Kanchanaburi lineage (PBS analysis, Table 1). However, the AMOVA analysis showed that 54% of bfibr variation is distributed among the six main geographic regions. This discrepancy between nuclear and mitochondrial genomes could result from several alternative phenomena. First, this could be explained by incomplete lineage sorting of alleles, given the low mutation rate and the low intraspecific variability of the bfibr gene and also the longer coalescent time of nuclear loci. The divergence of lineages within *L. neilli* would be too recent to allow a good phylogeographic resolution with the bfibr gene. A second explanation could be that a strong female philopatry existed as compared to male mediated nuclear gene flow. However this hypothesis seems fairly unlikely as bfibr alleles are not shared between the six mtDNA lineages (with the exception of one bfibr allele shared between the East and West Centre lineages). A third alternative explanation would be that a directional selection exists at the bfibr locus. However, this seems very unlikely as the nuclear fragment used in this study is a non-coding intron sequence and consequently is not under selection pressure.

The most divergent lineage includes the populations from Kanchanaburi province (western Thailand), which have been separated from the other genetic lineages since at least the Early Pleistocene. The other lineages are more closely related and have diverged since the Middle Pleistocene. These results are surprising as some populations are geographically very close to each other (e.g. less than 20 km between the East and West Centre lineages; Fig. 1). Such fine-scale population differentiation at the intraspecific level is very rare among rodents and small mammals. Phylogeographic studies available for some mammal species widespread in Southeast Asia have identified phylogeographic patterns with less genetic structure (Gorog et al. 2004; Campbell et al. 2004; Patou et al. 2009; Patou et al. 2010). Fine-scale population differentiation similar to that observed for *L. neilli* has been reported for very few mammal species strongly associated with a particular habitat, like the rodent species *Laonastes aenigmamus*, which is also highly endemic to Southeast Asian karst habitats (Rivière-Dobigny et al. 2011) and the fruit bat *Thoopterus nigrescens* endemic to Sulawesi and Mollucan Island chain and strongly associated with primary forest (Campbell et al. 2007). Fine scale population differentiation is more regularly observed among small vertebrates with low dispersal abilities, such as amphibians (e.g. Cabe

Author's personal copy

et al. 2007; Martinez-Solano et al. 2007; Garcia-Paris et al. 2000).

Within each lineage of *L. neilli*, the levels of haplotype and nucleotide diversities are very low for each gene. This low level of genetic diversity within lineages could be explained by small population sizes (in smaller populations genetic drift has a larger impact and balancing selection is less effective; Frankham et al. 2002), and/or by a prolonged demographic bottleneck in recent times (Avise 2000). The high level of divergence for mtDNA among *L. neilli* regional lineages, the low genetic diversity within lineages, and the relatively ancient separation of lineages, suggest a severe population fragmentation with similar effects than those observed in oceanic islands (small population sizes, restricted migration, genetic differentiation among populations, low genetic diversity). This fragmentation is correlated to the scattered distribution of the karst habitat of this species; the lowland areas between limestone karsts thus act as biogeographical barriers and prevent gene flow between lineages isolated on "karst islands".

Our results attest the high endemicity of *L. neilli* to limestone karsts and its high ecological specialization. Published data corroborate this conclusion since all specimens documented as *L. neilli* in the literature were caught in limestone karsts (Lekagul and McNeely 1988; Waengsothorn et al. 2007), with the exception of two specimens trapped in lowland bamboo forests in Kanchanaburi province (Wiles 1981). However, according to Waengsothorn et al. (2007), these samples have an external morphology similar to *L. sabanus* and could be misidentified. Field data of the CERoPath project (international project investigating the Southeast Asian rodent biodiversity and their pathogens; http://www.ceropath.org/) also confirm the endemicity of *L. neilli* to karst habitat. More than 2,500 Murinae rodents were caught in the framework of this project in the north, east and western part of Thailand in several habitat types and no *L. neilli* were trapped outside limestone karsts (S. Morand, pers. com.).

A similar phylogeographic pattern has also been observed for other taxa endemic to Southeast Asian limestone karsts, such as *Laonastes aenigmamus* endemic to Khammouan karst in Central Lao PDR (Rivière-Dobigny et al. 2011) and the mosquito, *Anophele scanloni* endemic to karsts of South Thailand (from Kanchanaburi at the north of its range to the Malaysian border in the south) (O'Loughlin et al. 2007). Important similarities, such as a strong genetic structure, several highly divergent lineages corresponding to particular karstic areas and restricted gene flow among these lineages, are observed for these three species.

Phylogeographic history of *Leopoldamys neilli*

The divergence of *L. neilli* and *L. edwardsi* was estimated to occur during the Late Pliocene, (around 2.7–3.2 Myr). The first splits within *L. neilli* occurred during the Early Pleistocene (around 2.2 Myr), separating the Kanchanaburi lineage from the other lineages of *L. neilli*. The divergence estimations obtained for the remaining lineages suggested that they appeared during the Middle Pleistocene (Table 4).

These divergence time estimates are to be treated with caution given the intrinsic inaccuracy of temporal estimates based on molecular data. Indeed, it has been demonstrated that the extrapolation of molecular rates across different evolutionary timescales can result in inaccurate age estimates (Heads 2005; Welch and Bromham 2005; Ho and Larson 2006). However, it was not possible to find more suitable calibration points than the split *Mus*/*Rattus* between 10 and 12 Myr and the divergence between *A. mystacinus* and (*A. flavicollis*–*A. sylvaticus*) at 7 Myr (younger paleontological data are absent). Therefore the divergence time estimates obtained in this study have to be considered in a relative time frame rather than as exact dating but they are consistent with the data available in the literature. Indeed, the Late Pliocene/Early Pleistocene is recognized as a period of intense speciation in Southeast Asia (Meijaard and Groves 2006). Verneau et al. (1998) described the near simultaneous generation of five Southeast Asian rat lineages around 2.7 Myr ago. During the Late Pliocene/Early Pleistocene, environmental changes may have lead to greater habitat diversity in Southeast Asia, with the isolation of tropical forest in Vietnamese and Laotian mountains, and the development of more open vegetation types in some regions (Meijaard and Groves 2006).

To explain the strong phylogeographic pattern of *L. neilli*, and the ancient separation of the Kanchanaburi lineage from the others, we hypothesise that palaeokarsts existed within Thailand; palaeokarsts are karsts currently buried by younger rocks (Simms 2005). These palaeokarsts would have allowed the dispersal of *L. neilli* throughout Thailand before they were buried, which consequently isolated lineages of *L. neilli*. The geological history of Thailand corroborates our hypothesis. Limestone of Thailand ranges in age from Ordovician to Triassic (Gillieson 2005). Throughout the Cenozoic Era, the collision of the Indian and Eurasian continental plates resulted in the tectonic deformation of Thailand (Dheeradilok 1995). From the Early Pleistocene to Holocene, sediments progressively buried the basement and Tertiary rocks of the Central Plain of Thailand. These Quaternary deposits of about 2000 m of thickness represent a complex sequence of alluvial, fluvial, and deltaic sediments and were overlaid in Lower Central

Author's personal copy

Plain with thick marine clay during Holocene sea transgression (Dheeradilok and Kaewyana 1983; Sinsakul 2000). In the Late Pleistocene, frequent floodings and progressive burring of the limestone karsts of the Central Plain would have isolated populations of *L. neilli* in western Thailand, which would explain the separation of the Kanchanaburi lineage from the other lineages. During the Quaternary, fluviatile deposits in the main river valley basins within the northern region and at the eastern margin of the Central Plain could have reduced migration between the northern and central lineages (Dheeradilok and Kaewyana 1983).

The biogeographical significance of the western region of Thailand has also been observed by Pagès et al. (2010) who found a divergent genetic lineage for the rodent genus *Berylmys* within this region. Moreover, White et al. (2004) found a difference in Early Holocene vegetation between northeast and northwest Thailand, which they suggested was due to the marked topographic differences between these two areas.

Another factor that could have influenced the phylogeographic history of *L. neilli* is the extent of dry forest in Thailand; a habitat to which this species appears to be particularly well adapted. The subtropical karst vegetation is a mixed evergreen and deciduous broad-leaved forest suffering from temporary drought events because of the shallow soils and highly porous limestone. Consequently, this vegetation is characterized by several morphological and physiological adaptations to resist drought stress and recover rapidly after rainfall (Liu et al. 2010). The typical vegetation of the subtropical karsts could be similar to the vegetation adapted to arid and seasonal climate of the cooling periods of the Pleistocene (White et al. 2004). We hypothesize that the dispersion of *L. neilli* throughout Thailand could have been facilitated by the expansion of dry forest during the cooling periods of the Pleistocene. Later, during the interglacial periods of the Quaternary era, dry forests were replaced by wet tropical forest in many areas of Thailand (Penny 2001; White et al. 2004), but were preserved on small isolated patches over limestone karst areas. These habitat changes would have reinforced the isolation and genetic differentiation of lineages due to the disappearance of karsts during the Quaternary.

Implications for the taxonomy and the conservation of *L. neilli*

Our study revealed the existence of several highly differentiated genetic lineages of *L. neilli* corresponding to different limestone karst regions throughout Thailand. The Kanchanaburi lineage is separated from the other lineages by a very high level of genetic divergence, which is similar to what has been observed among at least subspecies in other rodent species (e.g. Bradley and Baker 2001; Michaux et al. 2003, 2005; Mouline et al. 2008). This high differentiation could thus confer a sub-specific or specific status to this lineage. However, further studies based on other nuclear markers, morphological, ethological and ecological characters, are needed to assess the appropriate taxonomic status of the Kanchanaburi lineage.

Limestone karsts are currently highly threatened through quarrying, deforestation, hunting, and urbanization (Vermeulen and Whitten 1999; Clements et al. 2006). 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). As each limestone area of Thailand is characterized by a particular genetic lineage of *L. neilli*, the destruction of these karst habitats will lead to the disappearance of unique intraspecific strains not found elsewhere in Southeast Asia. Furthermore, the low genetic diversity within each lineage increases extinction risk, since populations with lower genetic diversity are poorer at coping with environmental changes and diseases than populations with higher genetic diversity (Frankham et al. 2002). As the six *L. neilli* lineages that we described exchange very few migrants and are geographically and genetically isolated, they should be considered at least as distinct MU and may require separate management and conservation plans (Moritz 1994; Avise 2000). Large undisturbed limestone karsts areas should be preserved in each karst region of Thailand in order to protect the genetic diversity of *L. neilli*. Moreover, as nuclear bfibr alleles are not shared between the six mtDNA lineages with the exception of one allele shared between the East and West Centre lineages, these six MUs could also be considered as distinct ESUs (Moritz 1994). Future investigation using more variable nuclear markers will help to confirm this last hypothesis.

Since *L. neilli* appears to be highly endemic to a strongly threatened habitat and to be distributed in a restricted area, we strongly recommend classifying this species as "Endangered" on the IUCN Red List (Lunde and Aplin 2008).

Conclusion

This study revealed an unexpected high level of genetic differentiation within the species *L. neilli*: six highly differentiated, allopatric genetic lineages were observed in our dataset. They exhibit a very high degree of genetic divergence, low gene flow among lineages and low levels of haplotype and nucleotide diversities within lineages. Our

Author's personal copy

results highlight the importance of protecting limestone habitats in Thailand to preserve both interspecific and intraspecific diversity. The study of the genetic structure of other species endemic to karst habitat will allow us to propose the best conservation and management measures.

Acknowledgments We thank Boonchai Tontan who worked hard in the field and help us to collect the samples used in this study. We are grateful to G. Vanhamel for his help in the lab. We express gratitude to S. Jittapalapong for his administrative help. We thank Drs D. Aigoin, C. Berthouly, C. Ek, S. Morand, M. Pagès for their great help and/or comments on the manuscript. We thank E. Latinne for her help in the drafting of the figures. We wish to acknowledge Dr Andy Jennings for English corrections and scientific comments. This work was supported by a Belgian FRS-FNRS (Fonds de la Recherche Scientifique) fellowship to A. Latinne (mandate "aspirant") and to J.R. Michaux (mandate "chercheur qualifié"), and a financial grant from the Belgian FNRS (crédits pour brefs séjours à l'étranger to A. Latinne), from the University of Liège (Patrimoine) and from the Communauté française de Belgique. 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.

Appendix 1

See Table 5.

Table 5 Sampling locality, samples used for sequencing (*N*) for each dataset and haplotype/allele distribution

Locality	Coord (Lat–Long)	Locality ID (Fig. 1)	Mitochondrial dataset		Nuclear dataset	
			N	Haplotype distribution (no. of individuals)	*N*	Allele distribution (no. of individuals)
Loei	17°05′ 101°47′	1	17	HM3 (5)–HM4 (1)–HM5 (2)–HM6 (6)–HM7 (2)–HM8 (1)	4	HBF2 (1)–HBF4 (3)
Loei	17°03′ 101°54′	2	14	HM1 (7)–HM2 (5)–HM3 (1)–HM9 (1)	4	HBF2 (3)–HBF4 (1)
Loei	17°06′ 101°56′	3	9	HM1 (7)–HM2 (2)	4	HBF1 (1)–HBF2 (2)–HBF3 (1)
Nan	19°23′ 100°36′	4	3	HM10 (1)–HM11 (2)	3	HBF5 (3)
Phrae	18°22′ 100°21′	5	6	HM12 (6)	5	HBF6 (3)–HBF7 (1)–HBF8 (1)
Saraburi	14°39′ 100°58′	6	4	HM16 (3)–HM18 (1)	4	HBF9 (4)
Saraburi	14°40′ 101°00′	7	1	HM17 (1)	1	HBF10 (1)
Lopburi	14°48′ 100°49′	8	1	HM19 (1)	1	HBF9 (1)
Saraburi	14°34′ 101°08′	9	8	HM13 (5)–HM14 (2)–HM15 (1)	4	HBF9 (4)
Saraburi	14°38′ 101°08′	10	1	HM13 (1)	1	HBF9 (1)
Nakhon Ratchasima	14°35′ 101°14′	11	1	HM13 (1)	1	HBF9 (2)
Nakhon Ratchasima	14°33′ 101°18′	12	7	HM13 (1)–HM14 (6)	5	HBF9 (5)
Kanchanaburi	14°14′ 99°03′	13	1	HM24 (1)	1	HBF11 (1)
Kanchanaburi	14°12′ 99°07′	14	8	HM25 (1)–HM26 (7)	5	HBF11 (3)–HBF13 (2)
Kanchanaburi	14°12′ 99°08′	15	2	HM24 (1)–HM27 (1)	2	HBF11 (1)–HBF13 (1)
Kanchanaburi	14°08′ 99°17′	16	7	HM20 (3)–HM24 (3)–HM25 (1)	4	HBF11 (4)
Kanchanaburi	14°03′ 99°25′	17	7	HM20 (3)–HM22 (2)–HM23 (1)–HM24 (1)	4	HBF11 (3)–HBF13 (1)
Kanchanaburi	13°54′ 99°18′	18	3	HM20 (2)–HM21 (1)	2	HBF11 (1)–HBF12 (1)
Kanchanaburi	14°30′ 99°24′	19	9	HM28 (3)–HM29 (6)	5	HBF11 (2)–HBF13 (2)–HBF16 (1)
Kanchanaburi	14°28′ 99°25′	20	6	HM28 (3)–HM29 (3)	5	HBF11 (3)–HBF14 (1)–HBF15 (1)

Author's personal copy

Appendix 2

See Tables 6 and 7.

Table 6 Genbank accession numbers for cytb and COI haplotypes and for bfibr alleles of *L. neilli*

cytb		COI		bfibr	
Haplotype	Accession no.	Haplotype	Accession no.	Sequence	Accession no.
HCB1	HM219591	HCOI1	HM219573	HBF1	HM219555
HCB2	HM219592	HCOI2	HM219574	HBF2	HM219556
HCB3	HM219593	HCOI3	HM219575	HBF3	HM219557
HCB4	HM219594	HCOI4	HM219576	HBF4	HM219558
HCB5	HM219595	HCOI5	HM219577	HBF5	HM219559
HCB6	HM219596	HCOI6	HM219578	HBF6	HM219560
HCB7	HM219597	HCOI7	HM219579	HBF7	HM219561
HCB8	HM219598	HCOI8	HM219580	HBF8	HM219562
HCB9	HM219599	HCOI9	HM219581	HBF9	HM219563
HCB10	HM219600	HCOI10	HM219582	HBF10	HM219564
HCB11	HM219601	HCOI11	HM219583	HBF11	HM219565
HCB12	HM219602	HCOI12	HM219584	HBF12	HM219566
HCB13	HM219603	HCOI13	HM219585	HBF13	HM219567
HCB14	HM219604	HCOI14	HM219586	HBF14	HM219568
HCB15	HM219605	HCOI15	HM219587	HBF15	HM219569
HCB16	HM219606	HCOI6	HM219588	HBF16	HM219570
HCB17	HM219607				
HCB18	HM219608				
HCB19	HM219609				
HCB20	HM219610				
HCB21	HM219611				
HCB22	HM219612				

Table 7 Vouchers specimens from the CERoPath tissue collection (Pagès et al. 2010) used in this study and their Genbank accession numbers

Number	Species	Locality	GenBank accession number		
			cytb	COI	bfibr
R4485	*Leopoldamys neilli*	Phrae	HM217459	HM217585	HQ454294
R4486	*Leopoldamys neilli*	Phrae	HM217460	HM217586	–
R4517	*Leopoldamys neilli*	Loei	HM217462	HM217699	–
R4527	*Leopoldamys neilli*	Loei	HM217463	HM217701	HQ454296
R4098	*Leopoldamys edwardsi*	Loei	HM217439	HM217566	HQ454289
R4222	*Leopoldamys edwardsi*	Loei	HM217444	HM217571	HQ454297
R4276	*Leopoldamys edwardsi*	Phrae	HM217449	HM217576	HQ454291
R4296	*Leopoldamys edwardsi*	Phrae	HM217450	HM217577	HQ454292
R4370	*Leopoldamys edwardsi*	Phrae	HM21745	HM217578	HQ454293
R3033	*Leopoldamys sabanus*	Kanchanaburi	HM217400	HM217531	HQ454288
R3111	*Leopoldamys sabanus*	Kanchanaburi	HM217404	HM217534	HQ454295
R4497	*Niviventer fulvescens*	Phrae	HM217461	HM217587	HQ454298
R4525	*Niviventer fulvescens*	Loei	HM217464	HM217589	–
R4223	*Maxomys surifer*	Loei	HM217445	HM217572	HQ454290
R3116	*Maxomys* sp.	Kanchanaburi	HM217405	HM217535	–

Author's personal copy

Appendix 3

See Tables 8 and 9.

Table 8 Mitochondrial haplotype definitions

Mitochondrial haplotypes (cytb + COI)	cytb haplotype	COI haplotype
HM1	HCB1	HCOI1
HM2	HCB2	HCOI1
HM3	HCB3	HCOI1
HM4	HCB4	HCOI2
HM5	HCB4	HCOI1
HM6	HCB5	HCOI1
HM7	HCB4	HCOI3
HM8	HCB6	HCOI1
HM9	HCB2	HCOI4
HM10	HCB7	HCOI5
HM11	HCB8	HCOI6
HM12	HCB9	HCOI7
HM13	HCB10	HCOI8
HM14	HCB11	HCOI8
HM15	HCB12	HCOI8
HM16	HCB13	HCOI9
HM17	HCB13	HCOI10
HM18	HCB14	HCOI11
HM19	HCB15	HCOI11
HM20	HCB16	HCOI12
HM21	HCB17	HCOI13
HM22	HCB18	HCOI14
HM23	HCB19	HCOI14
HM24	HCB16	HCOI14
HM25	HCB20	HCOI14
HM26	HCB17	HCOI15
HM27	HCB20	HCOI16
HM28	HCB21	HCOI14
HM29	HCB22	HCOI14

Table 9 Three combined genes haplotypes definitions

Three combined genes haplotype (cytb + COI + bfibr)	cytb haplotype	COI haplotype	bfibr allele
H1	HCB1	HCOI1	HBF1
H2	HCB1	HCOI1	HBF2
H3	HCB2	HCOI1	HBF2
H4	HCB1	HCOI1	HBF3
H5	HCB4	HCOI1	HBF4
H6	HCB5	HCOI1	HBF4
H7	HCB3	HCOI1	HBF4
H8	HCB5	HCOI1	HBF2
H9	HCB2	HCOI4	HBF2
H10	HCB7	HCOI5	HBF5
H11	HCB8	HCOI6	HBF5
H12	HCB9	HCOI7	HBF6
H13	HCB9	HCOI7	HBF7
H14	HCB9	HCOI7	HBF8
H15	HCB10	HCOI8	HBF9
H16	HCB11	HCOI8	HBF9
H17	HCB12	HCOI8	HBF9
H18	HCB13	HCOI9	HBF9
H19	HCB13	HCOI10	HBF10
H20	HCB14	HCOI11	HBF9
H21	HCB15	HCOI11	HBF9
H22	HCB16	HCOI12	HBF11
H23	HCB17	HCOI13	HBF12
H24	HCB18	HCOI14	HBF13
H25	HCB18	HCOI14	HBF11
H26	HCB16	HCOI14	HBF11
H27	HCB17	HCOI15	HBF13
H28	HCB17	HCOI15	HBF11
H29	HCB20	HCOI16	HBF11
H30	HCB16	HCOI14	HBF13
H31	HCB21	HCOI14	HBF11
H32	HCB22	HCOI14	HBF14
H33	HCB22	HCOI14	HBF15
H34	HCB21	HCOI14	HBF16
H35	HCB22	HCOI14	HBF11
H36	HCB22	HCOI14	HBF13

References

Achard F, Eva HD, Stibig HJ, Mayaux P, Gallego J, Richards T, Malingreau JP (2002) Determination of deforestation rates of the World's humid tropical forests. Science 297:999–1002

Aguilar JP, Michaux J (1996) The beginning of the age of Murinae (Mammalia: Rodentia) in southern France. Acta Zool Cracov 39:35–45

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge

Baker RH, DeSalle R (1997) Multiple sources of character information and the phylogeny of Hawaiian Drosophila. Syst Biol 46:654–673

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Mol Ecol 17:3754–3774

Bradley R, Baker R (2001) A test of the genetic species concepts: cytochrome-b sequences and mammals. J Mammal 82:960–973

Author's personal copy

Cabe PR, Page RB, Hanlon TJ, Aldrich ME, Connors L, Marsh DM (2007) Fine-scale population differentiation and gene flow in a terrestrial salamander (*Plethodon cinereus*) living in a continuous habitat. Heredity 98:53–60

Campbell P, Schneider CJ, Adnan AM, Zubaid A, Kunz TH (2004) Phylogeny and phylogeography of Old World fruit bats in the *Cynopterus brachyotis* complex. Mol Phylogenet Evol 33:764–781

Campbell P, Putnam A, Bonney C, Bilgin R, Morales JC, Kunz TH, Ruedas LA (2007) Contrasting patterns of genetic differentiation between endemic and widespread species of fruit bats (Chiroptera: Pteropodidae) in Sulawesi, Indonesia. Mol Phylogenet Evol 44:474–482

Clark AG (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. Mol Biol Evol 7:111–122

Clements R, Sodhi NS, Schilthuizen M, Ng PKL (2006) Limestone karsts of Southeast Asia: imperiled arks of biodiversity. Bioscience 56:733–742

Clements R, Ng PKL, Lu XX, Ambu S, Schilthuizen M, Bradshaw CJA (2008) Using biogeographical patterns of endemic land snails to improve conservation planning of limestone karsts. Biol Conserv 141:2751–2764

Corbet G, Hill J (1992) The mammals of the Indomalayan region: a systematic review. Oxford University Press, Oxford

Day MJ, Urich PB (2000) An assessment of protected karst landscapes in Southeast Asia. Cave Karst Sci 27:61–70

Dheeradilok P (1995) Quaternary coastal morphology and deposition in Thailand. Quat Int 26:49–54

Dheeradilok P, Kaewyana W (1983) On the quaternary deposits of Thailand. In: Conference on Geology and mineral resources of Thailand, Bangkok, pp 127–135

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22:1185–1192

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. Public Libr Sci Biol 4:699–710

Eldridge MDB, King JM, Loupis AK, Spencer PBS, Taylor AC, Pope LC, Hall GP (1999) Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock-wallaby. Conserv Biol 13:531–541

Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.0: an integrated software package for population genetics data analysis. Evol Biol Online 1:47–50

Frankham R (1997) Do island populations have less genetic variation than mainland populations? Heredity 78:311–327

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge

Garcia-Paris M, Good DA, Parra-Olea G, Wake DB (2000) Biodiversity of Costa Rican salamanders: implications of high levels of genetic differentiation and phylogeographic structure for species formation. Proc Natl Acad Sci USA 97:1640–1647

Gillieson D (2005) Karst in Southeast Asia. In: Gupta A (ed) The physical geography of Southeast Asia. Oxford University Press, Oxford, pp 157–176

Gorog AJ, Sinaga MH, Engstrom MD (2004) Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). Biol J Linn Soc 81:91–109

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hansen MC, Stehman SV, Potapov PV, Loveland TR, Townshend JRG, DeFries RS, Pittman KW, Arunarwati B, Stolle F, Steininger MK, Carroll M, DiMiceli C (2008) Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. Proc Natl Acad Sci USA 105:9439–9444

Hare MP, Palumbi SR (1999) The accuracy of heterozygous base calling from diploid sequence and resolution of haplotypes using allele-specific sequencing. Mol Ecol 8:1749–1752

Heads M (2005) Dating nodes on molecular phylogenies: a critique of molecular biogeography. Cladistics 21:62–78

Ho SYW, Larson G (2006) Molecular clocks: when times are a-changin'. Trends Genet 22:79–83

Jacobs LL, Downs WR (1994) The evolution of murine rodents in Asia. In: Tomida Y, Li CK, Setoguschi T (eds) Rodent and lagomorph families of Asian origins and their diversification, vol 8. National Science Museum Monograph, Tokyo, pp 149–156

Jacobs LL, Flynn LJ (2005) Of mice… again: the Siwalik rodent record, murine distribution, and molecular clocks. In: Lieberman D, Smith R, Kelley J (eds) Interpreting the past: essays on human primate and mammal evolution. Brill Academic, Leiden, pp 63–80

Jacobs LL, Pilbeam D (1980) Of mice and men: fossil-based divergence dates and molecular "Clocks". J Hum Evol 9:551–555

Jaeger JJ, Tong H, Denys C (1986) Age de la divergence *Mus-Rattus*: comparaison des données paléontologiques et moléculaires. C R Acad Sci II 302:917–922

Koh LP, Sodhi NS (2010) Conserving Southeast Asia's imperiled biodiversity: scientific, management, and policy challenges. Biodivers Conserv 19:913–917

Lekagul B, McNeely JA (1988) Mammals of Thailand. White Lotus Press, Bangkok

Librado P, Rozas J (2009) DnaSP v5: a software or comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Liu CC, Liu YG, Guo K, Zheng YR, Li GQ, Yu LF, Yang R (2010) Influence of drought intensity on the response of six woody karst species subjected to successive cycles of drought and rewatering. Physiol Plantarum 139:39–54

Lunde D, Aplin K (2008) *Leopoldamys neilli*. In: IUCN 2010 (ed) IUCN Red List of threatened species, version 2010.1

Martinez-Solano I, Jockusch EL, Wake DB (2007) Extreme population subdivision throughout a continuous range: phylogeography of *Batrachoseps attenuates* (Caudata: Plethodontidae) in western North America. Mol Ecol 16:4335–4355

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. Nature 14:114–116

Meijaard E, Groves CP (2006) The geography of mammals and rivers in mainland South-East Asia, chap 11. In: Lehman SM, Fleagle JG (eds) Primate biogeography. Springer, New York, pp 305–329

Michaux J, Aguilar JP, Montuire S, Wolff A, Legendre S (1997) Les Murinae (Rodentia, Mammalia) neogenes du Sud de la France: evolution et paleoenvironnements. Geobios 30:379–385

Michaux JR, Magnanou E, Paradis E, Nieberding C, Libois R (2003) Mitochondrial phylogeography of the woodmouse (*Apodemus sylvaticus*) in the western Palaearctic region. Mol Ecol 12:685–697

Michaux JR, Libois R, Filippucci MG (2005) So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicollis*) and the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. Heredity 94:52–63

Moritz C (1994) Defining 'evolutionarily significant units' for conservation. Trends Ecol Evol 9:373–375

Mouline K, Granjon L, Galan M, Tatard C, Doukary A, Solimane AA, Duplantier JM, Cosson JF (2008) Phylogeography of a Sahelian rodents species *Mastomys huberti*: a Plio-Pleistocene story of

Author's personal copy

emergence and colonization of humid habitats. Mol Ecol 17:1036–1053

Musser G, Carleton M (2005) Superfamily Muroidea. In: Wilson DE (ed) Mammal species of the World: a taxonomic and geographic reference, 3rd edn. Johns Hopkins University Press, Baltimore, pp 894–1531

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403:853–858

Nieberding C, Morand S, Libois R, Michaux JR (2006) Parasites and the island syndrome: the colonizationof the western Mediterranean islands by Heligmosomoides polygyrus (Dujardin, 1845). J Biogeogr 33:1212–1222

Nijman V (2010) An overview of international wildlife trade from Southeast Asia. Biodiver Conserv 19:1101–1114

O'Loughlin SM, Somboon P, Walton C (2007) High levels of population structure caused by habitat islands in the malarial vector *Anopheles scanloni*. Heredity 99:31–40

Pagès M, Chaval Y, Waengsothorn S, Cosson JF, Hugot JP, Morand S, Michaux J (2010) Refining the taxonomy of the Rattini tribe: a phylogeny-based delimitation of species boundaries. BMC Evol Biol. doi:10.1186/1471-2148-10-184

Patou ML, Chen J, Cosson L, Andersen DH, Cruaud C, Couloux A, Randi E, Zhang S, Veron G (2009) Low genetic diversity in the masked palm civet *Paguma larvata* (Viverridae). J Zool 278: 218–230

Patou ML, Wilting A, Gaubert P, Esselstyn JA, Cruaud C, Jennings A, Fickel J, G Veron (2010) Evolutionary history of the *Paradoxurus* palm civets—a new model for Asian biogeography. J Biogeogr. doi:10.1111/j.1365-2699.2010.02364.x

Peh KSH (2010) Invasive species in Southeast Asia: the knowledge so far. Biodiver Conserv 19:1083–1099

Penny D (2001) A 40,000 year palynological record from north-east Thailand; implications for biogeography and palaeo-environmental reconstruction. Palaeogeogr Palaeoclimatol Palaeoecol 171: 97–128

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality test against population growth. Mol Biol Evol 19: 2092–2100

Rivière-Dobigny T, Herbreteau V, Khamsavath K, Douangboupha B, Morand S, Michaux JR, Hugot JP (2011) Preliminary assessment of the genetic population structure of the enigmatic species *Laonastes aenigmamus* (Rodentia: Diatomyidae). J Mammal 92:620–628

Robins J, Hingston M, Matisoo-Smith E, Ross H (2007) Identifying *Rattus* species using mitochondrial DNA. Mol Ecol Notes 7:717–729

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

Rosenberg NA, Feldman MW (2002) The relationship between coalescence times and population divergence times. In: Slatkin M, Veuille M (eds) Modern developments in theoretical population genetics. Oxford University Press, New York, pp 130–164

Schilthuizen M, Liew TS, Bin Elahan B, Lackman-Ancrenaz I (2005) Effects of karst forest degradation on pulmonate and prosobranch land snail communities in Sabah, Malaysian Borneo. Conserv Biol 19:949–954

Schipper J, Chanson JS, Chiozza F, Cox NA, Hoffmann M, Katariya V, Lamoreux J et al (2008) The status of the world's land and marine mammals: diversity, threats and knowledge. Science 322: 225–230

Seddon JM, Santucci F, Reeve NJ, Hewitt GM (2001) DNA footprints of European hedgehogs, *Erinaceus europaeus* and *E. concolor*: Pleistocene refugia, postglacial expansion and colonization routes. Mol Ecol 10:2187–2198

Simms MJ (2005) Sedimentary processes, karsts and palaeokarst. In: Selley RC, Cocks LRM, Plimer IR (eds) Encyclopedia of geology. Elsevier, Oxford, pp 678–687

Sinsakul S (2000) Late Quaternary geology of the Lower Central Plain, Thailand. J Asian Earth Sci 18:415–426

Sodhi NS, Koh LP, Brook BW, Ng PKL (2004) Southeast Asian biodiversity: an impending disaster. Trends Ecol Evol 19: 654–660

Sodhi NS, Posa MRC, Lee TM, Bickford D, Koh LP, Brook BW (2010) The state and conservation of Southeast Asian biodiversity. Biodiver Conserv 19:317–328

Sorenson MD, Franzosa EA (2007) TreeRot, version 3. Boston University, Boston

Swofford DL (1998) PAUP*. Phylogenetic analysis using parsimony, (*and other methods), version 4. Sinauer Associates, Sunderland

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Taylor D (2010) Biomass burning, humans and climate change in Southeast Asia. Biodiver Conserv 19:1025–1042

Vermeulen J, Whitten T (1999) Biodiversity and cultural property in the management of limestone resources—lessons from East Asia. World Bank, Washington

Verneau O, Catzeflis F, Furano AV (1998) Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. Proc Natl Acad Sci USA 95:11284–11289

Waengsothorn S, Nabhitabhata J, Moochan T (2007) The ecological distribution of Thai endemic rodents with a new distributional range of *Niviventer hinpoon*. Thailand Nat Hist Mus J 2:31–42

Waengsothorn S, Kenthao A, Latinne A, Hugot JP (2009) Rodents within the Centre for Thai National Reference Collections (CTNRC), past, present and future. Kasetsart J (Nat Sci) 43: 118–124

Welch JJ, Bromham L (2005) Molecular dating when rates vary. Trends Ecol Evol 20:320–327

White JC, Penny D, Kealhofer L, Maloney B (2004) Vegetation changes from the late Pleistocene through the Holocene from three areas of archaeological significance in Thailand. Quat Int 113:111–132

Wiles GJ (1981) Abundance and habitat preferences of small mammals in southwestern Thailand. Nat Hist Bull Siam Soc 29: 44–54

World Bank (2004) Thailand Environment Monitor 2004, Biodiversity conservation. World Bank Report. http://go.worldbank.org/VPRE208MZ0

Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. J Hered 92:371–373