Online ISSN: 2588-3526
Journal of Wildlife and Biodiversity

Volume 8 (1): 203-212 (2024) (http://www.wildlife-biodiversity.com/)

Research Article

Characterization of a translocated Mitochondrial Cytochrome *b* pseudogene in *Meriones persicus* (Rodentia; Gerbillinae); a potential taxonomic pitfall

Ahmad Mahmoudi[1*], Ehsan Mostafavi[2], Boris Kryštufek[3]

[1]Department of Biology, Faculty of Science, Urmia University, Iran
[2]Department of Epidemiology and Biostatistics, Research Centre for Emerging and Reemerging Infectious Diseases, Pasteur Institute of Iran, Tehran, Iran
[3]Slovenian Museum of Natural History, Ljubljana, Slovenia
[*]Email: a.mahmoudi@urmia.ac.ir

Received: 18 March 2023// Revised: 28 June 2023 / Accepted: 15 July 2023/ Published online: 16 July. Ministry of Sciences, Research, and Technology, Arak University, Iran.

How to cite: Mahmoudi, A., Mostafavi, E., Kryštufek, B. (2024). Characterization of a translocated Mitochondrial Cytochrome b pseudogene in *Meriones persicus* (Rodentia; Gerbillinae); a potential taxonomic pitfall, Journal of Wildlife and Biodiversity, 8(1), 203-212. **DOI**: https://doi.org/10.5281/zenodo.8151941

Abstract

Due to its faster evolution rate compared to nuclear genes, haploid mitochondrial DNA (mtDNA) is a promising species identification tool. This has led to its significant use in taxonomic and phylogenetic studies. While mtDNA is subject to selective constraints that prevent the accumulation of deleterious mutations, the prevalence of nuclear-mitochondrial fragments (known as *NUMTs* or pseudogenes) in mammals has complicated the use of mtDNA for taxonomy. In the present study, a pseudogene of the mitochondrial cytochrome *b* (*Cytb*) was detected in *Meriones persicus*. This pseudogene differed from its mitochondrial counterpart at 235 out of 1140 sites and is characterized by frame-shift mutations, indels, and accumulation of non-synonymous substitutions. It is the first report of the *Cytb* pseudogene in a jird, highlighting the risk of misidentifying *NUMT*s as authentic mtDNA and the importance of addressing this potential pitfall in taxonomic studies.

Keywords: Species identification, Mitochondrial Cytochrome *b*, Jirds, Pseudogenes

Introduction

Rodents are known as the key components of every ecosystem as they impact the structure and processes of their ecosystems (Wilson et al., 2017). Rodents also influence in many ways human

well-being and lives in providing food, serving as pets, being agricultural pests, and the source of epidemiologic problems (Buckle & Smith, 1994; Plourde et al., 2017; Rabiee et al., 2018; Mahmoudi et al., 2022). As a result, a diverse range of disciplines and researchers regularly work with or on rodents, either directly or indirectly. Jirds of the genus *Meriones* are common inhabitants of the arid and semi-arid ecosystems including the Iranian plateau (Denys et al., 2017), posing a significant public health burden as they serve as reservoirs for numerous infectious diseases, including Plague, Leishmaniosis, Hepatic capillariasis, Trichinellosis, and Hymenolepsis (Rabiee et al., 2018). Although the species can be distinguished from the congeners through morphological characteristics (Darvish, 2011; Kryštufek & Vohralík, 2009), this approach is time-consuming and requires broad taxonomic expertise. Therefore, the analysis of a mitochondrial DNA sequence or DNA Barcoding has become a routine in taxonomy and related fields. The mtDNA evolves nearly four to ten times faster than nuclear genes in animals (Brown et al., 1979; Lynch, 1997; DeSalle et al., 2017), and is subject to different selective pressures (Shtolz & Mishmar, 2019). However, genomic analysis has revealed that most eukaryotes contain mtDNA fragments in their nucleus, referred to as *NUMTs* or nuclear-mitochondrial pseudogenes which particularly are prevalent in mammals (DeWoody et al., 1999; Zhang et al., 2004; Jaarola & Searle, 2004; Triant & DeWoody, 2007, 2008; Wei et al., 2022). Unlike mtDNA, nuclear copies are prone to the accumulation of deleterious mutations due to the absence of selective pressures (Li et al., 1980). Consequently, there is a high likelihood of finding i) premature stop codons or frame-shift mutations, especially when they are long enough, ii) an accumulation of non-synonymous mutations, and iii) a high rate of base call admixture (Jaarola & Searle, 2004; Triant & DeWoody, 2008; Dubey et al., 2009). Despite this fact, such signals could be diluted in various ways, so *NUMT*s can still introduce significant biases in species identification, systematics, and diversity estimation (Zhang & Hewitt, 1996, 2003; Triant & DeWoody, 2007; Dubey et al., 2009).

Material and methods

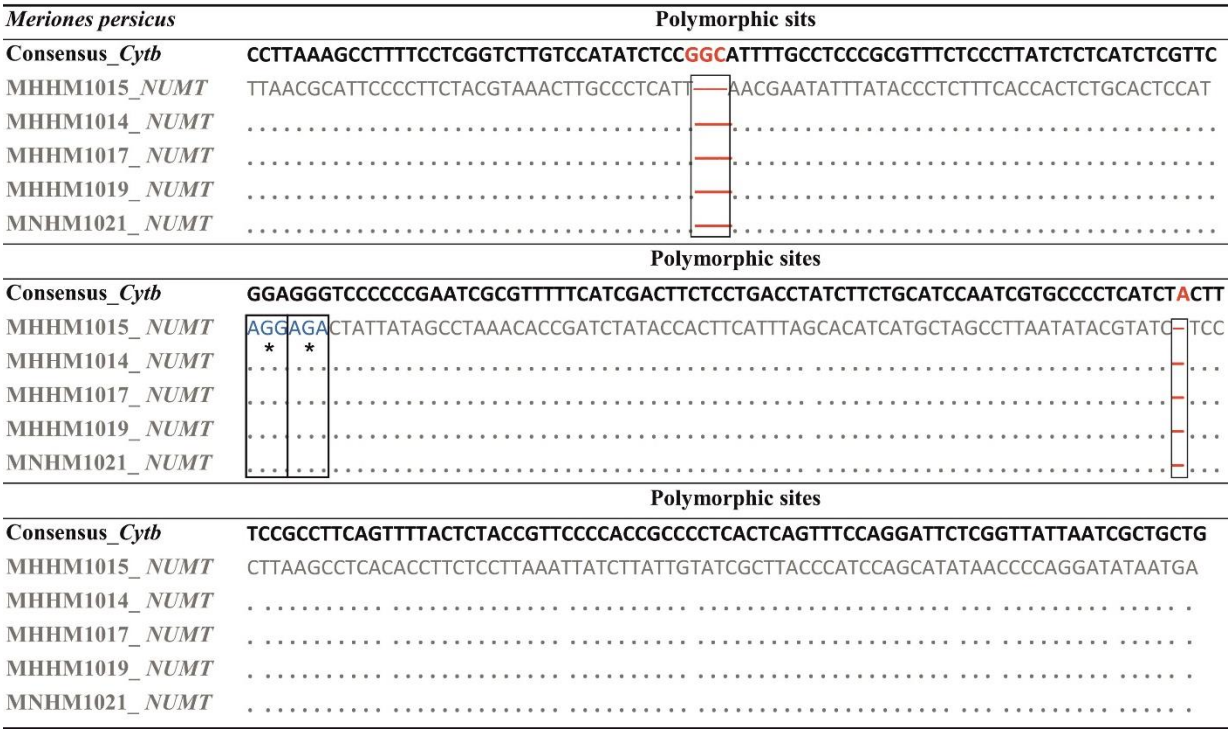
In our study on *Meriones* species in Iran, we utilized L7 and H6 (Montgelard et al., 2002), universal primers to amplify the entire cytochrome *b* gene (1140-bp). During electrophoresis, we did not observe any indication of *NUMT* contamination as there was only a single band in the 1100-1200 bp range. PCR products were purified using QIAquick PCR purification Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions and sequenced using dye-labeled dideoxy terminator cycle sequencing with Big Dye V.3.1 (Applied

Biosystems, Inc). Raw sequences were checked for ambiguous sites using CodonCode Aligner software (CodonCode Corp.). Nucleotide translation, premature stop codon, and the presence of indels were checked using MEGA v7.0 (Kumar et al., 2016). The phylogenetic tree was inferred by the Maximum likelihood (ML) approach in IQTREE (Nguyen et al., 2015) using 1000 bootstrap replications to obtain the nodal supports.

Results

In the final alignment (*n*=240 sequences with 1140-bp length), five samples from NW Iran showed a huge mismatch with all other *Cytb* sequences. Although morphologically identified as *Meriones persicus*, these specimens differed from other conspecific *Cytb* sequences at 235 sites (20.61%) out of 1140-bp. Two premature stop codons were also detected at positions 406-408 (AGG) and positions 424-426 (AGA), indicating that these sequences are pseudogenes. Moreover, a 3-bp deletion at 184-186 positions and a 1-bp deletion at position 747 were observed. (Table 1). The examination of the base compositions indicated that pseudogenes exhibit a greater amount of non-synonymous sites (NSS) accumulation compared to the authentic *Cytb* sequences in 136 *M. persicus* (NSS=76.64% versus NSS=58.39%).

Table 1. Polymorphic positions (235-bp) were observed between the consensus *Cytb* sequence of 136 samples and *NUMT* in *Meriones persicus*. A 3-bp and 1-bp deletions (highlighted in red) as well as two premature stop-codons were characterized in pseudogenes (highlighted in blue). see text for detailed information.


Meriones persicus
Polymorphic sits
Consensus_Cytb CCTTAAAGCCTTTTCCTCGGTCTTGTCCATATCTCCGGCATTTTGCCTCCCGCGTTTCTCCCTTATCTCTCATCTCGTTC
MHHM1015_NUMT TTAACGCATTCCCCTTCTACGTAAACTTGCCCTCATT-AACGAATATTTATACCCTCTTTCACCACTCTGCACTCCAT
MHHM1014_ NUMT
MHHM1017_ NUMT
MHHM1019_ NUMT
MNHM1021_ NUMT
Polymorphic sites
Consensus_Cytb GGAGGGTCCCCCCGAATCGCGTTTTTCATCGACTTCTCCTGACCTATCTTCTGCATCCAATCGTGCCCCTCATCTACTT
MHHM1015_ NUMT AGGAGACTATTATAGCCTAAACACCGATCTATACCACTTCATTTAGCACATCATGCTAGCCTTAATATACGTATC-TCC
MHHM1014_ NUMT
MHHM1017_ NUMT
MHHM1019_ NUMT
MNHM1021_ NUMT
Polymorphic sites
Consensus_Cytb TCCGCCTTCAGTTTTACTCTACCGTTCCCCACCGCCCCTCACTCAGTTTCCAGGATTCTCGGTTATTAATCGCTGCTG
MHHM1015_ NUMT CTTAAGCCTCACACCTTCTCCTTAAATTATCTTATTGTATCGCTTACCCATCCAGCATATAACCCCAGGATATAATGA
MHHM1014_ NUMT
MHHM1017_ NUMT
MHHM1019_ NUMT
MNHM1021_ NUMT

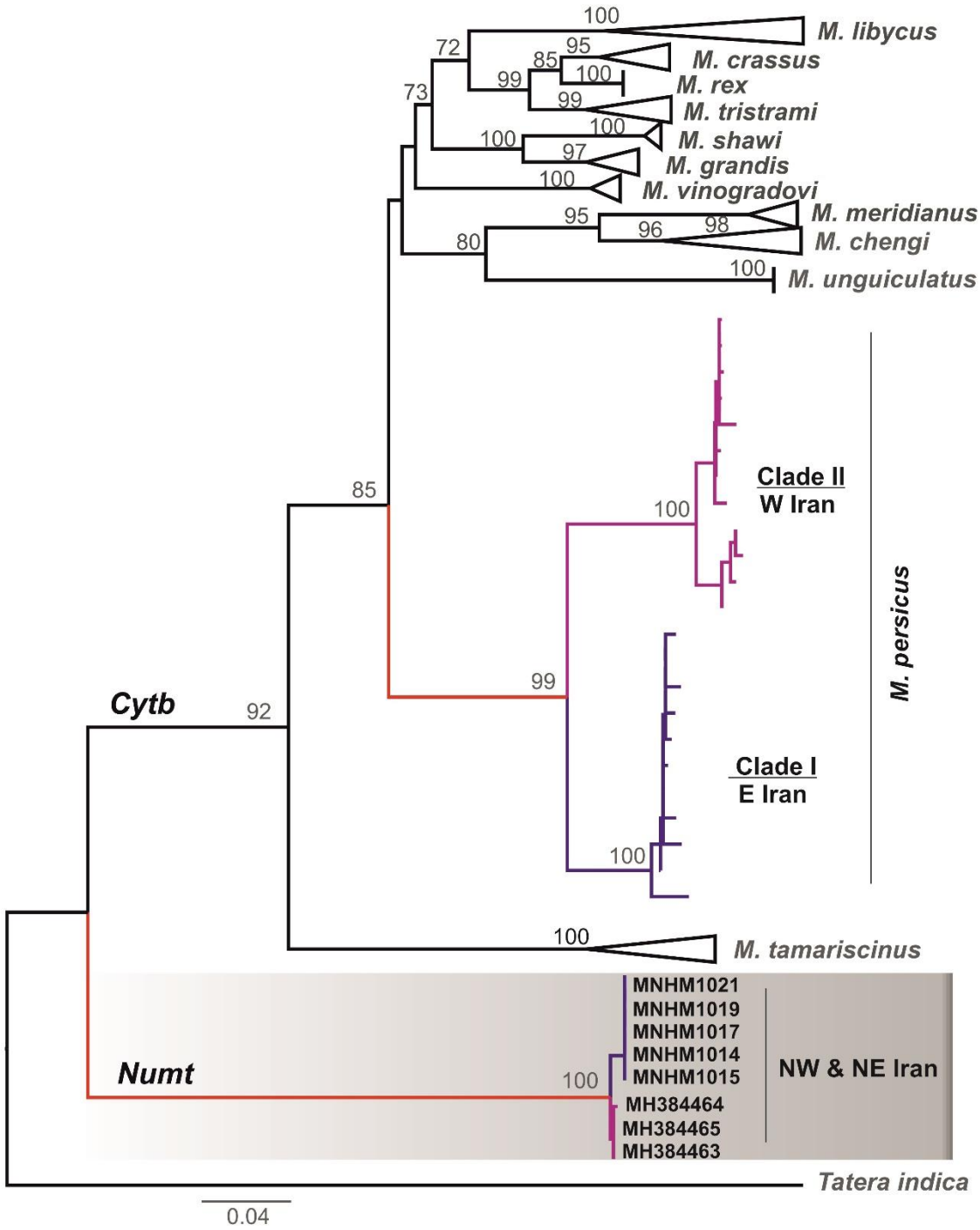


Figure 1. Maximum Likelihood tree based on *Cytb* sequences showing the relationship among the *Meriones* species. The remote position of the *NUMT* sequences is shown by a grey shadow.

We also observed that three short *Cytb* sequences (533-bp) (MH384463-5) retrieved from the GenBank database and labeled as *Meriones* sp. (Hosseini et al., unpublished) from NE Iran are distinct from the others with 101 unique mutations. Although no stop-codons or frame-shift

mutations were found in these sequences, a considerable accumulation of NSS (74.94%) is an indicator to its origin. Further analysis revealed that these cryptic *Cytb* pseudogenes closely aligned with the pseudogenes discovered in NW Iran as they formed a single highly divergent clade in the ML tree (Fig. 1), with significant genetic divergence from the *Meriones* species ranging from 24-31%.

Discussion

Based on *Cytb* data, *M. persicus* consists of two primary clades (I and II) that are mainly allopatric in western and eastern Iran with >9.3% K2P divergence (Dianat et al., 2017; see also Fig. 1). In terms of geographic distribution, the pseudogenes found in NW and NE Iran are belonging to clades I and II, respectively. However, their proximity suggests that mitochondrial translocation occurred once in their ancestor prior to the formation of clades I and II (~1.5 Mya cf. Dianat et al., 2017). However, there were three positions in which pseudogenes from NE were identical with authentic *Cytb* sequences but differed from pseudogenes from NW Iran. Moreover, the pseudogenes from NE Iran lack a 1-bp deletion at position 747 that was observed in pseudogenes from NW Iran (Fig. 2a). These variations can be explained under base-call admixture of the authentic *Cytb* and its nuclear copy. Finally, the sequence MH384464 from NE Iran has cytosine (C) at position 875, which is thymine (T) in all authentic *Cytb* sequences of the genus *Meriones* (n=235) and also pseudogenes from NW Iran (Fig. 2b). This variation could probably be originated from a bad quality sequence processing. The premature stop codons in our sequences from NW Iran resulted from the amino acid codons AGG and AGA which both encoded ‘Arginine’ in the universal genetic code, but function as stop codons in the vertebrate mitochondrial genome (Osawa et al., 1989). As a result, pseudogenes can produce false phylogenies (Dubey et al., 2009), particularly when the signals are masked by small sample size, the similarity in size of the authentic mtDNA and its nuclear copy, insufficient sequence length to detect frame-shift or stop codon mutations, and the formation of chimeric sequences due to the simultaneous amplification of both authentic and *NUMT* fragments (Fig. 3), as is the case with the pseudogenes from NE Iran.

The present study highlights the importance of careful utilization of mtDNA sequences, especially incomplete and partial sequences which might result in critical errors in taxonomy. Specifically, misidentification of cryptic pseudogenes might drive one to classify the divergent clade as either an unknown *Meriones* or, more likely, a new gerbil, a condition that once creates a false phylogeny among field mice (genus *Apodemus* cf. Dubey et al., 2009).

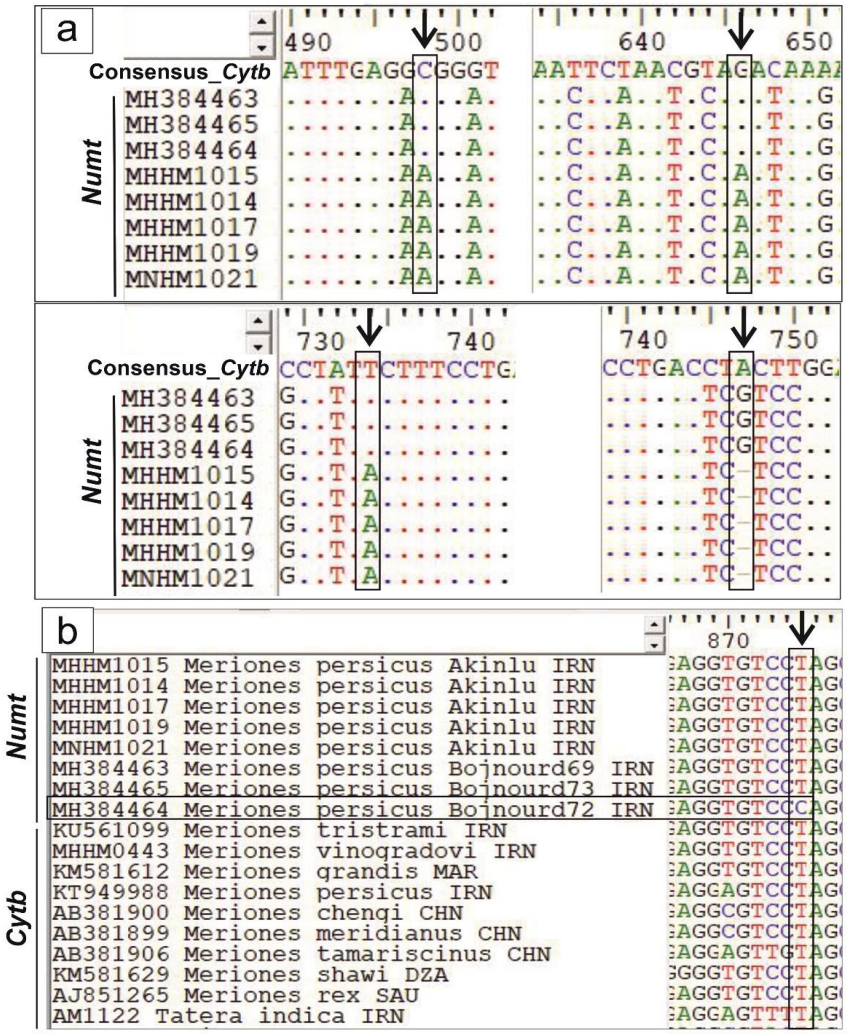


Figure 2. Variations among *NUMT* sequences of *Meriones persicus* from NW and NE Iran (a), and an example of artificially generated polymorphism in sample MH384464 in position 875, the position which is constant not only among 235 *Cytb* sequences belonging to seven *Meriones* species, but also among other *NUMT*s in *M. persicus* (b).

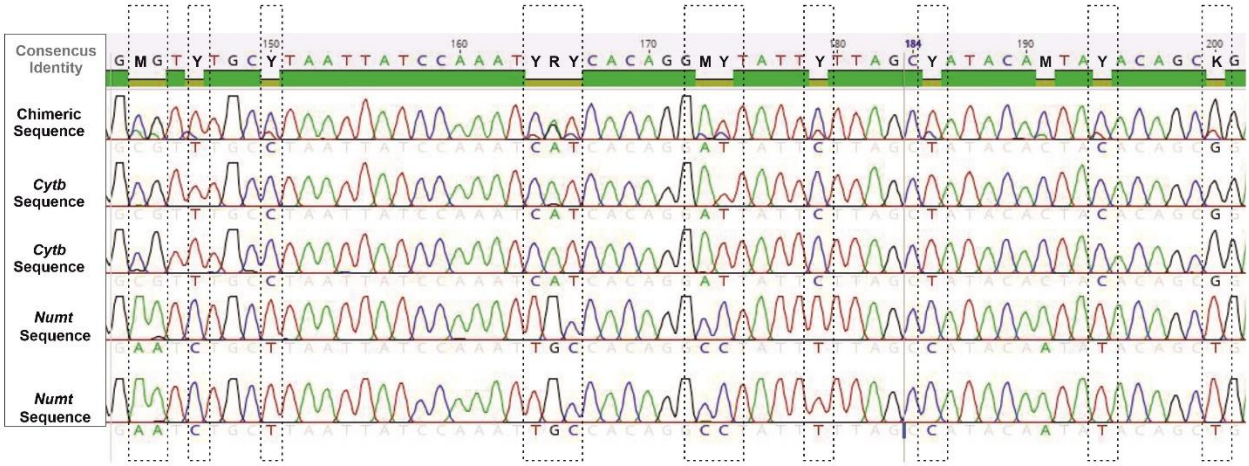


Figure 3. Chromatograms of one chimeric, two authentic *Cytb*, and two *NUMT* sequences in *Meriones persicus*.

Conclusion

This is the first report of a full-length *Cytb* pseudogene in gerbils (Gerbillinae; Rodentia) categorized as a "duplicated pseudogene", since it has not been subject to any selection and has accumulated frame-shift mutations and indels (Dainat et al., 2021). This study emphasizes the drawback of relying on incomplete mtDNA sequences for species identification, particularly in fields dealing with hosts and reservoirs of zoonotic infectious diseases. Consequently, this discovery highlights the need for combining morphology with genetic identification and extensive sampling to identify potential cryptic pseudogenes in gerbils, a significant rodent group in the arid and semi-arid ecosystems with poorly resolved phylogeny (Chevret & Dobigny, 2005; Alhajeri et al., 2015; Ding et al., 2022), and considerable public health burden.

Acknowledgements

This study was funded by a grant from Urmia University, Urmia, Iran (Project numbers: 001/4/00) to AM.

References

Brown, W. M., George, Jr-M., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences*, 76(4), 1967–1971. https://doi.org/10.1073/pnas.76.4.1967.

Buckle, A.P & Smith, R.H. (1994). *Rodent Pests and Their Control*. Cambridge University Press, Cambridge, United Kingdom.

Chevret, P., & Dobigny, G. (2005). Systematics and evolution of the subfamily Gerbillinae (Mammalia, Rodentia, Muridae). *Molecular phylogenetics and evolution*, *35*(3), 674–688. https://doi.org/10.1016/j.ympev.2005.01.001.

Dainat, J., & Pontarotti, P. (2021). Methods to identify and study the evolution of pseudogenes using a phylogenetic approach. In *Pseudogenes: Functions and Protocols* (pp. 21-34). New York, NY: Springer US.

Darvish, J. (2011). Morphological comparison of fourteen species of the genus *Meriones* Illiger, 1811 (Rodentia: Gerbillinae) from Asia and North Africa. *Iranian Journal of Animal Biosystematics*, 7(1).

Denys, C., Taylor, P., Burgin, C., Aplin, K., Fabre, P., Haslauer, R., et al. (2017). Species Account of Muridae. In: Wilson, D.E., Lacher, Jr, T. E., Mittermeier, R.A. (Eds) *Handbook of Mammals of the World*, Vol. 7, Rodents II, 599-884. Lynx Edicions. Barcelona. 1008 pp.

DeSalle, R., Schierwater, B., & Hadrys, H. (2017). MtDNA: The small workhorse of evolutionary studies. *Frontiers in Bioscience-Landmark*, 22(5), 873–887.

DeWoody, J. A., Chesser, R. K., & Baker, R. J. (1999). A translocated mitochondrial cytochrome b pseudogene in voles (Rodentia: *Microtus*). *Journal of Molecular Evolution*, 48, 380–382. https://doi.org/10.1007/PL00013154.

Dianat, M., Darvish, J., Cornette, R., Aliabadian, M., & Nicolas, V. (2017). Evolutionary history of the Persian Jird, *Meriones persicus*, based on genetics, species distribution modelling and morphometric data. *Journal of Zoological Systematics and Evolutionary Research*, 55(1), 29–45. https://doi.org/10.1111/jzs.12145.

Ding, L., Luo, G., Zhou, Q., Sun, Y., & Liao, J. (2022). Comparative mitogenome analysis of gerbils and the mitogenome phylogeny of Gerbillinae (Rodentia: Muridae). *Biochemical Genetics*, 60(6), 2226–2249. https://doi.org/10.1007/s10528-022-10213-8.

Dubey, S., Michaux, J., Brünner, H., Hutterer, R., & Vogel, P. (2009). False phylogenies on wood mice due to cryptic cytochrome-b pseudogene. *Molecular Phylogenetics and Evolution*, 50(3), 633–641. https://doi.org/10.1016/j.ympev.2008.12.008.

Jaarola, M., & Searle, J.B., (2004). A highly divergent mitochondrial DNA lineage of *Microtus agrestis* in southern Europe. *Heredity*, 92, 228–234. https://doi.org/10.1038/sj.hdy.6800400.

Kryštufek, B., & Vohralík, V. (2009). *Mammals of Turkey and Cyprus. Rodentia II: Cricetinae, Muridae, Spalacidae, Calomyscidae, Capromyidae, Hystricidae, Castoridae*. University of Primorska, Science and research centre Koper, Koper.

Kumar, S., Stecher, G, & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Li W-H, Gojobori T, & Nei, M. (1981) Pseudogenes as a paradigm of neutral evolution. *Nature*, 292, 237–239. https://doi.org/10.1038/292237a0.

Lynch, M. (1997). Mutation accumulation in nuclear, organelle, and prokaryotic transfer RNA genes. *Molecular Biology and Evolution*, 14(9), 914–925. https://doi.org/10.1093/oxfordjournals.molbev.a025834.

Mahmoudi, A., Kryštufek, B., Sludsky, A., Schmid, B. V., De Almeida, A. M., Lei, X., ... & Mostafavi, E. (2021). Plague reservoir species throughout the world. *Integrative Zoology*, 16(6), 820–833. https://doi.org/10.1111/1749-4877.12511.

Montgelard, C., Bentz, S., Tirard, C., Verneau, O., & Catzeflis, F. M. (2002). Molecular systematics of Sciurognathi (Rodentia): the mitochondrial cytochrome b and 12S rRNA genes support the Anomaluroidea (Pedetidae and Anomaluridae). *Molecular Phylogenetics and Evolution*, 22(2), 220–233.

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., & Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274. https://doi.org/10.1093/molbev/msu300.

Osawa, S., Ohama, T., Jukes, T. H., & Watanabe, K. (1989). Evolution of the mitochondrial genetic code I. Origin of AGR serine and stop codons in metazoan mitochondria. *Journal of Molecular Evolution*, 29, 202–207. https://doi.org/10.1007/BF02100203.

Plourde, B. T., Burgess, T. L., Eskew, E. A., Roth, T. M., Stephenson, N., & Foley, J. E. (2017). Are disease reservoirs special? Taxonomic and life history characteristics. *PloS one*, 12(7), e0180716. https://doi.org/10.1371/journal.pone.0180716.

Rabiee, M. H., Mahmoudi, A., Siahsarvie, R., Kryštufek, B., & Mostafavi, E. (2018). Rodent-borne diseases and their public health importance in Iran. *PLoS neglected tropical diseases*, 12(4), e0006256. https://doi.org/10.1371/journal.pntd.0006256.

Shtolz, N., & Mishmar, D. (2019). The mitochondrial genome–on selective constraints and signatures at the organism, cell, and single mitochondrion levels. *Frontiers in Ecology and Evolution*, 7, 342. https://doi.org/10.3389/fevo.2019.00342.

Triant, D. A., & DeWoody, J. A. (2007). The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. *Journal of Mammalogy*, 88(4), 908–920. https://doi.org/10.1644/06-MAMM-A-204R1.1.

Triant, D. A., & DeWoody, J. A. (2008). Molecular analyses of mitochondrial pseudogenes within the nuclear genome of arvicoline rodents. *Genetica*, 132, 21–33. https://doi.org/10.1007/s10709-007-9145-6.

Wei, W., Schon, K. R., Elgar, G., Orioli, A., Tanguy, M., Giess, A., ... & Chinnery, P. F. (2022). Nuclear-embedded mitochondrial DNA sequences in 66,083 human genomes. *Nature*, 611(7934), 105–114. https://doi.org/10.1038/s41586-022-05288-7.

Wilson, D. E., Lacher, T.E., & Mittermeier, R.A. (Eds.). (2017). *Handbook of the Mammals of the World: Rodents II*. Lynx Edicions, Barcelona.

Zhang, D. X., & Hewitt, G. M. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, 12(3), 563–584. https://doi.org/10.1046/j.1365-294X.2003.01773.x.

Zhang, D.X., & Hewitt, G.M. (1996). Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology and Evolution*, 11, 247–251. https://doi.org/10.1016/0169-5347(96)10031-8.

Zhang, Z., Carriero, N., & Gerstein, M. (2004). Comparative analysis of processed pseudogenes in the mouse and human genomes. Trends Genet 20, 62–67. https://doi.org/10.1016/j.tig.2003.12.005.