Functional genomics analysis reveals the evolutionary adaptation and demographic history of slow lorises

Ming-Li Li [1,3*], Sheng Wang [1,*], Penghui Xu [2,*], Hang-Yu Tian[1,4 *], Mixue Bai [2], Ya-Ping Zhang [1, 4], Yong Shao[1], Zi-Jun Xiong[1], Xiao-Guang Qi [6], David N. Cooper [10], Guojie Zhang [1, 5,7,11], He Helen Zhu [2,], Dong-Dong Wu [1, 5,7, 9,] †

[1.] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, 650223, China.

[2.] State Key Laboratory of Oncogenes and Related Genes, Renji-Med X Clinical Stem Cell Research Center, and Department of Urology, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200127, China

[3.] Department of Neurology, the First Affiliated Hospital of Zhengzhou University, Zhengzhou University, Zhengzhou, Henan, China

[4.] Kunming College of Life Science, University of the Chinese Academy of Sciences, Kunming, 650223, China

[5.] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, 650223, China

[6.] Shanxi Key Laboratory for Animal Conservation, College of Life Sciences, Northwest University, Xi'an, China

[7.] Villum Center for Biodiversity Genomics, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Denmark

[8.] National Resource Center for Non-Human Primates, Kunming Primate Research Center, and National Research Facility for Phenotypic & Genetic Analysis of Model Animals (Primate Facility), Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650107, China

[9.] Kunming Natural History Museum of Zoology, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

[10.] Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, CF14 4XN, UK.

[11.]China National Genebank, BGI-Shenzhen, Shenzhen, 518083, China

[*] These authors contributed equally to this work.

† Corresponding author.

Dong-Dong Wu: wudongdong@mail.kiz.ac.cn

Significance

Although strepsirrhines occupy a key node in primate phylogeny, our knowledge of this group remains quite limited. Here, we integrate comparative genomics analyses and functional assays to reveal the genetic underpinnings of evolutionary adaptation in the slow loris, a uniquely special strepsirrhine group. We identified a series of genes that have contributed to certain distinctive adaptive traits of the slow loris, namely *PITRM1* (low metabolic rate), *MYOF* (slow movement) and *PER2* (hibernation). Our findings serve to deepen our understanding of the adaptive evolution of the strepsirrhines, and may also provide useful information for future studies of human disorders related to abnormal metabolism, skeletal muscle development and circadian rhythms.

Abstract

Slow lorises are a group of globally threatened strepsirrhine primates that exhibit many unusual physiological and behavioral features, including a low metabolic rate, slow movement and hibernation, that are rarely seen in other primates. Here, we assemble a chromosome-level genome sequence of the pygmy slow loris (*Nycticebus pygmaeus*) and re-sequence 50 whole genomes from pygmy slow lorises and six whole genomes from Bengal slow lorises (*Nycticebus bengalensis*). We find that many gene families involved in detoxification have been specifically expanded in the slow lorises, including the *GSTA* gene family with many newly derived copies functioning specifically in the liver. We detect many genes displaying evolutionary convergence between slow loris and koala, including *PITRM1*. Significant decreases in PITRM1 enzymatic activity in these two species may have contributed to their characteristic low rate of metabolism. We also detect many evolutionarily convergent genes and positively selected genes in the slow loris that are involved in muscle development. Functional assays demonstrate the decreased ability of one positively selected gene, *MYOF,* to up-regulate the fast-type muscle fiber, consistent with the lower proportion of fast-twitch muscle fibers in the slow loris. The protein product of another positively selected gene in the slow loris, *PER2*, exhibits weaker binding to the key circadian core protein CRY, a finding which may be related to this species' unusual circadian rhythm. Finally, population genomics analysis reveals that these two extant slow loris species, which co-exist in the same habitat after sympatric speciation, have exhibited an inverse relationship in terms of their demography over the past 1 million years, implying strong inter-species competition after speciation.

Introduction

Strepsirrhines occupy a key node in primate phylogeny, and possess characteristics that are considered plesiomorphic to those of the haplorhines (monkeys, apes and humans) (**Fig. 1A**). Although some strepsirrhine genomes have been sequenced (1-4), high-quality genome sequence from this suborder is still rare, and this has impeded studies on the adaptive mechanisms underlying strepsirrhine evolution. Slow lorises (genus *Nycticebus*) are a group of globally threatened strepsirrhines that mainly inhabit the forests of Southeast Asia. Eight species are now recognized; most of which have been listed in the IUCN Red List as Vulnerable or Critically Endangered(5).

Slow lorises have evolved certain distinctive adaptive traits that are rarely seen in other primates(6). For example, slow lorises have a much lower metabolic rate relative to other eutherian species of similar body mass, which is thought to be related to the high concentrations of toxins and digestion inhibitors in their largely exudate diet(7-10). In addition, slow lorises display slow movement and locomotion among the strepsirrhines, and are characterized by a muscle fiber composition which differs from the general condition in mammals(11) which is characterized by a predominance of slow-twitch muscle fibers. These fibers are slow to contract and consume less energy than the white muscle fibers found in most primates, but are better suited to tasks that require endurance such as sustained locomotion(12-14). Slow lorises are the only primates that harbor toxins(15). The toxin in slow lorises is activated when oils exuded from a brachial gland are combined with saliva(16). Slow loris venom is known to function in ectoparasite control and as a protection against predators(17). It is one of the rare lineages that use venom for intraspecific competition(18), and can cause death in other slow lorises and small mammals as well as anaphylactic shock in humans(15). This notwithstanding, slow lorises display resistance to their own toxins. In addition, the pygmy slow loris (PSL, *Nycticebus pygmaeus*) has the capacity to hibernate, a very rare phenomenon in primates. Although these features are highly unusual for a primate, the precise mechanisms underlying their particular characteristics remain largely unknown.

Here, we assemble a chromosome-level genome sequence of the pygmy slow loris (*Nycticebus pygmaeus*) and re-sequence 50 whole genomes from pygmy slow lorises and six whole genomes from Bengal slow lorises (*Nycticebus bengalensis*). The genome sequence has enabled us to obtain unprecedented insights into the unique biology of the slow lorises, without having to harm or disturb an animal of conservation concern. Moreover, revealing the molecular genetics that underpin these unique primates may provide insights relevant to human genetic disease, including disorders of metabolism, skeletal muscle development and circadian rhythms(19, 20). The availability of the genome sequence should also empower a holistic, scientifically grounded approach to slow loris conservation.

Results

Slow loris chromosome-level assembly by long-read sequencing

We set out to study the genetics and genomics of slow lorises and to explore the underlying

evolutionary mechanisms of this highly unusual group of primates. We first sequenced the genome of a female pygmy slow loris (PSL) using single-molecule real-time sequencing technology (PacBio RSII) (Table S1-S3). From the long-read sequences, we assembled the PSL genome into 2 844 contigs, with a contig N50 of 4.73 Mb. An additional 60.82 Gb of Illumina-based short reads were generated to correct sequencing errors, and 254 Gb clean high-throughput chromosome conformation capture (Hi-C) data were used to order the contigs (Figure. S1B). The final assembly was 2.78 Gb distributed across 25 (2n = 50) chromosome-level pseudomolecules with a scaffold N50 of 129.9 Mb (Fig. 1B and Table S4). The PSL genome showed high quality and contiguity relative to other sequenced strepsirrhine genomes(1-4). For example, compared with the mouse lemur (*Microcebus murinus*) genome (1), the PSL genome contained fewer scaffolds (7 679 versus 2 844, respectively) and a higher scaffold N50 (103.22 Mb versus 129.9 Mb). Compared with the bushbaby (*Otolemur garnettii*), the PSL genome had 9 278-fold better contiguity and a 9.4-fold higher scaffold N50 (13.85 Mb versus 129.9 Mb, respectively). Gene annotation identified a total of 26 200 protein-coding genes in the PSL genome (Table S5; Figure S1C); this is somewhat higher than that seen in other primates(4, 21, 22) and may well result from false positives due to the use of *ab initio* predictions(23). Based on the well-annotated protein-coding genes, we performed comparative genomics, including analyses of the expansion/contraction of gene families, positive selection and convergent evolution, in an attempt to reveal some of the potential genetic mechanisms underlying adaptive evolution in slow lorises.

Enhanced detoxifying ability revealed by comparative genomics

Contrary to popular belief that slow lorises are largely frugivorous, slow lorises have recently been found to feed regularly on exudates that they obtain through active gouging, but are also capable of consuming insect prey (24) that contain toxins or digestive deterrents that are suspected to have negative effects on other mammals(7). For example, several of the secondary metabolic products that are commonly consumed by slow lorises contain high levels of heterocyclic compounds, such as phenolics and terpenoids(7). The sequencing of the PSL genome revealed substantial expansion of various gene families involved in the binding and metabolism of heterocyclic compounds (Figure S1D and Table S6). Considering the abundance of heterocyclic compounds that exist as secondary metabolites of the PSL diet (7), we propose that these expanded gene families may enhance the ability of slow lorises to detoxify secondary metabolites in their diet.

By integrating transcriptome data from eight PSL organs (i.e., heart, liver, spleen, stomach, kidney, brain, muscle and tongue), we identified a total of 29 expanded gene families which are highly and specifically expressed in the liver, a vital organ for xenobiotic detoxification and metabolism (Table S7 and Figure S1E). One of the most prominent of these gene families was represented by amplified copies of the glutathione *S*-transferase alpha (*GSTA*) gene (Fig. 2A). This gene encodes an enzyme that functions in phase II detoxification of electrophilic compounds (Fig. 2B), including carcinogens, therapeutic drugs, environmental toxins and products of oxidative stress, in conjugation with glutathione(25). We identified 14 copies of the *GSTA* gene family in the PSL haploid genome, considerably more than in other mammals, including humans, macaques and mice (Fig.2A). Because the 14 *GSTA* copies locate to different

chromosomes in the PSL genome (Figure 1B), they may not be due to tandem duplication. Transcriptome analysis across species revealed that the expression of *GSTA* in PSL liver and stomach was significantly higher than that in humans and mice (Fig. 2C, Table S8). Thus, based on the genome and transcriptome results, we propose that the *GSTA* gene family enhances the ability of PSLs to detoxify secondary metabolites from their diet. Interestingly, we found that the *GSTA* family is also significantly expanded in bushbaby. The majority of the bushbaby diet is rich in secondary metabolites, whilst the basal metabolic rate of bushbaby is 20-40% lower than the mammalian mass-specific standard(26). Thus, expansion of the *GSTA* gene family may also have helped the bushbaby to detoxify secondary metabolites in its diet.

We used the codeml package as implemented in PAML(11) to detect branch-specific positive selection in PSL, and identified a total of 610 genes displaying signals of positive selection (Table S9). From the transcriptome data, we also identified 453 genes specifically and highly expressed in the PSL liver compared with other tissues of this species (Table S10), hereafter termed liver-specific expressed genes. Among the liver-specific expressed genes, four (*FMO5*, *TRPC5*, *GDPD4* and *ABCC9*) exhibited positive selection in the PSL lineage (Figure S1F). Of these, flavin-containing monooxygenase 5 (FMO5) plays an important role in the phase I detoxification and oxygenation of a variety of xenobiotics (Fig.2B) (27). Interestingly, transcriptome analysis showed high expression of positively selected *FMO5* in the PSL liver, consistent with its role in detoxification (Fig. 2D). As is the case in PSL, *FMO5* is also a liver-specific expressed gene in both mouse and human (Figure S1G).

Convergent evolution of *PITRM1* between slow loris and koala may have contributed to their characteristic low metabolic rate

Owing to the high concentrations of toxic secondary compounds they encounter and their low calorific diet, slow lorises must maintain a low basal metabolic rate in order to detoxify the secondary compounds and to conserve energy(7, 28). This strategy is also employed by koalas(29). For this reason, we sought evidence for convergent evolution between the PSL and koala, and identified 226 genes containing amino acid substitutions that showed parallel evolutionary patterns specifically between the two species (Table S11). Of note, the *PITRM1* gene displayed three convergent amino acid substitutions in both the PSL and koala, i.e., K633R, T697S and D925E, which were conserved in the nine other mammals screened (Fig. 2E). Two of these substitutions (K633R, T697S) are also present in the Sunda slow loris (Figure S2A), which is a subspecies of slow loris that also exhibits a low metabolic rate like PSL(7). Our result suggests a similar adaptive evolutionary mechanism at the genetic level between Sunda slow loris and PSL.

PITRM1 is a highly conserved zinc metalloprotease, also known as presequence peptidase (PreP), which plays an important role in metabolism(30). Knockdown of *PITRM1* in mice is reported to decrease basal metabolism and oxygen consumption(31). To test the potential impact of the observed convergent changes in PITRM1 on metabolic function in PSL, we studied the functional consequences of the amino acid substitutions identified in PITRM1 (i.e., K633R, T697S and D925E). Our activation kinetics model showed that PSL PITRM1 exhibited a two-fold decrease in enzyme activity relative to that in humans (Fig. 2F-2G),

suggesting that the substitutions in the PSL PITRM1 sequence contribute to a decrease in basal metabolic rate, similar to *PITRM1* knockdown effect in mice(31).

Adaptive evolution of genes involved in muscle development

Slow lorises display slow movement and locomotion compared with many other mammals and the composition of skeletal muscle fiber in slow lorises differs from that of other primates(12, 14). Notably, slow lorises show a predominance of red (slow-twitch) muscle fibers, which are slow to contract and consume less energy than the white muscle fibers found in most primates(12-14). To explore the genetic mechanism(s) underlying the special skeletal muscle of slow lorises, we used our genomic sequence data to identify amino acid substitutions that correlated with muscle fiber development. Three genes (*MYOF*, *CHRNA1* and *HMCN2*) that are known to play important roles in muscle fiber development were found to be under positive selection in PSLs [$P < 0.05$ by branch site model in PAML] (Fig. 3A) (32). *MYOF*, which encodes the myoferlin protein, is associated with the plasma membrane(33). In humans, dysfunction of *MYOF* can cause abnormal muscle morphology and skeletal muscle fiber degeneration, which affects both proximal and distal muscles(34). 12 out of 24 amino acid positions in MYOF that display positive selection in PSL are also present in Sunda slow loris (Figure S2B), which also exhibits slow movement like PSL. To explore the functional consequences of positive selection on *MYOF*, we performed a functional assay by transfecting mouse myoblasts with PSL *MYOF*, human *MYOF* as well as an empty plasmid as a control. The assay revealed that the ability to up-regulate the expression of the *MyHC2a* (i.e. *MYH2*) and *MyHC2x* (i.e. *MYH1*) genes, which encode fast twitch myosin motor proteins, was weakened in the slow loris *MYOF* as compared to human *MYOF* (Fig. 3B). This finding is consistent with the lower proportion of fast-twitch muscle fibers found in slow lorises.

We also performed convergent evolution analysis between the PSL and the two-toed sloth (*Choloepus hoffmanni*), a mammalian species well known for its slow movement. A total of 20 protein-coding genes showing signals of convergent evolution were identified (Table S12). One of these genes, *NEB*, encodes nebulin, an actin-binding protein which is localized to the thin filament of the sarcomeres in skeletal muscle(35). Mutation of *NEB* is related to muscle weakness and nemaline myopathy in humans(36). Our results show that *NEB* was highly expressed in skeletal muscle in PSL (Figure. S1H). Further, we identified seven nonsynonymous amino acid substitutions which were common to PSL and sloth, that occurred at sites which were conserved in nine other mammals (Fig. 3C). Given its important role in muscle, we propose that these convergent substitutions in the *NEB* gene may well be related to slow movement in both PSL and sloth. Interestingly, we found that 5 of the 7 convergent amino acid positions also display substitutions in the Sunda slow loris (Figure S2C).

Genetic mechanism underlying abnormal circadian rhythm in slow lorises

Low temperatures during the winter months can result in high energetic costs for thermoregulation in slow lorises(37). During the winter, PSLs can enter a state of torpor, i.e., short hibernation, which can last up to 63 h(38). This ability to hibernate is unique in primates

outside of Madagascar and is in all likelihood associated with changes in circadian rhythmicity(39). Here, we identified five genes (*PER2*, *ENOX2*, *PTGER3*, *PROX1* and *SIX3*) that are involved in the regulation of circadian rhythm and which have also been under positive selection in PSL. Genes in the period (PER) gene family encode components of the circadian clock which regulates daily rhythms of locomotor activity, metabolism and behavior(40). PER2 is expressed in a circadian pattern in the suprachiasmatic nucleus, the primary circadian pacemaker in the mammalian brain(41). Mutations in PER2 can shorten, extend or even disrupt circadian rhythms(42). PER2 forms complexes with cryptochromes (CRY) which are required for photic entrainment of circadian rhythm and can negatively regulate the CLOCK/BMAL1-dependent transactivation of clock and clock-controlled genes(43). Here, we found multiple substitutions in PSL PER2, of which S1271A was predicted to affect the binding domain between PER2 and CRY based on 3D structure simulation (Fig. 3D-3E). As expected, a functional co-immunoprecipitation assay confirmed that the ability to bind with CRY was weakened in both PSL PER2 and S1271A-mutated human PER2 compared to human PER2 (35.181±3.867 in PSL; 63.269±1.20 in human, Fig. 3F). In conclusion, our findings suggest that the binding of PER2 to CRY1 has been weakened in PSL, and this may be associated with the unusual circadian rhythmicity in slow lorises.

Population genomics of slow loris

The slow loris is listed as either Vulnerable, Endangered or Critically Endangered by the International Union for Conservation of Nature (IUCN). Genomic variation is important for inferring demographic history and helpful for conservation management. Here, we re-sequenced the whole genomes of 50 PSLs and six Bengal slow lorises (BSLs, *N. bengalensis;* Table S13), which have an overlapping geographic distribution (Fig. 4A). By mapping to the reference genome, we identified 36.3 million high-quality single nucleotide polymorphisms (SNPs) across the two species. A number of analyses, including phylogenetic trees based on nuclear genomes and mtDNA (Figure S3A and Figure S3B), principal component analysis (PCA) (Figure S4A), and population structure by ADMIXTURE (Figure S4B), supported separation of the two species. Historical fluctuations in effective population size (*N*e) reconstructed by the pairwise sequential Markovian coalescent (PSMC) model(44) (Fig. 4B and Figure S4C) and relative cross-coalescence rate by the multiple sequentially Markovian coalescent (MSMC) model (Fig. 4C) showed that PSLs and BSLs diverged more than a million years ago. In addition, PSLs and BSLs have exhibited an inverse relationship in terms of the change in *N*e (effective population size) over the past 1 million years, which may have been caused by spatial competition for the same ecological niche and high-level habitat sharing in Southeast Asia (Fig. 4A). The effective population sizes of the two species have been dramatically reduced over the last 60 kyr (Fig. 4B), suggesting that the reduction in biological diversity has been ongoing for a considerable time, possibly due to human migration or the cold climate during the last glaciation.

Owing to high-level habitat sharing, speciation between PSLs and BSLs has proceeded sympatrically. Thus, we next investigated the pattern of genomic differentiation with regard to sympatric speciation. Genomic divergence between PSLs and BSLs was most pronounced

at the ends of the chromosomes, i.e., the telomeric regions (Figure S4D). This is similar to previous observations on the genomic landscape of species divergence in birds(45) and is probably due to the conserved role of telomeres during meiosis, which cluster on the nuclear envelope to form a 'telomere bouquet' enabling chromosome movements to promote homologous synapsis. Our findings suggest that rapid evolution of the telomeric region may have driven speciation in slow lorises.

Discussion

Our knowledge of strepsirrhines, as basal primates occupying a key node in primate phylogeny, still remains limited. Previous studies of strepsirrhines have been largely confined to macroscopic exploration (13, 24). Here, we present the first chromosome-scale genome assembly for the pygmy slow loris (*Nycticebus pygmaeus*). Comparative genomics combined with functional arrays disclosed the rapid evolution of *PITRM1*, *MYOF* and *PER2*, associated with the low metabolic rate, slow movement and unusual circadian rhythm of slow lorises, respectively.

Slow lorises have a much lower metabolic rate relative to other eutherian species of similar body mass, which is thought to be related to the high concentrations of toxins and digestion inhibitors in their largely exudate diet(7-10). Our findings show that the *GSTA* gene family expanded significantly in the pygmy slow loris, which is likely to be capable of enhanced detoxification of secondary metabolites from their diet. Whilst slow loris venom is harmful to other animals, slow lorises are resistant to their venom, probably due to the enhanced detoxification properties of the slow loris. We found that slow loris PITRM1 exhibited a decrease in enzyme activity relative to that in humans. PITRM1 is a highly conserved zinc metalloprotease which plays an important role in metabolism(30). Knockdown of *PITRM1* in mice is reported to decrease basal metabolism and oxygen consumption(31). Therefore, a decrease of PITRM1 is likely to be functionally associated with a low metabolic rate.

In addition to slow lorises displaying slow movement and locomotion in relation to other strepsirrhines, they have a muscle fiber composition which differs from the general condition evident in mammals(11), being characterized by a predominance of slow-twitch muscle fibers. These fibers are slow to contract and consume less energy than the white muscle fibers found in most primates; in this respect, they are better suited to tasks that require endurance such as sustained locomotion(12-14). We noted rapid evolution of *MYOF* in slow loris, and convergent evolution of the *NEB* gene between slow loris and sloth, which may well be responsible for the unique muscle fiber composition in slow loris.

Finally, it should be emphasized that slow lorises are a highly threatened group of Asian mammals, the threats being primarily due to forest loss and illegal trade(46). Genetic rescue through translocation and the introduction of variation into imperiled populations has been proposed as a powerful means to preserve, or even increase, the genetic diversity and hence the evolutionary potential of endangered species(47-49). Based on our large-scale population genomics analysis, we offer novel insights into both the historic and contemporary population dynamics of slow loris, which provide a theoretical framework to support conservation efforts

pertaining to this very special strepsirrhines species.

Methods

Sampling and DNA and RNA extraction

The skeletal muscle samples from an adult female PSL used for *de novo* genome assembly were provided by the Animal Branch of the Germplasm Bank of Wild Species, Kunming Institute of Zoology, Chinese Academy of Sciences. Eight different tissues, including liver, spleen, stomach, kidney, brain, muscle, heart and tongue, were used for transcriptome analysis. Sample collection was performed in accordance with the methods approved by the Institutional Animal Care and Use Committee (IACUC) at Kunming Institute of Zoology (ID: SMKX-20200118-06). Genomic DNA was extracted from skeletal muscle using a DNeasy Tissue Kit (Qiagen, Germany) according to the manufacturer's recommended instructions. The quality and quantity of the extracted DNA were measured using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, USA), Qubit 2.0 fluorometer (Thermo Fisher Scientific, USA), and agarose gel electrophoresis. Total RNA was extracted from the eight PSL tissues using a RNeasy Plus Universal Kit (Qiagen). The spectrophotometer and fluorometer were used to measure the quality and quantity of extracted RNA, respectively.

PacBio genome sequencing and assembly

All PSL libraries were generated with a SMRTbell Template Prep Kit (PacBio) and sequenced with 67 SMRT cells on the PacBio RSII platform. After filtering out adaptor sequences, low-quality reads and duplicate reads, a total of 136.61 Gb of PacBio sequence data were generated. Read N50 was approximately 14 136 bp, the longest read length was 169 536 bp, and the average length was 8 729 bp. We next used Canu (50) to assemble the genome with the PacBio reads and obtained a 2.89 Gb genome with a contig N50 of 1.51 Mb. We used WTDBG (https://github.com/ruanjue/wtdbg) to assemble the PSL genome, which was 2.82 Gb in size with a contig N50 of 3.84 Mb. The Canu and WTDBG assembly results were merged and optimized using Quickmerge (51). We further generated 60.82 Gb of short-read sequences by Illumina HiSeq 2000 and polished the assembly with Pilon v1.22 (parameters "--mindepth 10 --changes --threads 4 --fix bases.") (52). The final PSL genome was 2.89 Gb in size, with a contig N50 of 4.73 Mb.

Hi-C library preparation and sequencing

Hi-C libraries were prepared as described previously (53) using PSL skeletal muscle tissue from the above individual. Briefly, under adsorption of avidin magnetic beads, DNA with biotin was captured, and the Hi-C library was constructed through a series of steps, including the terminal repair of DNA fragments, addition of A, joint connection, evaluation of polymerase chain reaction (PCR) amplification, and library purification.

Qubit 2.0 was used for preliminary quantification. An Agilent 2100 bioanalyzer was used to check the insert size of the library. Libraries that passed inspection were sequenced using the Illumina HiSeq PE150 platform according to their effective concentrations. Thus, a total of 254 Gb sequence data were obtained.

The original PSL genome sequence draft was 2.89 Gb in length, with a contig N50 of 4.74 Mb and 3 858 contigs in total. After Hi-C-assisted genome assembly, the genome size was 2.78 Gb, contig N50 was 4.7 Mb, super scaffold N50 was 129.9 Mb, and anchor rate of contig length was 96.30%.

Evaluation of assembly results

Genome completeness was assessed by three methods: (1) Benchmarking Universal Single-Copy Orthologs (BUSCO) evaluation, resulting in a BUSCO-estimated PSL genome integrity of 89.06% (complete BUSCOs/total BUSCOs) (Table S14) (54); (2) Core Eukaryotic Genes Mapping Approach (CEGMA) evaluation (55), whereby assembly completeness was assessed by mapping 248 conserved core eukaryotic genes (CGEs) and 244 highly conserved CGEs to the genome (Table S15); (3) Comparative analysis with second-generation reads, with 98.21% of the reference genome mapped to the short sequences (56).

Repeat annotation

Owing to the relatively low level of conservation of repeat sequences between species, the prediction of repeat sequences for the PSL required the construction of a specific repeat sequence database. Therefore, based on structure and *de novo* predictions, we used LTR_FINDER v1.05 (57), MITE-Hunter (58), RepeatScout v1.0.5 (59), and PILER-DF v2.4 (60) to construct a repeat sequence database for the PSL genome. PASTE Classifier (61) was used to classify the database, which was then merged with the Repbase (62) as a final repeat sequence database. We next used RepeatMasker v4.0.6 (63) to predict repeat sequences based on the established repeat sequence database, yielding approximately 1.54 Gb of repeat sequence.

Genome annotation

Protein-coding genes were predicted using *ab initio* prediction, homologous species prediction, and UniGene prediction. GENSCAN (64), Augustus v2.4 (65), GlimmerHMM v3.0.4 (66), GeneID v1.4 (67), and SNAP v2006-07-28 (68) were used for *ab initio* prediction; GeMoMa v1.3.1 (69) was used for predictions based on related primate species, including *Callithrix jacchus*, *Rhinopithecus bieti*, *Homo sapiens* and *Macaca mulatta*; PASA v2.0.2 (70) was used to predict UniGene sequences based on full-length transcriptome data without reference assembly; EvidenceModeler (v2012-06-25) was used to integrate all predicted genes of different strategies (71) with parameters 'Mode: STANDARD, S-ratio: 1.13 score > 1000' and the following weight values: PROTEIN OTHER 50, PROTEIN GeMoMa 50, TRANSCRIPT assembler-PASA 50; ABINITIO PREDICTION Genscan 0.3, ABINITIO PREDICTION Augustus 0.3, ABINITIO PREDICTION GlimmerHMM 0.3, ABINITIO PREDICTION SNAP 0.3, ABINITIO_PREDICTION GeneID 0.3, and OTHER PREDICTION OTHER 100. PASA

was used to modify the final gene models(70). Finally, we obtained 26 200 genes (Table S6). Annotation of the predicted genes was performed by BLASTing the gene (and predicted protein) sequences against several nucleotide and protein sequence databases, including COG (72), KEGG (73), NCBI-NR, and Swiss-Prot (74), with an E value cutoff of $1e^{-5}$ used to assess orthology.

Based on the Rfam (75) and miRBase (76) databases, we used Infernal v1.1 (77) to predict ribosomal RNA (rRNA) and microRNA (miRNA) sequences and used tRNAscan-SE v1.3.1 (78) to identify transfer RNA (tRNA). We annotated 1 457 genes encoding long noncoding RNAs (lncRNAs) which were subdivided into 581 miRNAs, 539 rRNAs and 337 tRNAs.

Phylogenetic tree construction and evolution of gene families

Coding sequences of 11 species (human, chimpanzee, gorilla, orangutan, gibbon, rhesus macaque, marmoset, tree shrew, tarsier, mouse lemur, bush baby) were downloaded from the Ensembl genome browser and, together with that of the PSL, were used to construct gene families. Briefly, we removed CDS sequences with fewer than 30 encoded amino acids, premature stop codons, and nontriplet lengths. TreeFam v7.0 (http://www.treefam.org/) (79) was used to produce gene family categorie. The longest translation form was chosen to represent each gene. We obtained 32 250 gene families and 116 single-copy orthologs. Coding sequences from each single-copy family were aligned by MUSCLE v3.7 (http://www.ebi.ac.uk/Tools/msa/muscle/) (80). Then, RAxML was applied to these sequence sets to build phylogenetic trees (81). MCMCTREE in PAML v4.4 (32) was used to infer divergence times with approximate likelihood calculations using the 116 single-copy gene families (81). Expansion and contraction of gene families in the PSL were determined by CAFÉ v2.1 (82) . A phylogenetic tree based on the RAxML results was used in CAFÉ to infer changes in gene family size (83).

Positively selected genes

The 8 538 single-copy orthologous genes among human, chimpanzee, macaque, marmoset, mouse lemur, PSL, mouse and horse were identified by TreeFam v7.0 (http://www.treefam.org/) (79). The single-copy orthologous gene sequences across the nine species were aligned by MUSCLE (v3.8.31) (https://drive5.com/muscle5/), and low-quality aligned regions were further trimmed by Gblocks (v0.91b) with default parameters. The aligned genes with CDS lengths <100 bp were removed for downstream evolutionary analyses. Based on a reliable constructed species-guided tree topology, the branch-site model in PAML (v4.4) (http://abacus.gene.ucl.ac.uk/software/paml.html) and the likelihood rate test were utilized to detect positively selected genes in the PSL based on the single-copy

orthologous genes with $P < 0.05$ (χ^2 test, corrected for multiple testing by the false discovery rate (FDR)). The Bayes empirical algorithm was applied to calculate the posterior probabilities for inferred positively selected sites.

Convergent evolution

The one-to-one orthologous genes among human, chimpanzee, macaque, marmoset, mouse lemur, mouse, horse, cow, opossum, slow loris and koala (or sloth) were identified by TreeFam (79). The single-copy orthologous gene sequences across the eleven species were aligned by prank (v3.8.31)(84), and low-quality aligned regions were further trimmed by Gblocks (v0.91b) with default parameters(85). The aligned genes with CDS lengths <100 bp were removed for downstream analyses. The ancestral amino acid sequences were reconstructed for orthologs using the codeml program in the PAML package (version 4.7) (32) under the accepted phylogenetic tree. Convergent sites were identified where amino acid positions were identical in the slow loris and koala (or sloth) but were different from that of their respective most recent common ancestor. Then a PERL program that was developed by our laboratory was used to calculate the number of parallel amino acid replacements for a specified pair of branches(86). If the posterior probability of the reconstructed ancestral amino acid site was more than 90%, then they were retained and their states were deemed to be reliable.

Transcriptome analysis

A paired-end sequencing library was constructed from poly (A)+ RNA, as described in the Illumina manual, and sequenced on the Illumina Hiseq 2000 sequencing platform. For each sample, ~5 Gb of data were generated, which were deposited in the Genome Sequence Archive database (http://gsa.big.ac.cn/) under Accession ID CRA003461 for transcriptomes. RNA sequencing reads were aligned with HISAT2 v2.0.3 (87), and gene expression levels were quantified with Cufflinks 2.2.1(88). FPKM (FragmentsPer Kilobase per Million) values of a tissue that were two-fold higher than in other tissues and had an adjusted *P*-value < 0.05 were identified as tissue-specific highly expressed genes.

Measurement of PITRM1 enzyme activity

Total RNA from HEK293T cells was synthesized in order to clone the human *PITRM1* coding sequence (CDS). The *PITRM1* coding sequence of PSL was then derived by *in vitro* mutagenesis from the human *PITRM1* CDS (K642R, T707S, D963E) on the basis that the slow loris and human differ in terms of three PRTIM1 substitutions (K642R, T707S, D963E). The human and PSL *PITRM1* CDSs were flanked with *Xho*I and

*BamH*I restriction sites at their 5' and 3' ends via PCR using primers 5'-tgctcgagATGTGGCGCTGCGGCGGGCGGCA and 3'-ggatccTCATTGGATGATCCAGGATGGGTC. The CDSs were then cloned into the pTXB-6xHis expression vector to express 6xHis fusion proteins. Plasmids were transduced into *E. coli* BL21 (DE3) competent cells (Sangon, Shanghai), monoclone cultured to an OD (optical density) 0.6. Proteins were purified using a His-tag Protein Purification Kit (Beyotime #P2226, China) following the manufacturer's instructions. Proteins were stored in elution buffer (50% glycerol, NaCl 800 mM, Tris-HCl 25 mM, DTT 1 mM) at −80 °C and determined using the Bradford protein concentration assay (ThermoFisher #23200). The Aβ-derivative fluorogenic substrate Mca-KLVFFAEDK(Dnp)-OH was synthesized as described previously (89). FERT (fluorescence resonance energy transfer) assays for human PITRM1 and slow loris PITRM1 were performed as described previously (89). Briefly, each assay contained 2 nM enzyme and 10 μM substrate in 50 mM potassium phosphate buffer, pH7.0. The enzymatic assay was performed using a BioTek Synergy HT plate reader (excitation: 350 nm, emission: 405 nm, BioTek USA).

Functional validation of MYOF in unusual PSL musculature

Neonatal male mice (within 48 hours after birth) were sacrificed by cervical dislocation and disinfected with 70% ethanol, after which muscle tissue was taken for experimentation. Muscle tissues were cut into pieces and digested with collagenase at 37 °C for 60 min. Cell suspensions were successively passed through 100μm, 70μm and 40μm cell filters, and centrifuged at 300g for 5min. The precipitates were resuspended with complete medium and inoculated into 6-well plates for 30min at 37℃.

After the supernatant had been transferred to the new well and repeated three consecutive times, it was placed in petri dishes coated with poly-D-lysine and cultured at 37℃. The medium was changed every 2 days.

The plasmids CMV-MCS-polyA-EF1A-zsGreen-sv40-puromycin with slow loris and human MYOF coding sequences were transfected. At the logarithmic phase, the mixture was added to the culture dish with LIPO2000 transfection reagent according to the manufacturer's protocol. After incubation for 12 hours, fresh complete medium was replaced. The transfection efficiency was observed the next day. qPCR was used to detect the relative expression levels of MyHC2a, MyHC2b, MyHC2x and MyHC-slow before and after transfection.

Per2 functional assays

PER2(*Human*)-3×HA, PER2(*Human*,S1155A)-3×HA, and PER2(*Loris*)-3×HA were digested with *BamH*I and *Xho*I restriction enzymes. CRY1(*Human*)-3×FLAG was digested with *Kpn*I and *Not*I restriction enzymes. PER2(*Human*)-3×HA, PER2(*Human*,S1155A)-3×HA, PER2(*Loris*)-3×HA, and CRY1(*Human*)-3×FLAG were individually subcloned into the pCDNA3.1 vector.

The HEK293Tells were collected and lysed in RIPA buffer containing protease inhibitors (cocktail). In total, 5% of the cell extracts were retained for input, while the remainder were incubated with mouse IgG (ABclonal, #AC005) and anti-FLAG antibodies (ABclonal, #AE004) overnight at 4 °C. For immunoprecipitation of protein complexes, cell extracts were pre-cleared with protein-A/G beads (88803# Thermo Scientific) and incubated with the indicated antibodies for 2 h at 4 °C. The beads were washed three times with RIPA buffer, and the bound proteins were eluted by boiling in 1 × loading buffer and subjected to immunoblotting with the indicated antibodies.

Resequencing the slow loris genomes

We re-sequenced the whole genomes of 56 individuals, including fifty PSL and six BSL based on the existing sample information we have (Table S13). The majority of slow lorises came from Yunnan province, a southwest region of China. Only two of the slow lorises came from Vietnam. All the slow lorises in this study were sampled from the wild and taken after natural death; a few samples came from zoos. Samples were stored in the -80°C refrigerator. Combining information from 16S rRNA and COI (cytochrome c oxidase subunit I) genes as well as morphological feature of the slow loris, we distinguished the BSL and PSL species. The muscle and skin tissue of slow loris were used for genome resequencing. In total, 10 μg genomic DNA, prepared by the standard CTAB (cetyltrimethylammonium bromide) extraction protocol, were used to construct libraries with a 350-bp insert size. The sequence libraries were constructed according to the Illumina library preparation pipeline on the Illumina Hiseq 4000 platform, generating 150-bp paired-end reads.

All clean short reads were aligned against the assembled pygmy slow loris reference genome using BWA-MEM v0.7.12 (90). The aligned BAM files were sorted, and PCR duplicated reads were removed using SAMtools v1.3.1 (91). RealignerTargetCreator and IndelRealigner in the Genome Analysis Toolkit (GATK) v3.7.0 (92) were used for local realignment around indels. SNPs were genotyped using GenotypeGVCFs from the GATK package with default parameters after calling GVCFs (Genomic VCFs) in HaplotypeCaller. Finally, we applied the following criteria to all SNPs for hard filtering: QUAL < 30, QD < 2.0, MQ < 40.0, MQRankSum <−12.5, ReadPosRankSum < −8.0, and SOR > 4.0. Genotypes were imputed and phased with BEAGLE v4.1 (93).

Population genetic analysis

A neighbor-joining tree was constructed using the APE (94) package from the pair-wise identical-by-state (IBS) distance matrix calculated using PLINK 1.9 (95) and based on autosomal SNPs. After pruning the SNPs based on linkage disequilibrium using PLINK, we performed genetic structure clustering of autosomal SNPs using ADMIXTURE (96) and principal component analysis (PCA) with smartpca in EIGENSOFT v5.0.2 (97).

Demographic analysis

To reconstruct the demographic history of slow lorises, we applied the PSMC method based on high-depth (>20x) sequenced individuals. The parameters were set to "-N25 -t15 -r5 -p 4+25x+4+6". We applied 100 bootstraps by randomly sampling and replacing 5-Mb fragments of the consensus sequence. To infer more recent demographic history and divergence time between BSLs and PSLs, we applied MSMC v2.0 (98) to multiple individuals based on default parameters. MSMC input was created from all site VCF files genotyped from the GVCF files. MSMC was run on a four haplotype model by randomly selecting individuals. For all demographic analyses, the neutral mutation rate (μ) was set to 2.2×10^{-8} mutations per site per generation (99) and generation time (g) was set to 5 years.

Population divergence

We used the genetic differentiation index F_{ST} and absolute divergence Dxy to identify highly differentiated regions for the BSL and PSL species. We used vcftools v0.1.12 (100) to calculate average F_{ST} and $\theta\pi$ with 50-kb sliding windows and a step size of 25 kb. Dxy was calculated using an in-house python script. We defined potentially selective windows with the top 5% of values for F_{ST} and Dxy, and two-tailed top 5% of values for the $\theta\pi$-ratio. The outlier windows were annotated using snpEff (101), and Gene Ontology (GO) functional enrichment analysis was performed using the topGO package v2.40.0 (102) with Fisher's exact test.

Mitochondrial genome analysis

The Perl script NOVOPlasty v2.7.2 (103) was used to assemble individual mitochondrial genomes from paired-end short reads. The complete PSL and BSL mitochondrial genomes (GenBank accession numbers: KX397281 and MG515246) were used as seed and reference sequences for PSL and BSL mitochondrial genome assembly, respectively. The k-mer parameter was set to 35, unless assembly failure occurred, in which case it was reduced to 23–31. Only complete circularized mitochondrial genomes were kept for downstream analysis. We downloaded all published complete mitochondrial genomes for slow lorises, with the slender loris being used as the outgroup, from GenBank. Clustal Omega v1.2.0 (104) was applied for multiple sequence alignment after rotating the mitochondrial sequences using Cyclic DNA Sequence Aligner (105). Finally, we built the maximum-likelihood tree based on whole mitochondrial sequences using the Tamura-Nei model (106) in MEGA v7 (107).

References

1. P. A. Larsen *et al.*, Hybrid *de novo* genome assembly and centromere characterization of the gray mouse lemur (*Microcebus murinus*). *BMC Biology* **15**, 110 (2017).
2. M. T. R. Hawkins *et al.*, Genome sequence and population declines in the critically endangered greater bamboo lemur (*Prolemur simus*) and implications for conservation. *BMC Genomics* **19**, 445 (2018).
3. W. K. Meyer *et al.*, Evolutionary history inferred from the *de novo* assembly of a nonmodel organism, the blue-eyed black lemur. *Molecular Ecology* **24**, 4392-4405 (2015).
4. J. Schmitz *et al.*, Genome sequence of the basal haplorrhine primate *Tarsius syrichta* reveals unusual insertions. *Nat Commun* **7**, 12997-12997 (2016).
5. K. A. I. Nekaris, Extreme primates: Ecology and evolution of Asian lorises. *Evolutionary Anthropology: Issues, News, and Reviews* **23**, 177-187 (2014).
6. D. J. Chivers, Primates of the World: Distribution, Abundance and Conservation. Jaclyn H. Wolfheim, University of Washington Press, Seattle and London, 1983, $46·00. *Oryx* **18**, 252-253 (1984).
7. F. Wiens, A. Zitzmann, N. A. Hussein, Fast Food for Slow Lorises: Is Low Metabolism Related to Secondary Compounds in High-Energy Plant Diet? *Journal of Mammalogy* **87**, 790-798 (2006).
8. C. Xiao *et al.*, Energy metabolism and thermoregulation in pygmy lorises (*Nycticebus pygmaeus*) from Yunnan Daweishan Nature Reserve. *Acta Ecologica Sinica* **30**, 129-134 (2010).
9. C. Starr, K. A. I. Nekaris, Obligate exudativory characterizes the diet of the pygmy slow loris Nycticebus pygmaeus. *American Journal of Primatology* **75**, 1054-1061 (2013).
10. N. Swapna, S. Radhakrishna, A. K. Gupta, A. Kumar, Exudativory in the Bengal slow loris (*Nycticebus bengalensis*) in Trishna Wildlife Sanctuary, Tripura, northeast India. *American Journal of Primatology* **72**, 113-121 (2010).
11. T. Kimura, H. Kumakura, S. Inokuchi, H. Ishida, Composition of muscle fibers in the slow loris, using the m. biceps brachii as an example. *Primates* **28**, 525-532 (1987).
12. M. L. Boettcher, K. C. Leonard, E. Dickinson, A. Herrel, A. Hartstone-Rose, Extraordinary grip strength and specialized myology in the hyper-derived hand of *Perodicticus potto*? *Journal of Anatomy* **235**, 931-939 (2019).
13. A. Gyambibi, P. Lemelin, Comparative and Quantitative Myology of the Forearm and Hand of Prosimian Primates. *The Anatomical Record* **296**, 1196-1206 (2013).
14. Y. Higurashi, Y. Taniguchi, H. Kumakura, Density of Muscle Spindles in Prosimian Shoulder Muscles Reflects Locomotor Adaptation. *Cells Tissues Organs* **184**, 96-101 (2006).
15. N. B. Grow, Wirdateti, K. A. I. Nekaris, Does toxic defence in Nycticebus spp. relate to ectoparasites? The lethal effects of slow loris venom on arthropods. *Toxicon : official journal of the International Society on Toxinology* **95**, 1-5 (2015).
16. K. A.-I. Nekaris, R. S. Moore, E. J. Rode, B. G. Fry, Mad, bad and dangerous to know: the biochemistry, ecology and evolution of slow loris venom. *Journal of Venomous Animals and Toxins including Tropical Diseases* **19**, 21 (2013).
17. N. B. Grow, Wirdateti, K. A. I. Nekaris, Does toxic defence in Nycticebus spp. relate to ectoparasites? The lethal effects of slow loris venom on arthropods. *Toxicon* **95**, 1-5 (2015).
18. K. A. I. Nekaris *et al.*, Slow lorises use venom as a weapon in intraspecific competition. *Current Biology* **30**, R1252-R1253 (2020).
19. M. L. Benton *et al.*, The influence of evolutionary history on human health and disease.

Nature Reviews Genetics **22**, 269-283 (2021).

20. K. D. Reinhardt, V. V. Vyazovskiy, R. A. Hernandez-Aguilar, M. A. Imron, K. A.-I. Nekaris, Environment shapes sleep patterns in a wild nocturnal primate. http://dx.doi.org/10.1038/s41598-019-45852-2.
21. K. C. Worley *et al.*, The common marmoset genome provides insight into primate biology and evolution. *Nature Genetics* **46**, 850-857 (2014).
22. X. Zhou *et al.*, Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history. *Nature Genetics* **46**, 1303-1310 (2014).
23. J. Wang *et al.*, Vertebrate gene predictions and the problem of large genes. *Nature Reviews Genetics* **4**, 741-749 (2003).
24. U. Streicher, A. Wilson, R. L. Collins, K. A.-I. Nekaris, "Exudates and Animal Prey Characterize Slow Loris (*Nycticebus pygmaeus, N. coucang* and *N. javanicus*) Diet in Captivity and After Release into the Wild" in Leaping Ahead: Advances in Prosimian Biology, J. Masters, M. Gamba, F. Génin, Eds. (Springer New York, New York, NY, 2013), 10.1007/978-1-4614-4511-1_19, pp. 165-172.
25. B. F. Coles, F. F. Kadlubar, Detoxification of electrophilic compounds by glutathione S-transferase catalysis: Determinants of individual response to chemical carcinogens and chemotherapeutic drugs? *BioFactors* **17**, 115-130 (2003).
26. E. F. Müller, Basal metabolic rates in primates—the possible role of phylogenetic and ecological factors. *Comparative Biochemistry and Physiology Part A: Physiology* **81**, 707-711 (1985).
27. Y. Chen, N. R. Zane, D. R. Thakker, M. Z. Wang, Quantification of Flavin-containing Monooxygenases 1, 3, and 5 in Human Liver Microsomes by UPLC-MRM-Based Targeted Quantitative Proteomics and Its Application to the Study of Ontogeny. *Drug Metabolism and Disposition* **44**, 975 (2016).
28. C. Xiao *et al.*, Energy metabolism and thermoregulation in pygmy lorises (*Nycticebus pygmaeus*) from Yunnan Daweishan Nature Reserve. *Acta Ecologica Sinica* **30**, 129-134 (2010).
29. K. A. Nagy, R. W. Martin, Field Metabolic Rate, Water Flux, Food Consumption and Time Budget of Koalas, *Phascolarctos cinereus* (Marsupialia: Phascolarctidae) in Victoria. *Australian Journal of Zoology* **33**, 655-665 (1985).
30. L. Town *et al.*, The metalloendopeptidase gene Pitrm1 is regulated by hedgehog signaling in the developing mouse limb and is expressed in muscle progenitors. *Developmental Dynamics* **238**, 3175-3184 (2009).
31. C. J. Bult *et al.*, Mouse Genome Database (MGD) 2019. *Nucleic Acids Res* **47**, D801-D806 (2019).
32. Z. Yang, PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**, 1586-1591 (2007).
33. Y. Dong *et al.*, Myoferlin, a Membrane Protein with Emerging Oncogenic Roles. *BioMed Research International* **2019**, 7365913 (2019).
34. A. Kiselev *et al.*, Truncating Variant in Myof Gene Is Associated With Limb-Girdle Type Muscular Dystrophy and Cardiomyopathy. *Front Genet* **10**, 608-608 (2019).
35. M. W. Lawlor *et al.*, Novel mutations in *NEB* cause abnormal nebulin expression and markedly impaired muscle force generation in severe nemaline myopathy. *Skeletal Muscle* **1**, 23 (2011).

36. M. Scoto *et al.*, Nebulin (NEB) mutations in a childhood onset distal myopathy with rods and cores uncovered by next generation sequencing. *European Journal of Human Genetics* **21**, 1249-1252 (2013).
37. U. Streicher, Aspects of Ecology and Conservation of the Pygmy *Loris nycticebus pygmaeus* in Vietnam. (2004).
38. T. Ruf, U. Streicher, G. L. Stalder, T. Nadler, C. Walzer, Hibernation in the pygmy slow loris (*Nycticebus pygmaeus*): multiday torpor in primates is not restricted to Madagascar. *Scientific Reports* **5**, 17392 (2015).
39. K. D. Reinhardt, V. V. Vyazovskiy, R. A. Hernandez-Aguilar, M. A. Imron, K. A.-I. Nekaris, Environment shapes sleep patterns in a wild nocturnal primate. *Scientific Reports* **9**, 9939 (2019).
40. B. Zheng *et al.*, The mPer2 gene encodes a functional component of the mammalian circadian clock. *Nature* **400**, 169-173 (1999).
41. S. Masuda *et al.*, Mutation of a PER2 phosphodegron perturbs the circadian phosphoswitch. *Proc Natl Acad Sci U S A* **117**, 10888-10896 (2020).
42. P. L. Lowrey, J. S. Takahashi, Genetics of circadian rhythms in Mammalian model organisms. *Adv Genet* **74**, 175-230 (2011).
43. S. Langmesser, T. Tallone, A. Bordon, S. Rusconi, U. Albrecht, Interaction of circadian clock proteins PER2 and CRY with BMAL1 and CLOCK. *BMC Molecular Biology* **9**, 41 (2008).
44. H. Li, R. Durbin, Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493-496 (2011).
45. H. Ellegren *et al.*, The genomic landscape of species divergence in Ficedula flycatchers. *Nature* **491**, 756-760 (2012).
46. K. A. I. Nekaris, C. R. Starr, OVERVIEW: Conservation and ecology of the neglected slow loris: priorities and prospects. *Endangered Species Research* **28**, 87-95 (2015).
47. S. Feng *et al.*, The Genomic Footprints of the Fall and Recovery of the Crested Ibis. *Curr Biol* **29**, 340-349.e347 (2019).
48. A. R. Whiteley, S. W. Fitzpatrick, W. C. Funk, D. A. Tallmon, Genetic rescue to the rescue. *Trends in Ecology & Evolution* **30**, 42-49 (2015).
49. A. R. Weeks *et al.*, Genetic rescue increases fitness and aids rapid recovery of an endangered marsupial population. *Nat Commun* **8**, 1071 (2017).
50. S. Koren *et al.*, Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* **27**, 722-736 (2017).
51. M. Chakraborty, J. G. Baldwin-Brown, A. D. Long, J. J. Emerson, Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. *Nucleic Acids Res* **44**, e147-e147 (2016).
52. B. J. Walker *et al.*, Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**, e112963-e112963 (2014).
53. E. Lieberman-Aiden *et al.*, Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293 (2009).
54. F. Simão, R. Waterhouse, P. Ioannidis, E. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics (Oxford, England)* **31** (2015).
55. G. Parra, K. Bradnam, I. Korf, CEGMA: a pipeline to accurately annotate core genes in

eukaryotic genomes. *Bioinformatics* **23**, 1061-1067 (2007).
56. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv* **1303** (2013).
57. Z. Xu, H. Wang, LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res* **35**, W265-W268 (2007).
58. Y. Han, S. R. Wessler, MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res* **38**, e199-e199 (2010).
59. A. L. Price, N. C. Jones, P. A. Pevzner, *De novo* identification of repeat families in large genomes. *Bioinformatics* **21**, i351-i358 (2005).
60. R. C. Edgar, E. W. Myers, PILER: identification and classification of genomic repeats. *Bioinformatics* **21**, i152-i158 (2005).
61. T. Wicker *et al.*, A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* **8**, 973-982 (2007).
62. J. Jurka *et al.*, Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research* **110**, 462-467 (2005).
63. M. Tarailo-Graovac, N. Chen, Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics* **Chapter 4**, Unit 4.10 (2009).
64. C. Burge, S. Karlin, Prediction of complete gene structures in human genomic DNA11Edited by F. E. Cohen. *Journal of Molecular Biology* **268**, 78-94 (1997).
65. M. Stanke, S. Waack, Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215-ii225 (2003).
66. W. H. Majoros, M. Pertea, S. L. Salzberg, TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878-2879 (2004).
67. E. Blanco, G. Parra, R. Guigó, Using geneid to Identify Genes. *Current Protocols in Bioinformatics* **18**, 4.3.1-4.3.28 (2007).
68. I. Korf, Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
69. J. Keilwagen *et al.*, Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res* **44**, e89-e89 (2016).
70. M. A. Campbell, B. J. Haas, J. P. Hamilton, S. M. Mount, C. R. Buell, Comprehensive analysis of alternative splicing in rice and comparative analyses with *Arabidopsis*. *BMC Genomics* **7**, 327-327 (2006).
71. B. J. Haas *et al.*, Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol* **9**, R7-R7 (2008).
72. R. L. Tatusov *et al.*, The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**, 41 (2003).
73. M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).
74. B. Boeckmann *et al.*, The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* **31**, 365-370 (2003).
75. S. Griffiths-Jones *et al.*, Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res* **33**, D121-D124 (2005).
76. S. Griffiths-Jones, R. J. Grocock, S. van Dongen, A. Bateman, A. J. Enright, miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* **34**, D140-D144 (2006).
77. E. P. Nawrocki, S. R. Eddy, Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*

(Oxford, England) **29**, 2933-2935 (2013).

78. T. M. Lowe, S. R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**, 955-964 (1997).
79. H. Li *et al.*, TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res* **34**, D572-D580 (2006).
80. R. C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113-113 (2004).
81. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* **30**, 1312-1313 (2014).
82. T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269-1271 (2006).
83. S. B. Hedges, J. Marin, M. Suleski, M. Paymer, S. Kumar, Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution* **32**, 835-845 (2015).
84. A. Löytynoja, "Phylogeny-aware alignment with PRANK" in Multiple Sequence Alignment Methods, D. J. Russell, Ed. (Humana Press, Totowa, NJ, 2014), 10.1007/978-1-62703-646-7_10, pp. 155-170.
85. G. Talavera, J. Castresana, Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Systematic Biology* **56**, 564-577 (2007).
86. Y. Shao *et al.*, Genetic adaptations of the plateau zokor in high-elevation burrows. *Scientific Reports* **5**, 17262 (2015).
87. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**, 907-915 (2019).
88. A. Roberts, C. Trapnell, J. Donaghey, J. L. Rinn, L. Pachter, Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* **12**, R22-R22 (2011).
89. J. E. Smith-Carpenter, B. J. Alper, Functional requirement for human pitrilysin metallopeptidase 1 arginine 183, mutated in amyloidogenic neuropathy. *Protein Sci* **27**, 861-873 (2018).
90. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. https://doi.org/10.48550/arXiv.1303.3997 (2013)
91. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-2993 (2011).
92. M. A. DePristo *et al.*, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**, 491-+ (2011).
93. S. R. Browning, B. L. Browning, Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* **81**, 1084-1097 (2007).
94. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289-290 (2004).
95. C. C. Chang *et al.*, Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
96. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655-1664 (2009).

97. A. L. Price *et al.*, Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-909 (2006).
98. S. Schiffels, R. Durbin, Inferring human population size and separation history from multiple genome sequences. *Nat Genet* **46**, 919-925 (2014).
99. S. Kumar, S. Subramanian, Mutation rates in mammalian genomes. *Proc Natl Acad Sci U S A* **99**, 803-808 (2002).
100. P. Danecek *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
101. P. Cingolani *et al.*, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80-92 (2012).
102. A. Alexa, J. Rahnenfuhrer, T. Lengauer, Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600-1607 (2006).
103. N. Dierckxsens, P. Mardulyn, G. Smits, NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Res* **45**, e18 (2017).
104. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
105. F. Fernandes, L. Pereira, A. T. Freitas, CSA: an efficient algorithm to improve circular DNA multiple alignment. *BMC Bioinformatics* **10**, 230 (2009).
106. K. Tamura, M. Nei, Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10**, 512-526 (1993).
107. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870-1874 (2016).
108. J. L. Fribourgh *et al.*, Dynamics at the serine loop underlie differential affinity of cryptochromes for CLOCK:BMAL1 to control circadian timing. *Elife* **9**, e55275 (2020).

Acknowledgments

We thank the Primate Laboratory Animal Center from the Kunming Institute of Zoology, Chinese Academy of Sciences; Kunming Natural History Museum of Zoology, Kunming Institute of Zoology, Chinese Academy of Sciences and Animal Branch of the Germplasm Bank of Wild Species, Chinese Academy of Sciences (the Large Research Infrastructure Funding) for providing the sample of slow loris. This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDPB17), the National Natural Science Foundation of China (31822048), Yunnan Fundamental Research Project (2019FI010) and the Animal Branch of the Germplasm Bank of Wild Species of Chinese Academy of Science (the Large Research Infrastructure Funding).

Author contributions

D.D.W. led the project. D.D.W. and Z.H. designed and conceived the study. D.D.W., L.M.L., W.S., D.N.C. and G.J.Z. drafted the manuscript. L.M.L.,W.S., S.Y. and Z.J.X. performed the data analysis. P.H.X., H.Y.T., M.X.B. and L.M.L performed the functional experiments. X.G.Q. provided materials and performed some experiments.

Competing financial interests

The authors declare no competing financial interest.

Data availability: Genome assemblies, DNA sequencing data and RNA sequencing data have been deposited into the GSA database with project no. PRJCA003786. The *de novo* genome of PSLs is under Accession ID GWHBCHX00000000. The Re-sequenced genomes of PSLs and BSLs are under Accession ID CRA003477. Transcriptome data were deposited in the Genome Sequence Archive database (http://gsa.big.ac.cn/) under Accession ID CRA003461.

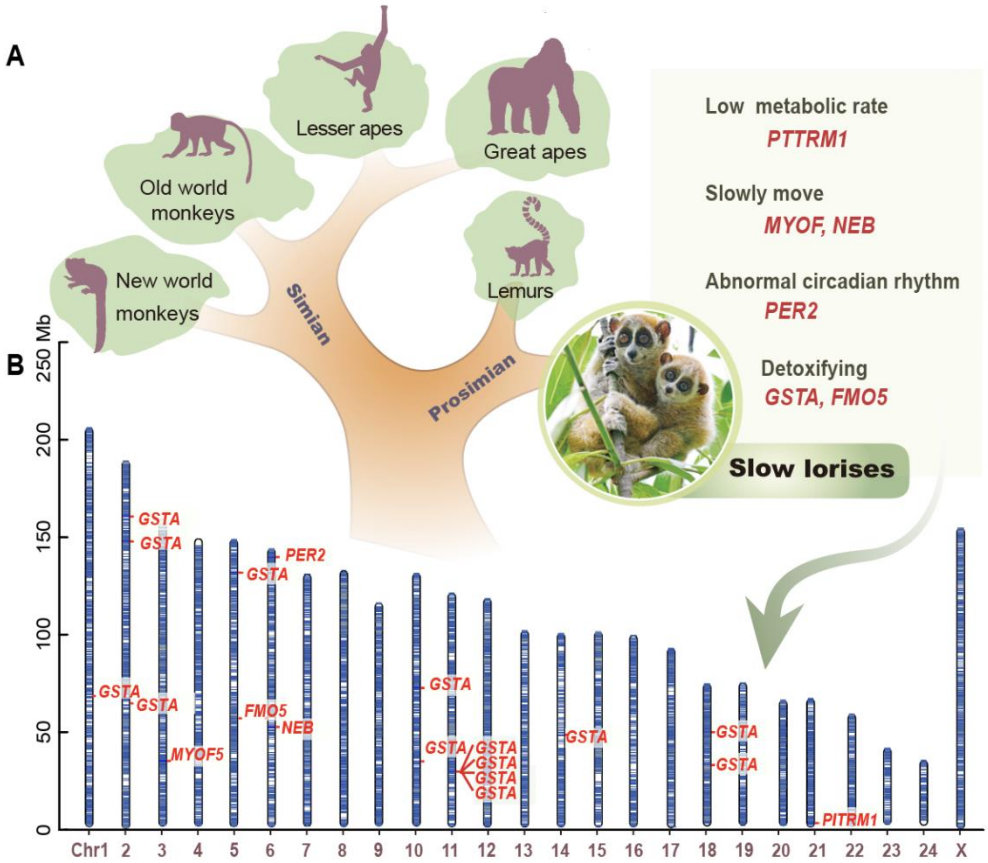

Fig.1. Study schematic. A Phylogenetic tree of primates and the crucial genes contribute to the specific phenotypes of slow loris. **B.** Karyotype of pygmy slow loris, and location of crucial genes at different chromosomes.

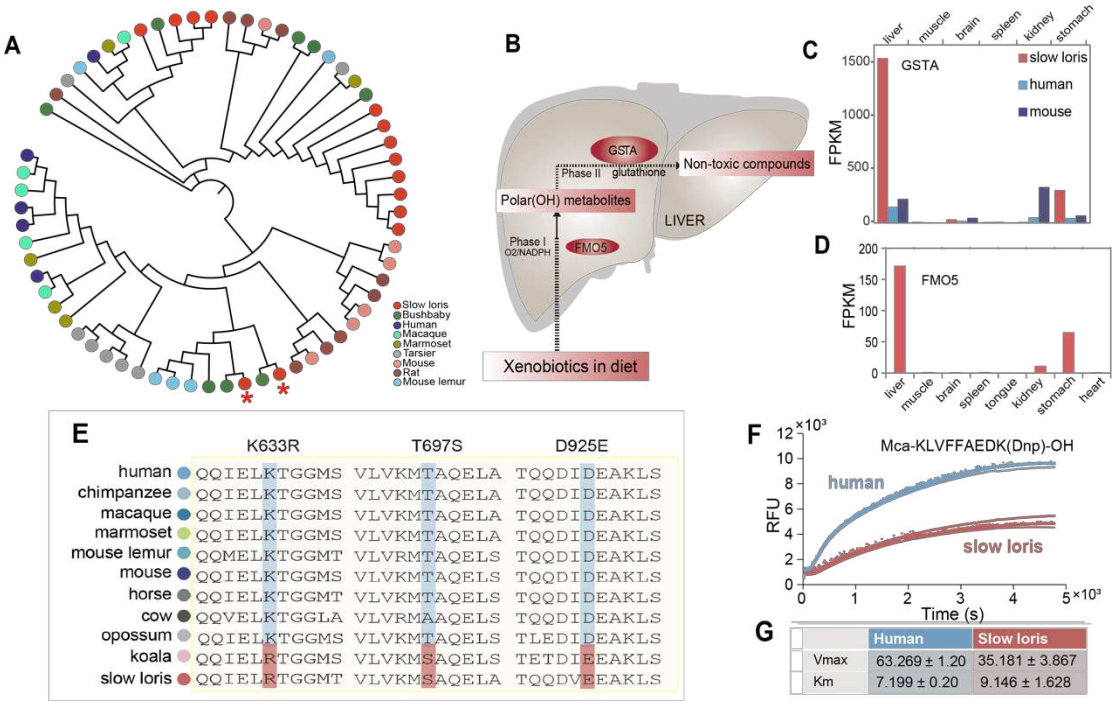

Fig.2. Functional genomics study of PSL. A. Unrooted neighbor-joining tree showing the *GSTA* gene families across species, with eight copies in *N. pygmaeus* and relatively few copies in other species. Red stars mark two *GSTA* genes showing highly expression levels in slow loris liver. **B.** *FMO5* and *GSTA* are involved in liver detoxification. **C.** Expression levels (FPKM) of *GSTA* in six tissues from humans, mice and slow lorises are presented. **D.** Expression levels (FPKM) of *FMO5* in eight tissues from slow loris. **E.** The alignment of mammalian PITRM1 amino acid sequences was suggestive of three convergent substitutions between koala and PSL. Amino acids unique to PSL and koala are shown in green; other amino acids at these positions are shown in orange. **F.** Reaction progress curves depicting sample fluorescence (RFU) as a function of time(s) for peptide hydrolysis assays conducted using Mca-KLVFFAEDK(Dnp)-OH, an established target of PITRM1 activity. **G.** Vmax and Km values of human and slow loris PITRM1 enzymes.

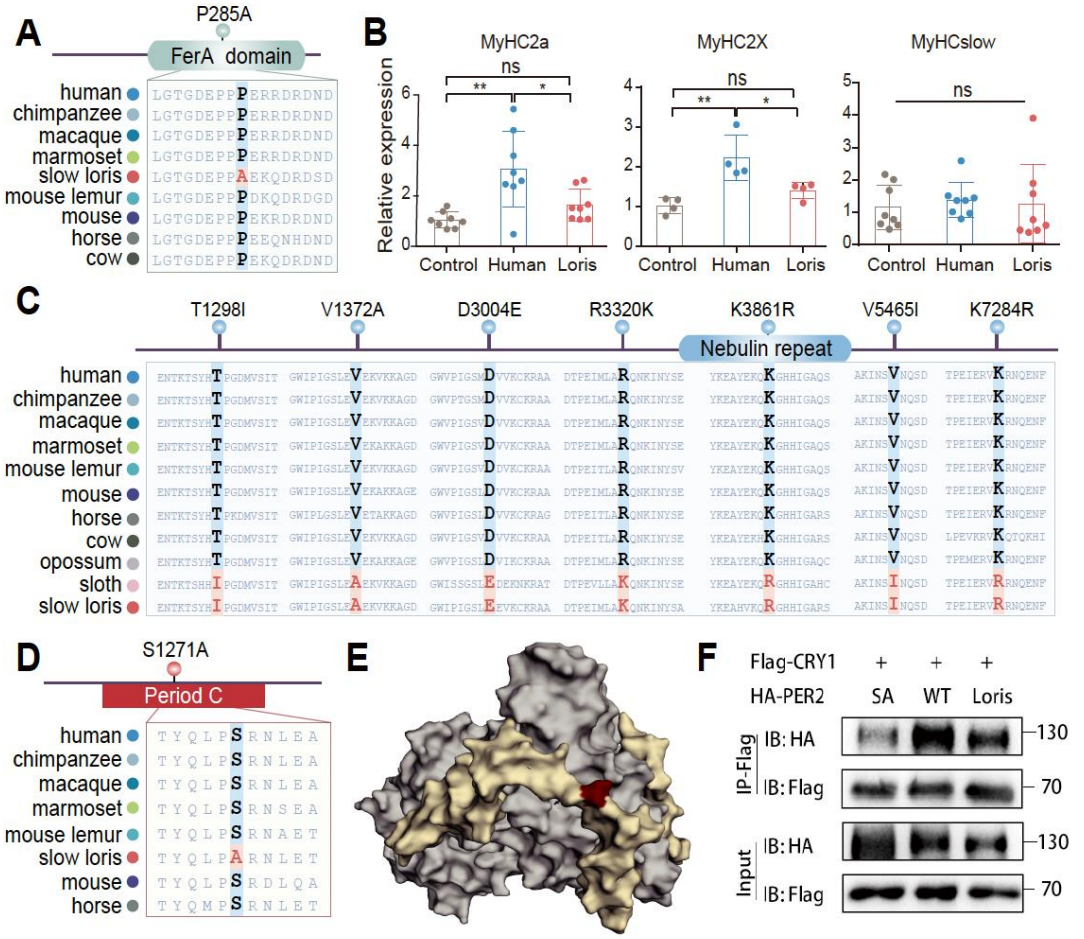

Fig. 3. Genetic mechanism underlying muscle evolution and abnormal circadian rhythm in PSL identified by functional genomics. A. Positive selection of *MYOF*, and alignment of protein sequence surrounding the positively selected site (P285A). **B.** PSL *MYOF* gene significantly inhibited the expression of the gene encoding the fast-twitch muscle fiber, as compared to the human *MYOF* gene. **P < 0.01, *P < 0.05. **C.** Convergent evolution of the NEB protein between PSL and two-toed sloth (*Choloepus hoffmanni*). The alignment of mammalian NEB amino acid sequences is presented. Amino acids unique to PSL and sloth are shown in green; alternative amino acids at these positions in other mammalian NEB orthologues are shown in orange. **D.** Alignment of mammalian amino acid sequences of period domain of PER2. Candidate amino acid substitution (S1271A) in PSL is shown in blue; other amino acids at these positions are shown in red. **E.** 3D structure model (PDB ID: 6OF7)(108) shows that the S1271A substitution in PER2 is located within the binding domain of the PER2-CRY complex. The grey shading represents PER2 whereas the light yellow shading represents CRY1. The site of this substitution is depicted in red. **F.** Co-immunoprecipitation assay identified the weakened binding ability of PSL PER2 with CRY1, as compared with human. WT is human Per2; SA is S1271A mutated human Per2, Loris is Per2 of PSL.

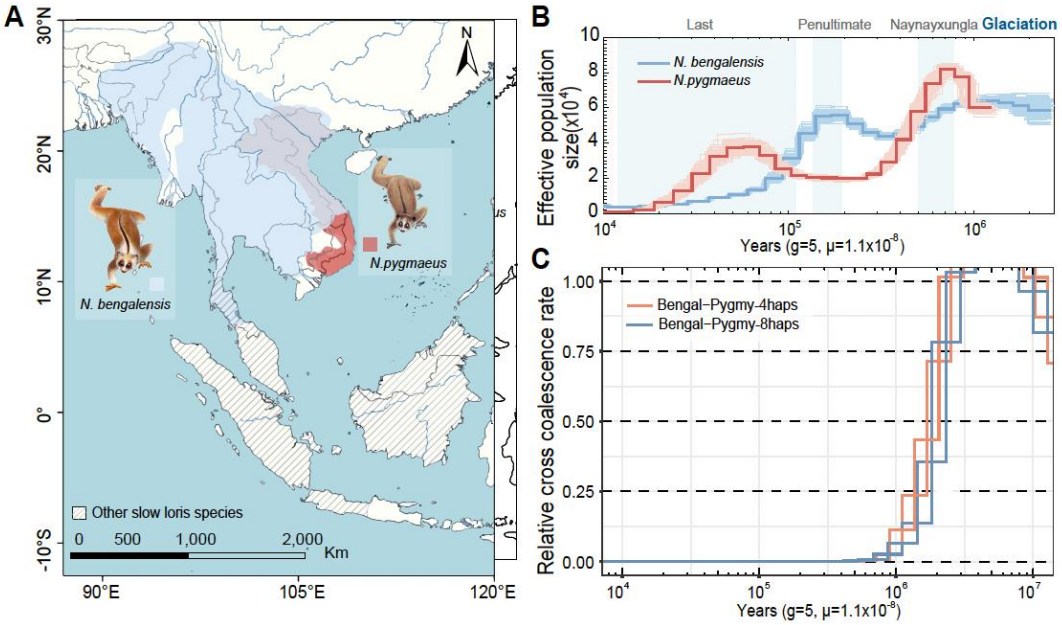

Fig.4. Evolutionary history of slow lorises. A. Geographic distribution of *N. pygmaeus* and *N. bengalensis*. **B.** Population demographic history inferred from sequences of *N. pygmaeus* and *N. bengalensis* using the pairwise sequential Markovian coalescent (PSMC) method. **C.** Divergence time between *N. pygmaeus* and *N. bengalensis* inferred by means of the multiple sequentially Markovian coalescent (MSMC) method.