DNA Research, 2022, 29(5), 1–13
https://doi.org/10.1093/dnares/dsac026
Research Article

OXFORD

Research Article

Genome assembly of the Pendlebury's roundleaf bat, *Hipposideros pendleburyi*, revealed the expansion of *Tc1/Mariner* DNA transposons in Rhinolophoidea

Wanapinun Nawae [1], Chutima Sonthirod[1], Thippawan Yoocha[1], Pitchaporn Waiyamitra[1], Pipat Soisook[2], Sithichoke Tangphatsornruang [1], and Wirulda Pootakham[1]*

[1]National Omics Center (NOC), National Science and Technology Development Agency, Thailand Science Park, Pathum Thani, Thailand, and [2]Princess Maha Chakri Sirindhorn Natural History Museum, Prince of Songkla University, Hat Yai, Thailand

*To whom correspondence should be addressed. Tel: +66 2 5646700 Ext 71445. Fax: +66 2 5646707.
Email: wirulda@alumni.stanford.edu

Received 11 January 2022; Editorial decision 12 July 2022

Abstract

Bats (Chiroptera) constitute the second largest order of mammals and have several distinctive features, such as true self-powered flight and strong immunity. The Pendlebury's roundleaf bat, *Hipposideros pendleburyi*, is endemic to Thailand and listed as a vulnerable species. We employed the 10× Genomics linked-read technology to obtain a genome assembly of *H. pendleburyi*. The assembly size was 2.17 Gb with a scaffold N50 length of 15,398,518 bases. Our phylogenetic analysis placed *H. pendleburyi* within the rhinolophoid clade of the suborder Yinpterochiroptera. A synteny analysis showed that *H. pendleburyi* shared conserved chromosome segments (up to 105 Mb) with *Rhinolophus ferrumequinum* and *Phyllostomus discolor* albeit having different chromosome numbers and belonging different families. We found positive selection signals in genes involved in inflammation, spermatogenesis and Wnt signalling. The analyses of transposable elements suggested the contraction of short interspersed nuclear elements (SINEs) and the accumulation of young *mariner* DNA transposons in the analysed hipposiderids. Distinct *mariners* were likely horizontally transferred to hipposiderid genomes over the evolution of this family. The lineage-specific profiles of SINEs and *mariners* might involve in the evolution of hipposiderids and be associated with the phylogenetic separations of these bats from other bat families.

Key words: Pendlebury's roundleaf bat, threatened animals, genome assembly, 10× linked-reads, transposable elements

1. Introduction

Bats (order Chiroptera) account for ~20% (>1,400 species[1]) of all living mammals and constitute the second largest order of the class Mammalia.[2] Compared to other mammals, bats show several unique characters, such as true self-powered flight, echolocating mechanisms and strong immunity.[2] Their habitats and diets are also diverse,

making them important pollinators and insect controllers in various ecosystems.[3] Bats have been originally classified into two groups, microbats and megabats, based on morphological features, such as body size and echolocating abilities.[4] Molecular data were later used to refine this classification, which led to the reclassification of bats into the suborders Yinpterochiroptera and Yangochiroptera.[4] Yinpterochiroptera comprises all frugivorous and nectarivorous bats in the family Pteropodidae and some echolocating bats in the superfamily Rhinolophoidea.[4] All other echolocating bats are members of Yangochiroptera.[4] This classification, which was supported by other molecular studies,[5,6] was controversial to the classification based on morphological features because the rhinolophoids, similar to the echolocating bats in the suborder Yangochiroptera, had a laryngeal echolocation ability and were insectivorous while pteropodids were frugivorous or nectarivorous and lacked the laryngeal echolocation ability but had a developed vision.[6]

Members of family Hipposideridae (superfamily Rhinolophoidea), also known as the Old World leaf-nosed bats, are widespread in tropical and subtropical areas.[7] *Hipposideros* is the largest genus of this family, with 88 species currently known.[7,8] They are members of laryngeal echolocating bats in the suborder Yinpterochiroptera.[7] The Pendlebury's roundleaf bat (*Hipposideros pendleburyi*) is endemic to Thailand[9] and being classified as vulnerable (facing a high risk of extinction in the wild) under the IUCN red list categories and criteria.[10] It has been estimated that there were only 4,700 of the *H. pendleburyi* bats remaining in Thailand.[11] They could be found in only 14 caves (3–800 individuals in each cave) in 7 provinces of peninsular Thailand.[11] The number of *H. pendleburyi* continues to decline due to habitat disturbances.[11]

Recent advance in the nuclear genome study has provided insights into the diversity and the evolution of Chiroptera members.[5,12,13] Currently, genome data from 45 species in the order Chiroptera were available for download at the NCBI Genomes database (https://www.ncbi.nlm.nih.gov/datasets/genomes/, last accessed: 25 March 2022). Two-thirds of these 45 genomes are from the echolocating bats in the suborder Yangochiroptera. Five genomes of the bats in the suborder Yinpterochiroptera are from the superfamily Rhinolophoidea (Supplementary Table S1) and only two of them are from the family Hipposideridae (*Hipposideros armiger* and *Hipposideros galeritus*). Most of the reported bat genomes have not been scaffolded to a chromosome level yet they provide valuable information to help improve our understanding of bat behaviours and evolution. For example, the analyses of the genomes of *H. armiger* and other bats have shown that the contraction of olfactory receptor genes might be associated with their diets.[12] Transposable element (TE) is another component of the genome that has been linked to the evolution of bats.[14] DNA transposons (class II TE), such as *Tc1/Mariner*, *hATs*, *piggyBac*s and *SPIN*, are active in bats[14–16] while some of these DNA transposons are inactive in other mammals such as humans, mice, rats and dogs.[17] Several pieces of evidence suggested that DNA transposons have been introduced into bat genomes via horizontal transfer.[14,18–20] The ability of bats to tolerate pathogen infection might facilitate the horizontal transfer in bats because viruses and parasitic protozoans were potential vectors for the horizontal introduction of DNA transposons.[14,15] The horizontal transfer of DNA transposons appears to be associated with the rapid diversification of bats in the family Vespertilionidae.[18]

In this study, we employed 10× Genomics linked-read technology to assemble the draft genome of *H. pendleburyi*. The analyses of *H. pendleburyi* genome allowed us to demonstrate unique features in the genome sequences of this bat species and other closely related bats in the family Hipposideridae (hipposiderids) and in the superfamily Rhinolophoidea (rhinolophoids). Our analyses demonstrated how TEs shaped the genome evolution of the analysed hipposiderids and rhinolophoids and provided a better understanding of their genetic diversity. To the best of our knowledge, our study is the first to report the profiles of TEs in hipposiderids and rhinolophoids. In addition, as *H. pendleburyi* is endemic to Thailand and its population continues to decline, the genome data reported here are valuable for future conservation programmes. Our analysis results also lay a foundation for comparative genomic studies in mammals.

2. Materials and methods

2.1. Sample collections and DNA extraction

A specimen of male *H. pendleburyi* was collected from Tham Le Stegodon Cave, Palian District, Trang Province, Thailand (7.141 N, 99.789 E) based on the guidelines for use of wild mammal species in research.[21] The collection was conducted under the permission by the Department of National Park, Wildlife and Plant Conservation (project number 6210306). The specimen was deposited in the Mammal Collection of the Princess Maha Chakri Sirindhorn Natural History Museum, Prince of Songkla University (PSU), Hat Yai, Songkhla, Thailand (https://nhm.psu.ac.th/, last accessed 21 July 2022) under the voucher number PSUZC-MM.2021.6.

To obtain high molecular weight DNA for 10× Genomics linked-read sequencing, DNA was extracted from *H. pendleburyi* blood using the QIAGEN Genomic-tip 100/G following the manufacturer's protocol (Qiagen, Germany). DNA quality was assessed using 0.75% pulsed-field gel electrophoresis and the concentration was tested with Qubit® dsDNA BR Assay Kits (Thermo Fisher Scientific) and NanoDrop (Thermo Fisher Scientific).

2.2. DNA library preparation and sequencing

For library preparation and sequencing, 1 ng of high quality and molecular weight DNA was used for the 10× Genomics linked-read library preparation using the Chromium Genome Library Kit & Gel Bead Kit v2, the Chromium Genome Chip Kit v2 and the Chromium i7 Multiplex Kit according to the manufacturer's instructions (10× Genomics, Pleasanton, USA). The library quality was assessed using Bioanalyzer DNA High Sensitivity DNA Assay (Agilent) and the concentration was tested with Qubit® dsDNA BR Assay Kits (Thermo Fisher Scientific). The 10× Genomics library was sequenced on the Illumina HiSeq X Ten (150 bp paired-end reads).

2.3. Genome assembly

The linked-read data were assembled using the Supernova assembler version 2.1.1 with the default parameter setting (https://support.10xgenomics.com/de-novo-assembly/software/pipelines/latest/using/running, last accessed 21 July 2022; 10× Genomics, Pleasanton, USA). For the quality assessment, short-read DNA sequences data obtained from this study were mapped back the final assembly sequences using minimap2[22] and the percentage of successful mapping was identified. We also employed the Benchmarking Universal Single-Copy Orthologues (BUSCO)[23] version 4.0.5 to evaluate the assembly by testing for the presence and completeness of the orthologues using the Laurasiatheria OrthoDB release 10.[24]

2.4. Repeat modelling and masking

We used RepeatModeler[25] (with default parameters) to construct libraries of the consensus sequences of TEs from the assemblies of *H. pendleburyi*, *H. armiger*, *H. galeritus*, *Rhinolophus ferrumequinum*, *Craseonycteris thonglongyai*, *Megaderma lyra*, *Rousettus aegyptiacus*, *Pteropus vampyrus, Myotis lucifugus*, *Phyllostomus discolor* and *Miniopterus natalensis*. The assemblies of species other than *H. pendleburyi* were retrieved from the NCBI Genomes database (Supplementary Table S1). The numbers of consensus sequences obtained from our RepeatModeler analyses are in Supplementary Table S2. We merged our repeat libraries with the sequence library generated and curated in previous study,[5] which contained the consensus sequences of TEs from several bat genera and the sequences from the RepBase database. The merged repeat library was used with RepeatMasker version 4.1.2[26] (with default parameters) to identify repeats in *H. pendleburyi* and other 10 species. The repeat landscapes of these 10 species were obtained by processing RepeatMasker primary results with calcDivergcenceFromAlign.pl and createRepeatLandscape.pl scripts, which were provided as utility scripts in the RepeatMasker packages.

2.5. Analysis of TEs

To construct the phylogenetic tree of the consensus sequences, we used CD-HIT software[27] with an identity cut-off of 95% to obtain representatives of the consensus sequences of *Tc1/Mariner*. The phylogenetic tree was then built from the representative sequences using automatic FastTree[28] and PhyML+SMS[29] pipelines on the NGPhylogeny.fr webserver.[30] To construct the phylogenetic tree of the full-length *Tc1/Mariner* elements, we performed BLASTN search by using the sequences of full-length *Tc1/Mariner* elements reported in other study[31] as query sequences and *H. pendleburyi* assembly as a subject sequence. The *Tc1/Mariner* sequences were extracted from *H. pendleburyi* assembly based on BLASTN results. Only the BLAST hits with the alignment coverage of more than 90% of the query length were considered. The phylogenetic tree was constructed in the same way as another tree. To identify transposase domains of the *Tc1/Mariner* elements, we performed TBLASTN search by using the amino acid sequences of transposase domains obtained from other study[32] as query sequences and the assemblies of *H. pendleburyi* and other bat species as subject sequences. We used an identity cut-off of 35% in TBLASTN search.[33]

2.6. Genome annotation

The repeat masked sequences of *H. pendleburyi* were annotated for gene models. For the evidence-based gene models, we used MAKER2 software[34] based on the assembled transcript sequences from *H. armiger* from the NCBI sequence read archive (ID: SRX1177145-SRX1177156) and the protein sequences from *H. armiger*, *R. ferrumequinum*, *Myotis brandtii*, *Myotis myotis*, *Eptesicus fuscus*, *Pipistrellus kuhlii*, *M. natalensis*, *Molossus molossus*, *P. discolor*, *Artibeus jamaicensis*, *Sturnira hondurensis*, *Desmodus rotundus*, *P. vampyrus*, *Pteropus giganteus* and *R. aegyptiacus* (from the NCBI Genomes database; Supplementary Table S1). The transcript sequences from *H. armiger* were additionally mapped to the assembled sequences using the PASA pipeline.[35] For the *ab initio* gene prediction, we used GeneMark-ET and AUGUSTUS via BRAKER2 pipeline[36] based on the mapped transcripts and the vertebrate orthologous protein sequences from the OrthoDB release 10 database[24] (odb10_vertebrata.fasta). EvidenceModeler was then used to compile all gene models to obtain final consensus gene models based on following weight for each evidence type: PASA = 10, protein2genome (maker2) = 5, est2genome (maker2) = 5 and BRAKER = 1. We employed the BUSCO[23] version 4.0.5 to evaluate the annotation results by testing for the presence and completeness of the orthologues using the Laurasiatheria OrthoDB release 10.[24]

2.7. Comparative genomics and phylogenetic analysis

We used OrthoFinder[37] to identify orthologous groups from the sequences of proteins of *H. pendleburyi* and the aforementioned species, which were downloaded from the NCBI database. The protein sequences from single-copy orthologous groups were aligned with MUSCLE software.[38] The alignments were further processed by trimming gap-rich regions with trimAl[39] (using the automated1 heuristic method) software and concatenating with catsequences software (https://github.com/ChrisCreevey/catsequences, last accessed: 21 July 2022). The final concatenated alignment was subjected to ModelTest-NG software[40] for identifying the substitution model of each alignment block. The RAxML-ng software[41] was used to construct a maximum-likelihood phylogenetic tree from the concatenated alignment and substitution models. Divergence times between *H. pendleburyi* and *H. armiger* and between *R. ferrumequinum* and the *Hipposideros* clade in the phylogenetic tree were estimated with the MCMCtree program (PAML4 package)[42] using the relaxed-clock model with the known divergence time obtained from the TimeTree database.[43] Synteny blocks of *H. pendleburyi* against the *R. ferrumequinum* and *P. discolor* genomes were obtained and visualized with SyMAP software.[44]

2.8. Ka/Ks analysis

The Ka/Ks ratio was estimated for each protein sequence in the single-copy orthologous groups using the Codeml software (PAML4 package).[42] We analysed branch-site selection[45] in two dataset systems. In the first set, all bat species were considered as foreground branches, and all other species were used as the background branches. In the second set, *Hipposideros* branch and *H. pendleburyi* branches were regarded as foreground branches and all other bat species were used as the background branches. The likelihood ratio test (LRT)[46] was used to test the alternative hypothesis (positive selection on the foreground branches; Ka/Ks > 1) compared with the null hypothesis (neutral selection; Ka/Ks = 1). Bayes empirical Bayes method was used to identify amino acid residues that have potentially evolved under the selection.[47] All aforementioned analysis steps were performed under VESPA pipeline.[48] The final results were manually inspected to find the genes that contained the positively selected sites only in the conserved regions of the sequence alignment. The LRT *P*-values of the finally selected genes were obtained using EasyCodeML software.[49] The gene ontology terms of the selected proteins were obtained by performed blast search against the reviewed protein data set from the uniprot database.[50]

3. Results and discussion

3.1. Genome assembly and annotations

We performed a whole-genome shotgun sequencing using the 10× Genomics linked-read strategy. We obtained a total of 100.85 Gb of Illumina paired-end 150 bp sequencing data from 336,176,017 raw reads. *De novo* assembly yielded a draft genome of 2.17 Gb, which was close to that of *H. armiger*.[12] Our assembly had a scaffold N50 length of 15,398,518 bases (L50 = 35 scaffolds; Table 1), which was

Table 1. Assembly statistics of the *H. pendleburyi* genome

Genomic features	10× genomics
N50 scaffold size (bases)	15,398,518
L50 scaffold number	35
N75 scaffold size (bases)	5,979,547
L75 scaffold number	90
N90 scaffold size (bases)	1,220,106
L90 scaffold number	208
Total (bases)	2,170,000,000
Number of scaffolds	28,685
Number of scaffolds ≥100 kb	492
Number of scaffolds ≥1 Mb	226
Number of scaffolds ≥10 Mb	60
Longest scaffold (bases)	98,956,843
GC content (%)	41.15
BUSCO	94.4%

about five times longer than that reported for the *H. armiger* assembly.[12] The analysis of *k*-mer distribution of the genome sequencing reads provided an estimated genome size of 2.28 Gb, which was close to the assembly size (Supplementary Fig. S1). For the quality assessment, 97% of the DNA reads could be mapped to the assembled genome. The assessment of gene content with the BUSCO software[23] showed that the gene prediction for *H. pendleburyi* recovered 94% of the 12,234 highly conserved orthologues in the Laurasiatheria lineage (Table 1). This percentage was similar to that estimated for *H. armiger* (93%) and *R. ferrumequinum* (99%).

3.2. Phylogenetic relationships

We analysed the interordinal relationships of *H. pendleburyi* with 15 bats and 12 mammals (Fig. 1). Our phylogenetic analysis showed that Chiroptera was placed as a sister group to Ferungulata within the Laurasiatheria superclade. This placement was consistent with several other molecular phylogenetic studies.[3,5,6] Two internal nodes within the Ferungulata clade showed a low bootstrap support value (Fig. 1) as also reported in the previous study,[6] suggesting the rapid speciation events in the Laurasiatheria superclade.[6]

For the order Chiroptera, the superfamily Rhinolophoidea was placed as a sister group to the family Pteropodidae within the suborder Yinpterochiroptera (Fig. 1), which was consistent with other phylogenetic studies.[4–6] *Hipposideros pendleburyi* formed a monophyletic group with *H. armiger* and *R. ferrumequinum* within the superfamily Rhinolophoidea (Fig. 1). As all species were placed with a 100% bootstrap value within the Chiroptera branch, we calculated the molecular divergence times only in this branch. The divergence between Yinpterochiroptera and Yangochiroptera occurred about 65 million years ago (Mya).[4] Within the Yinpterochiroptera branch, the divergence time between the superfamily Rhinolophoidea and the family Pteropodidae was estimated to be 58 Mya.[51] The divergence between *R. ferrumequinum* and the *Hipposideros* clade was estimated to be 33 Mya (with a 95% credibility interval of 21–48 Mya; Fig. 1). *H. pendleburyi* was estimated to diverge from *H. armiger* about 5 Mya (3–9 Mya; Fig. 1), which was in the same period when *P. vampyrus* and *P. giganteus* diverged from each other.[43] *H. pendleburyi* was formerly classified as a subspecies of the *Hipposideros turpis* complex.[9] The phylogenetic analysis based on *Cytochrome b* sequences and the analyses of morphological and ecological data later revealed that *H. pendleburyi* formed a monophyletic group with *H. armiger*.[9] Our phylogenetic tree constructed based on sequences of nuclear genes supported the close relationship between *H. pendleburyi* and *H. armiger*. However, the lack of the genome data of *H. turpis* has hindered further analysis of the relationship between this species and *H. pendleburyi*.

The topology of phylogenetic trees from this and previous studies suggested that pteropodids, which lack the ability of laryngeal echolocation, were nested among echolocating bat lineages, and their diversification appeared to be associated with a rising in global temperature in the early Eocene, which resulted in an increase in plant diversity and abundance and insect diversity.[4] The divergence timeline was also in agreement with the hypothesis stating that laryngeal echolocation ability might be originally evolved in bat ancestors and was inherited in both Yinpterochiroptera and Yangochiroptera lineages but was subsequently lost in pteropodids.[5]

3.3. Synteny analysis

The contiguity of our assembly enabled us to investigate the synteny between *H. pendleburyi* and other chromosome-level bat genomes. We identified synteny blocks in the *H. pendleburyi* (family Hipposideridae), *R. ferrumequinum* (Rhinolophidae) and *P. discolor* (Phyllostomidae) genome assemblies. The results showed 18,114 sequence matches (referred to as anchors[52]) from the alignment of the scaffolds of *H. pendleburyi* assembly to *R. ferrumequinum* chromosomes, while 17,604 anchors were obtained from the alignment with *P. discolor* chromosomes (Supplementary Table S3). The synteny blocks (generated from a set of high similarity scored anchors[52]) from the alignments with *R. ferrumequinum* and *P. discolor* genomes covered 91% and 89% of the *H. pendleburyi* genome, respectively (Supplementary Table S3). The longest alignment length between the synteny blocks of *H. pendleburyi* to the *R. ferrumequinum* chromosomes was 99 Mb, which was comparable to the longest alignment between the chromosomes of *P. discolor* and *R. ferrumequinum* (106 Mb). The analyses depicted that the synteny blocks (collinearity with rearrangements[52]) were almost as long as the chromosome of *R. ferrumequinum*. In addition, the synteny blocks were found to be conserved among the analysed bats, although they were from different families (Supplementary Fig. S2). Bat genomes were relatively small compared to other mammal genomes.[2] The small size of the genomes may restrict the reshuffling of chromosome segments and may be associated with the conservation of the synteny blocks among bats from different families.

The chromosome numbers of Chiroptera species (chiropterans) are highly varied.[53] *H. pendleburyi* is more closely related to *R. ferrumequinum* than to *P. discolor* yet the chromosome number of *H. pendleburyi* ($2n = 32$)[53] is equal to that of *P. discolor* and is noticeably lower than that of *R. ferrumequinum* ($2n = 58$). Despite the difference in chromosome number, the synteny block analyses showed that the sequences of *H. pendleburyi* scaffolds were more similar to the sequences of *R. ferrumequinum* chromosomes than to those of *P. discolor* chromosomes. The results suggested that the chromosome number might not be a good indicator to study the phylogenetic relationship and also highlighted the importance of genomic resources for studying the evolution in bats.

3.4. Genes under selection pressure

We performed the positive selection analysis with branch-site models[45] to identify positively selected sites in the sequences of orthologous genes. We first identified the selective sites that were specific to Chiroptera members. In this analysis, the foreground branch contained

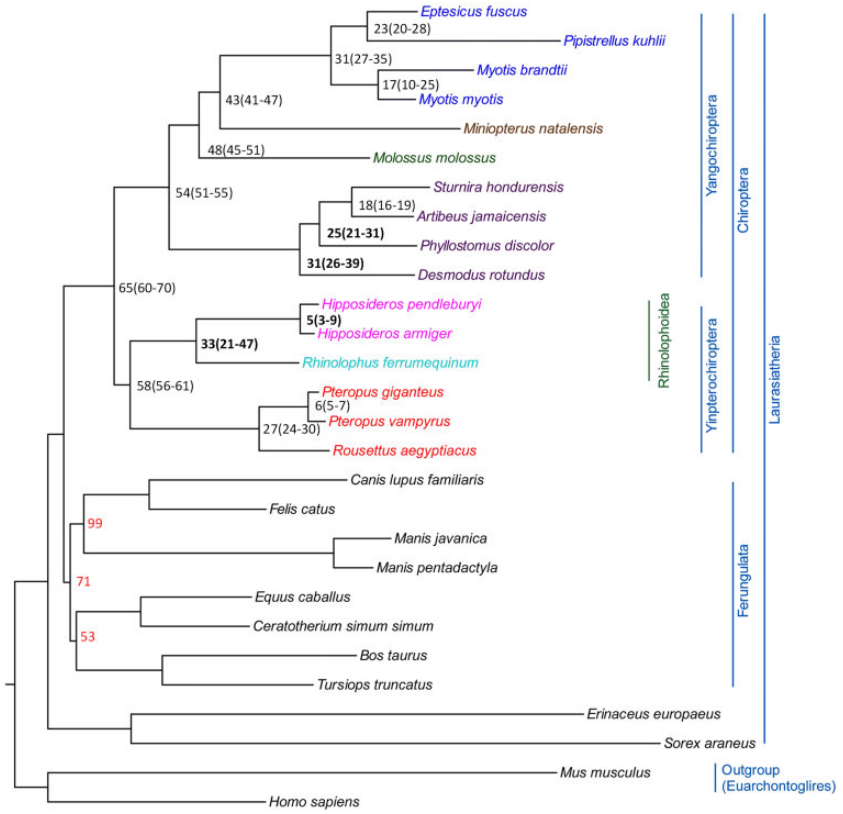

Figure 1. The maximum likelihood phylogenetic tree of bats and other mammals. All internal nodes receive a 100% bootstrap support value from the analysis (not shown in the figure for clarity), except for three internal nodes where bootstrap values are shown in red. Bats from different families are shown in different colours, Pteropodidae: red, Rhinolophidae: cyan, Hipposideridae: pink, Phyllostomidae: purple, Molossidae: green, Miniopteridae: brown and Vespertilionidae: blue. The superfamily Rhinolophoidea is shown in lime and other taxonomic ranges are in navy blue. The divergence times that are shown in a normal typeface at internal nodes are obtained from the TimeTree database. The values in the associated parentheses are ranges of divergence times compiled from several studies by the TimeTree database. These divergence times and the associated ranges are used as prior calibration time points for the estimation of unknown divergence times. The divergence times estimated in this study are displayed in bold typeface and the values in the associated parentheses show 95% credibility interval. (A color version of this figure appears in the online version of this article.)

all species within the Chiroptera clade of the phylogenetic tree (Fig. 1), and the background branch had other mammals. The analysis identified the selective sites in the sequences of interleukin (IL)-9 and Granzyme A (GzmA) encoding genes (Table 2), which were involved in inflammation.[54,55] The selection of these genes was consistent with the evolution of bat immunomodulation that limit virus-induced pro-inflammatory responses to reduce tissue damage while limiting virus propagation.[56] Previously, the positive selections of inflammation-related genes, caspase-8 and IL genes, were reported in *Rousettus* bats.[6] The positive selections of caspase-1 in *Pteropus* bats and IL-1β in *Myotis* bats have been reported to result in a dampened inflammation.[57] Similar to the roles of these two genes,[57] GzmA played a role in cleaving of the IL-1β precursor to produce functional IL-1β.[58] In addition to the Chiroptera-specific selections, we identified the selections specific to the analysed *Hipposideros* (*H. pendleburyi* + *H. armiger* in the foreground branch and all other Chiroptera members in the background branch). The results showed the positively selected sites in the sequences of the cell death regulator Aven (an inhibitor of caspase activation[59]), IL-5 receptor and IL-17F encoding genes at the *Hipposideros* branch level (Table 2). The manifestation of the positively selected sites in the sequences of these genes were also related to the immunomodulatory adaptation of bats.[2]

For *Hipposideros*, positively selected sites were additionally identified in the sequences of protein disulfide-isomerase-like protein of the testis (PDILT), SUN domain-containing protein 5 (Sun5) and ATP-dependent RNA helicase TDRD9 (Tdrd9) encoding genes (Table 2), which played roles in spermatogenesis. PDILT was shown in human to be induced in testis during puberty and essential for the spermatogenesis during the final stages of germ cell maturation.[60] Sun5 and Tdrd9 involve in spermatozoa integrity[61] and maturation arrest[62] and are associated with mice male infertility. The timing of spermatogenesis in bats depended on several conditions, such as climate, season, food availability and age, puberty onset age, genera and species of bats.[63,64] Positive selection of reproduction genes at the genus level may imply the formation of reproduction barriers between *H. pendleburyi* and *H. armiger* that diverged at about 5 Mya.

We also identified the positively selected sites in the sequence of the Secreted frizzled-related protein 5 (Sfrp5) encoding genes of *Hipposideros* (Table 2). The selection sites of Sfrp5 were not found in other Chiroptera members or other analysed mammals. Sfrp5 protein, a modulator in the Wnt signalling pathway,[65] involves in gut development in *Xenopus*[66] and mouse,[67] co-evolutionary patterning of teeth and taste buds in Cichlid,[68] and incisor renewal[69] and early trunk formation in mouse.[70] The positively selected sites in the Olfactory receptor 52M1 and the Cytochrome P450 8B1 encoding genes (Table 2) were additionally identified for *Hipposideros*. These genes play roles in regulating the intestinal absorption of dietary lipids in mouse.[71] These results suggested that the availability and type

Table 2. Selected bat protein-coding genes with positively selected sites in their sequences

ID	Uniprot proteins	Foreground	Background	lnL[a]	LRTs (2ΔLnL)	LRT *P*-value	Ka/Ks	Number of positively selected sites
OG0016899_Chiroptera	Granzyme A	Chiroptera	Other mammals	−7,094.87	8.73799	0.003116479	139.15457	3
OG0016905_Chiroptera	IL-9	Chiroptera	Other mammals	−3,790.17	5.454968	0.019512762	22.52779	3
OG0016989_Chiroptera	Secreted frizzled-related protein 5	*Hipposideros*	Other mammals	−3,928.83	5.108404	0.023810214	25.52567	3
OG0016989_Chiroptera	Secreted frizzled-related protein 5	*H. pendleburyi*	Other mammals	−3,929.51	7.89521	0.004956588	170.53445	1
OG0015897_Hipposideros	ATP-dependent RNA helicase TDRD9	*Hipposideros*	Other bats	−23,277.4	24.43744	0.000000768	78.08567	32
OG0016019_Hipposideros	Cytochrome P450 8B1	*Hipposideros*	Other bats	−6,775.04	8.79168	0.003026074	9.81229	10
OG0016047_Hipposideros	Protein disulfide-isomerase-like protein of the testis	*Hipposideros*	Other bats	−11,397	87.75038	0	119.58094	68
OG0016646_Hipposideros	SUN domain-containing protein 5	*Hipposideros*	Other bats	−4,558.43	19.26875	0.000011355	644.40457	7
OG0016719_Hipposideros	IL-5 receptor subunit alpha	*Hipposideros*	Other bats	−12,355.5	14.06578	0.000176527	65.42665	8
OG0016818_Hipposideros	IL-17F	*Hipposideros*	Other bats	−2,653.47	4.430288	0.03530655	10.11761	13
OG0016826_Hipposideros	Secreted frizzled-related protein 5	*Hipposideros*	Other bats	−2,562.66	5.142734	0.023343858	25.0218	3
OG0016836_Hipposideros	Cell death regulator Aven	*Hipposideros*	Other bats	−5,493.09	60.83971	0	137.77875	32
OG0016842_Hipposideros	Olfactory receptor 52M1	*Hipposideros*	Other bats	−3,269.07	6.784914	0.009193146	67.47649	3
OG0016826_H. pendleburyi	Secreted frizzled-related protein 5	*H. pendleburyi*	Other bats	−2,563.38	7.940124	0.004835052	184.41056	1

[a]LnL = the log-likelihood (lnL) values.

of diets, which varied in different ecological niches,[72,73] might exert selective pressure to these genes of *Hipposideros*.

3.5. Transposable elements

We obtained 12,212 consensus sequences of TEs from the assemblies of 11 bats, including three echolocating bats (*M. natalensis*, *M. lucifugus* and *P. discolor*) of the suborder Yangochiroptera and three hipposiderids (*H. pendleburyi*, *H. armiger* and *H. galeritus*), three other rhinolophoids (*R. ferrumequinum*, *C. thonglongyai* and *M. lyra*) and two pteropodids (*P. vampyrus* and *R. aegyptiacus*) of the suborder Yinpterochiroptera (Supplementary Table S2). For *H. pendleburyi*, TEs accounted for 38.6% (0.84 Gb) of the assembly, which was similar to the proportion of TEs in *H. armiger* (32.9%), *H. galeritus* (42.1%) and *R. ferrumequinum* (38.9%; Supplementary Table S4). For each species, the Kimura substitution levels (k-distances) of TE sequences were shown in the TE landscapes (Fig. 2). The k-distance values represented the divergence of the TE sequences from their consensus sequences. We found unique profiles of short interspersed nuclear elements (SINEs; class I TEs) and DNA transposons (class II TEs) within the k-distance range from three to seven in hipposiderids (Figs 2 and 3A). The TEs within this k-distance range were referred to as young TEs.

For class I TEs, we found fragments of MIR SINEs in all analysed species, while VES and MEG families of SINE class were found only in vespertilionids and pteropodids, respectively (Fig. 3B). These results were consistent with previous reports.[74,75] Rhin-1 SINEs were found only in hipposiderids and *R. ferrumequinum*, which was consistent with previous report.[76] The proportion of young Rhin-1 in *R. ferrumequinum* (0.385% of the assembly) was about 50 times higher than those in *H. pendleburyi* (0.008%), *H. armiger* (0.007%) and *H. galeritus* (0.006%; Fig. 3B). In addition, the average k-distance of all Rhin-1 in *H. pendleburyi* (15.4 ± 4.8), *H. armiger* (15.5 ± 4.8) and *H. galeritus* (16.7 ± 4.6) were significantly (Welch's *t*-test, $P < 0.01$) higher than that in *R. ferrumequinum* ($7.9 \pm 3.4\%$; Supplementary Table S5). These results suggested that Rhin-1 in *H. pendleburyi*, *H. armiger* and *H. galeritus* was older than those in *R. ferrumequinum* and implied that the amplification of Rhin-1 was likely suppressed in hipposiderids.

In most bat species, the proportion of SINE was observed to be correlated with the proportion of long interspersed nuclear element (LINE) because SINEs required reverse transcriptase encoded by LINEs for their amplification.[77,78] Intriguingly, hipposiderids had a slightly lower proportion of SINEs ($0.08 \pm 0.02\%$) than pteropodids did ($0.10 \pm 0.02\%$; Supplementary Table S5), even though its LINE proportion ($2.10 \pm 0.96\%$) was about two times higher than those of pteropodids ($0.96 \pm 0.38\%$), which lost LINE-1, the most prevalent LINE family, from their genomes.[79] The results suggested that the amplification of SINEs in the analysed hipposiderids might be suppressed by other mechanisms, such as the inhibition by APOBEC family proteins[78,80] and methylation.[78,81] Changes in the proportion of retroelements, including SINEs, have been shown to be associated with the evolution and radiation of pteropodids[79] and sigmodontine rodents.[82] In this study, we showed that the contraction of SINEs, particularly Rhin-1 and MEG families (Fig. 3B), might be linked to the phylogenetic separation of the hipposiderids from *R. ferrumequinum* (rhinolophid), other rhinolophoids and pteropodids.

For class II TEs, a high proportion of young DNA transposons was found in all analysed hipposiderids and rhinolophid (Fig. 3A). The exploration of DNA transposon families indicated that young *Tc1/Mariner* elements were the most prevalent in this group of

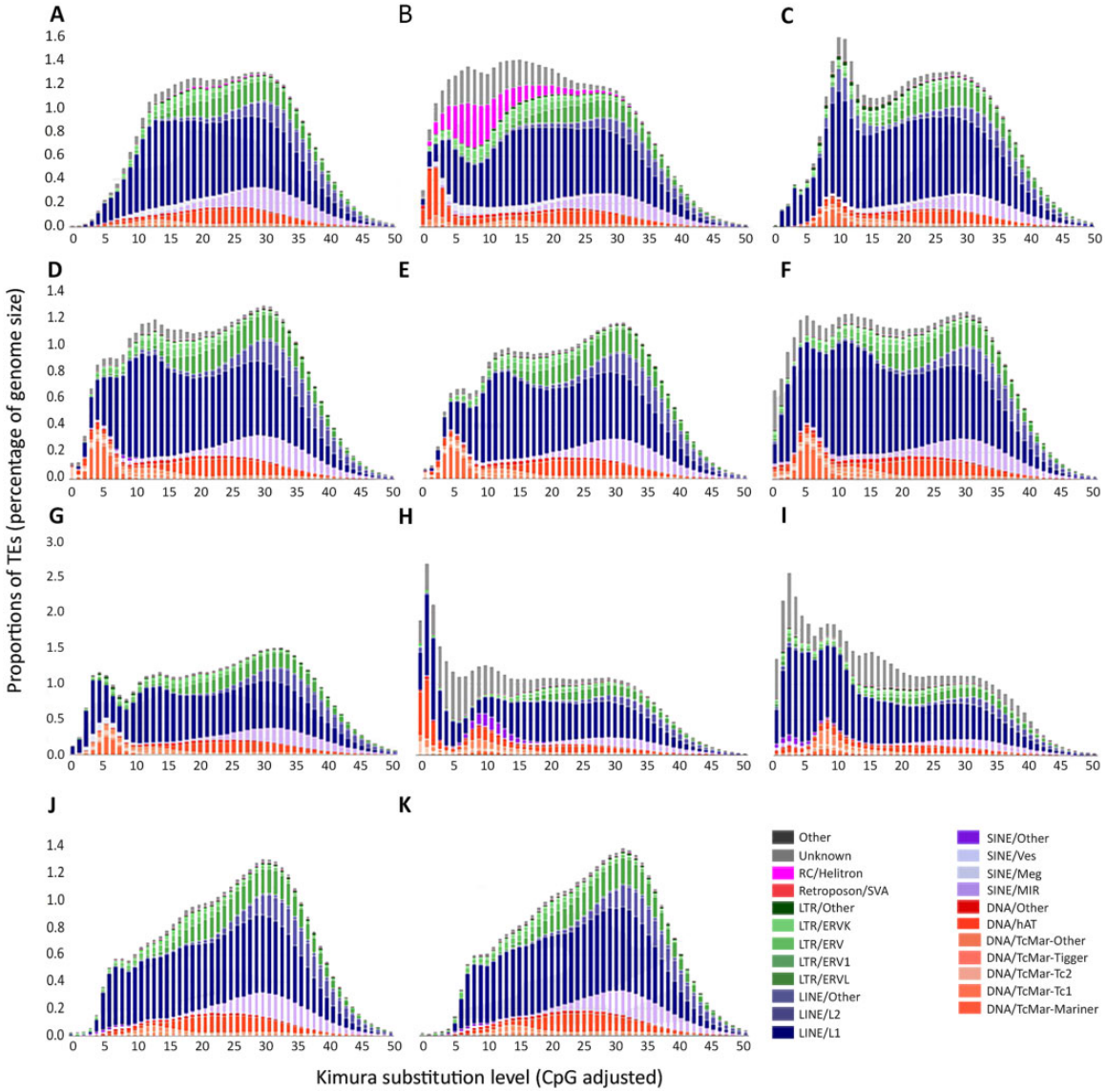

Figure 2. The TE profiles. The TE landscapes of (A) *M. natalensis*, (B) *M.lucifugus*, (C) *P. discolor*, (D) *H. pendleburyi*, (E) *H. armiger*, (F) *H. galeritus*, (G) *R. ferrumequinum*, (H) *C. thonglongyai*, (I) *M. lyra*, (J) *P. vampyrus* and (K) *R. aegyptiacus* are shown. Different TE families are shown with different colours. Younger TEs had a lower Kimura substitution level (k-distance) than the older TEs. The proportions of TEs at each k-distance level are expressed as a percentage of the genome size. The k-distances and the proportions of TEs are shown on the *x* and *y*-axes, respectively.

species (13.1 ± 0.1%; Fig. 3C and Supplementary Table S5). We did not find a similar profile in other bats, suggesting that a high accumulation of *Tc1/Mariner* was specific to the analysed hipposiderids and rhinolophid. The lineage-specific accumulation of particular DNA transposon families has been reported in Vespertilionidae and Miniopteridae echolocating bats.[13,14,18] To classify the *Tc1/Mariners*, the consensus sequences obtained from our RepeatModeler[25] analyses together with the *Tc1/Mariner* consensus sequences compiled in other study[5] were used to construct a phylogenetic tree (Fig. 4A). For rhinolophoids, we found four groups of *Tc1/Mariner* elements (Groups I–IV in Fig. 4). Compared to the TE sequences of *P. discolor* and *M. lucifugus* that each formed one group in the phylogenetic tree, rhinolophoids had a higher number of *Tc1/Mariner* groups, which was consistent with their high *Tc1/Mariner* proportions (Fig. 3C). RepeatModeler performed a *de novo* identification of consensus sequences from genome sequence.[25] As a support of RepeatModeler consensus-sequence identifications, we additionally identified *Tc1/Mariner* elements in the assemblies of *H. pendleburyi* and other species based on the amino acid sequences of transposase domains that are compiled in other study[32] (using TBLASTN search). Hits from the TBLASTN search included *Apismar1.1*, *Apismar2.1-2.2* (domains from *Acyrthosiphon pisum*), *Dnomar2.1-2.2*, *Dnomar3.1* (*Diuraphis noxia*) and *Mpmar1.1* (*Myzus persicae*) transposase domains (Supplementary Fig. S3). These domains belonged to the DD34D transposase family of *mariners*.[32] *Apismar2.1-2.2* and *Dnomar2.1-2.2* have also been found in *R. ferrumequinum*.[32] These results together indicated that the majority of young *Tc1/Mariner* in the analysed hipposiderids and rhinolophid were *mariner* elements, which were consistent with the results from RepeatModeler and RepeatMasker (Fig. 3C).

A phylogenetic tree based on full-length *mariner* sequences in *H. pendleburyi* showed four groups of *mariners* (Groups *i–iv* in

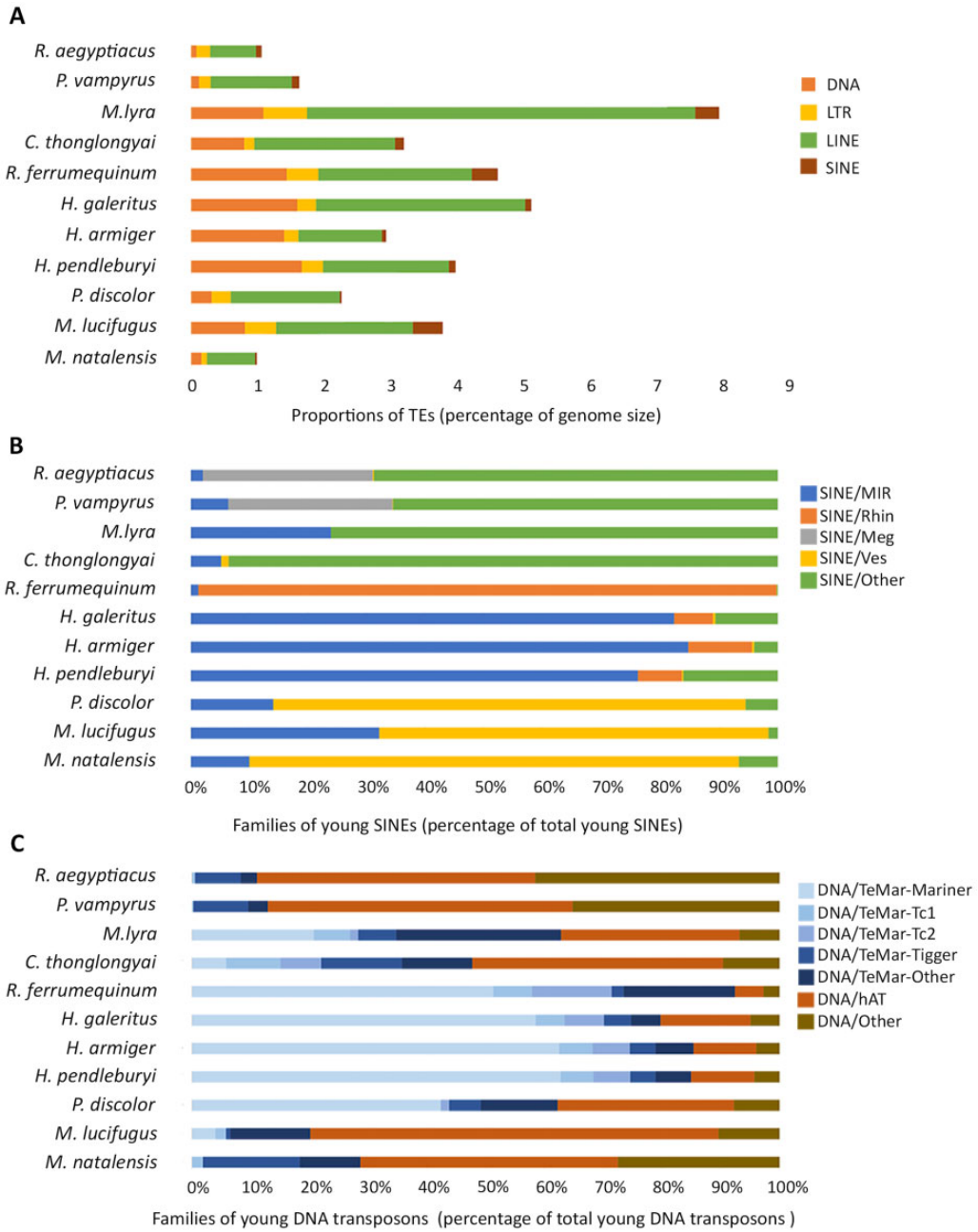

Figure 3. The proportion of young TEs. (A) The proportions of four main TE classes within the k-distance range between 3 and 7 (considered as young TE elements) in *H. pendleburyi* and other 10 species are shown. The proportions of each TE class are expressed as a percentage of the genome size. (B and C) Compositions of different families are shown as a percentage of total young SINEs or DNA transposons.

Fig. 4B), which was consistent with the phylogenetic tree of the *Tc1/Mariner* consensus sequences (Groups I–IV in Fig. 4A). The BLASTN results also showed that the *mariners* of Group *i* (corresponding to Group I in Fig. 4A) resembled that of the *mariners* of treeshrew (designated as *Tcmar* after *Tupaia chinensis*) and insect (*Camar1* after *Chymomyza amoena*) with average identities of $88 \pm 0.2\%$ and $71 \pm 0.3\%$, respectively (Fig. 4A and B; Supplementary Table S6). The *mariners* in Group *i* were found in all analysed hipposiderids but not in other rhinolophoids. In contrast, the *mariners* of Group *iv* (Group IV in Fig. 4A) were found in all rhinolophoids and resembled the *mariner* of caecilian (*Rbmar* after *Rhinatrema bivittatum*, referred to as Mari2_Pca in the previous study[5]; Fig. 4A and B). These data together indicated that the *mariner* elements found in hipposiderids and rhinolophoids were not detected in other bat groups but existed in species distantly related to bats. The presence of full-length copies of *mariner* elements suggested that the expansion of these elements is ongoing.[14] A similar observation has been described in vespertilionid bats.[14] In addition, the star-like topology of the *mariner* elements in Groups *i* and *iv* indicated the accumulation of distinct mutations after a single burst of transposition[15,19] (Fig. 4C and D). These results together suggested that *mariner* elements were introduced into hipposiderid

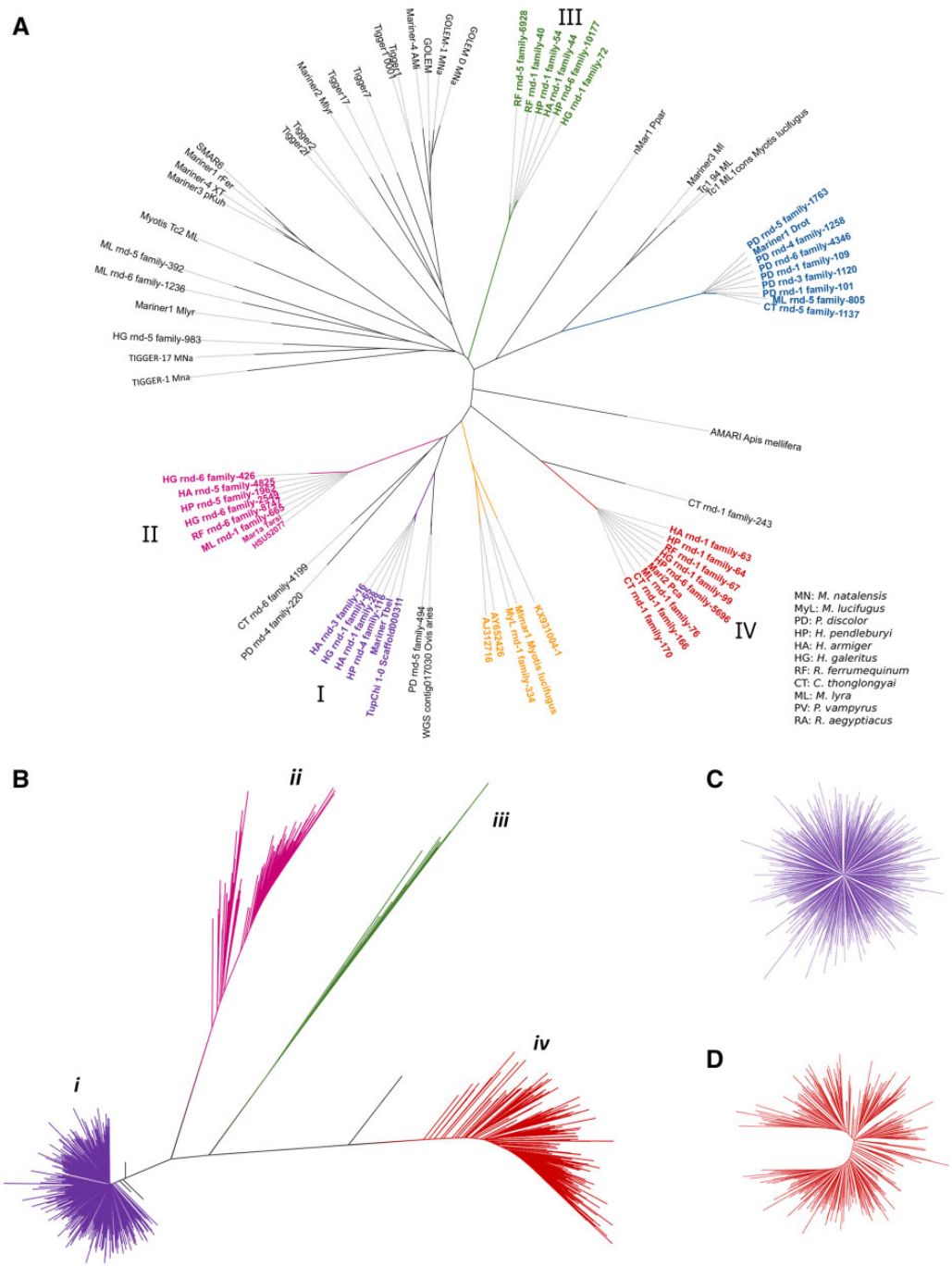

Figure 4. The phylogenetic trees of DNA transposons. (A) The unrooted phylogenetic tree of the consensus sequences of *Tc1*/*Mariner* elements is shown. (B) The unrooted phylogenetic tree of the full-length *mariner* sequences extracted from *H. pendleburyi* assembly is shown. (C and D) The subtrees of groups *i* and *iv* in the phylogenetic tree of the full-length *mariner* sequences are shown, respectively. Colours feature separate groups of DNA transposons and are synonymous between sub-panels.

and rhinolophoid genomes via horizontal transfer and are still active.[16,19,20]

3.6. Horizontal transfer of *mariner* elements

To further investigate the possibility of horizontal *mariner* acquisition in *H. pendleburyi*, we performed BLASTN search using the *mariner* consensus sequences of *H. pendleburyi* as query sequences against the NCBI nr/nt database. The BLAST results showed that only the *mariners* of groups I and IV (Fig. 4A) matched to non-mammalian sequences. The consensus sequences of Group I matched to the sequences of tree shrews (*T. chinensis* and *Tupaia belangeri*; Mammalia) and insects, including ant (*Tetramorium bicarinatum*; Hymenoptera), hoverfly (*Criorhina berberina*; Diptera) and hornet moth (*Sesia apiformis*; Lepidoptera), which were consistent with the results from another BLASTN (previous section). The consensus

sequences of Group IV matched to the sequences of rock hyrax (*Procavia capensis*; Mammalia) and two caecilians (*Geotrypetes seraphini* and *R. bivittatum*; Amphibia). These BLAST hits were obtained with the query coverages of >99% of the consensus sequence length (about 1,290 bases) and the identity of >86%. We also performed the BLASTN searches for other analysed bat species. We considered only the consensus sequences that were longer than about 50% of the full-length *mariner*s (550–1,300 bp) and only the BLAST hits with the query coverage of >90% and the identity of >80% (Supplementary Table S7). The results showed that the Group IV *mariners* of rhinolophoids matched to the sequences of *G. seraphini* and *R. bivittatum*, consistent with the results of *H. pendleburyi* (hipposiderid). For echolocating bats (Yangochiroptera), the *mariners* of *P. discolor* formed their own group (blue group in Fig. 4A) and were not related to any of the hipposiderid and rhinolophoid *mariners*. *P. discolor mariners* matched the sequences of *mariners* in other echolocating bats in the database (Supplementary Table S7). A *mariner* in *P. discolor* (PD_rnd-5_family-494) matched to the sequences of Artiodactyla mammals such as bovines, goats, sheep, deer, dolphins and whales. This particular *mariner* formed a separate clade on the branch, in which Group I resided (Fig. 4A). For pteropodid bats (Yinpterochiroptera), the *P. vampyrus mariner* did not show any significant match from the BLASTN search. These results together indicated that the BLAST hits of the *mariners* in groups I and IV were specific to the analysed hipposiderids and rhinolophids, respectively. These BLAST results also agreed with several examples of the horizontal transfer of *mariner* elements in mammals, insects and amphibians.[15,16,19,31,83]

The acquisition of *mariner* elements in rhinolophids and hipposiderids might be linked to their feeding behaviours. Rhinolophids and hipposiderids are both insectivorous and the insects in the order Lepidoptera, Diptera and Hymenoptera are their main preys.[84] The consumption of Hymenoptera ants by *Hipposideros* species has also been documented.[85] The interactions between mammals and insects, including the feeding on insects of mammals[31,83] and the transfer of pathogens from insects to mammal hosts,[15] were believed to increase the likelihood of TE horizontal transfer. It is therefore possible that hipposiderids have first received the *mariner* elements of Group I from their prey.

The timeline of *mariner* acquisition is also of interest. Even though rhinolophoids and hipposiderids fed on overlapping groups of prey, Group I *mariners* existed only in the analysed hipposiderids (Fig. 4A). Based on the assumption that these two groups of bats received the *mariner* elements from their prey, we hypothesized that hipposiderids acquired Group I *mariners* after they had diverged from other bats of the Rhinolophoidea superfamily. In contrast, Group IV *mariners*, which were found in both hipposiderids and rhinolophoids (Fig. 4A), might have been acquired before such divergence. This hypothesis was supported by the star-like topology of Group I *mariners* (Fig. 4C) that displayed a better proportioned shape than that of Group IV *mariners* (Fig. 4D). The difference of star-like topologies suggested that the *mariner* elements of Group IV, compared to those of Group I, might be accumulated more mutations, by the amplification of pre-existing mutated copies after vertical transfer events within Rhinolophoidea lineage. A higher k-distance value of the elements of Group IV (8.33 ± 1.22) than that of Group I (6.33 ± 0.91) also supported this hypothesis. Previously, the acquisition of *Tc1/Mariner* elements in *M. natalensis* and the vespertilionids were proposed to begin just before the divergence of the *Miniopterus* and the vespertilionid.[18] Sequence analyses indicated that four *mariners* in *M. natalensis* (GOLEM-1 MNa, GOLEM D MNa, TIGGER-1 Mna and TIGGER-17 MNa)[18] were similar to Tigger and GOLEM ancient *mariners*. The phylogenetic analysis additionally showed that these *mariners* were not in any *mariner* groups in hipposiderids and rhinolophoids (Fig. 4A). The difference of *mariner* profiles suggested that the acquisitions of *mariner* elements in vespertilionids[18] and rhinolophoids possibly occurred separately over the course of evolution.

The insertion of TEs introduces new nucleotide sequences to host genomes and can result in the emergence of new genetic features in host genomes by altering original coding or regulatory sequences.[86–89] The species with the insertions of TEs could adapt to new environment more quickly than those that relied solely on point mutations.[18] In primates, the fusion between *mariner* transposon and *SET* gene resulted in the emergence of SETMAR protein family, which had functions in DNA-double-strand break repair.[86,87] The insertion of TEs has been shown to alter transcriptional regulations and phenotypes in plants.[88] For bats, the lineage-specific accumulation of DNA transposons in vespertilionids was linked to the rapid and repeated introduction of novel p/miRNAs, which were involved in posttranscriptional regulation.[89] The expansion was also associated with the radiation of vespertilionids.[89] The insertion of TEs has been suggested to be related to the separation of hipposiderids from rhinonycterids (both belong to the superfamily Rhinolophoidea).[7] Based on this information, we proposed that the lineage-specific accumulation of *mariner* DNA transposons might be linked to the phylogenetic separation of hipposiderid and rhinolophid bats from other bats within the suborder Yinpterochiroptera.

4. Conclusions

We reported a draft genome assembly of Pendlebury's roundleaf bat (*H. pendleburyi*), endemic to Thailand and currently listed as a vulnerable species. The size of the assembled genome was 2.17 Gb, which was similar to those of other bat genomes. Our phylogenetic analysis placed *H. pendleburyi* within the clade of rhinolophoids in the suborder Yinpterochiroptera. The synteny analysis showed that genome structures of echolocating bats were highly conserved and shared contiguous chromosome segments, although they had different numbers of chromosomes and were from different families (Hipposideridae, Rhinolophidae and Phyllostomidae). We found positively selected sites in genes that might play roles in inflammation control, spermatogenesis and foregut morphogenesis. The positive selection of these genes implied the genetic adaptation of *H. pendleburyi* to its environments. From the analysis of TEs, we found a high accumulation of *mariner* DNA transposons in the analysed hipposiderid and rhinolophid genomes. The lineage-specific accumulation of this DNA transposon family might be involved in the evolution of hipposiderid and rhinolophid bats and could be related to the phylogenetic separation of this group of species from its related species. The analyses also suggested that some groups of *mariner* elements were horizontally transferred to hipposiderids after they had separated from rhinolophids and other rhinolophoids while other *mariner* groups had been transferred prior to the speciation. The contractions of Rhin-1 and other families of SINEs could also be linked to the separation of the hipposiderids from rhinolophid and other bats. The genome assembly data and its analysis results reported in this study are valuable for *H. pendleburyi* conservation and lay a foundation for mammal comparative genomic studies.

Supplementary data

Supplementary data are available at *DNARES* online.

Acknowledgements

We are grateful to the National Science and Technology Development Agency (NSTDA), Thailand for facilities. We thank Awatsaya Pimsai and Sakiya Morlor at Prince of Songkla University (PSU) for fieldwork and sample collection.

Accession numbers

Hipposideros pendleburyi genome assembly has been submitted to the NCBI database under the accession number JAHQIX000000000; BioSample SAMN19494061.

Funding

This work was supported by the National Science and Technology Development Agency (NSTDA), Thailand, grant number: P-19-52245 and Thailand Research Fund (TRF), grant number: DBG6180028.

Conflict of interest

None declared.

References

1. Wilson, D. E. and Mittermeier, R. A., eds. *Handbook of the Mammals of the World, Vol.* **9**. Barcelona: Lynx Edicions.
2. Teeling, E.C., Vernes, S.C., Dávalos, L.M., Ray, D.A., Gilbert, M.T.P. and Myers, E; Bat1K Consortium. 2018, Bat biology, genomes, and the Bat1K project: to generate chromosome-level genomes for all living bat species, *Ann. Rev. Anim. Biosci*., **6**, 23–46.
3. Sadier, A., Urban, D.J., Anthwal, N., Howenstine, A.O., Sinha, I. and Sears, K.E. 2021, Making a bat: the developmental basis of bat evolution, *Genet. Mol. Biol*., **43**, **1–14**.
4. Teeling, E.C., Springer, M.S., Madsen, O., Bates, P., O'brien, S.J. and Murphy, W.J. 2005, A molecular phylogeny for bats illuminates biogeography and the fossil record, *Science*, **307**, 580–4.
5. Jebb, D., Huang, Z., Pippel, M., et al. 2020, Six reference-quality genomes reveal evolution of bat adaptations, *Nature*, **583**, 578–84.
6. Nikaido, M., Kondo, S., Zhang, Z., et al. 2020, Comparative genomic analyses illuminate the distinct evolution of megabats within Chiroptera, *DNA Res*., **27**, 1–12.
7. Foley, N.M., Thong, V.D., Soisook, P., et al. 2015, How and why overcome the impediments to resolution: lessons from rhinolophid and hipposiderid bats, *Mol. Biol. Evol*., **32**, 313–33.
8. Monadjem, A., Soisook, P., Thong, V.D. and Kingston, T. Family Hipposideridae (old world leaf-nosed bats. In: Wilson, D. E. and Mittermeier, R. A., eds. *Handbook of the Mammals of the World*, *Vol.* **9**. Barcelona: Lynx Edicions, pp. 227–58.
9. Thong, V.D., Puechmaille, S.J., Denzinger, A., et al. 2012, Systematics of the Hipposideros turpis complex and a description of a new subspecies from Vietnam, *Mammal Rev*., **42**, 166–92.
10. The IUCN Species Survival Commission. 2012, *IUCN Red List Categories and Criteria, Version 3.1*, 2nd edition. Gland, Switzerland: IUCN (International Union for Conservation of Nature).
11. , Soisook, P. 2019. *Hipposideros pendlebury*. The IUCN Red List of Threatened Species 2019: e.T80224655A95642195. https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T80224655A95642195.en. last accessed 21 July 2022.
12. Dong, D., Lei, M., Hua, P., et al. 2017, The genomes of two bat species with long constant frequency echolocation calls, *Mol. Biol. Evol*., **34**, 20–34.
13. Ray, D.A., Pagan, H.J.T., Thompson, M.L. and Stevens, R.D. 2007, Bats with hATs: evidence for recent DNA transposon activity in genus *Myotis*, *Mol. Biol. Evol*., **24**, 632–9.
14. Ray, D.A., Feschotte, C., Pagan, H.J.T., et al. 2008, Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*, *Genome Res*., **18**, 717–28.
15. Gilbert, C., Schaack, S., Pace, J.K., II, Brindley, P.J. and Feschotte, C. 2010, A role for host–parasite interactions in the horizontal transfer of transposons across phyla, *Nature*, **464**, 1347–50.
16. Oliveira, S.G., Bao, W., Martins, C. and Jurka, J. 2012, Horizontal transfers of Mariner transposons between mammals and insects, *Mob. DNA*, **3**, 14.
17. Shen, D., Gao, B., Miskey, C., et al. 2020, Multiple invasions of visitor, a DD41D family of Tc1/Mariner transposons, throughout the evolution of vertebrates, *Genome Biol. Evol*., **12**, 1060–73.
18. Platt, R.N., Mangum, S.F. and Ray, D.A. 2016, Pinpointing the vesper bat transposon revolution using the *Miniopterus natalensis* genome, *Mob. DNA*, **7**, 12.
19. Pace, J.K., Gilbert, C., Clark, M.S. and Feschotte, C. 2008, Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods, *Proc. Natl. Acad. Sci. U S A*, **105**, 17023–8.
20. Thomas, J., Sorourian, M., Ray, D., Baker, R.J. and Pritham, E.J. 2011, The limited distribution of Helitrons to vesper bats supports horizontal transfer, *Gene*, **474**, 52–8.
21. Sikes, R.S; The Animal Care and Use Committee of the American Society of Mammalogists. 2016, 2016 Guidelines of the American Society of Mammalogists for the use of wild mammals in research and education, *J. Mammal*., **97**, 663–88.
22. Li, H. 2018, Minimap2: pairwise alignment for nucleotide sequences, *Bioinformatics*, **34**, 3094–100.
23. Seppey, M., Manni, M. and Zdobnov, E.M. BUSCO: assessing genome assembly and annotation completeness. In: Kollmar, M., ed., *Gene Prediction: Methods and Protocols*. New York, NY: Springer, 2019, pp.227–45.
24. Kriventseva, E.V., Kuznetsov, D., Tegenfeldt, F., et al. 2019, OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs, *Nucleic Acids Res*., **47**, D807–11.
25. Flynn, J.M., Hubley, R., Goubert, C., et al. 2020, RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. U S A*, **117**, 9451–7.
26. Tempel, S. Using and understanding RepeatMasker. In: Bigot, Y., ed. *Mobile Genetic Elements: Protocols and Genomic Applications*. Totowa, NJ: Humana Press, 2012, pp.29–51.
27. Li, W. and Godzik, A. 2006, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, *Bioinformatics*, **22**, 1658–9.
28. Price, M.N., Dehal, P.S. and Arkin, A.P. 2010, FastTree 2—approximately maximum-likelihood trees for large alignments, *PLoS One*, **5**, e9490.
29. Lefort, V., Longueville, J.-E. and Gascuel, O. 2017, SMS: smart model selection in PhyML, *Mol. Biol. Evol*., **34**, 2422–4.
30. Lemoine, F., Correia, D., Lefort, V., et al. 2019, NGPhylogeny.fr: new generation phylogenetic services for non-specialists, *Nucleic Acids Res*., **47**, W260–5.
31. Dupeyron, M., Baril, T., Bass, C. and Hayward, A. 2020, Phylogenetic analysis of the Tc1/Mariner superfamily reveals the unexplored diversity of pogo-like elements, *Mob. DNA*, **11**, 21.
32. Bouallègue, M., Filée, J., Kharrat, I., et al. 2017, Diversity and evolution of mariner-like elements in aphid genomes, *BMC Genomics*, **18**, 494.
33. Wang, H., Zhao, H., Sun, K., Huang, X., Jin, L. and Feng, J. 2020, Evolutionary basis of high-frequency hearing in the cochleae of echolocators revealed by comparative genomics, *Genome Biol. Evol*., **12**, 3740–53.

34. Holt, C. and Yandell, M. 2011, MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects, *BMC Bioinformatics*, **12**, 491.
35. Haas, B.J., Salzberg, S.L., Zhu, W., et al. 2008, Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments, *Genome Biol.*, **9**, R7.
36. Brůna, T., Hoff, K.J., Lomsadze, A., Stanke, M. and Borodovsky, M. 2021, BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database, *NAR Genom. Bioinform.*, **3**, lqaa108.
37. Emms, D.M. and Kelly, S. 2019, OrthoFinder: phylogenetic orthology inference for comparative genomics, *Genome Biol.*, **20**, 238.
38. Edgar, R.C. 2004, MUSCLE: a multiple sequence alignment method with reduced time and space complexity, *BMC Bioinformatics*, **5**, 113.
39. Capella-Gutiérrez, S., Silla-Martínez, J.M. and Gabaldón, T. 2009, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses, *Bioinformatics*, **25**, 1972–3.
40. Darriba, D., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B. and Flouri, T. 2020, ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models, *Mol. Biol. Evol.*, **37**, 291–4.
41. Kozlov, A.M., Darriba, D., Flouri, T., Morel, B. and Stamatakis, A. 2019, RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference, *Bioinformatics*, **35**, 4453–5.
42. Yang, Z. 2007, PAML 4: phylogenetic analysis by maximum likelihood, *Mol. Biol. Evol.*, **24**, 1586–91.
43. Kumar, S., Stecher, G., Suleski, M. and Hedges, S.B. 2017, TimeTree: a resource for timelines, timetrees, and divergence times, *Mol. Biol. Evol.*, **34**, 1812–9.
44. Soderlund, C., Bomhoff, M. and Nelson, W.M. 2011, SyMAP v3.4: a turnkey synteny system with application to plant genomes, *Nucleic Acids Res.*, **39**, e68.
45. Zhang, J., Nielsen, R. and Yang, Z. 2005, Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level, *Mol. Biol. Evol.*, **22**, 2472–9.
46. Anisimova, M., Bielawski, J.P. and Yang, Z. 2001, Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution, *Mol. Biol. Evol.*, **18**, 1585–92.
47. Yang, Z., Wong, W.S.W. and Nielsen, R. 2005, Bayes empirical Bayes inference of amino acid sites under positive selection, *Mol. Biol. Evol.*, **22**, 1107–18.
48. Webb, A.E., Walsh, T.A. and O'Connell, M.J. 2017, VESPA: very large-scale evolutionary and selective pressure analyses, *PeerJ Comput. Sci.*, **3**, e118.
49. Gao, F., Chen, C., Arab, D.A., Du, Z., He, Y. and Ho, S.Y.W. 2019, EasyCodeML: a visual tool for analysis of selection using CodeML, *Ecol. Evol.*, **9**, 3891–8.
50. The UniProt Consortium. 2019, UniProt: a worldwide hub of protein knowledge, *Nucleic Acids Res.*, **47**, D506–15.
51. Lavery, T.H., Leung, L.K.-P. and Seddon, J.M. 2014, Molecular phylogeny of hipposiderid bats (Chiroptera: Hipposideridae) from Solomon Islands and Cape York Peninsula, Australia, *Zool. Scr.*, **43**, 429–42.
52. Soderlund, C., Nelson, W., Shoemaker, A. and Paterson, A. 2006, SyMAP: a system for discovering and viewing syntenic regions of FPC maps, *Genome Res.*, **16**, 1159–68.
53. Sotero-Caio, C.G., Baker, R.J. and Volleth, M. 2017, Chromosomal Evolution In Chiroptera, *Genes (Basel)*, **8**, 272.
54. Lieberman, J. 2010, Granzyme A activates another way to die, *Immunol. Rev.*, **235**, 93–104.
55. Donninelli, G., Saraf-Sinik, I., Mazziotti, V., et al. 2020, Interleukin-9 regulates macrophage activation in the progressive multiple sclerosis brain, *J. Neuroinflammation*, **17**, 149.
56. Banerjee, A., Baker, M.L., Kulcsar, K., Misra, V., Plowright, R. and Mossman, K. 2020, Novel insights into immune systems of bats, *Front. Immunol.*, **11**, 26.
57. Goh, G., Ahn, M., Zhu, F., et al. 2020, Complementary regulation of caspase-1 and IL-1β reveals additional mechanisms of dampened inflammation in bats, *Proc. Natl. Acad. Sci. U S A*, **117**, 28939–49.
58. Irmler, M., Hertig, S., MacDonald, H.R., et al. 1995, Granzyme A is an interleukin 1 beta-converting enzyme, *J. Exp. Med.*, **181**, 1917–22.
59. Chau, B.N., Cheng, E.H.-Y., Kerr, D.A. and Hardwick, J.M. 2000, Aven, a novel inhibitor of caspase activation, *Mol. Cell.*, **6**, 31–40.
60. van Lith, M., Karala, A.-R., Bown, D., et al. 2007, A developmentally regulated chaperone complex for the endoplasmic reticulum of male haploid germ cells, *Mol. Biol. Cell.*, **18**, 2795–804.
61. Shang, Y., Zhu, F., Wang, L., et al. 2017, Essential role for SUN5 in anchoring sperm head to the tail, *eLife*, **6**, e28199.
62. Arafat, M., Har-Vardi, I., Harlev, A., et al. 2017, Mutation in TDRD9 causes non-obstructive azoospermia in infertile men, *J. Med. Genet.*, **54**, 633–9.
63. Pfeiffer, B. and Mayer, F. 2013, Spermatogenesis, sperm storage and reproductive timing in bats, *J. Zool.*, **289**, 77–85.
64. Linton, D.M. and Macdonald, D.W. 2020, Phenology of reproductive condition varies with age and spring weather conditions in male *Myotis daubentonii* and *M. nattereri* (Chiroptera: Vespertilionidae), *Sci. Rep.*, **10**, 6664.
65. Chang, J.T., Esumi, N., Moore, K., et al. 1999, Cloning and characterization of a secreted frizzled-related protein that is expressed by the retinal pigment epithelium, *Hum. Mol. Genet.*, **8**, 575–83.
66. Li, Y., Rankin, S.A., Sinner, D., Kenny, A.P., Krieg, P.A. and Zorn, A.M. 2008, Sfrp5 coordinates foregut specification and morphogenesis by antagonizing both canonical and noncanonical Wnt11 signaling, *Genes Dev.*, **22**, 3050–63.
67. Matsuyama, M., Aizawa, S. and Shimono, A. 2009, Sfrp controls apicobasal polarity and oriented cell division in developing gut epithelium, *PLoS Genet.*, **5**, e1000427.
68. Bloomquist, R.F., Parnell, N.F., Phillips, K.A., et al. 2015, Coevolutionary patterning of teeth and taste buds, *Proc. Natl. Acad. Sci. U S A*, **112**, E5954–62.
69. Juuri, E., Saito, K., Ahtiainen, L., et al. 2012, Sox2+ stem cells contribute to all epithelial lineages of the tooth via Sfrp5+ progenitors, *Dev. Cell.*, **23**, 317–28.
70. Satoh, W., Matsuyama, M., Takemura, H., Aizawa, S. and Shimono, A. 2008, Sfrp1, Sfrp2, and Sfrp5 regulate the Wnt/β-catenin and the planar cell polarity pathways during early trunk formation in mouse, *Genesis*, **46**, 92–103.
71. Li-Hawkins, J., Gåfvels, M., Olin, M., et al. 2002, Cholic acid mediates negative feedback regulation of bile acid synthesis in mice, *J. Clin. Invest.*, **110**, 1191–200.
72. Smirnov, D. and Vekhnik, V. 2014, Ecology of nutrition and differentiation of the trophic niches of bats (Chiroptera: Vespertilionidae) in floodplain ecosystems of the Samara Bend, *Biol. Bull. Russ. Acad.Sci.*, **41**, 60–70.
73. Machovsky-Capuska, G.E., Miller, M.G.R., Silva, F.R.O., et al. 2018, The nutritional nexus: linking niche, habitat variability and prey composition in a generalist marine predator, *J. Anim. Ecol.*, **87**, 1286–98.
74. Borodulina, O.R. and Kramerov, D.A. 1999, Wide distribution of short interspersed elements among eukaryotic genomes, *FEBS Lett.*, **457**, 409–13.
75. Gogolevsky, K.P., Vassetzky, N.S. and Kramerov, D.A. 2009, 5S rRNA-derived and tRNA-derived SINEs in fruit bats, *Genomics*, **93**, 494–500.
76. Borodulina, O.R. and Kramerov, D.A. 2005, PCR-based approach to SINE isolation: simple and complex SINEs, *Gene*, **349**, 197–205.
77. Dewannieux, M., Esnault, C. and Heidmann, T. 2003, LINE-mediated retrotransposition of marked Alu sequences, *Nat. Genet.*, **35**, 41–8.
78. Kramerov, D.A. and Vassetzky, N.S. 2011, Origin and evolution of SINEs in eukaryotic genomes, *Heredity. (Edinb).*, **107**, 487–95.
79. Cantrell, M.A., Scott, L., Brown, C.J., Martinez, A.R. and Wichman, H.A. 2008, Loss of LINE-1 activity in the megabats, *Genetics*, **178**, 393–404.
80. Koito, A. and Ikeda, T. 2013, Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases, *Front. Microbiol.*, **4**, 28.
81. Varshney, D., Vavrova-Anderson, J., Oler, A.J., Cowling, V.H., Cairns, B.R. and White, R.J. 2015, SINE transcription by RNA polymerase III is

suppressed by histone methylation but not by DNA methylation, *Nat. Commun*., **6**, 6569.
82. Yang, L., Scott, L. and Wichman, H.A. 2019, Tracing the history of LINE and SINE extinction in sigmodontine rodents, *Mob. DNA*, **10**, 22.
83. Novick, P., Smith, J., Ray, D. and Boissinot, S. 2010, Independent and parallel lateral transfer of DNA transposons in tetrapod genomes, *Gene*, **449**, 85–94.
84. Whitaker, J.O. and Black, H. 1976, Food Habits of Cave Bats from Zambia, Africa. *J. Mammal*., **57**, 199–204.
85. Myers, P., Espinosa, R., Parr, C.S., Jones, T., Hammond, G.S. and Dewey, T.A. The Animal Diversity Web (online), *Animal Diversity Web (ADW)* database. https://animaldiversity.org. 10 May 2022, date last accessed).
86. Lee, S.-H., Oshige, M., Durant, S.T., et al. 2005, The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair, *Proc. Natl. Acad. Sci. U S A*, **102**, 18075–80.
87. Cordaux, R., Udit, S., Batzer, M.A. and Feschotte, C. 2006, Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. *Proc. Natl. Acad. Sci. U S A*, **103**, 8101–6.
88. Butelli, E., Garcia-Lor, A., Licciardello, C., et al. 2017, Changes in anthocyanin production during domestication of citrus, *Plant Physiol*., **173**, 2225–42.
89. Platt, R.N., II, Vandewege, M.W., Kern, C., Schmidt, C.J., Hoffmann, F.G. and Ray, D.A. 2014, Large numbers of novel miRNAs originate from DNA transposons and are coincident with a large species radiation in bats, *Mol. Biol. Evol*., **31**, 1536–45.