



OPEN Mitogenomic analysis of Thai Sunda pangolins reveals regional phylogeography and informs conservation management

Nattapong Banterng^{1,2}, Kyle Ewart^{1,3,4}, Frankie Thomas Sitam⁵ & Rob Ogden^{1,4}✉

Pangolins are considered the most trafficked mammals in the world with all eight species listed on CITES Appendix I. Despite this pervasive threat to their survival, there remains a limited understanding of genetic diversity and connectivity among populations of Asian pangolin species, hampering effective conservation management. We analysed mitogenome sequences of the Sunda pangolin (*Manis javanica*) from across their Southeast Asia continental distribution, as well as Borneo. Phylogenetic reconstruction revealed six lineages, with clear separation north and south of the Kangar-Pattani biogeographic line in southern Thailand, revealing clear differentiation between Sundaland and Indochinese Sunda pangolin lineages. Further divergence across an east–west divide was observed in central and northern Thailand, extending northwards towards China. Our results provide new insights into the evolutionary relationships among Sunda pangolin populations in Southeast Asia, building on other recent research in this field and helping to establish the species' baseline phylogeography. These inferences will aid conservation planning and support the genetic traceability of the illegal pangolin trade.

Keywords Mitochondrial DNA, Conservation genetics, Phylogeography, Endangered species, Traceability

Pangolins are recognised as “the most heavily trafficked wild mammal in the world”^{1,2}; between 2000–2019, it is estimated that around one million pangolins were trafficked globally³. According to the International Union for Conservation of Nature Red List of Threatened Species⁴, all eight pangolin species in Africa and Asia are now threatened with extinction, with three of the Asian pangolins (Chinese pangolin (*Manis pentadactyla*), Philippine pangolin (*Manis culionensis*) and Sunda pangolin (*Manis javanica*)) considered to be Critically Endangered. Moreover, all pangolin species have been listed in CITES Appendix I since January 2017, which means international trade in pangolins is prohibited⁵. The Sunda pangolin is one of four Asian pangolin species distributed across Southeast Asia and southern China and living in varied habitats from primary forest to urban areas^{6,7} and the most trafficking pangolin in the pre-2000 period³. The illegal wildlife trade is driven by a demand for pangolin scales for use in traditional medicine and pangolin meat, which is considered a delicacy⁸ and Southeast Asia plays an important role in the pangolin trade as a source, hub, and destination⁹. As a result of increasing demand in Asia and the consequent decline of Asian pangolins, African pangolins are also now targeted for illegal trade into the Asian market^{10–13}; however, the pressure on Asian pangolins persists, they are still hunted and trafficked internationally¹⁴ and the number of individual pangolins in illegal trade is still increasing^{3,15}.

Over the last decade, pangolins have received considerable scientific attention, with researchers primarily focusing on pangolin conservation, but also on pangolin biology¹⁶. Despite this rise in interest, significant gaps in our knowledge of pangolins remains¹⁶, with the recent finding of a possible fifth species of Asian pangolin based solely on seized scale samples¹⁷ emphasizing our lack of fundamental zoological knowledge. Within widely distributed species there is relatively little understanding of geographic variation and population structure.

Mitochondrial DNA sequencing is well-recognised as a primary source of information for reconstructing phylogenetic relationships among species^{18,19}, to identify conservation units within species and to aid in the

¹Royal (Dick) School of Veterinary Studies and the Roslin Institute, University of Edinburgh, Easter Bush Campus, Edinburgh EH25 9RG, UK. ²Department of National Parks, Wildlife and Plant Conservation, Bangkok, Thailand. ³School of Life and Environmental Sciences, University of Sydney, Sydney, NSW 2050, Australia. ⁴TRACE Wildlife Forensics Network, Edinburgh EH12 6LE, UK. ⁵Department of Wildlife and National Parks (DWNP/PERHILITAN), National Wildlife Forensic Laboratory (NWFL), Kuala Lumpur, Malaysia. ✉email: rob.ogden@ed.ac.uk

forensic identification of wildlife evidence. For pangolins, it has been shown that mitochondrial DNA genes can be used for species assignment of scales or degraded samples^{11,20–26}. The ability to identify the geographic origin of seizures is more limited although seizures of white-bellied pangolin (*Phataginus tricuspis*) have been identified to one of six geographic lineages throughout the species range via mtDNA sequencing^{24,25}.

The Sunda pangolin is one of the most studied pangolin species¹⁶. A recent mtDNA genome phylogenetic study of the Sunda pangolin in Malaysia showed a clear distinction between samples from northern Borneo and those from west/south Borneo, Peninsular Malaysia and mainland Southeast Asia population²⁷ and consistent with the nuclear genome study that distinguished this species into two populations with five subpopulations²⁸, demonstrating the potential for regional geographic assignment. However, although the species distribution is arguably centered around Thailand, there is limited information on phylogeographic structure in this region. This is required both to inform conservation management and enable traceability of pangolin seizures for law enforcement in Thailand and the wider region. This study aims to characterize the phylogeographic structure and biogeography of the Sunda pangolin in Thailand to enhance our knowledge of the species biology, and to support tools for conservation genetic management and wildlife DNA forensics.

Materials and methods

Ethics statement

This study, in compliance with the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines for animal research and ethical approval to collect the samples was received from the University of Edinburgh Animal Welfare and Ethical Review Board (AWERB) (No. OS2-22). Permission to collect samples from Sunda pangolin, both dead and alive, from inside and outside-protected areas was granted under the Wildlife Protection and Preservation Act of 2562 B.E. from the Department of National Parks Wildlife and Plant Conservation, Thailand (Permit No.0909.204/9311), and the procedure in this study involve with animals were carried out in accordance to the Wildlife Protection and Preservation Act of 2562 B.E. and relevant regulations. The Sunda pangolin samples (blood and tissues) in this project were imported to the University of Edinburgh, UK, from Thailand under CITES export permit No. 22TH0902.2/521 and 23TH0902.2/2 (Thailand) and the CITES import permit No. 616968/01-03 and 624109/01 (UK).

Collection and DNA extraction

Sunda pangolin blood ($n=25$) and tissue ($n=6$) samples were collected opportunistically across Thailand (Sup. Figure S1); the samples were either from rescued or confiscated pangolins of known geographic origin. The geographical origin of confiscated specimens was based on enforcement records, as the animals were seized directly from local hunters who were apprehended within specific protected areas, and who confirmed the collection locations during their arrest. Given the nature of these enforcement operations, it is extremely unlikely that any of these samples derived from long-distance trafficking operations. The blood samples were collected in EDTA tubes and the tissue samples were collected in plastic tubes and stored at -20°C before analysis. DNA was extracted using the PureLink™ Genomic DNA Mini Kit (Cat. No. K182001) following the manufacturer's protocol. Existing Sunda pangolin sequence data were downloaded from NCBI GenBank (Sup. Table S1).

Genome sequencing and assembly

The 31 DNA samples were subject to whole genome sequencing using the Illumina NovaSeq platform (provided by Azenta Life Sciences). Paired-end 150 bp reads were generated to target an average sequencing depth of coverage of 10X across the genome. All paired-end reads were trimmed to remove low quality bases and adapter sequences using TrimGalore 0.6.6, with trimming parameters set to remove bases with Phred quality scores <30 and to discard reads shorter than 35 bp after trimming. The trimmed reads were then mapped against a reference mitogenome from NCBI (accession number NC026781)²⁹ with the BWA-MEM algorithm 0.7.17 and converted from SAM to BAM files with Samtools 1.9. BAM files were visualised and the mitogenome extracted in Geneious Prime 2022.2.2. Paired-end read data from NCBI (SRR9018664-5, SRR9018633, SRR25256520 and SRR25256582) were processed using the same method.

Data analysis

In addition to the 31 mitogenomes generated in this study, a further 27 sequences of Sunda pangolin were obtained from the NCBI database with known origins from Peninsular Malaysia ($n=14$), Borneo ($n=8$)²⁷, China (SRR9018664-5²⁸, MG196309³⁰ $n=3$), Myanmar (SRR9018633²⁸, $n=1$) and the only previous sequence from the south of Thailand (MG196302³⁰, $n=1$). One Chinese pangolin (*Manis pentadactyla*) sequence (MG196305)³⁰ was used as the outgroup. Additionally, two haplotypes from the recently proposed new Southeast Asian pangolin species, *Manis mystera*¹⁷, were included to assess their evolutionary relationship with the Thai Sunda pangolin samples. These 61 mitogenomes were aligned using MAFFT v7.490³¹ following the L-INS-i method and ambiguously aligned positions removed with Gblocks v0.91b^{32,33} with less stringent selection (allowing smaller final blocks, gap positions and fewer strict flanking positions).

We used PartitionFinder2³⁴ in PhyloSuite 1.2.3^{35,36} to determine the best-fit partitioning scheme and substitution model. We defined the 23 transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), d-loop and the 13 protein-coding genes as initial data blocks, and tested partitioning schemes based on the 'greedy' algorithm with branch lengths estimated as 'linked' with 'all' models to search for the optimal schemes. The corrected Akaike Information Criterion (AICc) was used to select the most optimal scheme. While other model selection criteria such as BIC and AIC are available, we selected AICc as it is particularly suitable for smaller sample sizes and provides a good balance between model fit and complexity³⁷.

Phylogenetic reconstruction was performed using both a Maximum Likelihood approach in IQ-TREE v1.6.12 (1000 bootstraps)³⁸, and a Bayesian approach performed with MrBayes v3.2.7a³⁹ (2 parallel runs, 2

M generations, 25% burn-in). Convergence was checked using Tracer v1.7⁴⁰ with two independent runs. The Effective Sample Sizes (ESS) values of 563.5 and 398.2 were well above the recommended threshold of 200⁴¹. Both methods used the best-fitting models for 14 partitions selected from PartitionFinder2. The resulting trees were visualised using FigTree v1.4.4⁴². Haplotype designations and networks were generated in PopART v1.7⁴³ using the Templeton-Crandall-Sing (TCS) algorithm using the sequence alignment after removing the hypervariable mtDNA control region and haplotype statistics using DnaSP v.6⁴⁴.

To examine geographic structuring within the data, we notionally divided a subset of the study samples into five groups according to sampling locality. Two groups were based on established forest complexes in Thailand: western forest complex (W. Forest, $n=5$), Khao Yai forest complex (KY forest, $n=11$), and the remaining three groups were defined by geographic regions: mid-southern Thailand (Mid-south, $n=7$), northern Borneo (N. Borneo, $n=7$) and the Thailand-Malaysia border region (TH-MY border, $n=20$) comprising samples from far south Thailand and peninsula Malaysia (Sup. Table S1). Occasional samples outside these areas were excluded from this analysis. Analysis of molecular variance (AMOVA) and F_{ST} statistics were calculated in Arlequin 3.5.2.2⁴⁵, using the same sequences as the haplotype network and phylogenetic tree analysis (2,000 AMOVA permutations; 200 pairwise F_{ST} permutations, $\alpha=0.05$).

Results

Phylogenetic reconstruction

A total of 52 novel Sunda pangolin mitogenome haplotypes were observed from the 31 samples sequenced in this study and the 27 retrieved from GenBank. A haplotype nomenclature system for Sunda pangolin mitogenomes including previously published sequences has been proposed (Sup. Table S1). Partition finder found highest support for fourteen blocks with differing substitution models (Sup. Table S4). The maximum likelihood and Bayesian trees show consistent clustering, with a principal division between northern Borneo and all other regions; primarily continental Southeast Asia (Peninsular Malaysia, Thailand, China, and Myanmar), but also a single sample from western Borneo ('PangSrwk'; Fig. 1). Note, only one mitogenome was available

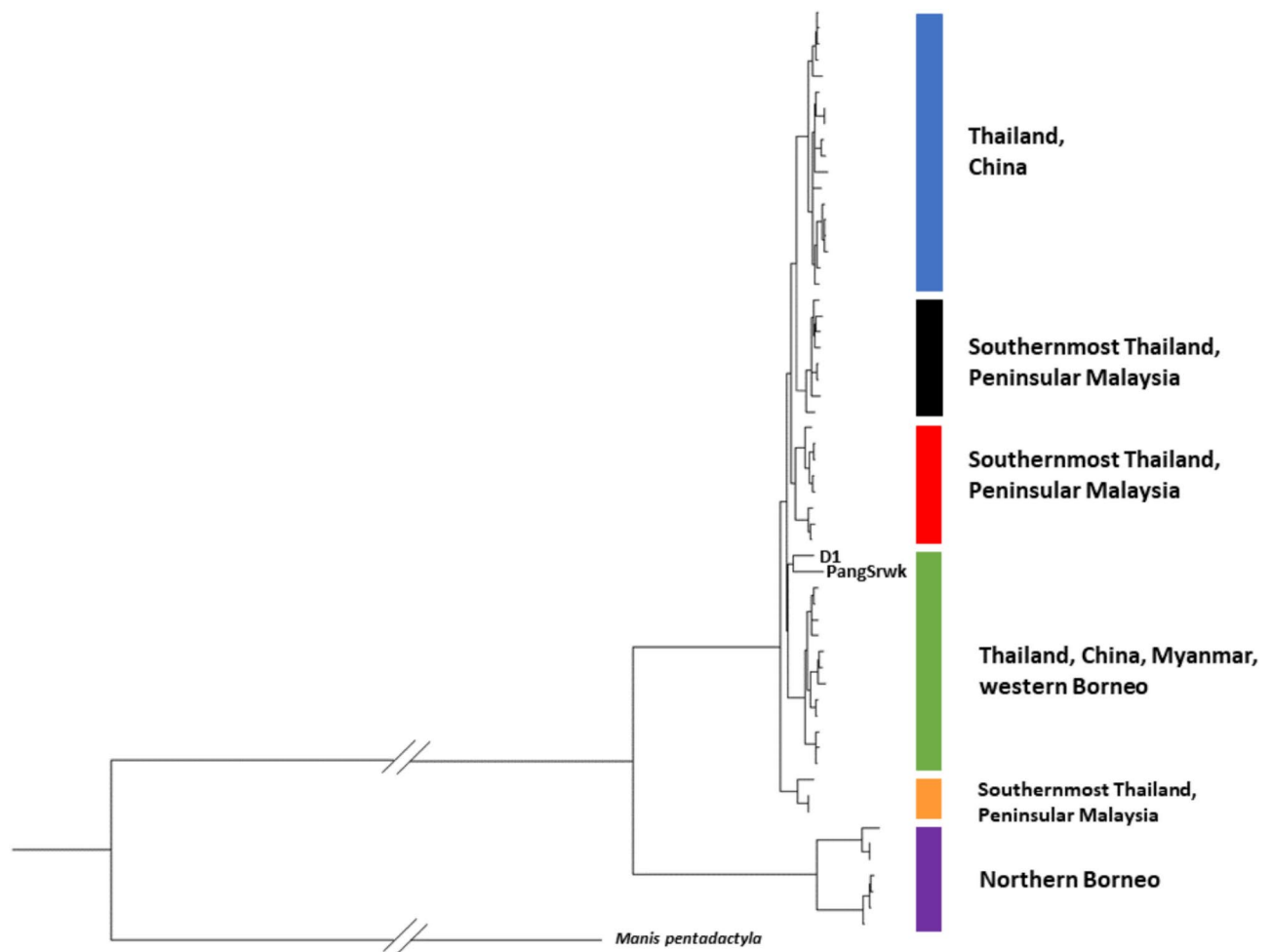


Fig. 1. Bayesian phylogenetic reconstruction of the Sunda pangolin based on mitogenome sequences (16,000 bp). Color-coding refers to clades with Bayesian posterior probabilities ≥ 0.99 . D1 sample was from the Southernmost Thailand and PangSrwk was from western Borneo.

from the western/southern Borneo lineage characterised in Sitam et al., (2023) (i.e., 'PangSrwk'); the other two Kalimantan samples analysed in this study have only partial mtDNA gene regions available. Within continental Southeast Asia, five distinct clades with high support (≥ 0.99 Bayesian posterior probability) can be identified (Blue, Black, Red, Green and Orange; Fig. 1).

The Peninsular Malaysia and southernmost Thailand samples grouped predominantly into three clades (black, red and orange; Fig. 1), with overlapping sympatric genetic distributions. Interestingly, these Thai-Malay border clades did not form a phylogeographically cohesive group with respect to the blue and green clades observed further north in Thailand, China, and Myanmar, which exhibited a much wider geographic distribution (> 1000 km north-south; Fig. 3). One sample from southernmost Thailand (D1) formed an identifiable cluster with one from western Borneo (PangSrwk), with relatively longer branch lengths (Fig. 1), but with bootstrap support in the maximum likelihood tree (45%) and the Bayesian posterior probabilities (0.5) were both low (Sup. Figure S2 and S3), resulting in them being indistinguishable from the green clade under both phylogenetic analyses.

The two mitogenomes from the recently proposed new Asian pangolin species, *Manis mysteria*¹⁷, form a clade this is sister to all of the known Sunda pangolins included in this study (Sup. Figure S5). This *Manis mysteria* clade is genetically more similar to the Sunda pangolin than to the Chinese pangolin, *Manis pentadactyla*.

Haplotype network

The haplotype network suggests seven genetic clusters, with the two samples from western Borneo and southernmost Thailand more clearly distinguished than under phylogenetic analysis (Fig. 2, yellow circle). The remaining six clusters correspond to the clades observed in the phylogenetic trees, although the network reveals considerable haplotype diversity within many of these, notably the red and orange clusters from the Thai-Malay border and purple cluster from northern Borneo.

In terms of geographic distribution, pangolin samples from the mid-south area and Khao Yai forest complex were grouped together mostly with one sample from central area and one sample from Guanxi, China (Blue clade). Haplotypes from the western forest complex group together along with one from mid-south of Thailand and other from southern China and northern Myanmar (green clade) (Figs. 2 and 3). All haplotypes observed in Peninsular Malaysia and southernmost Thailand were placed within the red, black, and orange clusters, but these showed significant differentiation from one another in terms of mutation steps.

Analysis of haplotype diversity revealed high genetic variation across most groups. When analysed by phylogenetic clades, six of the seven groups showed high haplotype diversity ($H_d > 0.95$), only the orange clade showed relatively lower diversity ($H_d = 0.67$) (Sup. Table S2). Geographic analysis showed similarly high diversity across regions, with the Thai-Malaysia border region containing 17 haplotypes among 20 samples ($H_d = 0.98$), and maximum diversity ($H_d = 1.0$) in both mid-south and western regions (Sup. Table S3).

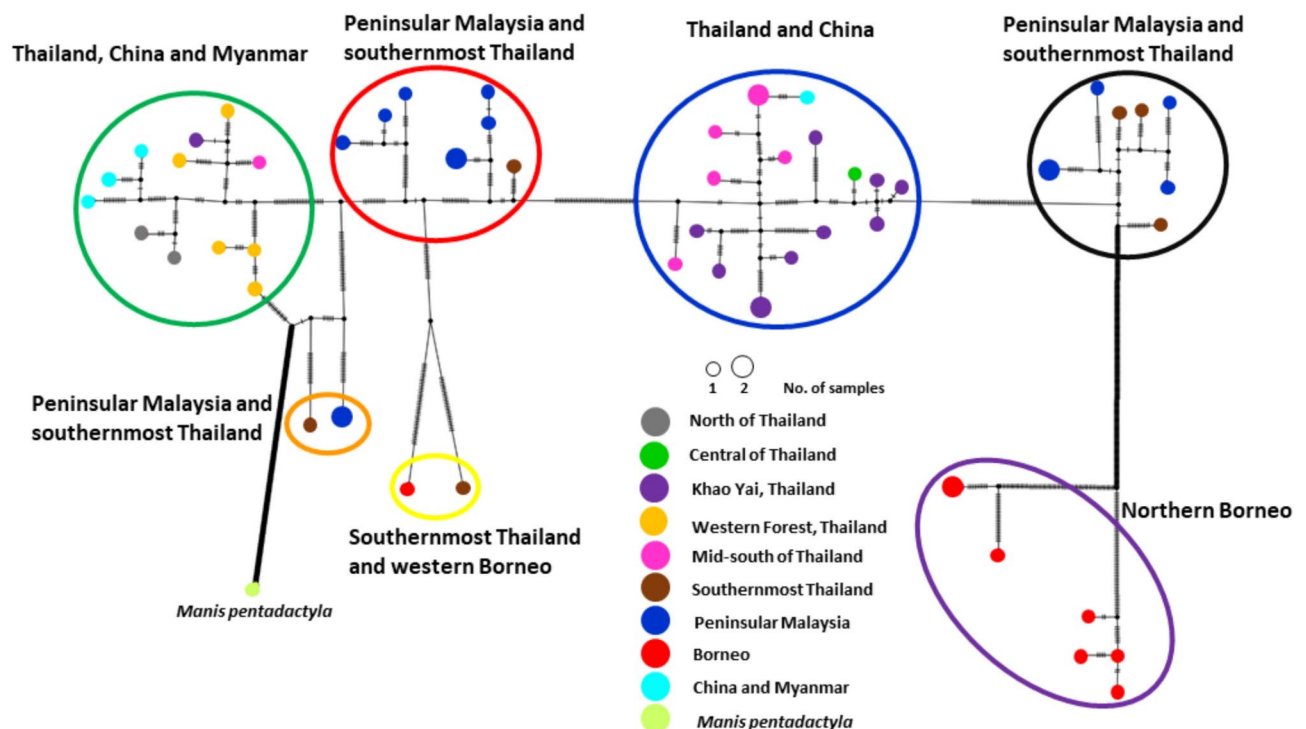


Fig. 2. Haplotype network based on mitogenome (15,408 bp) of the samples from Thailand, Peninsular Malaysia, Borneo, China, and Myanmar. The diameter of haplotype circle reflects the total number of samples carrying this haplotype, the small black circle represents the missing haplotype and hatch-marks across lines represent mutational steps between haplotypes. Ellipse colours match those used in the phylogenetic tree (Fig. 1).

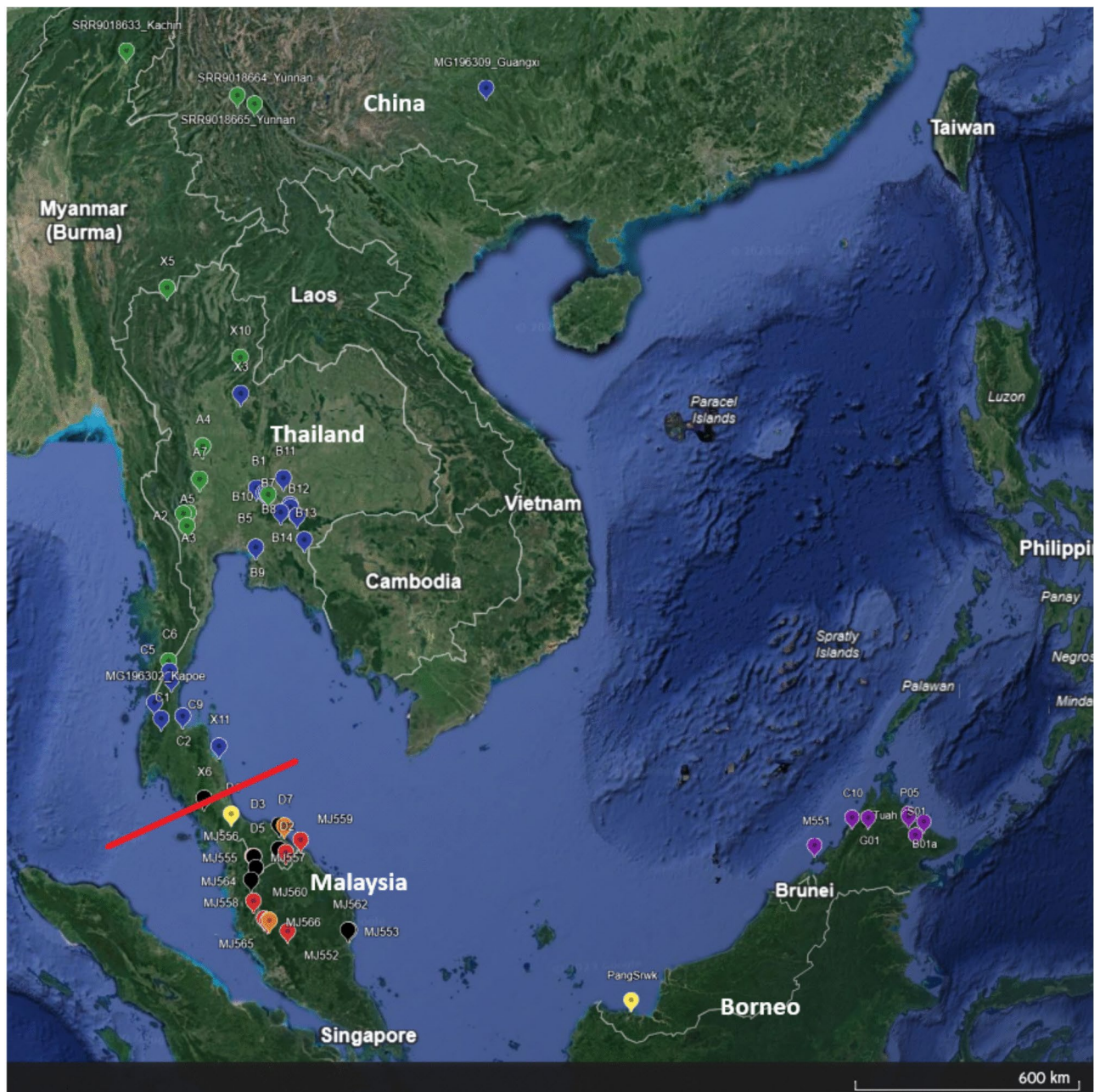


Fig. 3. Distribution map of Sunda pangolin samples/sequence data used in this study, from Thailand, Peninsular Malaysia, Bornean Malaysia, China, and Myanmar. Colours represent clusters according to the haplotype network and the phylogenetic tree, and the red line represents the Kangar-Pattani Line (KPL). The map created using Google Earth, v. 10.69.0.1 (<https://earth.google.com/>).

Geographic structure

AMOVA analysis revealed significant genetic variation distributed among the five pangolin sampling localities (76% of variance, P -value < 0.01, Table 1), while the proportion of variance within populations was relatively low (23%). The pairwise F_{ST} matrix indicated significant differentiation of mtDNA haplotypes between all localities except one: the mid-south area and Khao Yai forest complex ($F_{ST} = 0.13$). The highest continental differentiation occurred between western forest complex and mid-south area ($F_{ST} = 0.63$), followed by the western forest complex and Khao Yai forest complex ($F_{ST} = 0.61$) (Fig. 4), supporting a historic east–west split within Thailand, visible in Fig. 3. One striking result, given the lack of shared haplotypes across the KPL, was the relatively low F_{ST} values for haplotype differentiation between the Thailand–Malaysia border region, and the western forest complex, Khao Yai forest complex and mid-South Thailand regions (pairwise $F_{ST} = 0.30$ to 0.43). Consistent with previous research, the northern Borneo population exhibited the greatest pairwise F_{ST} values, ranging from 0.87 to 0.89.

Source of variation	Sum of squares	Variance components	Percentage variation	P-value
Among populations	5696.83	75.50	76.72	P < 0.01
Within populations	2177.01	22.91	23.28	
Total	7873.84	98.41		

Table 1. Analysis of molecular variance (AMOVA) results with the six populations according to the phylogenetic tree and the haplotype network result.

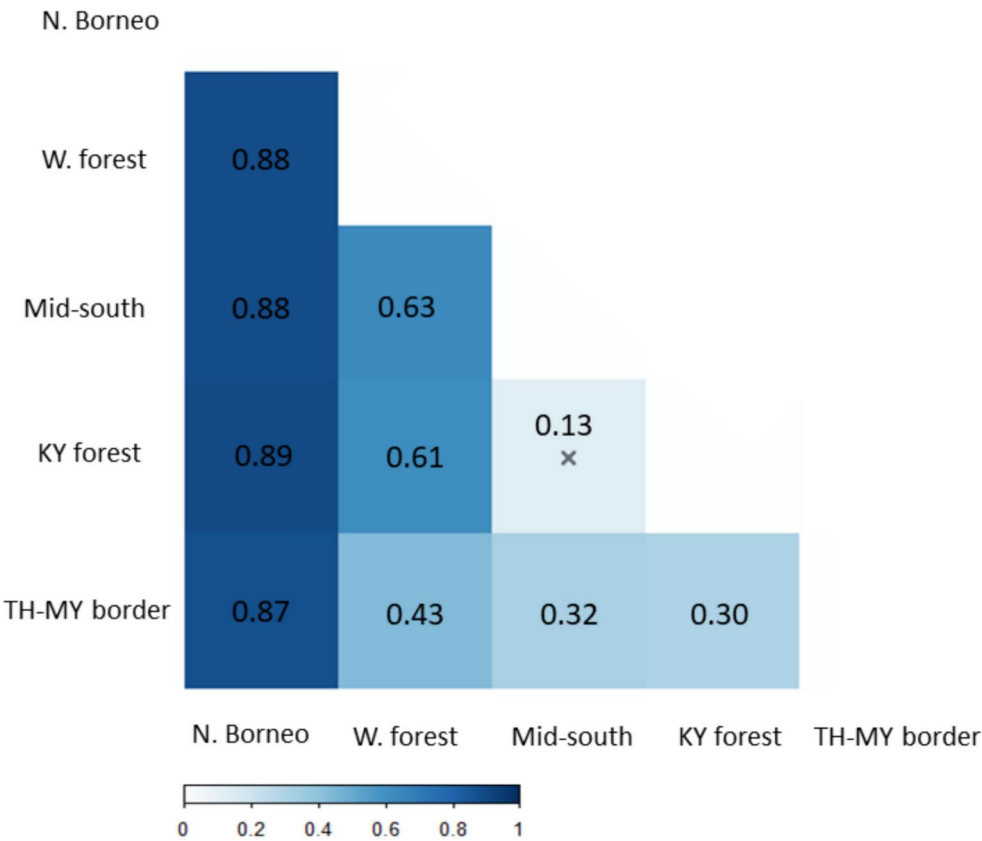


Fig. 4. Genetic distance matrix based on F_{ST} among five geographic areas selected according to the sampling locations within discrete existing pangolin forest habitats. Pairwise F_{ST} values values are significant ($P < 0.05$), except the box marked 'x'.

Discussion

Our results provide the first detailed assessment of evolutionary genetic diversity in the Sunda pangolin throughout Thailand, and reveal multiple previously unknown mitochondrial DNA lineages. The most prominent pattern in our data was the clear phylogeographic structuring within continental Southeast Asia, where we observed five distinct clades with notable geographic distributions. Three clades were concentrated in the southern Thai-Malay border region (red, black and orange) and two (green and blue) were distributed further north in Thailand, stretching up to Myanmar and China, respectively. This north–south separation corresponds to the Kangar-Pattani Line (KPL), the biogeographical barrier which divides the Indochinese and Sundaic bioregions^{46,47}. This division is thought to be driven by a climate-mediated shift in forest habitat type^{48,49}. This result may explain the findings of a previous study²⁸ that relied primarily on seized Sunda pangolins, but which postulated the separation of the species into two main continental groups: Yunnan / Myanmar and Malaysia, based on nuclear DNA data. The KPL has typically been observed to separate flora, in relation to climatic variation in Malaysia at 6°N to 7°N⁴⁶, while the Isthmus of Kra (11°N to 13°N) is better known as a zoogeographic boundary⁴⁷. However, our results indicate separation around the KPL, similar to other mammalian species such as the grey-bellied (*Callosciurus caniceps*) and Asian red-cheeked (*Dremomys rufigenis*) squirrels⁵⁰, and invertebrates such as the fruit fly (*Zeugodacus cucurbitae*)⁵¹. The Sunda pangolin phylogeographic patterns we have inferred suggest a complex evolutionary history potentially involving multiple migration events across the KPL.

North of the KPL, the geographic distributions of the two main clades show increasing divergence, with an east–west split emerging and becoming more defined towards northern Indo-China, although further sampling in Myanmar, Lao PDR and Vietnam would be required to verify this pattern. The distribution of the

western (green) clade, predominant in Thailand's western forest complex and other western sampling localities in Thailand and Myanmar, relative to the eastern (blue) clade found in the Khao Yai forest complex, mid-south Thailand and eastern China, suggests a historical parapatric distribution of Sunda pangolin lineages. The haplotype network analysis revealed further structuring within these two clades, with individual haplotypes grouping largely according to geographic origin, reinforcing the phylogeographic distribution patterns of Sunda pangolin in Thailand.

Within central Thailand, genetic divergence of pangolin lineages between the western forest complex and the Khao Yai forest complex is analogous to the speciation pattern observed within and between other species. For example, Tickell's brown hornbill (*Anorrhinus tickelli*) and Austen's brown hornbill (*Anorrhinus austeni*), originally classified as conspecific subspecies, show a similar east–west partitioning, with *A. tickelli* occupying the western forest complex, while *A. austeni* is found in the Khao Yai forest complex⁵². Hybridization between white-handed (*Hylobates lar*) and pileated (*Hylobates pileatus*) gibbons in a contact zone in Khao Yai National Park reflects another east–west phylogenetic split among closely related taxa in this region⁵³, indicating the possible presence of an east–west biogeographic division in Thailand, which could be explained by the Chao Phraya River Basin acting as a natural barrier between populations in some species, such as the gibbons and squirrel^{54,55}.

In contrast, Sunda pangolin clades to the south of the KPL (red, orange, and black) were not observed to exhibit geographic structuring and appear to be distributed sympatrically around the Thai–Malay border. Within each clade, there was no apparent association between haplotype and sampling locality, while the positioning of samples within the orange and yellow clusters in the network emphasized the degree of mitochondrial genetic divergence in this region. The results of pairwise F_{ST} analysis for the Thai–Malay border region were curious, suggesting closer genetic affinity to both the western and Khao Yai forest complexes than these forest complexes had for each other, despite sharing some haplotypes. We interpret this as being due to the breadth of haplotype variation within the Thai–Malay border region (red, orange, black clusters) and the shorter network distances between these clusters and each of the western (green) and Khao Yai (blue) forest complex clusters, compared to the genetic divergence between the forest complexes. This pattern is particularly notable given that Sunda pangolin has relatively limited capacity for large-scale migration^{56,57}.

The results also provide further evidence of the division between the two phylogenetic groups lineages previously recorded on Borneo²⁷, whereby the pangolins in west/south Borneo appear to be associated with a mitogenome clade that includes pangolins from Peninsular Malaysia, Thailand, Myanmar, and China, rather than the neighbouring northern Borneo clade. In this study, the finding of a mitogenome haplotype in southern Thailand that is much more closely related to the west/south Borneo haplotype (with “PangSrwk”), compared to haplotypes previously observed, significantly strengthens the phylogenetic link between these geographic regions and reinforces the deep phylogenetic division within Borneo²⁷. This pattern aligns with mitogenome phylogenies of shrews⁵⁸, colugo⁵⁹ and sun bears⁶⁰, which also show shared haplotypes between western Borneo and the Southeast Asian continent.

Overall, the levels of genetic differentiation observed among the major Sunda pangolin mitogenome clades was generally lower than intraspecific divergences characterized in the Chinese pangolin²⁸. Based on 782 bp of CYTB (the region utilised by²⁸), the largest divergence within the Sunda pangolin, between the northern Borneo and other clades, were separated by 17 substitutions, while the two major Chinese pangolin clades were separated by 37 substitutions (Sup. Figure S6). Intraspecific diversity in the other two Asian pangolin species is not well-established.

Conservation genetic management

The strong regional genetic differentiation and distribution of lineage diversity observed in this species has important implications for conservation management. The geographic patterns of mitochondrial DNA divergence on the continent, either side of the Kangar–Pattani Line, while strong, do not warrant designation as ESUs due to the lack of reciprocal monophyly. However, the high diversity of Sunda pangolin mtDNA haplotype clades observed in this relatively small area straddling northern peninsular Malaysia and southern Thailand (Sup. Figures S4 and Sup. Table S3), highlighted by the network analysis, identifies this region as an evolutionary genetic hotspot for the species. In terms of current conservation management, Malaysia–Thailand border area should be prioritised to preserve the species' genetic diversity and evolutionary potential. Despite existing transboundary protected areas, increasing habitat fragmentation threatens these unique populations, highlighting the urgent need for enhanced connectivity between protected areas in both countries.

Within Thailand, the western forest complex and Khao Yai forest complex may also warrant separate Management Unit status, based on their high genetic differentiation ($F_{ST}=0.61$) (but not entirely distinct) and the fact that these two habitats are separated by a large waterway (Chao Phraya River) and more recently by a large urban development north of Bangkok. As pangolins are known to swim and survive in peri-urban environments, neither factor necessarily comprises an absolute barrier to gene flow, and further research to characterise nuclear genetic differentiation is required; however, it seems likely that considering separate Sunda pangolin management units between eastern and western Thailand is appropriate.

The primary phylogeographic split separating the northern Borneo population from Sundaland and Indochina has been previously observed²⁷, albeit with a slightly smaller dataset, and the current findings reinforce this distinction that supports the recognition of the north Borneo pangolin as a separate Evolutionary Significant Unit (ESU)^{61,62}. Further sampling, analyses of nuclear DNA markers and evaluation of morphological traits in the northern Borneo are required to determine whether or not the northern Borneo pangolin population merits further taxonomic elevation, as has been suggested²⁷.

Implications for traceability and repatriation

Beyond informing regional in situ conservation management of the Sunda pangolin and its habitat, this study can support efforts to trace the geographic origin of seized pangolins or pangolin scales, as well as helping to reconstruct routes of illegal wildlife trade^{63,64}. The discrete distribution of many haplotypes observed in this study may provide clues as to the geographic origin of the trafficked pangolins for forensic intelligence⁶⁵. While this approach is disrupted by human-mediated animal translocation and limited by the broad geographic scale of certain haplotypes, mitochondrial DNA data has the advantage of being readily generated in most wildlife DNA forensic laboratories and readily shared among them, enabling traceability, for example in African elephants⁶⁶ and lions⁶⁷. Our study adds to several existing mitochondrial DNA studies on the geographic provenance of pangolins^{25,27,28,68,69}, providing a broad scale geographic provenance information across an increasingly comprehensive range of species and distributions. For the Sunda pangolin, additional samples from Indonesia, in particular Sumatra and Java, identified as a major origin of pangolin trafficking between 2010–2015⁷⁰, alongside Borneo, Vietnam, Lao PDR, and Cambodia, would enable a more complete characterization of genetic variation to identify the geographic origins across the species' range. Given the relatively deep intra-specific phylogenetic divisions observed in this study, traceability is also important to ensure rescued or confiscated pangolins can be returned to appropriate locations that match their genetic provenance, to optimise their chance of survival and to maintain genetically cohesive populations and evolutionary processes in their native locations^{71,72}.

Conclusion

This study has provided the first conservation genetic management data on the Sunda pangolin in Thailand, revealing mitochondrial DNA diversity hotspots and phylogeographic structure that can be used to aid in species and habitat management and trade traceability. Further genetic studies that expand geo-referenced sample coverage and employ nuclear genomic markers will enhance our understanding of fine-scale population genetic processes.

Data availability

The sequences produced in this work can be found in the GenBank database with the accession numbers PP266608–PP266638.

Received: 8 December 2023; Accepted: 1 April 2025

Published online: 23 April 2025

References

- Challender, D., Waterman, C., Baillie, J. Scaling up pangolin conservation. IUCN SSC pangolin specialist group conservation action plan. Zoological Society of London, London, UK (2014).
- Nuwer, R.L. Poached: inside the dark world of wildlife trafficking: Hachette UK (2018).
- Challender, D. W. S., Heinrich, S., Shepherd, C. R. & Katsis, L. K. D. Chapter 16 - International trade and trafficking in pangolins, 1900–2019. In *pangolins* (eds Challender, D. W. S. et al.) 259–276 (Academic Press, 2020).
- IUCN. The IUCN Red List of Threatened Species 2021 [Available from: <https://www.iucnredlist.org>].
- Challender, D. W. S. & O'Cruidain, C. Chapter 19 - Addressing trade threats to pangolins in the convention on international trade in endangered species of wild fauna and flora (CITES). In *Pangolins* (eds Nash, H. C. et al.) 305–320 (Academic Press, 2020).
- Challender, D., Willcox, D., Panjang, E., Lim, N., Nash, H., Heinrich, S., Chong, J. *Manis javanica*. The IUCN red list of threatened species 2019: e. T12763A123584856. (2019).
- Chong, J. L. et al. Chapter 6 - Sunda pangolin *Manis javanica* (Desmarest, 1822). In *Pangolins* (eds Challender, D. W. S. et al.) 89–108 (Academic Press, 2020).
- Wong, R.W.Y. *The Illegal Wildlife Trade in China Understanding The Distribution Networks* / by Rebecca W. Y. Wong. 1st ed. 2019. ed. Cham: Springer International Publishing, (2019).
- Heinrich, S., Wittman, T., Ross, J., Shepherd, C., Challender, D. Cassey, P. The global trafficking of pangolins: A comprehensive summary of seizures and trafficking routes from 2010–2015. Selangor, Malaysia; 2017. Report No.: 978-983-3393-76-3.
- Challender, D. W. & Hywood, L. African pangolins under increased pressure from poaching and intercontinental trade. *Traffic Bull.* **24**(2), 53–55 (2012).
- Zhang, H. et al. Molecular tracing of confiscated pangolin scales for conservation and illegal trade monitoring in Southeast Asia. *Global Ecol. Conserv.* **4**, 414–422 (2015).
- Heinrich, S. et al. Where did all the pangolins go? International CITES trade in pangolin species. *Global Ecol. Conserv.* **8**, 241–253 (2016).
- Shepherd, C. R., Connelly, E., Hywood, L. & Cassey, P. Taking a stand against illegal wildlife trade: the Zimbabwean approach to pangolin conservation. *Oryx* **51**(2), 280–285 (2017).
- Vigne, L. & Nijman, V. Elephant ivory, rhino horn, pangolin and helmeted hornbill products for sale at the Myanmar–Thailand–China border. *Environ. Conserv.* **49**(3), 187–194 (2022).
- UNODC. World Wildlife Crime Report 2020: Trafficking in Protected Species. United Nations Office on Drugs and Crime (2020).
- Heighton, S. P. & Gaubert, P. A timely systematic review on pangolin research, commercialization, and popularization to identify knowledge gaps and produce conservation guidelines. *Biol. Cons.* **256**, 109042 (2021).
- Gu, T.-T. et al. Genomic analysis reveals a cryptic pangolin species. *Proc. Natl. Acad. Sci.* **120**(40), e2304096120 (2023).
- Frankham, R., Ballou, J. D. & Briscoe, D. A. *Introduction to conservation genetics* 2nd edn. (Cambridge University Press, 2010).
- Lee, J. C. I., Hsieh, H. M. & Tsai, L. C. DNA and endangered species. In *Encyclopedia of forensic sciences* 2nd edn (eds Siegel, J. A. et al.) 409–412 (Academic Press, 2013).
- Mwale, M. et al. Forensic application of DNA barcoding for identification of illegally traded African pangolin scales. *Genome* **60**(3), 272–284 (2017).
- Kumar, V. P., Rajpoot, A., Shukla, M., Nigam, P. & Goyal, S. P. Inferring the molecular affinity of Indian pangolin with extant Manidae species based on mitochondrial genes: A wildlife forensic perspective. *Mitochondrial DNA Part B Res.* **3**(2), 640–644 (2018).
- Singh, A. et al. Pangolin Indexing System: Implications in forensic surveillance of large seizures. *Int. J. Legal Med.* **134**(5), 1613–1618 (2020).
- Priyambada, P. et al. Digging out the keys in the heap of seized pangolin scales: up scaling pangolin conservation using wildlife forensics. *Forensic Sci. Int.* **323**, 110780 (2021).

24. Gaubert, P. et al. Phylogeography of the heavily poached African common pangolin (*Pholidota*, *Manis tricuspis*) reveals six cryptic lineages as traceable signatures of Pleistocene diversification. *Mol. Ecol.* **25**(23), 5975–5993 (2016).
25. Ewart, K. M. et al. DNA analyses of large pangolin scale seizures: Species identification validation and case studies. *Forensic Sci. Int. Animals Environ.* **1**, 100014 (2021).
26. Zhang, H. et al. Genetic identification of African pangolins and their origin in illegal trade. *Global Ecol. Conserv.* **23**, e01119 (2020).
27. Sitam, F. T. et al. Phylogeography of the Sunda pangolin, *Manis javanica*: Implications for taxonomy, conservation management and wildlife forensics. *Ecol. Evol.* **13**(8), e10373 (2023).
28. Hu, J.-Y. et al. Genomic consequences of population decline in critically endangered pangolins and their demographic histories. *Natl. Sci. Rev.* **7**(4), 798–814 (2020).
29. Hassanin, A., Hugot, J.-P. & van Vuuren, B. J. Comparison of mitochondrial genome sequences of pangolins (Mammalia, Pholidota). *C.R. Biol.* **338**(4), 260–265 (2015).
30. Gaubert, P. et al. The complete phylogeny of pangolins: Scaling up resources for the molecular tracing of the most trafficked mammals on earth. *J. Hered.* **109**(4), 347–359 (2018).
31. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**(4), 772–780 (2013).
32. Dereeper, A. et al. Phylogeny. fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–W469 (2008).
33. Dereeper, A., Audic, S., Claverie, J.-M. & Blanc, G. BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol. Biol.* **10**(1), 1–6 (2010).
34. Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T. & Calcott, B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **34**(3), 772–773 (2017).
35. Zhang, D. et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **20**(1), 348–355 (2020).
36. Xiang, C. Y. et al. Using PhyloSuite for molecular phylogeny and tree-based analyses. *iMeta.* **2**(1), e87 (2023).
37. Hurvich, C. M. & Tsai, C.-L. A corrected akaike information criterion for vector autoregressive model selection. *J. Time Ser. Anal.* **14**(3), 271–279 (1993).
38. Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**(1), 268–274 (2015).
39. Ronquist, F. et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**(3), 539–542 (2012).
40. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**(5), 901–904 (2018).
41. Lanfear, R., Hua, X. & Warren, D. L. Estimating the effective sample size of tree topologies from bayesian phylogenetic analyses. *Genome Biol. Evol.* **8**(8), 2319–2332 (2016).
42. Rambaut, A. FigTree v1.4.4, a Graphical Viewer of Phylogenetic Trees (2018).
43. Leigh, J. W. & Bryant, D. POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**(9), 1110–1116 (2015).
44. Rozas, J. et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**(12), 3299–3302 (2017).
45. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**(3), 564–567 (2010).
46. Woodruff, D. S. Biogeography and conservation in Southeast Asia: How 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodiv. Conserv.* **19**, 919–941 (2010).
47. Hughes, A. C., Satasook, C., Bates, P. J., Bumrungsri, S. & Jones, G. Explaining the causes of the zoogeographic transition around the Isthmus of Kra: Using bats as a case study. *J. Biogeogr.* **38**(12), 2362–2372 (2011).
48. Baltzer, J. L., Davies, S. J., Bunyavechewin, S. & Noor, N. S. M. The role of desiccation tolerance in determining tree species distributions along the Malay–Thai Peninsula. *Funct. Ecol.* **22**(2), 221–231 (2008).
49. Phillips, O. L. Drought, dispersal, and distribution in the inner tropics. *J. Biogeogr.* **34**(11), 1846–1847 (2007).
50. Hinckley, A., Hawkins, M. T., Maldonado, J. E. & Leonard, J. A. Evolutionary history and patterns of divergence in three tropical east Asian squirrels across the Isthmus of Kra. *J. Biogeogr.* **50**(6), 1090–1102 (2023).
51. Boontop, Y. et al. Population structure in Zeugodacus cucurbitae (Diptera: Tephritidae) across Thailand and the Thai–Malay peninsula: Natural barriers to a great disperser. *Biol. J. Lin. Soc.* **121**(3), 540–555 (2017).
52. Trisurat, Y. et al. An assessment of the distribution and conservation status of hornbill species in Thailand. *Oryx* **47**(3), 441–450 (2013).
53. Markviriyi, D., Asensio, N., Brockelman, W. Y., Jeratthitikul, E. & Kongrit, C. Genetic analysis of hybridization between white-handed (*Hylobates lar*) and pileated (*Hylobates pileatus*) gibbons in a contact zone in Khao Yai National Park, Thailand. *Primates*. 1–13 (2022).
54. Geissmann, T. Sympatry between white-handed gibbons (*Hylobates lar*) and pileated gibbons (*H. pileatus*) in Southeastern Thailand. *Primates* **32**(3), 357–363 (1991).
55. Hinckley, A., Maldonado, J. E., Tamura, N., Leonard, J. A. & Hawkins, M. T. Lost in synonymy: Integrative species delimitation reveals two unrecognized species of Southern Asian tree squirrels (Rodentia: Sciuridae: Callosciurinae). *Vertebrate Zoology*, 74 (2024).
56. Lim, N. T. & Ng, P. K. Home range, activity cycle and natal den usage of a female Sunda pangolin *Manis javanica* (Mammalia: Pholidota) in Singapore. *Endangered Species Res.* **4**(1–2), 233–240 (2008).
57. Gray, R. J., Van Le, D., Thi Thanh Nguyen, H., Nhat Cau, L., Van Nguyen, T. & Van Pham, T. et al. Home ranges and activity patterns of Sunda pangolins *Manis javanica* (Pholidota: Manidae) in Vietnam. *J. Asia-Pacific Biodiver.* (2023).
58. Hinckley, A. et al. Evolutionary history of Sundaland shrews (Eulipotyphla: Soricidae: Crocidura) with a focus on Borneo. *Zool. J. Linn. Soc.* **194**(2), 478–501 (2021).
59. Mason, V. C., Helgen, K. M. & Murphy, W. J. Comparative phylogeography of forest-dependent mammals reveals paleo-forest corridors throughout Sundaland. *J. Hered.* **110**(2), 158–172 (2019).
60. Kunde, M. N. et al. First mitogenome phylogeny of the sun bear *Helarctos malayanus* reveals a deep split between Indochinese and Sundic lineages. *Ecol. Evol.* **13**(4), e9969 (2023).
61. Fraser, D. J. & Bernatchez, L. Adaptive evolutionary conservation: Towards a unified concept for defining conservation units. *Mol. Ecol.* **10**(12), 2741–2752 (2001).
62. Moritz, C. Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas* **130**(3), 217–228 (1999).
63. Phelps, J., Biggs, D. & Webb, E. L. Tools and terms for understanding illegal wildlife trade. *Front. Ecol. Environ.* **14**(9), 479–489 (2016).
64. Nijman, V. An overview of international wildlife trade from Southeast Asia. *Biodivers. Conserv.* **19**(4), 1101–1114 (2010).
65. Ogden, R. & Linacre, A. Wildlife forensic science: A review of genetic geographic origin assignment. *Forensic Sci. Int. Genet.* **18**, 152–159 (2015).
66. Zhao, K. et al. Loxodonta Localizer: A software tool for inferring the provenance of African elephants and their ivory using mitochondrial DNA. *J. Hered.* **110**(7), 761–768 (2019).
67. Au, W. C., Dures, S. G., Ishida, Y., Green, C. E., Zhao, K., Ogden, R. & Roca, A. L. Lion Localizer: a software tool for inferring the provenance of lions (*Panthera leo*) using mitochondrial DNA. *J. Heredity*, esad072 (2023).

68. Nash, H. C. et al. Conservation genomics reveals possible illegal trade routes and admixture across pangolin lineages in Southeast Asia. *Conserv. Genet.* **19**(5), 1083–1095 (2018).
69. Wiradeti, R. T. P. N., Yulianto, & Semiadi, G. Identification of confiscated pangolin for conservation purposes through molecular approach. *J. Threatened Taxa.* **14**(6), 21127–21139 (2022).
70. Gomez, L., Leupen, B., Krishnasamy, K. Heinrich S. Scaly nexus: mapping Indonesian pangolin seizures (2010–2015). TRAFFIC. Southeast Asia Regional Office, Petaling Jaya, Selangor. (2017).
71. Kawecki, T. J. & Ebert, D. Conceptual issues in local adaptation. *Ecol. Lett.* **7**(12), 1225–1241 (2004).
72. Weeks, A. R. et al. Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evol Appl.* **4**(6), 709–725 (2011).

Acknowledgements

The research was funded through a Royal Thai Government Scholarship awarded to NB at the University of Edinburgh. We would like to thank Emily Humble, Melissa Marr and Marc-Alexander Gose for bioinformatic consultation, and Prateep Duengkae from Kasetsart University for ecological advice. We also thank all staff at DNP Thailand who helped collect the samples in this project.

Author contributions

NB performed laboratory work and data analysis. RO conceived the study, supervised lab work and data analysis. NB, RO and KE wrote the manuscript. FS helped interpret the results and revised the manuscript. All authors read and approved the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-97182-1>.

Correspondence and requests for materials should be addressed to R.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025