Genetic diversity and matrilineal structure in Chinese tree shrews inhabiting Kunming, China

CHEN Shi-Yi 1, XU Ling 1,3, LÜ Long-Bao 2, YAO Yong-Gang 1,*

(1. Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming Yunnan 650223, China; 2. Experimental Animal Core Facility & Kunming Primate Research Center, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming Yunnan 650223, China; 3. Graduate School of the Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Due to their special phylogenetic position in the Euarchontoglires and close affinity to primates, tree shrews have been proposed as an alternative experimental animal to primates in biomedical research. However, the population genetic structure of tree shrews has largely remained unknown and this has hindered the development of tree shrew breeding and selection. Here we sampled 80 Chinese tree shrews (Tupaia belangeri chinensis) in Kunming, China, and analyzed partial mtDNA control region sequence variation. Based on our samples and two published sequences from northern tree shrews (T. belangeri), we identified 29 substitutions in the mtDNA control region fragment (~ 604 bp) across 82 individuals and defined 13 haplotypes. Seventeen samples were selected for sequencing of the cytochrome b (Cyt b; 1134 bp) gene based on control region sequence variation and were analyzed in combination with 34 published sequences to solidify the phylogenetic pattern obtained from control region data. Overall, tree shrews from Kunming have high genetic diversity and present a remarkable long genetic distance to the two reported northern tree shrews outside China. Our results provide some caution when using tree shrews to establish animal models because of this apparent genetic difference. In addition, the high genetic diversity of Chinese tree shrews inhabiting Kunming suggests that systematic genetic investigations should be conducted before establishing an inbred strain for medical and biological research.

Key words: Chinese tree shrews; mtDNA; Control region; Cytochrome b; Genetic diversity

昆明城郊中国树鼩群体线粒体DNA遗传多样性

陈仕毅1，许凌1,3, 吕龙宝2, 姚永刚1,*

(1. 中国科学院和云南省动物模型与人类疾病机理重点实验室, 中国科学院昆明动物研究所, 云南 昆明 650223；2. 中国科学院昆明动物研究所 实验动物中心, 云南 昆明 650223；3. 中国科学院研究生院, 北京 100049)

摘要：由于树鼩是灵长类动物的近亲，且具有体型小、繁殖周期短、饲养管理成本低等优点，长期以来被认为有望替代灵长类动物用于人类疾病的动物模型研究。然而，目前对树鼩的群体遗传结构还知之甚少，这极大地限制了其在疾病动物模型研究的应用，也是其品系资源创制的瓶颈。本研究通过分析 80 只采自于云南省昆明周边地区的野生树鼩 (Tupaia belangeri chinensis) 线粒体DNA (mtDNA) 多态性，结合国外报道的 2 个树鼩 (Tupaia belangeri) 序列比较后发现，在 604 bp 的mtDNA控制区片段中共检测到 29 个核苷酸替代变异。这些变异共界定了 13 种单倍型，表现在较高的群体遗传多样度。另外，昆明地区的树鼩与国外报道的 2 个树鼩间存在较大的遗传分化，mtDNA控制区单倍型间的核苷酸替换数大于 18 个，远高于昆明地区树鼩群体内部不同单倍型之间的差异，选择含有代表性的mtDNA控制区单倍型的 17 个昆明地区树鼩个体进一步测定了细胞色素 b 基因片段 (~ 1134 bp)，结合前人报道的数据分析，结果进一步支持mtDNA控制区内数据反映的遗传格局及揭示的昆明地区树鼩与国外报道树鼩间的明显差异。本研究结果提示，昆明地区的树鼩与国外树鼩之间存在较大遗传差异，在将树鼩用于人类疾病动物模型研究中要注意这些遗传差别。昆明城郊的树鼩群体具有较高的遗传多样性，在开展近交系建立等工作时须考虑选取群体内部具有代表性的mtDNA世系。
Tree shrews (*Tupaia belangeri*) are small mammals of a squirrel-like appearance formerly placed in the Primates order despite a lack of derived features characteristic of primate species (Sargis, 2004). It is currently placed in the family of Tupaiidae, Scandentia, which together with Dermoptera and Primates, form the group Euarchonta (Murphy et al, 2001; Helgen, 2005). Although the exact phylogenetic position of tree shrews remains resolved, this animal is still regarded as a close relative to primates (Nie et al, 2008; Peng et al, 1991; Seiffert et al, 2003). Given that tree shrews are easy to raise, cheap to maintain, have a high reproductive rate, small body size, and close affinity to primates, this animal may be an alternative to the use of primates in biomedical research. For example, tree shrews have been used in creating animal models for hepatitis B virus and hepatitis C virus infection (Amako et al, 2010; Cao et al, 2003; Köck et al, 2001; Ren & Nassal, 2001; Su, 1987; Xu et al, 2007; Yan et al, 1996). Tree shrews were also successfully used in the study of pathogenesis of human myopia and psychosocial stress diseases (Cao et al, 2003; Fuchs, 2005; Norton et al, 2006). In addition, tree shrews have the highest brain to body ratio of all mammals, including humans (Peng et al, 1991; Previc, 2009), which raises the possibility of investigating brain function and creating animal models for human neuro-degenerative diseases.

There are four genera within the family Tupaiidae, consisting of *Tupaia*, *Anathana*, *Urogale*, and *Dendrogale* (Helgen, 2005). In genus *Tupaia*, a total of 15 species are recognized and broadly distributed across Southeast Asia including southern China, India, Philippines, Java, Borneo, Sumatra and Bali (Olson et al, 2005; Peng et al, 1991). Chinese tree shrews (*Tupaia belangeri*) are distributed across southwest China. According to geographical distribution and morphological characteristics, Wang (1987) proposed that Chinese tree shrews could be divided into six subspecies: *Tupaia belangeri chinensis* distributed in most of parts of Yunnan, Yun-Gui plateau, and southwest Sichuan; *T. b. gaoligongensis* distributed in central and north Gaoligong Mountain; *T. b. modesta* on Hainan Island; *T. b. tonquinia* distributed in southwest Guangxi; *T. b. yuoshanensis* distributed in southwest Guangxi; and *T. b. yunalis* distributed in central and south Yunnan, northwest Guangxi and southwest Guizhou. Recently, Jia et al (2009) studied the morphometrics of the skull and mandible of *T. b. chinensis* collected in Luquan, Jianchuan, Lijiang, and Yunlong Counties of Yunnan and morphological differences between localities. In combination these studies suggest that Chinese tree shrews show regional differentiation.

Despite growing attention and the strong potential of tree shrews as an animal model for human disease, scant biological data about these species have been collected, especially at the molecular level (Peng et al, 1991). This deficit has constrained the utilization of tree shrews as an animal model in biomedical studies. Although the critical problem regarding laboratory breeding and domestication of wild tree shrews was solved more than two decades ago (Peng et al, 1991), an inbred strain of tree shrews has not been established because, among other factors, of a lack of genetic data.

Here we aimed to analyze the genetic structure of tree shrews collected from urban Kunming, Yunnan, China to provide useful information to establish several inbreeding lines from these captured animals. We used mitochondrial DNA (mtDNA) as a genetic marker as it is useful in revealing the origin and demographic history and overall genetic diversity of domestic and wild species (He et al, 2008).

1 Materials and Methods

1.1 Sampling, DNA amplification and sequencing

We collected ear or muscle tissue from 80 Chinese tree shrews (*T. b. chinensis*) inhabiting the suburb of Kunming, Yunnan, China. Animals were raised at the experimental animal core facility of the Kunming Institute of Zoology and procedures were approved by the Institutional Animal Care and Use Committee of the Kunming Institute of Zoology, the Chinese Academy of Sciences.

Genomic DNA was extracted from tissue samples using the standard phenol/chloroform method. A fragment with a size of 1488 bp that covers the entire mtDNA control region of Chinese tree shrews was amplified using primer pairs L15356: 5'-CTCAAGGAA GAAGAACAATA-3' / H50: 5'-TGTATGTTTATGGATGAG TCTATG-3'. Primers were named relative to their
positions in the complete mtDNA sequence of northern tree shrew (GenBank accession number NC_002521). PCR amplification was performed in a volume of 50 µL reaction mixture containing 100 ng of DNA, 10 mmol/L Tris-HCl (pH 8.3), 2.5 mmol/L MgCl₂, 50 mmol/L KCl, 10 pmol/L of each primer, and one unit of LA Taq polymerase (Takara, Dalian, China) following 35 cycles of 50 s at 94°C, 50 s at 55°C and 60 s at 72°C. The complete mitochondrial Cyt b gene was amplified and sequenced in 17 Chinese tree shrews representing different haplotypes as identified by mtDNA control region data. We used primer pair L14058: 5'-ACAAAAACACAATAAGGCA -3' / H15298: 5'-CTCCGT TTCTGGTTTACAAAG -3' to amplify the Cyt b gene, applying the same amplification conditions and procedures as used for mtDNA control region. PCR products were purified on spin columns and directly sequenced using Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, USA) on an ABI 3730 DNA sequencer according to the manufacturer's manual. Two inner sequencing primers (H16093: 5'-CAAGTTAAGTCCAGCTACAAT-3', for mtDNA control region; L15059: 5'-TCCCATTCCTTACACGTC-3', for Cyt b gene) and primers L15356, L14058, and H15298 were used for overlapping sequencing of the analyzed region. The mtDNA sequences of Chinese tree shrew have been deposited in GenBank with accession numbers HQ836245- HQ836324 (for mtDNA control region sequences) and HQ836325-HQ836341 (for Cyt b gene sequences).

1.2 Data Analysis

We analyzed a fragment of mtDNA control region (~604 bp) in 80 Chinese tree shrews, corresponding to 15416th–16017th region in mitochondrial genome of T. belangeri (GenBank accession number NC_002521). Two reported individuals of northern tree shrews (Tupaia belangeri) outside China were retrieved from GenBank and were aligned with those of Chinese tree shrews (Tab. 1 and Fig. 1). Sequence variations were exported using Mega 4.0 (Tamura et al, 2007). The available Cyt b gene sequences of tree shrews in GenBank, including 32 Chinese tree shrews and two northern tree shrews outside China (Tab. 1), were also analyzed together with 17 newly sequenced samples. Relationships among the lineages were explored using the median-joining networks (Bandelt et al, 1999). Networks were constructed manually and were checked using Network 4.1 (http://www.fluxus-engineering.com/sharenet.htm), in which transitions and transversions were equally weighted. Nucleotide diversity (π) and haplotype diversity (h) were estimated according to Nei (1987) and using DnaSP 5.0 (Librado & Rozas, 2009).

To further determine the phylogenetic position of Chinese tree shrews we retrieved the Cyt b gene sequences of Tupaia glis (AY321645), T. longipes (AY321651), T. salatana (AY321654), and Dendrogale melanura (AY321634, as the outgroup) from GenBank and constructed a neighbor-joining (NJ) tree using Mega 4.0 (Tamura et al, 2007). The maximum composite likelihood nucleotide substitution model was employed.

The rate variation among sites, with four rate categories for the nucleotide gamma shape parameter, was

![Tab. 1 mtDNA sequence information for tree shrews analyzed in this study](image-url)

<table>
<thead>
<tr>
<th>Species</th>
<th>GenBank accession number</th>
<th>No.</th>
<th>h *</th>
<th>π *</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese tree shrews</td>
<td>HQ836245- HQ836324</td>
<td>80</td>
<td>0.790 ± 0.032</td>
<td>0.00436 ± 0.00017</td>
<td>This study</td>
</tr>
<tr>
<td>Northern tree shrews outside</td>
<td>AF217811 b</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>Schmitz et al, 2000</td>
</tr>
<tr>
<td>China</td>
<td>AJ421453</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td><strong>Cyt b gene</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese tree shrews</td>
<td>AY321635-AY321637</td>
<td>3</td>
<td>0.926 ± 0.035</td>
<td>0.00338 ± 0.00029</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>EU376531, EU3531749-EU3531776</td>
<td>29</td>
<td>—</td>
<td>—</td>
<td>From GenBank</td>
</tr>
<tr>
<td>Northern tree shrews outside</td>
<td>AF217811 b</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>Jia et al, 2008</td>
</tr>
<tr>
<td>China</td>
<td>AJ421453</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

* Nucleotide diversity (π) and haplotype diversity (h) were estimated according to Nei (1987). We pooled all Cyt b sequences of Chinese tree shrews together to estimate these two parameters.

b This sequence has been curated by NCBI staff and has the accession number NC_002521.
Sequence variation and haplotype distribution

The mtDNA control region sequences of tree shrews presented length mutations, which were caused by insertion/deletion (indels) of tandem repeat or length mutations of a stretch of duplicated nucleotides in region 15449–15456. One sample (TS07) had an extra insertion of CATT repeat in region 15557–15564 which contained two tandem repeats in the analyzed fragment (~604 bp). After the indels were excluded, we found 29 sequence variations in the analyzed fragment of mtDNA control region, which defined 13 haplotypes (H1–H13; Fig. 1). Among them, haplotype H5 was dominant and occurred in 30 individuals, followed by haplotypes H1 (in 14 individuals), H2 (in 12 individuals) and H12 (in 8 individuals). The other haplotypes had a low frequency and were shared by no more than four samples. The two reported northern tree shrews outside China shared haplotype H13, and presented strong differences to the haplotypes identified in Chinese tree shrews (Fig. 1). The overall genetic diversity of 49 Chinese tree shrew Cyt b gene sequences was 0.00333 ± 0.00032.

2.2 Network profiles of mtDNA haplotypes in Chinese tree shrews

The network profile of the Cyt b haplotypes followed a similar pattern observed for mtDNA control region sequences (Fig. 3b). Because we selected the samples for Cyt b sequencing based on the haplotypes identified by control region sequence information, an estimation of the frequency of each Cyt b haplotype may be biased, despite a fact that we took 32 published Chinese tree shrews samples into consideration. Nonetheless, the phylogenetic relationship as indicated by the network of Cyt b haplotypes might offer further information for us to discern the genetic difference between different lineages. Consistent with the pattern of mtDNA control region sequence data, haplotype h3 was also located in the central location of the network, suggesting its ancestral status. The difference of Cyt b gene nucleotide changes between Chinese tree shrews and the reported northern tree shrews (h26) was estimated using jModelTest 0.1.1 (Guindon & Gascuel, 2003) which finally yielded the alpha value 0.64.
Fig. 2  Sequence variation of mtDNA cytochrome b gene fragments (1134 bp) of 17 Chinese tree shrews newly sequenced in this study and 34 individuals from GenBank. The detailed sources of these reported sequences are listed in Tab. 1. Dots (·) denotes identical sites. Substitutions were numbered according to the mitochondrial genome of northern tree shrew (GenBank accession number NC_002521). The number of individuals sharing certain mtDNA control region haplotypes (in parentheses) is listed in the right column.

Fig. 3  Network profiles of the mtDNA control region (a) and Cyt b gene (b) haplotypes of tree shrews. The links are labeled by nucleotide positions to designate variations. The order of the mutations on a branch is arbitrary. The area of each circle is proportional to haplotype frequency. White color denotes individuals analyzed for both mtDNA control region and Cyt b gene sequence variation. Chinese tree shrew Cyt b gene sequences from GenBank are marked black.
remarkable, with a ratio (89/1134) even higher than that of the mtDNA control region (18/604).

2.3 Phylogenetic position of Chinese tree shrews

We constructed an NJ tree of the 26 Cyt b haplotypes of Chinese tree shrews, together with four species from GenBank (Tupaia glis, Tupaia longipes, Tupaia salatana, and Dendrogale melanura), to further demonstrate the affinity of Chinese tree shrews to other species (Fig. 4). All Chinese tree shrews haplotypes were clustered together with high bootstrap support. Chinese tree shrews are clearly separated from northern tree shrews outside China and other species, but presented an overall closer relationship to northern tree shrews.

Fig. 4 The neighbor-joining (NJ) tree of the Cyt b gene of 26 Tupaia belangeri haplotypes and four other species (T. glis, T. longipes, T. salatana, and Dendrogale melanura)

Tree was rooted by Dendrogale melanura. The values on the branch are bootstrap support based on 1000 replications and values lower than 50% were omitted.

3 Discussion

Tree shrews are received more and more attention due to their potential use as animal models for human disease and as alternatives for primates in biomedical research (Cao et al, 2003; Peng et al, 1991). Although the exact phylogenetic position of tree shrews within the Euarchontoglires remains of debate (Arnason et al, 2002; Murphy et al, 2001; Nie et al, 2008; Olson et al, 2005; Peng et al, 1991; Sargis, 2004; Schmitz et al, 2000), available evidence has undoubtedly suggested a closer affinity of tree shrews to well-known primates such as rhesus monkeys (Macaca mulatta) and crab-eating macaques (Macaca fascicularis) than mice or rats. Given that tree shrews are also small animals and easy to raise, it may be promising to develop them as replacements for monkeys and other endangered non-human primate species.

There are 15 recognized species and 40 subspecies in genus Tupaia (http://www.bucknell.edu/msw3/; Helgen, 2005). However, to our knowledge an insufficient number of molecular studies have been done and necessary information for biomedical research and breeding as an experimental animal remain unknown. Work on gene expression in the hippocampal tissue of tree shrews treated with cortisol (Alfonso et al, 2004) and proteome analysis of aflatoxin B1-induced hepatocarcinogenesis (Li et al, 2008) has yielded useful information, but compared to the available knowledge for mice and rats there is still much unknown about these animals.

Here we focused on mtDNA sequence variation in Chinese tree shrews captured from urban Kunming. Our initial aim was to better understand the genetic structure of this species and to establish a matrilineal marker for tracing lineages in a subsequent tree shrews breeding project. In addition, we wanted to better understand overall genetic diversity of this animal resource. By comparing the mtDNA sequence differences between Chinese tree shrews and northern tree shrews, we found that they had a large divergence (Figs.1, 2). Network analyses showed that Chinese tree shrews and northern tree shrews are clearly separated from each other (Fig. 3). Our results raise caution over the routine use of tree shrews as animal models and genetic differences are ignored, especially when studies simply named their subjects as Tupaia belangeri.

The divergence of Chinese tree shrews and northern tree shrews was unexpected and implies that a systematic analysis of the six proposed subspecies of Chinese tree shrew (Wang, 1987) requires verification in future studies. Moreover, capturing and domesticating wild animals with the representative lineages should be encouraged during our efforts in establishing inbred tree shrew strains for medical and biological research. The observation that the tree shrew population inhabiting Kunming has a considerably high mtDNA genetic diversity is consistent with the result of a recent study of
microsatellite markers in which He et al. (2009) found that nine out of 11 loci exhibited polymorphisms in 30 Chinese tree shrews. Data such as that presented here will be helpful in selecting founding individuals for the establishment of an inbred strain of Chinese tree shrews.

Acknowledgement: We thank YANG Zhao-Qing for collecting some tree shrews samples and JIANG Xue-Long for comments on an early version of the manuscript.

References:


