Sequencing and analysis of the complete mitochondrial genome of the Eurasian least shrew (*Sorex minutissimus*) from China

Liu Zhu, Dang Yun-Qi and Li Jiao-Jiao

College of Life Science and Technology, Mudanjiang Normal University, Mudanjiang, PR China

ABSTRACT

The complete mitogenome sequence of the Eurasian least shrew (*Sorex minutissimus*) was determined using long PCR. The genome was 17,061 bp in length and contained 13 protein-coding genes, 2 ribosomal RNA genes, 22 transfer RNA genes, 1 origin of L strand replication and 1 control region. The overall base composition of the heavy strand is A (32.6%), C (24.7%), T (29.1%) and G (13.6%). The base compositions present clearly the A–T skew, which is most obviously in the control region and protein-coding genes. The extended termination-associated sequence domain, the central conserved domain and the conserved sequence block domain are defined in the mitochondrial genome control region of the Eurasian least shrew. Mitochondrial genome analyses based on MP, ML, NJ and Bayesian analyses yielded identical phylogenetic trees. The ten *Sorex* species formed a monophyletic group with the high bootstrap value (100%) in all examinations. The control region of the Eurasian least shrew mitochondrial genome control region (Zhang et al. 2009): the extended termination-associated sequence (ETAS) domain, the central conserved domain (CD) and the conserved sequence block (CSB) domain. Three CSBs were found in the CSB domain and they were located in positions 16,284–16,319, 16,714–16,748 and 16,764–16,802. Also only one repetitive sequence region (RS) was found, which was located between the CSB1 and CSB2, and was rich in A and C. The repetitive pattern of segments in the RS was 5'-TG-(TACACGTACACGTA)n-CA-3' (n = 12).

The total length of the protein-coding gene sequences was 11,475 bp. Most protein-coding genes initiate with ATG except for ND2, ND3, ND5 and ND6, which began with ATA, ATT or TAC. Six protein-coding genes terminated with TAA whereas the Cyt b gene terminated with AGG. The incomplete stop codons (T—— or TA——) were used in ND1, ND2, COX3, ND3 and ND6. A strong bias against A at the third codon position was observed in the protein-coding genes. The frequencies of CTA (Leu), ATT (Ile), TTA (Leu) and ATA were higher than those of other codons. The length of tRNA genes varied from 59 to 75 bp. Twenty-one of them could be folded into the typical cloverleaf secondary structure except the tRNA-Ser (AGY), whose complete dihydrouridine arm was lacking.

Most the Eurasian least shrew mitochondrial genes were encoded on the H strand, except for the ND6 gene and eight tRNA genes, which were encoded on the L strand. Some reading frame intervals and overlaps were found. One of the most typical was between ATP8 and ATP6. The L-strand replication origin (OL) was located within the WANCY region.
containing five tRNA genes (tRNATrp, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr). This region was 37 bp long and had the potential to fold into a stable stem-loop secondary structure.

The total base composition of the Eurasian least shrew mitochondrial genome was A (32.6%), C (24.7%), T (29.1%) and G (13.6%). The base compositions clearly present the A-T skew, which was most obviously in the control region and protein-coding genes.

In order to explore the evolution of Insectivora shrews which include Soricidae and Talpidae, especially the evolution of genus Sorex from China, here, we investigate the molecular phylogenetics of Chinese the Eurasian least shrew using complete mitochondrial genome sequence of 31 species. All sequences generated in this study have been deposited in the GenBank (Figure 1).

Mitochondrial genome analyses based on MP, ML, NJ and Bayesian analyses yielded identical phylogenetic trees, indicating a close phylogenetic affinity of shrews. The phylogram obtained from Maximum Parsimony method is shown in Figure 1. It shows that two major phyletic lineages were present in Insectivora: Soricidae and Talpidae. Soricidae comprised Crocidura lasiura, Crocidura attenuata, Crocidura russula, Episoricus macrurus, Episoricus caudatus, Neomys fodiens, Nectogale elegans, Anourosorex squamipes, Blarinella quadricauda, Soriculus fumidus, Suncus murinus, Sorex araneus, Sorex tundrensis, Sorex caecutiens, Sorex minutissimus, Talpa europaea, Urotrichus talpoides, Uropsilus soricipes, Galemys pyrenaicus, Scapanulus oweni, Mogera robusta, Mogera wogura, Mogera wagleri (AB099483), Mogera robusta (KT934322), Condylura cristata (KU144678), Galemys pyrenaicus (AY833419), Scapanulus oweni (KMS06754) and Erinaceus europaeus (NC002080).

Figure 1. Phylogenetic tree generated using the Maximum Parsimony method based on complete mitochondrial genomes. Crocidura lasiura (KR007669), Crocidura shantungensis (JX968507), Crocidura attenuata (KP120863), Crocidura russula (AY789264), Episoricus macrurus (KU246040), Episoricus caudatus (KMS03097), Neomys fodiens (KM092492), Nectogale elegans (KC503902), Anourosorex squamipes (KJ545899), Blarinella quadricauda (KJ131179), Suncus murinus (KJ920196), Soriculus fumidus (AF348081), Sorex araneus (KT210896), Sorex cylindrica (KF696672), Sorex unguculatus (AB061527), Sorex tundrensis (KMS07275), Sorex caecutiens (MF374796), Sorex roboreatus (KY930906), Sorex isodon (MG983792), Sorex gracilis (MF426913), Sorex mirabilis (MF426265), Sorex minutissimus (MH823669), Talpa europaea (Y19192), Urotrichus talpoides (AB099483), Uropsilus soricipes (JQ658979), Uropsilus gracilis (KM379136), Mogera wagleri (AB099482), Mogera robusta (KT934322), Condylura cristata (KU144678), Galemys pyrenaicus (AY833419), Scapanulus oweni (KMS06754) and Erinaceus europaeus (NC002080).
Sorex roboratus, Sorex isodon, Sorex gracillus, Sorex mirabilis, Sorex cylindricauda, Sorex unguiculatus and Sorex minutissimus was supported by bootstrap values of 100%. Talpidae comprised Talpa europaea, Urotrichus talpoides, Mogera wogura, Condylura cristata, Uropsilus soricipes, Mogera robusta, Galemys pyrenaicus, Uropsilus gracilis and Scapanulus oweni was supported by bootstrap values of 100%. The ten Sorex species formed a monophyletic group with the high bootstrap value (100%) in all examinations.

Disclosure statement
The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding
This research was supported by the Heilongjiang Provincial Department of Education Filing Project [1351MSYYB009], Heilongjiang Provincial Natural Funds [C2017065] and the Doctoral Scientific Research Foundation Project of Mudanjiang Normal University [MNUB201404].

References
Fontanillas P, Depraz A, Giorgi MS, Perrin N. 2005. Nonshivering thermogenesis capacity associated to mitochondrial DNA haplotypes and gender in the greater white-toothed shrew, Crocidura russula. Mol Ecol. 14:661–670.
Huang T, Yan CC, Tan Z, Tu FY, Yue BS, Zhang XY. 2014. Complete mitochondrial genome sequence of Nectogale elegans. Mitochondrial DNA. 25:253.
Huang T, Dang X, An M, Chen L, Zhang J. 2016. The complete mitochondrial genome of the Sorex araneus. Mitochondrial DNA. 27:3655–3656.
Jin ZM, Liu Z, Ma JZ. 2017. Sequencing and analysis of the complete mitochondrial genome of the masked shrew (Sorex caecutiens) from China. Mitochondrial DNA Part B. 2:486–488.
Kim HR, Park JK, Cho JY, Chul Park Y. 2013. Complete mitochondrial genome of an Asian Lesser White-toothed Shrew, Crocidura shantungensis (Soricidae). Mitochondrial DNA. 24:202–204.
Kim TW, Kim YK, Oh DJ, Park JH, Kim D, Adhikari P, Kim G, Park SM, Lee JW, Jung YH, et al. 2017. Complete mitochondrial genome of the Ussuri white-toothed shrew Crocidura lasiura (Insectivora, Soricidae). Mitochondrial DNA Part A. 28:216–217.
Liu Z, Bai W, Wang AN, Tian XM, Li DW. 2018. Sequencing and analysis of the complete mitochondrial genome of the taiga shrew (Sorex isodon) from China. Mitochondrial DNA Part B. 3:466–468.
Liu Z, Tian XM, Jin JL, Jin ZM, Li DW, Zhang JS. 2017. Sequencing and analysis of the complete mitochondrial genome of the slender shrew (Sorex gracilinus) from China. Mitochondrial DNA Part B. 2:642–644.
Liu Z, Tian XM, Jin ZM, Dong M, Zhang JS. 2017. Sequencing and analysis of the complete mitochondrial genome of the Ussuri shrew (Sorex mirabilis) from China. Mitochondrial DNA Part B. 2:645–647.
Liu Z, Wang AN, Zhang JS, Yang X, Liu H. 2017. Sequencing and analysis of the complete mitochondrial genome of flat-skulled shrew (Sorex roboratus) from China. Mitochondrial DNA Part B. 2:369–371.
Liu Z, Zhao W, Liu P, Li S, Xu C. 2016. The complete mitochondrial genome of Eurasian water shrew (Neomys fodiens). Mitochondrial DNA Part A. 27:2381–2382.
Meganathan PR, Pagan HJT, McCulloch ES, Stevens RD, Ray DA. 2012. Complete mitochondrial genome sequences of three bats species and whole genome mitochondrial analyses reveal patterns of codon bias and lend support to a basal split in Chiroptera. Gene. 492:121–129.
Nikaido M, Kawai K, Cao Y, Harada M, Tomita S, Okada N, Hasegawa M. 2001. Maximum likelihood analysis of the complete mitochondrial genomes of eutherians and a reevaluation of the phylogeny of bats and insectivores. J Mol Evol. 53:506–508.
Xu CZ, Zhang HH, Ma JZ. 2013. The complete mitochondrial genome of sable, Martes flavigula. Mitochondrial DNA. 24:240–242.
Xu CZ, Zhang HH, Ma JZ, Liu ZH. 2012. The complete mitochondrial genome of sable, Martes zibellina. Mitochondrial DNA. 23:167–169.
Xu CZ, Zhao S, Wu HL, Wu SY, Zhang ZW, Wang B, Dou HS. 2016. Sequencing analysis of the complete mitochondrial genome of tundra shrew (Sorex tundrensis) from China. Mitochondrial DNA. 27:2354–2355.
Yoon KB, Kim HR, Kim JY, Jeon SH, Park YC. 2013. The complete mitochondrial genome of the Ussurian tube-nosed bat Murina ussuriensis (Chiroptera: Vespertilionidae) in Korea. Mitochondrial DNA. 24:397–399.
Zhang HH, Xu CZ, Ma JZ. 2009. Structure of the mtDNA control region and phylogeny of the Mustelidae species. Acta Ecol Sin. 29:3585–3592.