



# Mitochondrial DNA Diversity of *Mesocricetus auratus* and Other Cricetinae Species among Cricetidae Family

Ruijing Xuan<sup>1</sup> · Jiping Gao<sup>1</sup> · Qiang Lin<sup>2</sup> · Wenbin Yue<sup>3</sup> · Tianfu Liu<sup>1</sup> · Songnian Hu<sup>2</sup> · Guohua Song<sup>1</sup> 

Received: 11 June 2021 / Accepted: 25 January 2022 / Published online: 5 February 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Unique anatomical and physiological features have made hamster species desirable research models. Comparative genomics and phylogenetic analysis of the hamster family members to clarify their evolution and genetic relationship, can provide a genetic basis for the comprehension of the variable research results obtained using different hamster models. The Syrian golden hamster (*Mesocricetus auratus*) is the most widely used species. In this study, we sequenced the complete mitochondrial genome (mitogenome) of *M. auratus*, compared it with the mitogenome of other Cricetinae subfamily species, and defined its phylogenetic position in the Cricetidae family. Our results show that the mitogenome organization, gene arrangement, base composition, and genetic analysis of the protein coding genes (PCGs) of *M. auratus* are similar to those observed in previous reports on Cricetinae species. Nonetheless, our analysis clarifies some striking differences of *M. auratus* relative to other subfamily members, namely distinct codon usage frequency of TAT (Tyr), AAT (Asn), and GAA (Glu) and the presence of the conserved sequence block 3 (CSB-3) in the control region of *M. auratus* mitogenome and other hamsters (not found in Arvicolinae). These results suggest the particularity of amino acid codon usage bias of *M. auratus* and special regulatory signals for the heavy strand replication in Cricetinae. Additionally, Bayesian inference/maximum likelihood (BI/ML) tree shows that Cricetinae and Arvicolinae are sister taxa sharing a common ancestor, and Neotomiinae split prior to the split between Cricetinae and Arvicolinae. Our results support taxonomy revisions in *Cricetulus kamensis* and *Cricetulus migratorius*, and further revision is needed within the other two subfamilies. Among the hamster research models, *Cricetulus griseus* is the species with highest sequence similarity and closer genetic relationship with *M. auratus*. Our results show mitochondrial DNA diversity of *M. auratus* and other Cricetinae species and provide genetic basis for judgement of different hamster models, promoting the development and usage of hamsters with regional characteristics.

Extended author information available on the last page of the article

**Keywords** *Mesocricetus auratus* · Mitochondrial genome · Cricetinae · Mitogenomic comparison · Phylogenetic relationship

## Introduction

Hamsters, due to several unique anatomical and physiological features, are desirable research models. The Syrian hamster (golden), *Mesocricetus auratus*; the Chinese hamster (striped-back), *Cricetulus griseus*; the Armenian (gray), *Cricetulus migratorius*; and the European hamster, *Cricetus cricetus* are the species majorly used as research models (Miedel et al. 2015). Among them, the golden hamster is the most widely used species (Tchitchek et al. 2014), and it was recently reported as a useful small animal model that can be used to evaluate the effects of vaccines, immunotherapy, and antiviral drugs (Chan et al. 2020; Imai et al. 2020; Kreye et al. 2020; Rosenke et al. 2020; Sia et al. 2020).

Due to local and regional developmental restrictions, the growth of laboratory animals in China is unbalanced and hamster resources are scarce and expensive. Contradictory results can often be observed when using different hamster models. To overcome this problem, comparative genomics and phylogenetic analysis can be used to clarify the evolution and genetic relationship between different hamster species, enlightening researchers on the usage of hamster resources, and improving their utilization for biomedical research and human health. The mitogenome has been widely used in comparative genomics research (Yang et al. 2019) and the detailed comparison of the mitogenome from different species helps in understanding of their sequence characteristics and structural diversity. *M. auratus* from inbred strains show expected lower diversity values (YİĞİT et al. 2007) compared with the diversity of the outbred group, an observation that requires more attention. We thus aimed to analyze the complete mitogenome information of different hamster species to clarify the phylogenetic and evolutionary relationship between them.

In the present study, the complete mitogenome of *M. auratus* was sequenced, and compared with the mitogenome of other members of the Cricetinae subfamily. Their phylogenetic position in the Cricetidae family was defined. The results of this study enlighten the evolution of and genetic relationship between several hamster species commonly used as research models, facilitating the comprehension of research results obtained using different models and playing a positive role in the development of hamster resources.

## Materials and Methods

### Sample Source and DNA Extraction

*M. auratus*, an outbred hamster, was obtained from Beijing Vital River Laboratory Animal Technology Co., Ltd. in China. The tail tip (1–2 cm) was cut and stored in absolute ethanol. SDS/Protease K was used to lyse the tissues. Total genomic DNA

was obtained using phenol–chloroform extraction method (Lu 1999) and analyzed using agarose gel electrophoresis.

### Primer Design and PCR Amplification

The complete mitogenome sequences from KK mouse (EF108339), Wistar Rat (AC000022), *Cavia porcellus* (AJ222767), Taiwan voles (NC003041), Southern Voles (NC008064), and other hamsters were aligned using CLUSTAL X 1.8 software. Twenty seven pairs of primers covering the complete mitochondrial genome of *M. auratus* were designed using Oligo 6.0 software (Table S1).

Touchdown PCR program was used with the following procedure and cycle parameters: (1) 94 °C pre-denaturation for 5 min (min); (2) two cycles of 94 °C denaturation for 30 s (s), 48–45 °C annealing for 1 min, 72 °C extension for 1 min 30 s; (3) 30 cycles of 94 °C denaturation for 30 s, 44 °C annealing for 1 min, 72 °C extension for 1 min 30 s; and (4) 72 °C extension for 10 min. The PCR product was analyzed using 1.0% agarose gel electrophoresis.

### Product Purification, Sequencing, and Sequence Splicing

The PCR products were purified with SAP enzymes, and two-way sequencing was performed with the ABI 3730 DNA sequence automatic analyzer (Waltham, USA). DNASTAR and Chromas 2.22 were used to perform sequence alignment and concatenation. Subsequently, the complete mitogenome of *M. auratus* was obtained.

### Gene Identification and Genome Analysis

BLAST online service and Clustal W 1.8 were used to search the homologous sequence and complete the multiple sequences alignment among the mitogenomes of Cricetidae species. The protein coding genes (PCGs), transfer-RNA (tRNA) genes, ribosomal-RNA (rRNA) genes, and control region (CR, also D-loop region) were analyzed and their location was obtained using DNASTAR, RNASTructure 4.5, and tRNA-scan SE web server (Lowe and Chan 2016). The mitogenome sequence of *M. auratus* was also submitted to MITOS Web Server and annotated by referring the vertebrate genetic code. In addition to the confirmation of the initial annotations for PCGs, tRNA, rRNA and D-loop, O<sub>L</sub> (origin for the light strand replication) was also annotated and the secondary structure of tRNAs was acquired. The Mtv online tool was used to draw the circular mitogenome chart of *M. auratus* (Ding et al. 2016a).

The base composition of the mitogenome was determined with MEGA 5.2.2, and the skewness of nucleotide composition was calculated according to the following formulas: AT skew =  $[A - T]/[A + T]$  and GC skew =  $[G - C]/[G + C]$  (Perna and Kocher 1995). Codon usage was determined with MEGA 6.0. For each of the PCGs from Cricetinae mitogenomes described in NCBI, the synonymous substitution rate (Ks) and the nonsynonymous substitution rate (Ka) were calculated with DnaSP v6.12.03 (Rozas 2017). Genetic distances of the 13 PCGs were calculated with MEGA 7.0 using the Kimura-2-parameter model (K2P) (Tamura et al. 2013).

## Phylogenetic Tree Construction

Heavy chain PCGs from 32 species (30 Cricetidae species including *M. auratus* and two outgroups) were used to construct the phylogenetic tree of Cricetidae. *Mus musculus* [NC\_005089] (Bayona-Bafaluy et al. 2003) and *Rattus norvegicus* [KF011917] were used as the outgroups. The FASTA files of 12 PCGs from 32 species were downloaded from GenBank (Table 1). Multiple alignments of the complete nucleotide sequences of the 12 PCGs individually were conducted with MEGA 7.0. SequenceMatrix 1.8 was used to concatenate the individual gene alignment and create a combined sequence matrix (Li et al. 2018). Paup 4.0 was used to transform the file format to nex., so that it could be recognized by Bayesian. Bayesian inference (BI) (Yang and Rannala 1997) and maximum likelihood (ML) methods (Felsenstein 1981) were used to construct the phylogenetic tree based on the concatenated sequence matrix. MrModeltest 2.3 (Nylander 2004) was used to determine the best-fit evolutionary model for the concatenated sequence matrix with the Akaike Information Criterion (AIC) (Posada and Buckley 2004). Based on the test results, the GTR+I+G was considered as the appropriate model for the BI and ML analysis. The ML analysis was performed with IQ-TREE v1.6.2, and the node support was assessed by the bootstrap values calculated after 1,000 replicates. The BI analysis was executed using MrBayes v3.2.6 (Ronquist et al. 2012), with the following command program: Lset nst=6 rates=invgamma; Prset statefreqpr=dirichlet (1, 1, 1, 1); mcmc ngen=2,000,000 printfreq=1,000 samplefreq=100 nchains=4 nruns=2 checkfreq=5,000; sumt burninfrac=0.25. A good indication of convergence was considered to be reached when the average standard deviation of split frequencies was below 0.01. FigTree v1.4.3 was used to view the phylogenetic trees.

## Results and Discussion

### Genome Organization and Base Composition

*M. auratus* mitogenome is 16,264 base-pair (bp) in length, with a typical circular structure (Fig. 1), including 37 genes and two non-coding regions. The 37 genes include 13 PCGs, 22 tRNA genes, and two rRNA genes, with nine genes located in light chain and the other 28 genes in heavy chain. The non-coding regions are  $O_L$  and CR. Both gene overlap and spacing are observed in *M. auratus* mitogenome, with the overlapped length being longer than the spacing (Table S2). Such organization is common in Cricetinae species (Nylander 2004, 1981, 2007; Ding et al. 2016a, 2016b). *Cricetulus longicaudatus* is an exception, with 13 bp of intergenic nucleotides, differing from the overlaps observed in the other six species (Table S3). Overlapping genes may compose a compact genome organization to facilitate gene regulation efficiency (Ho et al. 2012).

The base composition of the mitogenome of *M. auratus* is biased towards A and T (Table S4). In all the surveyed mitogenomes, the absolute values of GC skews (negative) were always higher than AT skews (positive) (Table 1), underlining the stronger GC skew. The GC skew of *M. auratus* is closest to the one of *C. griseus*.

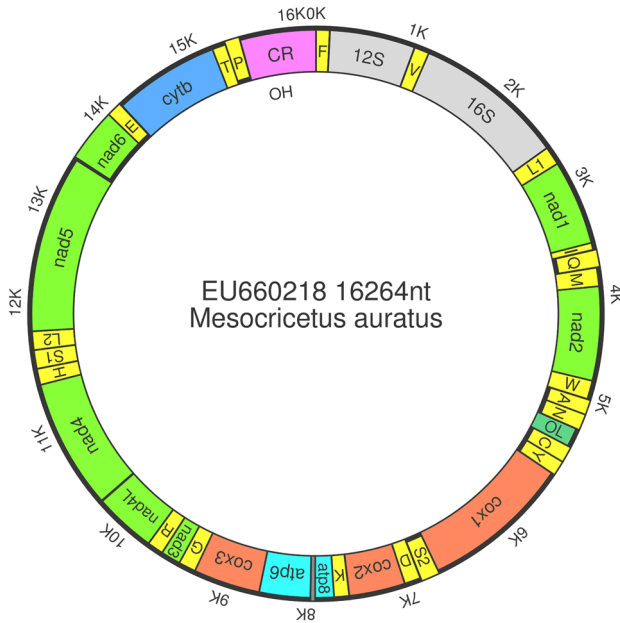
**Table 1** Sequences data used in the phylogenetic analysis

Family	Subfamily	Genus	species	GBAN	References	A + T%	AT skew	GC skew
Cricetidae	Cricetinae	Mesocricetus	<i>Mesocricetus auratus</i>	EU660218	This study	63.71	0.024	-0.280
Cricetidae	Cricetinae	Cricetulus	<i>Cricetulus griseus</i>	KX576660	Partridge et al. (2007)	64.24	0.050	-0.274
Cricetidae	Cricetinae	Cricetulus	<i>Cricetulus kamensis</i>	KJ680375	Kang et al. (2016)	60.01	0.106	-0.362
Cricetidae	Cricetinae	Cricetulus	<i>Cricetulus longicaudatus</i>	KM067270	Zhang et al. (2016a)	63.54	0.035	-0.258
Cricetidae	Cricetinae	Cricetulus	<i>Cricetulus migratorius</i>	KT918407	Ding et al. (2016a)	60	0.024	-0.322
Cricetidae	Cricetinae	Cricetulus	<i>Cricetulus cricetus</i>	MF405145	Ding et al. (2020)	59.13	0.041	-0.320
Cricetidae	Cricetinae	Phodopus	<i>Phodopus roborovskii</i>	KU885975	Ding et al. (2016b)	61.37	0.106	-0.312
Cricetidae	Neotominae	Peromyscus	<i>Peromyscus pectoralis</i>	KY707309	Sullivan et al. (2017)	63.11	0.111	-0.335
Cricetidae	Neotominae	Peromyscus	<i>Peromyscus atwateri</i>	KY707299	Sullivan et al. (2017)	64.09	0.099	-0.324
Cricetidae	Neotominae	Peromyscus	<i>Peromyscus crinitus</i>	KY707308	Sullivan et al. (2017)	62.18	0.101	-0.312
Cricetidae	Neotominae	Peromyscus	<i>Peromyscus melanophrys</i>	KY707303	Sullivan et al. (2017)	62.88	0.122	-0.323
Cricetidae	Neotominae	Peromyscus	<i>Peromyscus megalops</i>	KY707305	Sullivan et al. (2017)	62.85	0.120	-0.333
Cricetidae	Neotominae	Peromyscus	<i>Peromyscus aztecus</i>	KY707306	Sullivan et al. (2017)	63.08	0.093	-0.303
Cricetidae	Neotominae	Peromyscus	<i>Peromyscus polionotus</i>	KY707301	Sullivan et al. (2017)	62.85	0.110	-0.309
Cricetidae	Neotominae	Neotomodon	<i>Neotomodon alstoni</i>	KY707310	Sullivan et al. (2017)	62.51	0.119	-0.343
Cricetidae	Neotominae	Neotoma	<i>Neotoma fuscipes</i>	KU745736	Bendova et al. (2016)	62.19	0.132	-0.349
Cricetidae	Neotominae	Neotoma	<i>Neotoma magister</i>	NC_039670	Schofield et al. (2018)	60.88	0.123	-0.351
Cricetidae	Neotominae	Podomys	<i>Podomys floridanus</i>	KY707302	Sullivan et al. (2017)	64.69	0.099	-0.303
Cricetidae	Neotominae	Habromys	<i>Habromys taylori</i>	KY707304	Sullivan et al. (2017)	62.31	0.102	-0.321
Cricetidae	Neotominae	Isthmomyys	<i>Isthmomyys pirrensis</i>	KY707312	Sullivan et al. (2017)	62.55	0.141	-0.341
Cricetidae	Arvicolinae	Myodes	<i>Myodes glareolus</i>	KF918859	Bendova et al. (2016)	60.13	0.100	-0.343
Cricetidae	Arvicolinae	Myodes	<i>Eothenomys regulus</i>	JN629046	Ds	59.71	0.105	-0.345
Cricetidae	Arvicolinae	Myodes	<i>Myodes rufocanus</i>	KT725595	Lu et al. (2017)	59.93	0.103	-0.343
Cricetidae	Arvicolinae	Eothenomys	<i>Eothenomys chinensis</i>	F1483847	Yang et al. (2012)	59.86	0.108	-0.335
Cricetidae	Arvicolinae	Eothenomys	<i>Eothenomys melanogaster</i>	KP997311	Yang et al. (2012)	59.48	0.104	-0.335
Cricetidae	Arvicolinae	Eothenomys	<i>Eothenomys inez</i>	KU200225	Yu et al. 2016	60.24	0.104	-0.348

Table 1 (continued)

Family	Subfamily	Genus	species	GBAN	References	A + T%	AT skew	GC skew
Cricetidae	Arvicolinae	Neodon	<i>Neodon sikimensis</i>	KU891252	Zhang et al. (2016b)	59.61	0.118	-0.349
Cricetidae	Arvicolinae	Neodon	<i>Neodon fuscus</i>	NC_040138	Li et al. (2019)	58.45	0.092	-0.309
Cricetidae	Arvicolinae	Ondatra	<i>Ondatra zibethicus</i>	KX377613	Zhao et al. (2018)	62.05	0.107	-0.330
Cricetidae	Arvicolinae	Microtus	<i>Microtus rossiaemeridionalis</i>	DQ015676	Triant and Dewoody (2006)	60.24	0.094	-0.316
Muridae	Murinae	Mus	<i>Mus musculus</i>	NC_005089	outgroup	63.26	0.092	-0.328
Muridae	Murinae	Rattus	<i>Rattus norvegicus</i>	KF011917	outgroup	61.48	0.108	-0.361

\*Ds: Direct submission, GBAN: Genbank accession numbers



**Fig. 1** Circular chart of *M. auratus* mitochondrial genome. The single-letter abbreviation of the tRNA-carried amino acid was used to indicate the corresponding tRNA gene. The thick black lines outside the ring indicate that the genes are located in the heavy chain and that inside the ring indicate that the genes are located in the light chain. Different gene families were represented with different colors: gray for rRNAs, yellow for tRNAs, light green for nad (1–6), dark green for OL, cyan for atp (6 and 8), orange for cox (1–3), blue for cytb and pink for CR

**PCGs**

The PCGs sequence accounted for 70% of *M. auratus* mitogenome. We could observe that the shortest gene is *atp8*, and that *nad5* is the longest one. From the 13 PCGs of the mitogenome, 12 are located in heavy chain, and only *nad6* is localized in light chain. For most PCGs, ATN (N=T/G) is the start codon, and TAN (N=A/G) is the stop codon. As exceptions, *nad1* start from GTG, which is found in several *Cricetulus* species (Partridge et al. 2007; Jiang et al. 2012; Ding et al. 2016a, 2016b) and four PCGs stop as T-. Incomplete stop codon is common in vertebrate mitogenomes and is thought to regulate mRNA maturation via post-transcriptional modification (Ojala et al. 1981; Jiang et al. 2012).

The codon usage analysis of *M. auratus* mitogenome revealed that TAT (Tyr) and AAT (Asn) are the most frequently used codons (Fig. S1), while in other *Cricetulus* species, AUU (Ile) and CUA (Leu) are the mostly used (Ding et al. 2016a, 2016b), suggesting that *M. auratus* has a particular bias for amino acid codon usage. In *M. auratus* mitogenome, most amino acids have more than one synonymous codon, with the exception of Glutamate, which has only one codon, GAA; this means that Glutamate and its corresponding proteins are at greater mutation risk.

Figure S2 shows the genetic analysis of 13 PCGs from seven Cricetinae mitogenomes. The gene *nad2* has the largest genetic distance (K2P) followed by *atp8* and *nad3*, suggesting these genes vary largely between species, while *cox1* and *cox3* are relatively conserved genes. The rich phylogenetic information and high degree of conservation of *cox1* allows for this gene to be used as a bar-code in species identification of multiple taxa. The lowest Ka/Ks of *cox1*, when compared to other mitogenome genes, suggests the weakest natural selective pressure of this gene, while *atp8* shows the opposite tendency. Ka/Ks lower than 1 indicates that the 13 PCGs are undergoing purification selection.

## tRNA

In our analysis, we could find 22 tRNAs with variable length (Table S2) spread over the entire *M. auratus* mitogenome. They can recognize all amino acid codons and transfer the 20 common amino acids. Fourteen tRNAs are in heavy chain and eight in light strand (Fig. 1). Concerning the formation of secondary structures, *trnS1* (GCT) is the only gene that cannot be folded into the typical cloverleaf structure due to lack of dihydrouridine (DHU) and acceptor stems (Fig. S3). This feature is common in vertebrates (Jiang et al. 2012).

The ratio of mitochondrial tRNAs to encoded amino acids is approximately 1:1. If a specific mitochondrial tRNA gene has a mutation with functional consequences, other tRNAs are not able to compensate for it. Therefore, many diseases related to defects in oxidative phosphorylation are caused by mutations in mitochondrial tRNA genes. Sequencing of genes of experimental animals coding for mitochondrial tRNA will facilitate the research of mitochondrial diseases originating from tRNA gene mutations.

## CR and OL

We also analyzed the location, size, and base composition of CR (Tables S2 and S4), as well as the internal organization of CR included ETAS (extended termination associated sequences), central domain, and conserved sequence blocks (CSBs) (Table S5). These sequences were identified through the alignment with sequences of dozens of mammalian species already reported, including several *Cricetulus* species (Sbisà et al. 1997; Ding et al. 2016a, 2016b). The CSB and ETAS domains represent the initiation and stop region of heavy chain synthesis, respectively. CSB domains have been proposed to be regulatory signals of RNA primer for the heavy chain replication. Figure S4 shows the conservative sequences in CSB domains. CSB-1 is present in all the organisms considered in this study, while CSB-2 and CSB-3 could not be found in the mitogenome of three species of genus *Microtus* (Arvicolinae, Rodentia) (Jiang et al. 2012). The hairpin structures formed by the complementary sequences in ETAS-1 and ETAS-2 may be the recognition sites for synthesis conclusion (Saccone et al. 1991).

O<sub>L</sub> region is located between *trnN* and *trnC* in the WANCY cluster, with expected length of 34 bp in *M. auratus*, similar to *C. griseus* (Partridge et al. 2007). Figure S5



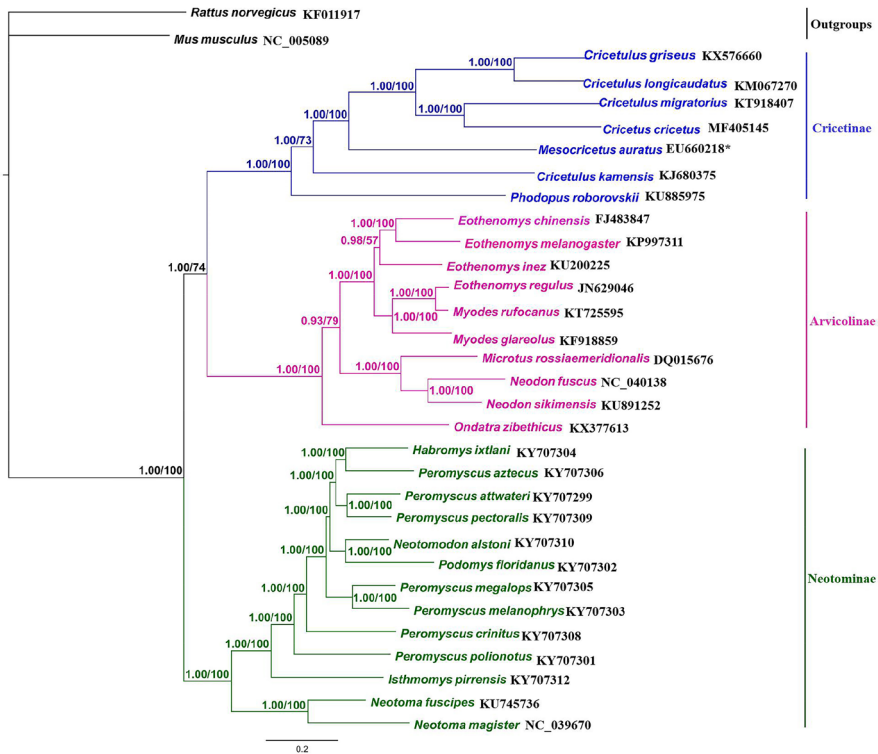
shows its secondary structure with a stable stem-loop. This origin is highly conserved among the conveyed seven Cricetinae species (Fig. S6). Its T-rich sequence is of great significance for initiating light chain synthesis (Hixson et al. 1986).

## Phylogenetic Analysis

The 30 Cricetidae species used in the phylogenetic tree analysis were selected from the NCBI blast results, based on the similarity of their mitogenome with the one of *M. auratus*. Three subfamilies were included: Neotominae, Arvicolinae, and Cricetinae with seven species from the most latter (Table 1). In the blast analysis, *M. auratus* mitogenome shared highest sequence similarity with Cricetinae species. Following traditional taxonomy, our seven Cricetinae samples included four *Cricetulus* species (*Cricetulus kamensis*, *C. griseus*, *C. longicaudatus*, *C. migratorius*), one *Cricetus* species (*C. cricetus*), one *Mesocricetus* species (*M. auratus*) and one *Phodopus* species (*Phodopus roborovskii*). Previous evolution studies at both the genes and the chromosomal levels indicated that *Cricetus*, *Mesocricetus* and *Phodopus* represented a monophyletic clade (Neumann et al. 2006; Romanenko et al. 2007). However, various studies have suggested that *Cricetulus* is not monophyletic (Neumann et al. 2006; Lebedev et al. 2018; Ding et al. 2020), such that Lebedev et al. (2018) formally proposed a new monotypic genus *Nothocricetulus* for *C. migratorius*, and suggested elevating the subgenus *Urocricetus* to the rank of a full genus. Our BI/ML tree also showed that *C. (Urocricetus) kamensis* was phylogenetically distant from *Cricetulus* (Fig. 2), which would support the taxonomic revisions proposed by Lebedev et al. (2018) and indicate that names on public databases need updating.

The evolutionary time of the three Cricetidae subfamilies can be analyzed according to the horizontal length of different branches of the tree, i.e. the order of their occurrence was: Neotominae, (Arvicolinae + Cricetinae). Although, according to the cumulative distance of branch length and the NCBI blast analysis, Neotominae species have higher sequence similarity with Cricetinae, both BI and ML analysis showed that Cricetinae and Arvicolinae are sister taxa. They share a common ancestor, and split at the same time. Neotominae split prior to the split between Arvicolinae and Cricetinae. These observations are consistent with ML analysis results by Jiang et al. (2012) and emphasizing that high similarity is not equivalent to close genetic relationship. According to the horizontal length of Cricetinae branches, the order of occurrence of the seven Cricetinae species was as follows: (1) *Phodopus roborovskii*, (2) *C. kamensis* + *M. auratus*, (3) *C. griseus* + *C. longicaudatus*, and (4) *C. cricetus* + *C. migratorius*.

Among the several hamster species used as research models, *C. griseus* (also known as the Chinese hamster or the striped-back hamster) shares the highest sequence similarity with *M. auratus*. At the same time, *C. griseus* has specific characteristics that encourage its use in biomedical research: small size, short gestation period, polyestrous cycle, and low chromosome number. Although the use of this species in research is greatly overshadowed by the extensive use of cell lines derived from its ovarian cells (Miedel et al. 2015), there is still room



**Fig. 2** A consensus tree constructed using BI and ML inferred from 12 concatenated PCGs of 30 Cricetidae species. The number before and after the slash were the Bayesian posterior probabilities and bootstrap values, representing the support values of BI and ML, respectively. Muridae species (*Mus musculus* and *Rattus norvegicus*) were used as outgroups. GeneBank accession numbers were listed after species names; \**Mesocricetus auratus* was sequenced in the present study. Different colors represent different subfamilies: blue for Cricetinae, pink for Arvicolinae, green for Neotominae and black for outgroups

for development of their use in tumor transplantation and, like the golden hamster (*M. auratus*) (Chan et al. 2020; Imai et al. 2020; Kreye et al. 2020; Rosenke et al. 2020; Sia et al. 2020), in new coronavirus infection animal model to evaluate vaccines, immunotherapy, and antiviral drugs effects.

## Conclusions

Although the mitogenome organization, gene arrangement, base composition, and genetic analysis of PCGs of *M. auratus* are similar to previous reports on Cricetinae species, in this study we could identify special characteristics for this species. In *M. auratus* mitogenome, TAT (Tyr) and AAT (Asn) were the most frequently used codons, while GAA (Glu) was the least used one, indicating the particularity of amino acid codon usage bias of this species. The CSB-3 in control region, which is not found in

Arvicolinae species, can be observed in the mitogenome of *M. auratus* and other hamsters, suggesting that the regulatory signals for the heavy chain replication is specific for Cricetinae species. The BI/ML tree showed that Cricetinae and Arvicolinae are sister taxa sharing a common ancestor, and Neotominae split prior to the split between Cricetinae and Arvicolinae. Our results support taxonomy revisions in *C. kamensis* and *C. migratorius*, and further revision is needed within the other two subfamilies. Among the hamster species used as research models, *C. griseus* shares the highest sequence similarity with *M. auratus*, as well as closer genetic relationship, allowing for a better understanding of distinct experimental results between the two models. We consider there is still a variety of possible approaches for the improvement of usage of these species of hamster as research models.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10528-022-10195-7>.

**Acknowledgements** This research was funded by Special Funds for Experimental Animal Technologies of Shanxi Province, China (NO. 2009K02), Shanxi Natural Science Foundation of China (No. 2012011034-1), PhD Start-up Fund of Shanxi Medical University (No. 03201552), and Youth Fund of Shanxi Medical University (No. 02201626).

**Author Contributions** Conceptualization, RX; Data curation, RX, JG, QL and GS; Methodology, RX, JG, QL and GS; Resources, WY, TL, SH and GS; Software, RX, JG and QL; Writing—original draft, RX; Writing—review & editing, RX and JG; Funding acquisition, GS and RX; Supervision, WY, TL and SH; Project administration, GS.

**Funding** This research was funded by Special Funds for Experimental Animal Technologies of Shanxi Province, China (No. 2009K02), Shanxi Natural Science Foundation of China (No. 2012011034–1), PhD Start-up Fund of Shanxi Medical University (No. 03201552), and Youth Fund of Shanxi Medical University (No. 02201626).

**Data Availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflicts of Interest** The authors declare that they have no conflicts of interest to this work.

## References

- Bayona-Bafaluy MP, Acin-Perez R, Mullikin JC, Park JS, Moreno-Loshuertos R, Hu P, Perez-Martos A, Fernandez-Silva P, Bai Y, Enriquez JA (2003) Revisiting the mouse mitochondrial DNA sequence. *Nucleic Acids Res* 31(18):5349–5355. <https://doi.org/10.1093/nar/gkg739>
- Bendova K, Markova S, Searle JB, Kotlik P (2016) The complete mitochondrial genome of the bank vole *Clethrionomys glareolus* (Rodentia: Arvicolinae). *Mitochondrial DNA A DNA Mapping Sequencing, and Analysis* 27(1):111–112. <https://doi.org/10.3109/19401736.2013.873927>
- Brown SK, Blois JL (2016) The complete mitochondrial genome of the dusky-footed woodrat (*Neotoma fuscipes*) (Rodentia, Cricetidae). *Mitochondrial DNA Part B-Resources* 1(1):728–729. <https://doi.org/10.1080/23802359.2016.1233464>
- Chan JF, Zhang AJ, Yuan S (2020) Simulation of the clinical and pathological manifestations of Coronavirus Disease 2019 (COVID-19) in golden Syrian hamster model: implications for disease pathogenesis and transmissibility. *Clinical Infections Diseases*. <https://doi.org/10.1093/cid/ciaa325>

- Chen S, Chen G, Wei H, Wang Q (2016) Complete mitochondrial genome of the Pere David's Vole, *Eothenomys melanogaster* (Rodentia: Arvicolinae). Mitochondrial DNA A DNA Mapping Sequencing, and Analysis 27(4):2496–2497. <https://doi.org/10.3109/19401736.2015.1036246>
- Ding L, Li W, Liao J (2016a) Mitochondrial genome of *Cricetulus migratorius* (Rodentia: Cricetidae): Insights into the characteristics of the mitochondrial genome and the phylogenetic relationships of *Cricetulus* species. Gene 595:121–129. <https://doi.org/10.1016/j.gene.2016.10.003>
- Ding L, Li W, Liao J (2016b) Characterization of the complete mitochondrial genome of *Phodopus roborovskii* (Rodentia: Cricetidae) and systematic implications for Cricetinae phylogenetics. Biochem Syst Ecol 69:226–236
- Ding L, Zhou Q, Sun Y, Feoktistova NY, Liao J (2020) Two novel cricetine mitogenomes: Insight into the mitogenomic characteristics and phylogeny in Cricetinae (Rodentia: Cricetidae). Genomics 112(2):1716–1725. <https://doi.org/10.1016/j.ygeno.2019.09.016>
- Emily LM, Hankenson FC (2015) Biology and diseases of hamsters. Laboratory Animal Medicine, Third Edition, Academic Press. <https://doi.org/10.1016/B978-0-12-409527-4.00005-5>
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376
- Fonseca MM, Harris DJ (2008) Relationship between mitochondrial gene rearrangements and stability of the origin of light strand replication. Genet Mol Biol 30:566–574
- Hixson JE, Wong TW, Clayton DA (1986) Both the conserved and divergent 5'-flanking sequences are required for initiation at the human mitochondrial origin of light strand replication. J Biol Chem 261:2384–2390 (PMID: 3944140)
- Ho MR, Tsai KW, Lin WC (2012) A unified framework of overlapping genes: towards the origination and endogenous regulation. Genomics 100(4):231–239 (PMID: 22766524)
- Imai M, Iwatsuki-Horimoto K, Hatta M (2020) Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. Proc Natl Acad Sci USA 117(28):16587–16595. <https://doi.org/10.1073/pnas.2009799117>
- Jiang X, Gao J, Ni L, Hu J, Li K, Sun F, Xie J, Bo X, Gao C, Xiao J, Zhou Y (2012) The complete mitochondrial genome of *Microtus fortis calamorum* (Arvicolinae, Rodentia) and its phylogenetic analysis. Gene 498:288–295. <https://doi.org/10.1016/j.gene.2012.02.022>
- Kang C, Yue H, Liu M, Huang T, Liu Y, Zhang X, Yue B, Zeng T, Liu S (2016) The complete mitochondrial genome of *Cricetulus kamensis* (Rodentia: Cricetidae). Mitochondrial DNA A 27(2):976–977. <https://doi.org/10.3109/19401736.2014.926513>
- Kreye J, Reincke SM, Kornau HC (2020) A SARS-CoV-2 neutralizing antibody protects from lung pathology in a COVID-19 hamster model. Preprint Biorxiv. <https://doi.org/10.1101/2020.08.15.252320>
- Lebedev VS, Bannikova AA, Neumann K, Ushakova MV, Ivanova NV, Surov AV (2018) Molecular phylogenetics and taxonomy of dwarf hamsters *Cricetulus Milne-Edwards*, 1867 (Cricetidae, Rodentia): description of a new genus and reinstatement of another. Zootaxa 4387(2):331–349. <https://doi.org/10.11646/zootaxa.4387.2.5>
- Li Q, Yang M, Chen C, Xiong C, Jin X, Pu Z, Huang W (2018) Characterization and phylogenetic analysis of the complete mitochondrial genome of the medicinal fungus *Laetiporus sulphureus*. Sci Rep 8:9104. <https://doi.org/10.1038/s41598-018-27489-9>
- Li JQ, Li L, Fu BQ, Yan HB, Jia WZ (2019) Complete mitochondrial genomes confirm the generic placement of the plateau vole, *Neodon fuscus*. Bioscience Reports 39(8):BSR20182349. <https://doi.org/10.1042/BSR20182349>
- Liu Z, Qin KS, Li BQ, Yu CW, Cai H, Zhang JS, Wang Z (2019) The complete mitochondrial genome of the striped hamster (*Cricetulus barabensis*) China and its phylogenetic analysis. Mitochondrial DNA Part B-Resources 4(2):2593–2595. <https://doi.org/10.1080/23802359.2019.1641440>
- Lowe TM, Chan PP (2016) tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–57. <https://doi.org/10.1093/nar/gkw413>
- Lu S (1999) Modern molecular biology experimental technology (Second Edition) [M]. Medium National Union Medical University Press, Beijing
- Lu T, Zhu M, Yi C, Si C, Yang C, Chen H (2017) Complete mitochondrial genome of the gray red-backed vole (*Myodes rufocanus*) and a complete estimate of the phylogenetic relationships in Cricetidae. Mitochondrial DNA A DNA Mapp Seq Anal 28(1):62–64. <https://doi.org/10.3109/19401736.2015.1110799>

- Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ (1997) Two novel gene orders and the role of light-strand replication in rearrangements of the vertebrate mitochondrial genome. *Mol Biol Evol* 14:91–104. <https://doi.org/10.1093/oxfordjournals.molbev.a025706>
- Neumann K, Michaux J, Lebedev V, Yigit N, Colak E, Ivanova N, Poltoraus A, Surov A, Markov G, Maak S (2006) Molecular phylogeny of the Cricetinae subfamily based on the mitochondrial cytochrome b and 12S rRNA genes and the nuclear vWF gene. *Mol Phylogenet Evol* 39:135–148. <https://doi.org/10.1016/j.ympev.2006.01.010>
- Nylander J (2004) MrModeltest v2 program distributed by the author 2. Uppsala University, Evolutionary Biology Centre
- Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290(5806):470–474. <https://doi.org/10.1038/290470a0>
- Partridge MA, Davidson MM, Hei TK (2007) The complete nucleotide sequence of Chinese hamster (*Cricetus griseus*) mitochondrial DNA. *DNA Seq* 18(5):341–346. <https://doi.org/10.1080/10425170601101287>
- Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol* 41:353–358. <https://doi.org/10.1007/BF00186547>
- Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *System Biology* 53:793–808. <https://doi.org/10.1080/10635150490522304>
- Romanenko SA, Volobouev VT, Perelman PL, Lebedev VS, Serdukova NA, Trifonov VA, Biltueva LS, Nie W, O'Brien PC, Bulatova NS (2007) Karyotype evolution and phylogenetic relationships of hamsters (Cricetidae, Muroidea, Rodentia) inferred from chromosomal painting and banding comparison. *Chromosome Res* 15:283–298. <https://doi.org/10.1007/s10577-007-1124-3>
- Romanenko S, Perelman P, Trifonov V, Graphodatsky A (2012) Chromosomal evolution in Rodentia. *Heredity* 108:4–16. <https://doi.org/10.1038/hdy.2011.110>
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542. <https://doi.org/10.1093/sysbio/sys029>
- Rosenke K, Jarvis MA, Feldmann F (2020) Hydroxychloroquine proves ineffective in hamsters and macaques infected with SARS-CoV-2. Preprint Biorxiv. <https://doi.org/10.1101/2020.06.10.145144>
- Rozas J (2017) DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Mol Biol Evol* 34(12):3299–3302. <https://doi.org/10.1093/molbev/msx248>
- Saccone C, Pesole G, Sbisà E (1991) The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *J Mol Evol* 33:83–91. <https://doi.org/10.1007/BF02100199>
- Sbisà E, Tanzariello F, Reyes A, Pesole G, Saccone C (1997) Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* 205:125–140. [https://doi.org/10.1016/s0378-1119\(97\)00404-6](https://doi.org/10.1016/s0378-1119(97)00404-6)
- Schofield M, Duchamp J, Larkin JL, Smyser TJ, Doyle JM (2018) Mitochondrial genome of an Allegheny Woodrat (*Neotoma magister*). *Mitochondrial DNA B Resource* 3(1):256–258
- Sullivan KAM, Platt RN, Bradley RD, Ray DA (2017) Whole mitochondrial genomes provide increased resolution and indicate paraphyly in deer mice. *BMC Zoology* 2:11. <https://doi.org/10.1186/s40850-017-0020-3>
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Tchitchek N, Safronetz D, Rasmussen AL, Martens C, Virtaneva K, Porcella SF, Feldmann H, Ebihara H, Katze MG (2014) Sequencing, annotation and analysis of the Syrian hamster (*Mesocricetus auratus*) transcriptome. *PLoS One* 9(11):e112617. <https://doi.org/10.1371/journal.pone.0112617>
- Triant DA, Dewoody JA (2006) Accelerated molecular evolution in *Microtus* (Rodentia) as assessed via complete mitochondrial genome sequences. *Genetica* 128(1–3):95–108. <https://doi.org/10.1007/s10709-005-5538-6>
- Wei SJ, Shi M, Chen XX, Sharkey MJ, Van CA, Ye GY, He JH (2010) New views on strand asymmetry in insect mitochondrial genomes. *PLoS ONE* 5(9):e12708. <https://doi.org/10.1371/journal.pone.0012708>
- Yang Z, Rannala B (1997) Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. *Mol Biol Evol* 14:717–724. <https://doi.org/10.1093/sysbio/syx087>

- Yang C, Hao H, Liu S, Liu Y, Yue B, Zhang X (2012) Complete mitochondrial genome of the Chinese oriental vole *Eothenomys chinensis* (Rodentia: Arvicolinae). *Mitochondrial DNA* 23(2):131–133. <https://doi.org/10.3109/19401736.2012.660926>
- Yang T, Xu G, Gu B, Shi Y, Mzuka HL, Shen H (2019) The complete mitochondrial genome sequences of the *Philomycus bilineatus* (Stylommatophora: Philomycidae) and phylogenetic analysis. *Genes* 10:198–210. <https://doi.org/10.3390/genes10030198>
- Yiğit N, Kankiliç T, Çolak R et al (2007) Allozyme variations and genetic differentiation in *Mesocricetus brandti* Nehring, 1898 and *Mesocricetus auratus* (Waterhouse, 1839) (Mammalia: Rodentia). *Turkish J Zool* 31(3):219–227
- Yu P, Kong LM, Li YC, Cong H, Li Y (2016) Analysis of complete mitochondrial genome and its application to phylogeny of *Caryomys inez* (Rodentia: Cricetidae: Arvicolinae). *Mitochondrial DNA Part B Resource* 1(1):343–344. <https://doi.org/10.1080/23802359.2016.1172275>
- Zardoya R, Garrido-Pertierra A, Bautista JM (1995) The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*. *J Mol Evol* 41:942–951. <https://doi.org/10.1007/BF00173174>
- Zhang Q, He T, Wei H, Li F, Feng Y, Zong H, Chen S (2016a) Characterization of the complete mitochondrial genome and phylogenetic relationship of *Neodon sikimensis* (Rodentia: Arvicolinae). *Mitochondrial DNA B Resource* 1(1):445–446
- Zhang Z, Sun T, Kang C, Liu Y, Liu S, Yue B, Zeng T (2016b) The complete mitochondrial genome of lesser long-tailed Hamster *Cricetulus longicaudatus* (Milne-Edwards, 1867) and phylogenetic implications. *Mitochondrial DNA A DNA Mapping, Sequencing and Analysis* 27(2):1303–1304. <https://doi.org/10.3109/19401736.2014.945567>
- Zhao H, Qi X, Li C (2018) Complete mitochondrial genome of the muskrat (*Ondatra zibethicus*) and its unique phylogenetic position estimated in Cricetidae. *Mitochondrial DNA Part B-Resources* 3(1):296–298. <https://doi.org/10.1080/23802359.2017.1390402>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Ruijing Xuan<sup>1</sup> · Jiping Gao<sup>1</sup> · Qiang Lin<sup>2</sup> · Wenbin Yue<sup>3</sup> · Tianfu Liu<sup>1</sup> · Songnian Hu<sup>2</sup> · Guohua Song<sup>1</sup> 

✉ Guohua Song  
ykdsgh@sxmu.edu.cn

<sup>1</sup> Laboratory Animal Center, Shanxi Medical University, Taiyuan 030001, China

<sup>2</sup> Key Laboratory of Genome Information and Sciences, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China

<sup>3</sup> College of Animal Science and Technology, Shanxi Agricultural University, Taigu 030801, China