Positively selected genes of the Chinese tree shrew (*Tupaia belangeri chinensis*) locomotion system

Yu FAN[1, 2], Dan-Dan YU[1], Yong-Gang YAO[1,2,3,*]

1. *Key Laboratory of Animal Models and Human Disease Mechanisms of Chinese Academy of Sciences & Yunnan Province*, *Kunming Institute of Zoology*, *Kunming, Yunnan* 650223, *China*
2. *Kunming College of Life Science*, *University of Chinese Academy of Sciences*, *Kunming, Yunnan* 650223, *China*
3. *Kunming Primate Research Center*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences, Kunming* 650223, *China*

Abstract: While the recent release of the Chinese tree shrew (*Tupaia belangeri chinensis*) genome has made the tree shrew an increasingly viable experimental animal model for biomedical research, further study of the genome may facilitate new insights into the applicability of this model. For example, though the tree shrew has a rapid rate of speed and strong jumping ability, there are limited studies on its locomotion ability. In this study we used the available Chinese tree shrew genome information and compared the evolutionary pattern of 407 locomotion system related orthologs among five mammals (human, rhesus monkey, mouse, rat and dog) and the Chinese tree shrew. Our analyses identified 29 genes with significantly high ω (Ka/Ks ratio) values and 48 amino acid sites in 14 genes showed significant evidence of positive selection in the Chinese tree shrew. Some of these positively selected genes, e.g. *HOXA6* (homeobox A6) and *AVP* (arginine vasopressin), play important roles in muscle contraction or skeletal morphogenesis. These results provide important clues in understanding the genetic bases of locomotor adaptation in the Chinese tree shrew.

Keywords: Chinese tree shrew; Locomotion system; Positively selected genes

The Chinese tree shrew (*Tupaia belangeri chinensis*), currently placed in the order Scandentia, has attracted wide attention as a viable alternative animal model for biomedical research due to a variety of unique characteristics, e.g., small adult body size, high brain-to-body mass ratio, short reproductive cycle and life span, low cost of maintenance, and most importantly, a closer affinity to primates, particularly humans. These physical factors make the tree shrew an attractive alternative experimental animal to primates, while the closer affinity to humans make them preferable to rats or mice (Cao et al, 2003; Fuchs & Corbach-Söhle, 2010; Peng et al, 1991; Xu et al, 2013). The Chinese tree shrew is widely distributed in the tropical forests of South Asia, Southeast Asia and Southwest China (Peng et al, 1991). As an arboreal mammal living in the wild, tree shrew needs to cope with unfavorable circumstances (e.g., insufficient food and the attack of predators), and accordingly has developed an important adaption ability, that is, the ability of quick climbing and jumping movement (Fuchs & Corbach-Söhle, 2010; Peng et al, 1991). To adapt to the inevitably dangerous circumstances, the ancestor of tree shrew gradually developed typical arboreal adaptation patterns of the muscular system, including well-developed deltoideus and acromion of the scapula, existence of post-scapular fossa as in many arboreal mammals, the triceps brachii and the teres major with large bundles in favor of forelimb flexion, and strong and massive structure of the biceps femoris. These features may play a role in knee flexion during the locomotion on the branches of trees (Endo et al, 1999). Therefore, tree shrew has a significant advantage in locomotion ability of climbing and jumping, and in strong explosive power.

The evolution of these typically arboreal traits poses several interesting questions, the most pertinent being

Received: 09 January 2014; Accepted: 17 March 2014
Foundation items: This study was supported by the National 863 Project of China (2012AA021801, 2012AA022402) and grants from Chinese Academy of Sciences (KSCX2-EW-R-11, KSCX2-EW-J23) and Yunnan Province (2013FB071)
*Corresponding author, E-mail: ygyaozh@gmail.com; yaoyg@mail.kiz.ac.cn

whether or not there any selective signals in the tree shrew's locomotion system. During every evolutionary process, a series of sequence changes occurs on multiple genes to induce phenotypic changes in morphology and physiology, allowing species to adapt to their novel environments (Hoekstra & Coyne, 2007). Because positive selection could leave an imprint on genes, the identification of positively selected genes via genome-wide scanning may help explain the genetic bases underlying adaptive evolution in mammals (e.g. Zeng et al, 2013). The currently available high-quality genome sequence of Chinese tree shrew (Fan et al, 2013) provides an opportunity to more fully characterize the potential molecular underpinnings of the tree shrew's locomotion ability, as well as give potentially significant clues to how the tree shrew evolved and differentiated itself from other related mammals, many of which are also used as animal models in biomedical research.

In this paper, we searched for locomotion system related genes against the Gene Ontology (GO) database (http://www.geneontology.org/) and then reconstructed gene sets of six mammalian species, including the Chinese tree shrew, human, rhesus monkey, mouse, rat and dog, to identify genes under positive selection along the Chinese tree shrew lineage. We also investigated rapidly evolving genes of locomotion system related GO category in Chinese tree shrew. Collectively, the results of these analyses provide some important clues to understanding the genetic bases of locomotion ability and adaptation of Chinese tree shrew.

METHODS

Source of gene dataset and identification of orthologs

A total of 848 human locomotion system related genes were obtained from the Gene Ontology database (GO:0007626 locomotory behavior, GO:0040017 positive regulation of locomotion, GO:0003012 muscle system process, GO:0001501 skeletal system development) for comparative study of the Chinese tree shrew's locomotive system. To determine orthologous relationships between Chinese tree shrew and human, we downloaded human protein sequence data from Ensembl (release 64; http://www.ensembl.org/index.html) for comparison. The longest transcript was chosen to represent each gene with alternative splicing variants. We aligned the orthologous human protein sequences of these 848 genes onto the Chinese tree shrew genomes using tblastn (Mcginnis & Madden, 2004), and then the best hit regions of each gene with a 5kb flanking sequence were cut down and re-aligned using GeneWise (Birney et al, 2004) in order to define the detailed exon-intron structure of each gene. The identified orthologous sequences that may have potential errors such as frame-shift and premature termination were further eliminated, and ortholog sequences from the other surveyed species were identified using a two-step method, as follows. First, using human genes as a reference, we obtained orthologous sequences of other species from Ensembl one2one orthology gene list, and downloaded the corresponding coding region nucleotide sequences and amino acid sequences according to IDs from Ensembl (release 64). Next, we identified the sequences of these genes not included in Ensembl one2one orthology gene list using computational gene prediction, as described in the Chinese tree shrew genome project (Fan et al, 2013).

Sequence alignment and filtering

All coding sequences (CDS) from the six surveyed species were aligned via MUSCLE 3.7 (Edgar, 2004) with the guidance of aligned protein sequences. To reduce the occurrence of false positive prediction, we carried out a series of filtering processes. First, we deleted all gaps and "N" from the alignments. Second, the aligned regions with more than 4 non-synonymous mutations in 7 continuous amino acids were filtered. Finally, the entire alignment was discarded if the remaining sequences after alignment were shorter than 100 bp.

Detection of positively selected genes

To detect potential candidate genes under positive selection along the Chinese tree shrew lineage, we applied CODEML from the PAML4 package, which is based on the Maximum-Likelihood method of molecular evolution (Yang, 2007), to the human, rhesus monkey, Chinese tree shrew, mouse, rat and dog ortholog gene sets. The same guide tree present in our previous study (Fan et al, 2013) was used to detect the positive selection genes using the following models: (a) branch model, which allows different ω values (ω is the ratio of non-synonymous to synonymous substitution rates, i.e., the Ka/Ks ratio) between the foreground branch (ω_2) and background branch (ω_1) while its corresponding null model assumes all branches have an identical ω_0 value; and (b) branch-site model with fixed foreground branch ω_2=1 and non-fixed foreground branch ω_2, which is used to test whether a gene has undergone positive selection on a foreground branch. Finally, likelihood ratio test (LRT) was performed on following model pairs: (a) Method 1, to test whether the foreground branches ω ratio were significantly different from that of background branches; and (b) Method 2, to test whether a proportion of sites in the sequence provided statistically significant support for ω>1 on foreground branches.

Phylogenetic analysis

Maximum Likelihood (ML) trees were reconstructed based on amino acid sequence using MEGA5.0

(Tamura et al, 2011). For the phylogenetic analysis, dog sequence obtained served as an outgroup to root the phylogenetic tree. Accuracy of phylogenetic tree were measured with 1 000 bootstrap replicates.

Analysis of rapidly evolving function categories

Based on GO annotation of human, we sought to detect rapidly evolving function categories of genes between Chinese tree shrews and humans, or between humans and rats. If a given GO category contained 20 or more genes from our ortholog dataset, it was selected for further analysis. The values of *K*a and *K*s and the *K*a/*K*s ratio (ω) were estimated for each gene using the KaKs_Calculator program (Zhang et al, 2006). The average values of *K*a and *K*s for all genes annotated to a given GO were calculated using the following equations (Cho et al, 2013).

$$k_a = \frac{\sum_{ieT} a_i}{\sum_{ieT} A_i}, \quad k_s = \frac{\sum_{ieT} s_i}{\sum_{ieT} S_i}$$

Where a_i is the number of non-synonymous substitutions and A_i is the number of non-synonymous sites in gene i; s_i is the number of synonymous substitutions and S_i is the number of synonymous sites in gene i; T is the number of genes annotated by GO. The expected proportion of non-synonymous substitutions to all substitutions P_A in a GO category was then estimated as previously described in Cho et al (2013):

$$P_A = \frac{k_a \sum_{ieC} A_i}{k_a \sum_{ieC} A_i + k_s \sum_{ieC} S_i}$$

Finally, for a given GO category, we used binomial distributions (Chimpanzee Sequencing and Analysis Consortium, 2005) to estimate the divergence of the proportion of non-synonymous substitutions and synonymous sites between the observed and expected values. The pbinom package in R (http://www.r-project.org/) was used to calculate the *P*-value. Rapidly evolving function categories were defined as $P<0.05$, with the parameter of lower.tail=FALSE.

RESULTS

Filtered ortholog genes

From the gene list of the GO database, we identified 604 locomotion system related orthologs between the genomes of the Chinese tree shrew, mouse, rat, dog, rhesus monkey and human. After alignment and filtering, 197 genes were eliminated (see Methods) leaving 407 genes for further positive selection tests.

Genes under positive selection

To detect the difference in selective pressure between Chinese tree shrew and other species, each aligned gene was evaluated in terms of its ratio of non-synonymous to synonymous substitution rates (ω value) using CODEML in the PAML package (Yang, 2007) under the guide tree previously described by Fan et al (Fan et al, 2013). Testing under the branch model (Method 1) showed 29 genes (7.1 % of the total number of genes) had a significantly higher ω value ($P<0.05$) in the Chinese tree shrew (Table 1) than in the other five surveyed species. Next, we used the branch-site test of positive selection (Method 2) to detect signals of positive selection on each alignment, identifying 48 amino acid sites in 14 genes (3.4 % of the total number of genes) that exhibited significant evidence of positive selection ($P<0.05$) in the Chinese tree shrew lineage (Table 2). While the final gene lists retrieved from these two methods were largely different, two genes, *HOXA6* (homeobox A6) and *AVP* (arginine vasopressin) that showed positive selection signals were identified by both of these methods (Figure 1).

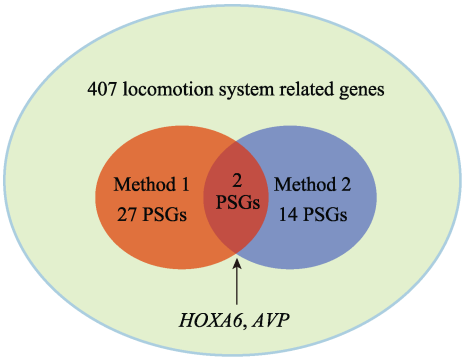

Figure 1 Summary of positively selected genes identified by two different detection methods

Method 1: likelihood ratio test under branch model; Method 2: likelihood ratio test under branch-site model (positive selection). PSGs: positively selected genes.

Phylogenetic analysis

To gain a clearer picture of these two potential locomotive related genes showing positive selection, two maximum likelihood (ML) trees were reconstructed based on the amino acid sequences of HOXA6 and AVP, respectively. Both gene trees showed a clustering pattern inconsistent with the recognized species tree (Fan et al, 2013; Lindblad-Toh et al, 2011; Murphy et al, 2004) and had a long branch for the Chinese tree shrew (Figure 2). Several other gene trees were constructed for the other positively selected genes, and only 9 trees showed a clustering pattern that was consistent with the species tree (data not shown). The observed discrepancy between the gene tree and species tree was in agreement with results of the above prediction of selection signal.

Table 1 29 positively selected genes detected by method 1 in Chinese tree shrew

Gene	GO category	*P*-value
ACTC1	GO:0030049 muscle filament sliding; GO:0060048 cardiac muscle contraction; GO:0030049 muscle filament sliding; GO:0003012 muscle system process	0.031449
ACTN2	GO:0030049 muscle filament sliding; GO: 0006936 muscle contraction	0.013104
ARRB2	GO:0007628 adult walking behavior	0.016337
AVP	GO:0007626 locomotory behavior	0.005769
AXIN2	GO:0003413 chondrocyte differentiation involved in endochondral bone morphogenesis	0.00714
BSX	GO:0007626 locomotory behavior	0.026649
CALM1	GO:0006936 muscle contraction; GO: 0010881 regulation of cardiac muscle contraction by regulation of the release of sequestered calcium ion; GO: 0055117 regulation of cardiac muscle contraction	5.26E-05
FGF1	GO:0030335 positive regulation of cell migration; GO:0040017 positive regulation of locomotion	0.001138
GJC1	GO:0006936 muscle contraction	0.037772
GNB2L1	GO:0030335 positive regulation of cell migration	0.005448
HOXA6	GO:0048704 embryonic skeletal morphogenesis	0.009921
HSBP1	GO:0006936 muscle contraction; GO:0003012 muscle system process	0.025508
KCNJ10	GO:0007628 adult walking behavior; GO:0007626 locomotory behavior	0.049375
LAMB1	GO:0030335 positive regulation of cell migration	0.039006
OPRD1	GO:0008344 adult locomotory behavior	2.92E-05
PDGFB	GO:0014911 positive regulation of smooth muscle cell migration; GO: 0030335 positive regulation of cell migration; GO: 2000591 positive regulation of metanephric mesenchymal cell migration	0.002046
PDGFRB	GO:0014911 positive regulation of smooth muscle cell migration; GO: 0030335 positive regulation of cell migration; GO: 0048705 skeletal morphogenesis	0.012328
PLN	GO:0003012 muscle system process; GO: 0010881 regulation of cardiac muscle contraction by regulation of the release of sequestered calcium ion	0.016657
PRKRA	GO:0048705 skeletal morphogenesis	0.047973
PRRX1	GO:0048701 embryonic cranial skeleton morphogenesis; GO: 0048705 skeletal morphogenesis	0.000916
RRAS2	GO:0030335 positive regulation of cell migration	0.008652
SATB2	GO:0048704 embryonic skeletal morphogenesis	0.043399
SRSF1	GO:0060048 cardiac muscle contraction	0.048949
TBX20	GO:0006936 muscle contraction	0.012294
TGFB3	GO:0048702 embryonic neurocranium morphogenesis; GO: 0060364 frontal suture morphogenesis	0.040708
TGFBR1	GO:0048701 embryonic cranial skeleton morphogenesis; GO: 0048705 skeletal morphogenesis	0.018154
TNNC2	GO:0003009 skeletal muscle contraction; GO: 0006937 regulation of muscle contraction; GO: 0030049 muscle filament sliding	0.017249
VIPR1	GO:0006936 muscle contraction	0.009035
WNT11	GO:0030335 positive regulation of cell migration	0.024982

Method 1 is likelihood ratio test under branch model.

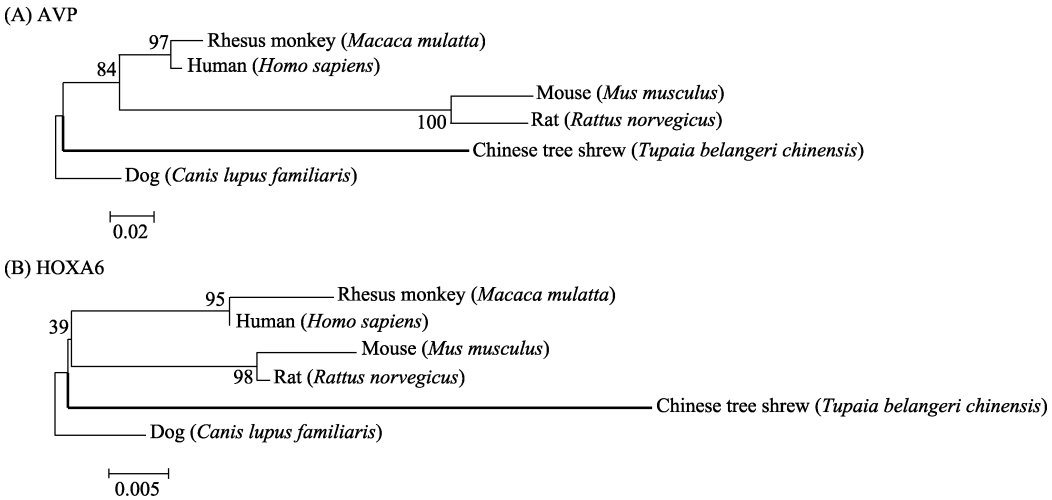

Figure 2 Maximum likelihood (ML) trees based on amino acid sequences of AVP (A) and HOXA6 (B)

Numbers on branches denote ML bootstrap values.

Table 2 14 Positively selected genes detected by method 2 in Chinese tree shrew

Gene	GO category	*P*-value
APLP2	GO:0007626 locomotory behavior	0.006183
AVP	GO:0007626 locomotory behavior	0.024482
C1QBP	GO:0040017 positive regulation of locomotion	0.000929
CACNA1E	GO:0007626 locomotory behavior	0.005731
CAV1	GO:0003012 muscle system process	8.85E-05
CBS	GO:0001958; GO:0048705 skeletal morphogenesis	0.000192
CCR7	GO:0040017 positive regulation of locomotion	0.014505
CHRNA1	GO:0003009 skeletal muscle contraction; GO:0003012 muscle system process	0.010052
FOXC2	GO:0040017 positive regulation of locomotion	0.017599
HOXA6	GO:0048704 embryonic skeletal morphogenesis	6.50E-10
HOXC8	GO:0048705 skeletal morphogenesis	0.032523
MMP14	GO:0030335 positive regulation of cell migration; GO:0040017 positive regulation of locomotion	0.002603
MYH4	GO:0030049 muscle filament sliding; GO:0006936 muscle contraction; GO:0003012 muscle system process	0.004462
MYOT	GO:0006936 muscle contraction; GO:0003012 muscle system process	0.001562

Method 2 is likelihood ratio test under branch-site model.

Rapidly evolving function categories

To further determine the rapidly evolution of locomotion system related GO category in Chinese tree shrew, we calculated the average *K*a and *K*s values for all GO category containing 20 genes or less. When compared to humans, Chinese tree shrews had 23 GO categories showing rapid evolvement ($P<0.05$), of which 7 categories were closely related to locomotion (GO: 0040012 regulation of locomotion; GO:0040017 positive regulation of locomotion; GO:0006937 regulation of muscle contraction; GO:0007626 locomotory behavior; GO:0003012 muscle system process and GO:0030534 adult behavior) (Table 3). Compared to rats, the Chinese tree shrew had 26 GO categories showing a rapid evolvement ($P<0.05$). Among these GO categories, 10 were closely related to locomotion (GO:0060537 muscle tissue development; GO:0014706 striated muscle tissue development; GO:0040012 regulation of locomotion; GO:0042692 muscle cell differentiation; GO:0040017 positive regulation of locomotion; GO:0006941 striated muscle contraction; GO:0003012 muscle system process; GO:0007517 muscle organ development; GO:0006936 muscle contraction; GO:0006937 regulation of muscle contraction) (Table 4). Comparisons between GO categories revealed four rapidly evolving categories relevant to locomotion (GO:0040012 regulation of locomotion, GO:0040017 positive regulation of locomotion, GO: 0003012 muscle system process, GO:0006937 regulation of muscle contraction), which were commonly observed in the Chinese tree shrew, human, and rat.

DISCUSSION

As a small mammal with a close affinity to primates, tree shrews are increasingly considered a viable alternative to primates in biomedical research (Cao et al, 2003; Fuchs & Corbach-Söhle, 2010; Peng et al, 1991). While the recently published genome of the Chinese tree shrew greatly extends the necessary knowledge needed to make it an effective animal model, there are still many aspects of Chinese tree shrew evolution and genetics that we still do not entirely understand. Take for example locomotion: previous studies have shown that to adapt to dangerous living environments in the wild, the ancestor of tree shrew gradually developed a faster rate of speed and stronger jumping ability that partially differentiated it from other mammals (Fuchs & Corbach-Söhle, 2010; Peng et al, 1991). While these changes are well known and fit well with both the extant empirical observations and theoretical models, relatively little is known about the underlying genetic mechanisms that partly distinguish the tree shrew from other related mammalian species—in particular those mechanisms that may serve to make the Chinese tree shrew a more valuable animal model.

In the present study, we used the recently published Chinese tree shrew genome as a foundation to begin understanding the development of the tree shrew's locomotive system by examining related orthologs in six mammals. We identified several genes under positive selection in Chinese tree shrew that may help explain its significant advantage in locomotion ability of explosive power and jumping (Fuchs & Corbach-Söhle, 2010; Peng et al, 1991), but *HOXA6* and *AVP* were the most attractive targets, being detected commonly by two distinct detection methods. *HOXA6* was previously reported to be involved in skeletal morphogenesis (Wellik, 2007), but little is known about the function of *AVP*, though future research on the function of *AVP* may uncover putatively unknown function of *AVP* related to locomotion.

Among the other 39 genes under positive selection, 13 genes have been reported to be involved in muscle contraction, especially with fast-twitch fibers contraction. Fast-twitch fibers have several unique characteristics, including shorter contraction time, more powerful and more advantage in explosive power, all of which would likely prove useful to the tree shrew in its natural environment (Eberstein & Goodgold, 1968). *TNNC2* and *MYH4* were detected to be under positive selection by both Method 1 and Method 2. Previously, both genes were shown to play important roles in fast-twitch fibers contraction (Davoli et al, 2003; Farah & Reinach, 1995). As a subunit of Troponin C, *TNNC2* is expressed exclusively in fast-twitch skeletal muscle in human and plays a critical role in regulating fast-twitch skeletal muscle contraction (Farah & Reinach, 1995). *MYH4* (myosin, heavy chain 4, skeletal muscle), an isoform of myosin, is highly expressed in adult fast-twitch fibers in human (Davoli et al, 2003). Myosins, which are composed of a family of ATP-dependent motor proteins, are best known for their role in muscle contraction and are responsible for actin-based motility (Sellers, 2000). The positive selection of this gene may potentially contribute to the improvement in explosive power of Chinese tree shrew, which is consistent with a previous observation (Schmidt & Schilling, 2007) that tree shrews possess a high content of fast-twitch fibers in infraspinatus muscle. Together, the rapid evolution of *TNNC2* and *MYH4* in Chinese tree shrew may have increased fast-twitch muscle content and strength, thus improving explosive power and jumping ability.

Another identified gene, *ACTN2* (actinin, alpha 2), which encodes alpha-actinin protein (a member of family of the actin-binding proteins), was previously reported to be closely related to explosive power (Norman et al, 2009). ACTN2 is the major structural components of the sarcomeric Z-line involved in anchoring together the actin-containing thin filaments (Tiso et al, 1999). ACTN2 also takes part in the anaerobic muscle metabolism and compensates for the function of ACTN3, an actin protein associated with explosive power (Macarthur & North, 2004). The detected rapid evolution of *ACTN2* in Chinese tree shrew supports the view that this species has evolved an advantage in explosive power. In addition, another positively selected gene in Chinese tree shrew, *MYOT* gene that encodes myotilin, is a skeletal muscle protein found within the Z-disc of sarcomeres, and it was previously found that myotilin could induce the formation of actin bundles (Salmikangas et al, 2003).

While we found several interesting candidate genes that may explain the locomotive system of the tree shrew, previous studies have effectively argued that some complex function is determined not only by one gene, but also by a series of genes or even by evolution of the whole pathway or network (Huynen et al, 2005). In both our key comparisons of the Chinese tree shrew to humans and rats, we observed many rapidly evolving GO categories related to the locomotion system, especially regarding muscle contraction. These rapidly evolving functions may suggest that Chinese tree shrew has evolved an arboreal locomotion ability to adapt to its living circumstance at a broader level.

While our results may greatly expand the knowledge of the tree shrews locomotive system, the current study has several limitations worth noting. First, some key genes may not have been detected due to the limited power of the statistical methods (Kosakovsky Pond et al, 2011). Second, the current study utilizes the list of locomotor genes in human as reference genes, and this may lead to a partial coverage of all possible locomotion related genes in Chinese tree shrew, meaning there may still be more as of yet unidentified genes at work contributing to the tree shrew locomotive system. Third, no functional assay was performed to validate the findings. Future study taking these limitations into account will likely facilitate greater in-depth understandings of the Chinese tree shrew's locomotion ability.

In summary, our screening for putative locomotor genes under positive selection provided some important clues for understanding the locomotion adaptation in Chinese tree shrew. Further study is indispensable to characterize the exact function of these positively selected locomotor genes.

Acknowledgements: We are grateful to Dr. Dong WANG for helpful discussion.

Table 3 GO categories showing a rapid evolution in Chinese tree shrew as compared with human

GO ID	Gene number	GO name	*Ka*/*Ks*	Amino Acid divergence	*P*-value
GO:0040012	48	Regulation of locomotion	0.083451	0.049934	2.25E-23
GO:0040017	38	Positive regulation of locomotion	0.078734	0.049505	9.93E-18
GO:0030335	35	Positive regulation of cell migration	0.075789	0.048236	2.46E-14
GO:0043085	57	Positive regulation of catalytic activity	0.072898	0.044996	6.25E-14
GO:0051270	46	Regulation of cell motion	0.075461	0.046457	3.75E-13
GO:0030334	41	Regulation of cell migration	0.07441	0.04616	4.62E-12
GO:0051272	38	Positive regulation of cell motion	0.073068	0.046683	8.31E-12
GO:0007610	88	Behavior	0.067098	0.042765	8.81E-09
GO:0001558	24	Regulation of cell growth	0.07434	0.049782	2.07E-08
GO:0008283	44	Cell proliferation	0.06937	0.044489	1.96E-07
GO:0006937	30	Regulation of muscle contraction	0.065469	0.046684	2.48E-06
GO:0007626	73	Locomotory behavior	0.065127	0.042458	4.95E-06
GO:0006928	63	Cell motion	0.06289	0.039346	0.00096
GO:0030097	23	Hemopoiesis	0.065893	0.040171	0.001051
GO:0016477	42	Cell migration	0.065758	0.038147	0.001705
GO:0048870	43	Cell motility	0.063748	0.038383	0.004006
GO:0048534	28	Hemopoietic or lymphoid organ development	0.064639	0.038657	0.004228
GO:0016044	27	Membrane organization	0.068598	0.041368	0.004382
GO:0014070	21	Response to organic cyclic substance	0.061588	4.17E-02	0.00834
GO:0003012	74	Muscle system process	0.057302	0.040496	0.008362
GO:0030534	29	Adult behavior	0.063637	0.04071	0.009514
GO:0007010	33	Cytoskeleton organization	0.061117	0.043273	0.013121
GO:0030155	21	Regulation of cell adhesion	0.068809	0.036182	0.025276

Table 4 GO categories showing a rapid evolution in Chinese tree shrew as compared with rat

GO ID	Gene number	GO name	*Ka*/*Ks*	Amino Acid divergence	*P*-value
GO:0030155	21	Regulation of cell adhesion	0.097633	0.119016	3.55E-49
GO:0048514	44	Blood vessel morphogenesis	0.082959	0.108724	1.92E-46
GO:0043085	57	Positive regulation of catalytic activity	0.084787	0.107279	4.26E-43
GO:0001568	48	Blood vessel development	0.079933	0.106575	1.32E-42
GO:0001944	49	Vasculature development	0.080965	0.105684	2.00E-42
GO:0060537	31	Muscle tissue development	0.077758	0.117899	3.24E-37
GO:0014706	29	Striated muscle tissue development	0.073017	0.116237	1.87E-31
GO:0040012	48	Regulation of locomotion	0.08633	0.10316	1.10E-27
GO:0030334	41	Regulation of cell migration	0.083986	0.101519	2.35E-23
GO:0042692	22	Muscle cell differentiation	0.065632	0.112218	1.37E-18
GO:0051270	46	Regulation of cell motion	0.080463	0.096135	1.32E-15
GO:0040017	38	Positive regulation of locomotion	0.0823	0.098041	1.48E-15
GO:0030335	35	Positive regulation of cell migration	0.080413	0.096529	1.47E-13
GO:0008015	39	Blood circulation	0.07289	0.097362	4.87E-13
GO:0006941	21	Striated muscle contraction	0.055351	0.103401	1.45E-10
GO:0051272	38	Positive regulation of cell motion	0.078057	0.092413	3.92E-10
GO:0001558	24	Regulation of cell growth	0.079741	0.100792	7.15E-09
GO:0001666	23	Response to hypoxia	0.072469	0.088672	1.44E-08
GO:0016044	27	Membrane organization	0.078575	0.094073	5.07E-07
GO:0003012	74	Muscle system process	0.06129	0.086659	9.54E-06
GO:0007517	46	Muscle organ development	0.057935	0.088743	1.23E-05
GO:0006936	68	Muscle contraction	0.057845	0.086065	0.000125
GO:0006937	30	Regulation of muscle contraction	0.069204	0.084317	0.00053
GO:0048870	43	Cell motility	0.075263	0.081279	0.000704
GO:0016477	42	Cell migration	0.07596	0.081453	0.00111
GO:0051094	48	Positive regulation of developmental process	0.06674	0.076614	0.004085

References

Birney E, Clamp M, Durbin R. 2004. GeneWise and Genomewise. *Genome Research,* **14**(5): 988-995.

Cao J, Yang EB, Su JJ, Li Y, Chow P. 2003. The tree shrews: adjuncts and alternatives to primates as models for biomedical research. *Journal of Medical Primatology,* **32**(3): 123-130.

Cho YS, Hu L, Hou H, Lee H, Xu J, Kwon S, Oh S, Kim HM, Jho S, Kim S, Shin YA, Kim BC, Kim H, Kim CU, Luo SJ, Johnson WE, Koepfli KP, Schmidt-Kuntzel A, Turner JA, Marker L, Harper C, Miller SM, Jacobs W, Bertola LD, Kim TH, Lee S, Zhou Q, Jung HJ, Xu X, Gadhvi P, Xu P, Xiong Y, Luo Y, Pan S, Gou C, Chu X, Zhang J, Liu S, He J, Chen Y, Yang L, Yang Y, He J, Liu S, Wang J, Kim CH, Kwak H, Kim JS, Hwang S, Ko J, Kim CB, Kim S, Bayarlkhagva D, Paek WK, Kim SJ, O'brien SJ, Wang J, Bhak J. 2013. The tiger genome and comparative analysis with lion and snow leopard genomes. *Nature Communications,* **4**: 2433.

Davoli R, Fontanesi L, Cagnazzo M, Scotti E, Buttazzoni L, Yerle M, Russo V. 2003. Identification of SNPs, mapping and analysis of allele frequencies in two candidate genes for meat production traits: the porcine myosin heavy chain 2B (MYH4) and the skeletal muscle myosin regulatory light chain 2 (HUMMLC2B). *Animal Genetics,* **34**(3): 221-225.

Eberstein A, Goodgold J. 1968. Slow and fast twitch fibers in human skeletal muscle. *The American Journal of Physiology,* **215**(3): 535-541.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research,* **32**(5): 1792-1797.

Endo H, Rerkamuaychoke W, Kimura J, Sasaki M, Kurohmaru M, Yamada J. 1999. Functional morphology of the locomotor system in the northern smooth-tailed tree shrew (*Dendrogale murina*). *Annals of Anatomy-Anatomischer Anzeiger,* **181**(4): 397-402.

Fan Y, Huang ZY, Cao CC, Chen CS, Chen YX, Fan DD, He J, Hou HL, Hu L, Hu XT, Jiang XT, Lai R, Lang YS, Liang B, Liao SG, Mu D, Ma YY, Niu YY, Sun XQ, Xia JQ, Xiao J, Xiong ZQ, Xu L, Yang L, Zhang Y, Zhao W, Zhao XD, Zheng YT, Zhou JM, Zhu YB, Zhang GJ, Wang J, Yao YG. 2013. Genome of the Chinese tree shrew. *Nature Communications,* **4**: 1426.

Farah CS, Reinach FC. 1995. The troponin complex and regulation of muscle contraction. *The FASEB Journal,* **9**(9): 755-767.

Fuchs E, Corbach-Söhle S. 2010. Tree Shrews. *In*:The UFAW Handbook on the Care and Management of Laboratory and Other Research Animals. Oxford: Wiley-Blackwell, 262-275.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution,* **61**(5): 995-1016.

Huynen MA, Gabaldon T, Snel B. 2005. Variation and evolution of biomolecular systems: searching for functional relevance. *FEBS Letters,* **579**(8): 1839-1845.

Kosakovsky Pond SL, Murrell B, Fourment M, Frost SD, Delport W, Scheffler K. 2011. A random effects branch-site model for detecting episodic diversifying selection. *Molecular Biology and Evolution,* **28**(11): 3033-3043.

Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Mauceli E, Ward LD, Lowe CB, Holloway AK, Clamp M, Gnerre S, Alfoldi J, Beal K, Chang J, Clawson H, Cuff J, Di Palma F, Fitzgerald S, Flicek P, Guttman M, Hubisz MJ, Jaffe DB, Jungreis I, Kent WJ, Kostka D, Lara M, Martins AL, Massingham T, Moltke I, Raney BJ, Rasmussen MD, Robinson J, Stark A, Vilella AJ, Wen J, Xie X, Zody MC, Baldwin J, Bloom T, Chin CW, Heiman D, Nicol R, Nusbaum C, Young S, Wilkinson J, Worley KC, Kovar CL, Muzny DM, Gibbs RA, Cree A, Dihn HH, Fowler G, Jhangiani S, Joshi V, Lee S, Lewis LR, Nazareth LV, Okwuonu G, Santibanez J, Warren WC, Mardis ER, Weinstock GM, Wilson RK, Delehaunty K, Dooling D, Fronik C, Fulton L, Fulton B, Graves T, Minx P, Sodergren E, Birney E, Margulies EH, Herrero J, Green ED, Haussler D, Siepel A, Goldman N, Pollard KS, Pedersen JS, Lander ES, Kellis M. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature,* **478**: 476-482.

Macarthur DG, North KN. 2004. A gene for speed? The evolution and function of alpha-actinin-3. *BioEssays,* **26**(7): 786-795.

Mcginnis S, Madden TL. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research,* **32**(2): W20-W25.

Murphy WJ, Pevzner PA, O'brien SJ. 2004. Mammalian phylogenomics comes of age. *Trends in Genetics,* **20**(12): 631-639.

Norman B, Esbjornsson M, Rundqvist H, Osterlund T, Von Walden F, Tesch PA. 2009. Strength, power, fiber types, and mRNA expression in trained men and women with different ACTN3 R577X genotypes. *Journal of Applied Physiology,* **106**(3): 959-965.

Peng YZ, Ye ZZ, Zou RJ, Wang YX, Tian BP, Ma YY, Shi LM. 1991. Biology of Chinese Tree Shrews (*Tupaia belangeri chinensis*). Kunming, China: Yunnan Science and Technology Press.

Salmikangas P, Van Der Ven PF, Lalowski M, Taivainen A, Zhao F, Suila H, Schroder R, Lappalainen P, Furst DO, Carpen O. 2003. Myotilin, the limb-girdle muscular dystrophy 1A (LGMD1A) protein, cross-links actin filaments and controls sarcomere assembly. *Human Molecular Genetics,* **12**(2): 189-203.

Schmidt M, Schilling N. 2007. Fiber type distribution in the shoulder muscles of the tree shrew, the cotton-top tamarin, and the squirrel monkey related to shoulder movements and forelimb loading. *Journal of Human Evolution,* **52**(4): 401-419.

Sellers JR. 2000. Myosins: a diverse superfamily. *Biochimica et Biophysica Acta,* **1496**(1): 3-22.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution,* **28**(10): 2731-2739.

The Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature,* **437**: 69-87.

Tiso N, Majetti M, Stanchi F, Rampazzo A, Zimbello R, Nava A, Danieli GA. 1999. Fine mapping and genomic structure of ACTN2, the

human gene coding for the sarcomeric isoform of alpha-actinin-2, expressed in skeletal and cardiac muscle. *Biochemical and Biophysical Research Communications,* **265**(1): 256-259.

Wellik DM. 2007. Hox patterning of the vertebrate axial skeleton. *Developmental Dynamics,* **236**(9): 2454-2463.

Xu L, Zhang Y, Liang B, Lu LB, Chen CS, Chen YB, Zhou JM, Yao YG. 2013. Tree shrews under the spot light: emerging model of human diseases. *Zoological Research*, **34**(2): 59-69. (in Chinese)

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution,* **24**(8): 1586-1591.

Zeng YN, Shen YY, Zhang YP. 2013. Genome-wide scan reveals the molecular mechanisms of functional differentiation of *Myotis lucifugus* and *Pteropus vampyrus*. *Zoological Research,* **34**(3): 221-227. (in Chinese)

Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics Proteomics Bioinformatics,* **4**(4): 259-263.