

Mitogenome characterization and phylogenetic relationships of Turkish horses (*Equus caballus*)

© Saffet Teber^{1,2*}

¹Erciyes University Faculty of Agriculture, Department of Agricultural Biotechnology, Kayseri, Türkiye

²Erciyes University, Genome and Stem Cell Center (GenKök), Kayseri, Türkiye

Cite this article as: Teber, S. (2025). Mitogenome characterization and phylogenetic relationships of Turkish horses (*Equus caballus*) *Trakya University Journal of Natural Sciences*, 27(1), xx–xx.

Abstract

Background: Horses (*Equus caballus*) have been selectively bred for numerous purposes since their domestication, leading to various breeds and increased genetic diversity within the species.

Aims: This study focused on the mitogenome characterization and phylogenetic relationships of Turkish feral (yıkı) and domestic horse breeds.

Methods: In this study, the whole mitogenomes of Turkish feral (yıkı) and domestic horse breeds were first amplified using long-range polymerase chain reaction, sequenced with ~4000× coverage on the Illumina MiSeq platform, and their phylogenetic relationships were subsequently analyzed.

Results: The mitogenomes of Turkish horses were 16657 base pairs in length, encompassing 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes, two ribosomal RNA (rRNA) genes, one origin of replication (O_L), and one non-coding control region (displacement-loop). In the mitogenomes, the *ND6* gene and eight tRNAs (tRNA^{Gln}, tRNA^{Ala}, tRNA^{Asn}, tRNA^{Cys}, tRNA^{Tyr}, tRNA^{Ser(UCN)}, tRNA^{Glu}, and tRNA^{Pro}) were encoded on the light strand (L), while the origin of replication (O_L), 12 PCGs, 14 tRNAs, and two rRNA genes were encoded on the heavy strand (H). A total of 45 gaps and 68 overlaps were identified in the intergenic regions of the assembled mitogenomes. In the Maximum Likelihood phylogenetic tree, the feral (yıkı) horse clustered within haplogroup A together with breeds from North America, the Middle East, Southern Europe, and Central Asia, while the domestic horse was clustered in haplogroup B together with breeds from Southern Europe and Central Europe.

Özet

Dayanak: Atlar (*Equus caballus*), evcilleştirmenin bir sonucu olarak birçok farklı görev için yetiştirilmiş ve bununla birlikte çeşitli ırklar elde edilerek tür içi genetik çeşitlilik artmıştır.

Amaçlar: Bu çalışma, Türk yıkı ve evcil at ırklarının mitogenom karakterizasyonu ve filogenetik ilişkilerine odaklanmıştır.

Yöntemler: Bu çalışmada Türkiye yıkı ve yerli at ırklarının tüm mitogenomu ilk kez uzun parçalı polimeraz zincir reaksiyonu ile çoğaltılmış, Illumina MiSeq platformu ile ~4000× kapsamla dizilenecek şekilde karakterize edilmiş ve filogenetik ilişkileri incelenmiştir.

Bulgular: Türk atlarına ait mitogenomlar 16.657 baz çifti uzunluğunda olup 13 protein kodlayan gen (PCG), 22 transfer RNA (tRNA) geni, iki ribozomal RNA (rRNA) geni, bir replikasyon başlangıç bölgesi (O_L) ve bir kodlamayan kontrol bölgesi (displacement-loop) içermektedir. Mitogenomlarda *ND6* geni ile sekiz tRNA (tRNA^{Gln}, tRNA^{Ala}, tRNA^{Asn}, tRNA^{Cys}, tRNA^{Tyr}, tRNA^{Ser(UCN)}, tRNA^{Glu} ve tRNA^{Pro}) hafif iplik (L) üzerinde kodlanırken; replikasyon başlangıç bölgesi (O_L), 12 PCG, 14 tRNA ve iki rRNA geni ağır iplik (H) üzerinde kodlanmıştır. Dizilenmiş mitogenomların intergenik bölgelerinde toplam 45 boşluk (gap) ve 68 bindirme (overlap) belirlenmiştir. Maksimum Olabilirlik filogenetik ağacında yıkı atı; Kuzey Amerika, Orta Doğu, Güney Avrupa ve Orta Asya'dan ırklarla birlikte A haplogrubunda, evcil at ise Güney Avrupa ve Orta Avrupa'dan ırklarla birlikte B haplogrubunda kümelenmiştir.

Sonuç: Bu çalışma, Türk atları için ilk kapsamlı mitogenom verilerini sunarak evrimsel biyoloji, genetik çeşitlilik ve koruma

Edited by: Coşkun Tez

*Corresponding Author: Saffet Teber, E-mail: saffetteber@gmail.com

ORCID iDs of the author(s): ST. 0000-0002-9987-3287



Received: 07 August 2025, Accepted: 03 November 2025, Epub: 26 November 2025



Copyright© 2025 The Author(s). Published by Galenos Publishing House on behalf of Trakya University. Licensed under a Creative Commons Attribution (CC BY) 4.0 International License.



Conclusion: This study significantly contributes to evolutionary biology, genetic diversity, and conservation by providing the first comprehensive mitogenome data for Turkish horses, establishing a foundational resource for future comparative and evolutionary genomic research.

çalışmalarına önemli katkı sağlamaktadır. Ayrıca gelecekte yapılacak karşılaştırmalı ve evrimsel genomik araştırmalar için temel bir kaynak oluşturmaktadır.

Keywords: *Equus caballus*, horse, mitogenome, next-generation sequencing, Türkiye

Introduction

Horses, belonging to the genus *Equus* (Mammalia: Perissodactyla: Equidae), were domesticated later than many other mammals. Although not definitive, archaeological and genetic evidence suggests that domestication occurred on the Eurasian steppes approximately 5,000–6,000 years ago (Lippold et al., 2011; Outram et al., 2009). Throughout history, horses have been primarily utilized for agriculture, transportation, sports, and military purposes, thereby enhancing mobility and trade, influencing warfare, and profoundly transforming human civilization (Levine, 2005; Outram et al., 2009). Selective breeding for traits such as strength, size, appearance, temperament, and endurance has driven intraspecific genetic differentiation and increased phenotypic variation among modern breeds (Lippold et al., 2011).

Due to its unique geographical position, Anatolia has served as a critical center for species domestication, animal migration, and transportation routes (MacHugh & Bradley, 2001; Naderi et al., 2008; Rezaei et al., 2010; Zeder, 2008). In Türkiye, the principal horse breeds include Arabian, Anatolian Native, Ayvacık Pony, Canik, Çamardı Kula, Cirit, Çukurova, Eastern Anatolian, Hınıs, Karacabey, Karakacan, Malakan, Thrace (Rumelia), Rahvan, Turkish-Arab, Turkish Nonius, and Uzunyayla (Bayram, 2005; Ozbeyaz & Akcapinar, 2010). In addition to these, there are also horses known as “*yılkı*,” which have been released into the wild, with their use declining due to increased mechanization in Anatolia (Hacan et al., 2018). Despite the rich diversity of native horse populations, the genetic variability of Anatolian horses remains insufficiently explored, and comprehensive studies are needed to uncover their evolutionary and historical contributions to global horse diversity.

Although numerous studies have analyzed complete mitochondrial DNA (mitogenomes) to investigate genetic diversity and population history in horses (Achilli et al., 2012; Ahlawat et al., 2025; Lippold et al., 2011; Sheikh et al., 2019; 2023), complete mitogenome data for horses in Türkiye are still lacking. Previous research on Turkish horses has mainly focused on morphological traits (Celik et al., 2015; Yilmaz, 2013), partial mtDNA regions (particularly the D-loop) (Koban et al., 2012; Köseman et al., 2019), and SSR-based analyses (Aksu et al., 2010; Koban et al., 2012; Koseman et al., 2020). Within the scope of the TURKHAYGEN-1 project, designated to genotype Anatolian domestic animals in Türkiye, 21 SSR markers and mtDNA D-loop sequences were analyzed across four horse breeds. This study revealed high allelic diversity and significant sequence variation (Koban et al., 2012). Similarly, a partial mtDNA (D-loop) study of 28 Turkish horses identified 42

polymorphic sites and 10 haplogroups (Köseman et al., 2019). The presence of repeat sequences in the mitochondrial D-loop region can affect the topology of phylogenetic trees; therefore, whole mitogenome sequencing provides a more reliable and comprehensive approach for assessing genetic diversity and elucidating phylogenetic relationships (Achilli et al., 2012). Sixteen mitogenome haplogroups (A–R) have been identified among 83 horse individuals distributed across Asia, Europe, the Middle East, and the Americas (Achilli et al., 2012).

The advent of next-generation sequencing technologies has enabled rapid and accurate characterization of mitochondrial genomes, which, being smaller and less complex than nuclear genomes, serves as highly effective molecular markers (Brankovics et al., 2017; Yong et al., 2015). The mitochondrial genome is widely employed in studies involving evolutionary analysis, genetic diversity, phylogeny, population structure, and taxonomy (Carpi et al., 2016; Delsuc et al., 2003; Hassanin et al., 2013; Olivieri et al., 2017; Ramos et al., 2018). Moreover, mitochondrial genomes provide critical insights for comparative and evolutionary genomic research (Carpi et al., 2016; Du et al., 2025; Olivieri et al., 2017; Ramos et al., 2018). In mammals, mitochondrial DNA typically ranges from 15 to 18 kilobases in length (Galtier et al., 2009; Jenuth et al., 1996; Saccone et al., 1999) and contains 37 genes, including 22 tRNAs, 2 rRNAs, and 13 protein-coding genes (PCGs), and 1 non-coding control region (Galtier et al., 2009; Iborra et al., 2004).

Accordingly, the objectives of this study were to i) characterize the mitogenomes of Turkish domestic and feral (*yılkı*) horses, ii) establish a reference mitogenome database for Turkish horse populations, and iii) elucidate phylogenetic relationships among Turkish horses.

Materials and Methods

Sample Collection and Genomic DNA (gDNA) Isolation

Hair samples from one female feral (*yılkı*) and one domestic horse from Kayseri, Türkiye, were collected and preserved in absolute alcohol at –20 °C until genomic DNA extraction. gDNA was isolated using the DNeasy Blood and Tissue Kit (Cat. No: 69504, QIAGEN) following the manufacturer’s protocol. The concentration of extracted gDNA was quantified using the dsDNA Quantification Broad Range kit (Cat. No: Q32850, Thermo Fisher Scientific). DNA purity and integrity were assessed by electrophoresis on a 1% agarose gel and spectrophotometric absorbance measurements. The extracted gDNA samples were subsequently used to amplify the entire mitochondrial genome in two overlapping fragments via long-range polymerase chain reaction (PCR).

PCR, Library Preparation, and Sequencing

The complete mitochondrial genome was amplified in two overlapping fragments using the long-range PCR method with NEB LongAmp® Taq 2X Master Mix (Cat. No: M0287S, NEB). Primer pairs AL1_2024L–BH1_13002H (~11,000 base pairs [bp]) and LuLu_12690L–LuLu_2503H (~6,500 bp) were used, as described by Ibis (2019). PCR reactions and cycling conditions were performed according to Ibis (2019). The concentration of the resulting amplicons was determined using the QuBit dsDNA HS Kit (Cat. No: Q32851, Thermo Fisher Scientific) and diluted to 0.2 ng/μL for sequencing library preparation. Sequencing libraries were prepared using the Illumina Nextera XT DNA Library Preparation Kit (Cat. No: FC-131-1096) and the Nextera XT Index Kit (Cat. No: FC-131-1002) (Illumina), according to the manufacturer's instructions. Library normalization was performed using a bead-based approach, and sequencing was conducted on the Illumina MiSeq platform (Genom and Stem Cell Center, GENKOK, Erciyes University, Kayseri, Türkiye) with the MiSeq Reagent Kit v2.

Bioinformatics Analyses

The raw sequencing reads were imported into Geneious Prime (<https://www.geneious.com>) for quality assessment, filtering, and subsequent analyses. Adapter sequences were removed from the raw reads, and both low-quality reads (quality score <25) and short reads (<50 bp) were discarded using the BBDuk trimming tool within Geneious Prime. The remaining high-quality reads were mapped against the reference mitochondrial genome (GenBank accession number NC_001640) using the Geneious Mapper algorithm, configured with highest sensitivity/medium parameters and fine-tuning for up to 25 iterations). Annotation information was transferred from the reference mitogenome and applied through

Geneious Prime. To further validate the data, the raw reads were also subjected to *de novo* assembly using GetOrganelle software (Jin et al., 2020). Annotation boundaries were manually inspected and confirmed with MITOS2 (Donath et al., 2019) and tRNAscan-SE 2.0 Web Server (Lowe & Chan, 2016). Visualization of tRNA secondary structures was performed using the VARNA Java program (Darty et al., 2009), and circular representations of the mitogenome were generated with CG View web software (Grant & Stothard, 2008). The base composition of the mitogenomes was analyzed using Geneious Prime. Strand asymmetries were determined according to the following formulas: AT skew = $[A - T] / [A + T]$ and GC skew = $[G - C] / [G + C]$.

Phylogenetic Analyses

In addition to the newly obtained mitogenomes, a comparative dataset was assembled, comprising 31 horse mitogenomes and two outgroup mitogenomes from the genus *Equus*, all obtained from the GenBank database (*National Center for Biotechnology Information*) (Table 1). Sequence alignments was generated using the MAFFT algorithm (Katoh et al., 2002) implemented in Geneious Prime with default settings. Following alignment, the D-loop region characterized by a high mutation rate, repetitive motifs, and alignment ambiguities in mammals, was excluded from subsequent phylogenetic reconstruction. Ambiguous regions that introduced uncertainty into the dataset were further removed using the GBlocks v.0.91b program (Castresana, 2000) with default parameters. The optimal base substitution model was identified using jModelTest 2.1.10 (Darriba et al., 2012) based on the AIC criterion. Phylogenetic inference was performed using the Maximum Likelihood (ML) method in MEGA11 (Tamura et al., 2021) with 10,000 bootstrap replicates to assess node support.

Table 1. List of horse mitogenomes used in phylogenetic analyses, including two outgroups.

Name	Species	Accession number	Haplogroup	Reference
Feral-TR	<i>E. caballus</i>	PQ573015	A	In this study
JN398377_Efc-m_A	<i>E. caballus</i>	JN398377	A	Achilli et al. (2012)
JN398378_Efc-m_A	<i>E. caballus</i>	JN398378	A	Achilli et al. (2012)
JN398382_Efc-m_A	<i>E. caballus</i>	JN398382	A	Achilli et al. (2012)
JN398384_Efc-m_A	<i>E. caballus</i>	JN398384	A	Achilli et al. (2012)
JN398380_Efc-m_A	<i>E. caballus</i>	JN398380	A	Achilli et al. (2012)
JN398385_Efc-m_A	<i>E. caballus</i>	JN398385	A	Achilli et al. (2012)
JN398381_Efc-m_A	<i>E. caballus</i>	JN398381	A	Achilli et al. (2012)
JN398383_Efc-m_A	<i>E. caballus</i>	JN398383	A	Achilli et al. (2012)
JN398379_Efc-m_A	<i>E. caballus</i>	JN398379	A	Achilli et al. (2012)
Domestic-TR	<i>E. caballus</i>	PQ573014	B	In this study
JN398391_Efc-m_B	<i>E. caballus</i>	JN398391	B	Achilli et al. (2012)
JN398390_Efc-m_B	<i>E. caballus</i>	JN398390	B	Achilli et al. (2012)
JN398386_Efc-m_B	<i>E. caballus</i>	JN398386	B	Achilli et al. (2012)
JN398388_Efc-m_B	<i>E. caballus</i>	JN398388	B	Achilli et al. (2012)

Table 2. Sequencing read statistics for Turkish horse mitogenomes.

Isolate	Location	Raw sequences		Filtered sequences		Coverage (X)	Mitogenome length (bp)
		Read count	Average sequence length (bp)	Read count	Average sequence length (bp)		
EqCabTR1 (PQ573014)	Kayseri	495458	224.5	357324	195.9	4183.4	16657
EqCabTR2 (PQ573015)	Kayseri	487640	219.4	357336	191.4	4019.1	16657

Table 3. Annotation of genomic regions in the mitogenomes of Turkish horses.

Region	Start–stop		Length	Direction	Intergenic region	Start codon	Stop codon	Anti-codon
tRNA ^{Phe}	1	70	70	H	0			GAA
s-rRNA	71	1046	976	H	0			
tRNA ^{Val}	1047	1113	67	H	0			TAC
l-rRNA	1114	2693	1580	H	0			
tRNA ^{Leu}	2694	2768	75	H	2			TAA
ND1	2771	3727	957	H	−1	ATG	TAG	
tRNA ^{Ile}	3727	3795	69	H	−3			GAT
tRNA ^{Gln}	3793	3865	73	L	2			TTG
tRNA ^{Met}	3868	3936	69	H	0			CAT
ND2	3937	4977	1041	H	−2	ATA	TAG	
tRNA ^{Trp}	4976	5045	70	H	5			TCA
tRNA ^{Ala}	5051	5119	69	L	1			TGC
tRNA ^{Asn}	5121	5193	73	L	0			GTT
rep origin	5195	5227	33	H	−2			
tRNA ^{Cys}	5226	5291	66	L	0			GCA
tRNA ^{Tyr}	5292	5358	67	L	1			GTA
COX1	5360	6904	1545	H	−3	ATG	TAA	
tRNA ^{Ser}	6902	6970	69	L	8			TGA
tRNA ^{Asp}	6979	7045	67	H	0			GTC
COX2	7046	7729	684	H	3	ATG	TAA	
tRNA ^{Lys}	7733	7800	68	H	1			TTT
ATP8	7802	8005	204	H	−43	ATG	TAG	
ATP6	7963	8643	681	H	−1	ATG	TAA	
COX3	8643	9426	784	H	0	ATG	T--	
tRNA ^{Gly}	9427	9495	69	H	0			TCC
ND3	9496	9842	347	H	0	ATA	TA-	
tRNA ^{Arg}	9843	9911	69	H	1			TCG
ND4L	9913	10209	297	H	−7	ATG	TAA	
ND4	10203	11580	1378	H	0	ATG	T--	

Table 3. Continued.

Region	Start–stop		Length	Direction	Intergenic region	Start codon	Stop codon	Anti-codon
tRNA ^{His}	11581	11649	69	H	0			GTG
tRNA ^{Ser}	11650	11709	60	H	1			GCT
tRNA ^{Leu}	11711	11780	70	H	6			TAG
<i>ND5</i>	11787	13601	1815	H	−17	ATG	TAA	
<i>ND6</i>	13585	14112	528	L	0	ATG	TAA	
tRNA ^{Glu}	14113	14181	69	L	4			TTC
<i>CYTB</i>	14186	15325	1140	H	0	ATG	AGA	
tRNA ^{Thr}	15326	15398	73	H	1			TGT
tRNA ^{Pro}	15400	15465	66	L	0			TGG
D-loop	15466	16657	1192	H	0			

Table 4. Nucleotide compositions of the mitogenomes of Turkish horses.

Feral (domestic)	Length	%	A%	T%	G%	C%	A + T%	G + C%	AT skew	GC skew
Total length	16657	100	32.20	25.90 (25.80)	13.40	28.60	58	42	0.10	-0.36
PCGs	11401	68.42	31.30 (31.40)	26.10	12.10	30.40	57.50	42.50	0.09	-0.43
rRNAs	2556	15.34	36.70	23.80 (27.70)	16.90	22.60 (22.70)	60.50 (60.40)	39.50 (39.60)	0.21	-0.14
tRNAs	1517	9.10	34.70	27.20	15.60	22.50	62	38	0.12	-0.18
D-loop	1192	7.15	27.10 (27.30)	26.30 (26.20)	15.70 (15.50)	30.90 (31.00)	53.40	46.60	0.01	-0.32

The combined length of the 13 PCG in the Turkish horses mitogenomes was 11401 bp, representing 68.42% of the total mitogenome length. These genes encoded 3,799 amino acids. Among the encoded amino acids, 4.3% were acidic, 6.9% basic, 11.2% charged, 29.3% polar uncharged, and 61.9% hydrophobic. Of the PCGs, 11 genes initiated with the “ATG” codon, whereas *ND2* and *ND3* began with ‘ATA’. The stop codon “TAA” was present in six genes, while “TAG” appeared in *ND1*, *ND2*, and *ATP8*, and “AGA” was found in *CYTB*. Incomplete stop codons were detected as “T—” in *COX3* and *ND4*, and “TA-” in *ND3* (Table 3). Codon usage frequencies in the Turkish horse mitogenomes are provided in Table 5. The most frequently used codons were “CUA” (285), “AUC” (208), “AUA” (186), and “UUC” (160), while the least frequent codon was “CGG” (3).

The total length of rRNA genes was 2556 bp comprising the 12S-rRNA (s-rRNA) was 976 bp and the 16S-rRNA (l-rRNA) was 1580 bp, located between tRNA^{Phe} and tRNA^{Leu}(UUR), and separated by tRNA^{Val}. The tRNA genes ranged in length from 60 to 75 bp. All tRNAs exhibited the typical cloverleaf secondary structure, except for tRNA^{Ser}(AGY), which lacked a dihydrouridine (DHU) arm (Figure 2). The replication origin, 32 bp in length, was situated within the WANCY tRNA cluster. The control region

(D-loop) measured 1,192 bp, making it the longest non-coding sequence in the mitogenome.

Phylogenetic Analyses

The mitogenomes of Turkish horses, together with those representing other horse haplogroups available in the GenBank database, were analyzed using the GTR + I substitution model under the ML method with 10,000 bootstrap replications (Table 1, Figure 3). In the resulting phylogenetic tree, the Turkish feral horse (yılki) clustered with strong bootstrap support within Haplogroup A, grouping with other A haplogroup horses from various geographic regions, including North America (Chincoteague Pony, JN398377), the Middle East (Caspian Pony, Arab, and unidentified Syrian and Iranian breeds, JN398378, JN398384, JN398380, and JN398383), Southern Europe (Maremmano breed, JN398379, JN398381, and JN398382), and Central Asia (Akhal-Teke breed, JN398385). In contrast, the domestic Turkish horse was positioned with high bootstrap support within Haplogroup B, together with undefined Syrian breeds from the Middle East (JN398391 and JN398389), Maremmano and undefined Italian breeds from Southern Europe (JN398390, JN398387, and JN398388), and the Westphalian breed from Central Europe (JN398386).

Table 5. Codon usage in the mitogenomes of Turkish horses.

		Domestic horse		Feral horse	
Amino acid	Codon	Count	Fraction	Count	Fraction
Ala	GCG	5.00	0.02	5.00	0.02
	GCA	76.00	0.32	76.00	0.32
	GCT	51.00	0.22	51.00	0.22
	GCC	103.00	0.44	103.00	0.44
Cys	TGT	5.00	0.22	5.00	0.22
	TGC	18.00	0.78	18.00	0.78
Asp	GAT	26.00	0.40	26.00	0.39
	GAC	39.00	0.60	40.00	0.61
Glu	GAG	19.00	0.20	20.00	0.21
	GAA	78.00	0.80	77.00	0.79
Phe	TTT	78.00	0.33	77.00	0.32
	TTC	160.00	0.67	161.00	0.68
Gly	GGG	30.00	0.14	30.00	0.14
	GGA	99.00	0.46	99.00	0.46
	GGT	21.00	0.10	22.00	0.10
	GGC	64.00	0.30	63.00	0.29
His	CAT	22.00	0.21	22.00	0.21
	CAC	81.00	0.79	81.00	0.79
Ile	ATT	131.00	0.39	132.00	0.39
	ATC	208.00	0.61	207.00	0.61
Lys	AAG	10.00	0.10	10.00	0.10
	AAA	87.00	0.90	87.00	0.90
Leu	TTG	21.00	0.03	21.00	0.03
	TTA	61.00	0.10	61.00	0.10
	CTG	38.00	0.06	38.00	0.06
	CTA	285.00	0.46	285.00	0.46
	CTT	68.00	0.11	68.00	0.11
	CTC	140.00	0.23	140.00	0.23
Met	ATG	45.00	0.19	45.00	0.20
	ATA	186.00	0.81	185.00	0.80
Asn	AAT	40.00	0.27	40.00	0.27
	AAC	110.00	0.73	109.00	0.73
Pro	CCG	7.00	0.04	7.00	0.04
	CCA	76.00	0.40	76.00	0.40
	CCT	36.00	0.19	36.00	0.19
	CCC	72.00	0.38	72.00	0.38
Gln	CAG	8.00	0.09	8.00	0.09
	CAA	83.00	0.91	83.00	0.91
Arg	CGG	3.00	0.05	3.00	0.05
	CGA	37.00	0.58	37.00	0.58
	CGT	8.00	0.13	7.00	0.11

Table 5. Continued.

Amino acid	Codon	Domestic horse		Feral horse	
		Count	Fraction	Count	Fraction
Ser	CGC	16.00	0.25	17.00	0.27
	AGT	12.00	0.04	12.00	0.04
	AGC	53.00	0.17	53.00	0.17
	TCG	9.00	0.03	9.00	0.03
	TCA	121.00	0.40	121.00	0.40
Thr	TCT	36.00	0.12	36.00	0.12
	TCC	73.00	0.24	73.00	0.24
	ACG	11.00	0.03	11.00	0.03
	ACA	144.00	0.46	144.00	0.46
Val	ACT	57.00	0.18	57.00	0.18
	ACC	103.00	0.33	103.00	0.33
	GTG	25.00	0.13	25.00	0.13
	GTA	78.00	0.42	79.00	0.42
Trp	GTT	31.00	0.17	31.00	0.16
	GTC	53.00	0.28	53.00	0.28
	TGG	8.00	0.08	8.00	0.08
Tyr	TGA	96.00	0.92	96.00	0.92
	TAT	47.00	0.37	47.00	0.37
	TAC	81.00	0.63	81.00	0.63

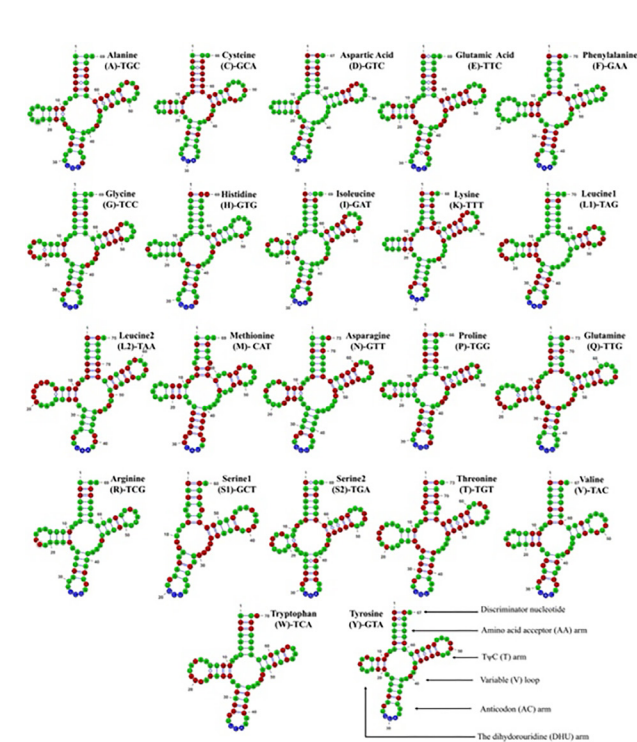


Figure 2. Predicted secondary structures of the 22 tRNAs identified in the mitochondrial genome of Turkish horses.

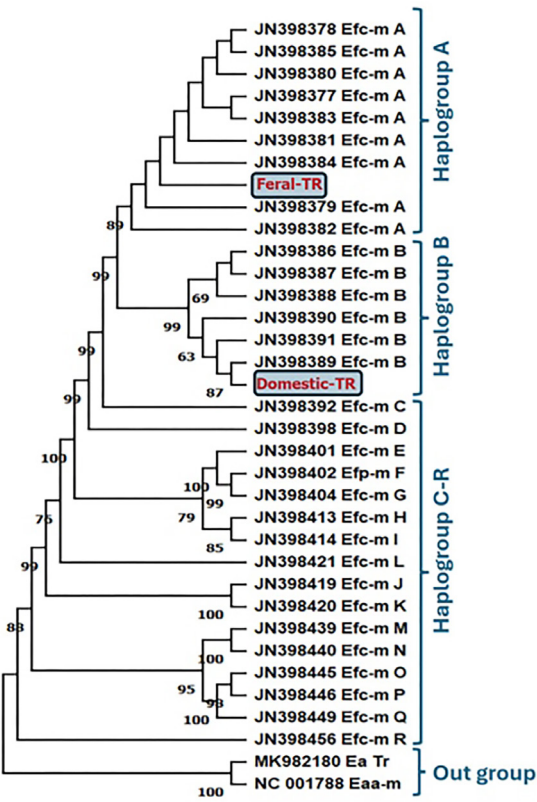


Figure 3. Phylogenetic tree constructed using the ML method, showing Turkish horses and reference samples from GenBank.

Discussion

The two horse mitogenomes obtained in the present study exhibited structural features similar to previously reported horse mitogenomes (Achilli et al., 2012; Lippold et al., 2011; Luo et al., 2011; Sarkissian et al., 2015; Vilstrup et al., 2013), consisting of a circular molecule containing 37 gene regions (Figure 1, Table 2). According to the GenBank database, a total of 535 complete mitogenomes are currently available under the name *Equus caballus*, representing horse breeds from various regions worldwide, with lengths ranging from 16,403 to 16,681 bp ([https://www.ncbi.nlm.nih.gov/nuccore/txid9796\[Organism:noexp\]](https://www.ncbi.nlm.nih.gov/nuccore/txid9796[Organism:noexp])): Access date: 02.11.2024). The reference horse mitogenome (*E. caballus*) (NC_001640/X79547) has a total length of 16,660 bp (Xiufeng & Árnason, 1994), which closely corresponds to that of the Turkish horse breeds (16,657 bp). The variation in sequence length observed among existing horse mitogenomes is primarily attributed to differences in the length of the control region sequence. Although both modern and ancient horse mitogenomes have been reported, the mitogenomes of Turkish horses were previously unavailable and are presented here for the first time.

The general structural characteristics of the Turkish horse mitogenomes were consistent with those of other vertebrates. The nucleotide composition displayed a distinct bias toward Adenine and Thymine (58% A + T), a common feature in vertebrate mitogenomes. Similarly, the presence of numerous intergenic gaps (45 bp) and overlaps (68 bp) reflects the compact organization typical for horse and other vertebrate mitogenomes (Guo et al., 2016; Ibis, 2019; Krause et al., 2008; Rohland et al., 2007; Wada et al., 2010). The occurrence of incomplete stop codons (“T--” in *COX3* and *ND4*, and “TA--” in *ND3*) aligns with the reference horse mitogenome and other vertebrates (Ibis, 2019; Xiufeng & Árnason, 1994) where these codons function equivalently to complete stop codons (Ojala et al., 1981).

The codon usage pattern in Turkish horses, with “CUA,” “AUC,” and “AUA” being the most frequently used codons, was highly similar to that of the donkey mitogenome from Türkiye, suggesting a conserved codon usage pattern within the *Equus* genus (Ibis, 2019). The structural features of the tRNA genes in Turkish horses were also consistent with those typically observed in mammalian mitogenomes. The 22 identified tRNAs ranged from 60 to 75 bp in length, and all but one displayed the canonical cloverleaf secondary structure. The single exception, tRNA^{Ser}(AGY), lacked the DHU arm, a variation that is known to be functionally acceptable (Steinberg & Cedergren, 1994). Furthermore, the lengths of the ribosomal RNA genes (976 bp for s-rRNA and 1,580 bp for l-rRNA) and the D-loop region (1,192 bp), were nearly identical to those reported in the reference horse mitogenome (975 bp, 1,581 bp and 1,192 bp, respectively), indicating a high degree of structural conservation (Xiufeng & Árnason, 1994). The observed length differences among horse mitogenomes are mainly attributed to variations in the repeat regions within the D-loop, and it has been noted that variation in the number of repeat motifs within the

D-loops, a region known for its heteroplasmic nature (Xiufeng & Árnason, 1994).

In the ML phylogenetic tree reconstructed in this study, the Turkish horses mitogenomes clustered in accordance with the haplogroups defined by Achilli et al. (2012). The mitogenome of the Turkish feral horse grouped within Haplogroup A, whereas the domestic Turkish horse clustered within Haplogroup B. The findings indicate that the Turkish feral horse shares a common maternal lineage with other horses in Haplogroup A, which is widely distributed across global horse populations and may represent an ancient or broadly dispersed lineage. In contrast, the domestic Turkish horse’s placement in Haplogroup B suggests a distinct maternal lineage, likely reflecting differences in breeding history or geographical origin. Consistent with previous research, a previous study reported that the Przewalski horse in Eurasia possesses only the F and J-K lineages, with the greatest haplogroup diversity observed on the Asian continent (Achilli et al., 2012). The broad distribution of Haplogroup A across North America, the Middle East, and Central Asia supports the hypothesis of a shared evolutionary history among geographically distinct horse populations. Similarly, the inclusion of the Turkish domestic horse in Haplogroup B indicates connections with Central and Southern European populations. The identification of Haplogroup B in both Italy and East Asia (Achilli et al., 2012; Cardinali et al., 2016) further highlights its historical dissemination and emphasizes the pivotal role of Anatolia in horse domestication history. The domestication process, which originated in the Eurasian steppes and subsequently spread across various regions (Lippold et al., 2011; Outram et al., 2009), underscores Türkiye’s position as an integral part of this evolutionary and cultural pathway.

Conclusion

In conclusion, the complete mitochondrial genomes of Turkish horses, presented here for the first time, provide a valuable foundation for future genetic studies, including assessments of genetic diversity and conservation efforts within Türkiye. Further investigation of mitogenomic and genomic data from additional Turkish horse breeds, combined with comparative analyses across global datasets, will enhance understanding of horse evolutionary history and contribute to conservation strategies.

Ethics

Ethics Committee Approval: This study was approved by the Erciyes University Local Ethics Committee for Animal Experiments (approval number: 23/061, dated: 05.04.2023).

Data Sharing Statement: All data are available within the study.

Footnotes

Conflict of Interest: The author(s) have no conflicts of interest to declare.

Funding: This study was supported by The Scientific Research Projects (BAP) of Erciyes University (Project No.: FBAÜ-2023-12262).

References

- Achilli, A., Olivieri, A., Soares, P., Lancioni, H., Kashani, B. H., Perego, U. A., & Torroni, A. (2012). Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proceedings of the National Academy of Sciences*, 109(7), 2449–2454. <https://doi.org/10.1073/pnas.1111637109>
- Ahlatwat, S., Sharma, U., Niranjani, S. K., Chhabra, P., Arora, R., Sharma, R., Singh, K. V., Vijh, R. K., & Mehta, S. C. (2025). Unraveling the maternal heritage: Identifying the complex origins of indigenous Indian horse and pony breeds through mitochondrial genome analysis. *Mammalian Genome*, 36(1), 118–128. <https://doi.org/10.1007/s00335-024-09963-0>
- Aksu, S., Aktoprakligil, D., Koban, E., & Aslan, O. (2010). Genetic characterization of native horse breeds within the scope of protection of gene resources and biodiversity. *Research Journal of Biological Sciences*, 3, 121–130. <http://www.aensiweb.com/RJBS/>
- Bayram, D. (2005). The phenotypic characteristics of horses raised in Van region. *Van Veterinary Journal*, 16(1), 85–88. <https://dergipark.org.tr/en/pub/vanvet/issue/40274/487006>
- Brankovics, B., Dam, P., Rep, M., Hoog, G. S., Lee, J. A., Waalwijk, C., & Diepeningen, A. D. (2017). Mitochondrial genomes reveal recombination in the presumed asexual *Fusarium oxysporum* species complex. *BMC Genomics*, 18, 1–14. <https://doi.org/10.1186/s12864-017-4066-7>
- Cardinali, I., Lancioni, H., Giontella, A., Capodiferro, M. R., Capomaccio, S., Buttazzoni, L., Biggio, G. P., Cherchi, R., Albertini, E., Olivieri, A., Cappelli, K., Achilli, A., & Silvestrelli, M. (2016). An overview of ten Italian horse breeds through mitochondrial DNA. *PLOS ONE*, 11(4), e0153004. <https://doi.org/10.1371/journal.pone.0153004>
- Carpi, G., Kitchen, A., Kim, H. L., Ratan, A., Drautz-Moses, D. I., McGraw, J. J., & Schuster, S. C. (2016). Mitogenomes reveal diversity of the European Lyme borreliosis vector *Ixodes ricinus* in Italy. *Molecular Phylogenetics and Evolution*, 101, 194–202. <https://doi.org/10.1016/j.ympev.2016.05.028>
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17(4), 540–552. <https://doi.org/10.1093/oxfordjournals.molbev.a026334>
- Celik, S., Coskun, F., & Yilmaz, O. (2015). Investigation by different orthogonal methods of body sizes of Turkish spotted horses according to age group. *COMU Journal of Agriculture Faculty*, 3(1), 81–87. <https://dergipark.org.tr/tr/pub/comuagri/issue/27156/285456>
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and high-performance computing. *Nature Methods*, 9(8), 772. <https://doi.org/10.1038/nmeth.2109>
- Darty, K., Denise, A., & Ponty, Y. (2009). VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, 25(15), 1974–1975. <https://doi.org/10.1093/bioinformatics/btp250>
- Delsuc, F., Stanhope, M. J., & Douzery, E. J. P. (2003). Molecular systematics of armadillos (Xenarthra, Dasypodidae): Contribution of maximum likelihood and Bayesian analyses of mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution*, 28(2), 261–275. [https://doi.org/10.1016/S1055-7903\(03\)00036-8](https://doi.org/10.1016/S1055-7903(03)00036-8)
- Donath, A., Jühling, F., Al-Arab, M., Bernhart, S. H., Reinhardt, F., Stadler, P. F., & Bernt, M. (2019). Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Research*, 47(20), 10543–10552. <https://doi.org/10.1093/nar/gkz833>
- Du, W., Sun, Q., Hu, S., Yu, P., Kan, S., & Zhang, W. (2025). *Equus* mitochondrial pangenome reveals independent domestication imprints in donkeys and horses. *Scientific Reports*, 15(1), 6803. <https://doi.org/10.1038/s41598-025-65526-6>
- Galtier, N., Nabholz, B., Glémin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology*, 18(22), 4541–4550. <https://doi.org/10.1111/j.1365-294X.2009.04380.x>
- Grant, J. R., & Stothard, P. (2008). The CGView server: A comparative genomics tool for circular genomes. *Nucleic Acids Research*, 36(suppl_2), 181–184. <https://doi.org/10.1093/nar/gkn179>
- Guo, X., Pei, J., Chu, M., Wu, X., Bao, P., Ding, X., & Yan, P. (2016). The complete mitochondrial genome of Hequ horse. *Mitochondrial DNA Part A*, 27(6), 4657–4658. <https://doi.org/10.3109/19401736.2015.1101548>
- Hacan, Ö., Koçak, S., Çelikeoğlu, K., Bozkurt, Z., Erdoğan, M., & Tekerli, M. (n.d.). The independent spirit of Turkey: Wild horse.
- Hassanin, A., An, J., Ropiquet, A., Nguyen, T. T., & Couloux, A. (2013). Combining multiple autosomal introns for studying shallow phylogeny and taxonomy of Laurasiatherian mammals: Application to the tribe Bovini. *Molecular Phylogenetics and Evolution*, 66(3), 766–775. <https://doi.org/10.1016/j.ympev.2012.11.003>
- Ibis, O. (2019). Türkiye Anadolu Eşeği (*Equus asinus*)’nin mitogenom karakterizasyonu ve filogenetik ilişkileri. *Türkiye Tarımsal Araştırmalar Dergisi*, 6(3), 257–267. <https://doi.org/10.19159/tutad.460381>
- Iborra, F. J., Kimura, H., & Cook, P. R. (2004). The functional organization of mitochondrial genomes in human cells. *BMC Biology*, 2, 1–14. <https://doi.org/10.1186/1741-7007-2-9>
- Jenuth, J. P., Peterson, A. C., Fu, K., & Shoubridge, E. A. (1996). Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. *Nature Genetics*, 14(2), 146–151. <https://doi.org/10.1038/ng1096-146>
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., DePamphilis, C. W., Yi, T. S., & Li, D. Z. (2020). GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology*, 21, 1–31. <https://doi.org/10.1186/s13059-020-02154-5>
- Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30(14), 3059–3066. <https://doi.org/10.1093/nar/gkf436>
- Koban, E., Denizci, M., Aslan, O., Aktoprakligil, D., Aksu, S., Bower, M., & Arat, S. (2012). High microsatellite and mitochondrial diversity in Anatolian native horse breeds shows Anatolia as a genetic conduit between Europe and Asia. *Animal Genetics*, 43(4), 401–409. <https://doi.org/10.1111/j.1365-2052.2011.02284.x>
- Koseman, A., Ozsensoy, Y., Erdogan, M., Yarali, C., Toprak, B., Zengin, K., & Seker, I. (2020). Investigation of genetic variations using microsatellite markers in colored horses in Turkey. *Russian Journal of Genetics*, 56, 592–602. <https://doi.org/10.1134/S1022795420040073>
- Köseman, A., Ozsensoy, Y., Erdogan, M., Yarali, C., Toprak, B., Zengin, K., & Şeker, İ. (2019). Investigation of genetic structures of coloured horses by mtDNA D-loop sequence analysis in Turkey. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 25(6), 769–778. <https://doi.org/10.9775/kvfd.2019.22056>
- Krause, J., Unger, T., Noçon, A., Malaspinas, A. S., Kolokotronis, S. O., Stiller, M., & Hofreiter, M. (2008). Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene–Pliocene boundary. *BMC Evolutionary Biology*, 8, 1–12. <https://doi.org/10.1186/1471-2148-8-224>
- Levine, M. A. (2005). Domestication and early history of the horse. In *The domestic horse: The origins, development and management of its behaviour*. Cambridge University Press.
- Lippold, S., Matzke, N. J., Reissmann, M., & Hofreiter, M. (2011). Whole mitochondrial genome sequencing of domestic horses reveals incorporation of extensive wild horse diversity during domestication. *BMC Evolutionary Biology*, 11, 1–10. <https://doi.org/10.1186/1471-2148-11-147>
- Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE online: Integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research*, 44(W1), 54–57. <https://doi.org/10.1093/nar/gkw413>

- Luo, Y., Chen, Y., Liu, F., Jiang, C., & Gao, Y. (2011). Mitochondrial genome sequence of the Tibetan wild ass (*Equus kiang*). *Mitochondrial DNA*, 22(1–2), 6–8. <https://doi.org/10.3109/10425171003618702>
- MacHugh, D. E., & Bradley, D. G. (2001). Livestock genetic origins: Goats buck the trend. *Proceedings of the National Academy of Sciences*, 98(10), 5382–5384. <https://doi.org/10.1073/pnas.101126598>
- Naderi, S., Rezaei, H. R., Pompanon, F., Blum, M. G., Negrini, R., Naghash, H. R., & Taberlet, P. (2008). The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proceedings of the National Academy of Sciences*, 105(46), 17659–17664. <https://doi.org/10.1073/pnas.0804782105>
- 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>
- Olivieri, A., Sidore, C., Achilli, A., Angius, A., Posth, C., Furtwängler, A., & Torrioni, A. (2017). Mitogenome diversity in Sardinians: A genetic window onto an island's past. *Molecular Biology and Evolution*, 34(5), 1230–1239. <https://doi.org/10.1093/molbev/msx086>
- Outram, A. K., Stear, N. A., Bendrey, R., Olsen, S., Kasparov, A., Zaibert, V., & Evershed, R. P. (2009). The earliest horse harnessing and milking. *Science*, 323(5919), 1332–1335. <https://doi.org/10.1126/science.1168594>
- Ozbeyaz, C., & Akcapinar, H. (2010). *Course notes of horse breeding*. Ankara University Faculty of Veterinary Medicine.
- Ramos, B., González-Acuña, D., Loyola, D. E., Johnson, W. E., Parker, P. G., Massaro, M., & Vianna, J. A. (2018). Landscape genomics: Natural selection drives the evolution of mitogenome in penguins. *BMC Genomics*, 19, 1–17. <https://doi.org/10.1186/s12864-018-4840-7>
- Rezaei, H. R., Naderi, S., Chintauan-Marquier, I. C., Taberlet, P., Virk, A. T., Naghash, H. R., & Pompanon, F. (2010). Evolution and taxonomy of the wild species of the genus *Ovis* (Mammalia, Artiodactyla, Bovidae). *Molecular Phylogenetics and Evolution*, 54(2), 315–326. <https://doi.org/10.1016/j.ympev.2009.09.008>
- Rohland, N., Malaspina, A. S., Pollack, J. L., Slatkin, M., Matheus, P., & Hofreiter, M. (2007). Proboscidean mitogenomics: Chronology and mode of elephant evolution using mastodon as outgroup. *PLoS Biology*, 5(8), 207. <https://doi.org/10.1371/journal.pbio.0050207>
- Saccone, C., Giorgi, C., Gissi, C., Pesole, G., & Reyes, A. (1999). Evolutionary genomics in Metazoa: The mitochondrial DNA as a model system. *Gene*, 238(1), 195–209. [https://doi.org/10.1016/S0378-1119\(99\)00270-X](https://doi.org/10.1016/S0378-1119(99)00270-X)
- Sarkissian, C., Vilstrup, J. T., Schubert, M., Seguin-Orlando, A., Eme, D., Weinstock, J., & Orlando, L. (2015). Mitochondrial genomes reveal the extinct *Hippidion* as an outgroup to all living equids. *Biology Letters*, 11(3), 20141058. <https://doi.org/10.1098/rsbl.2014.1058>
- Sheikh, A., Ahmed, M. M. M., Mutawakil, M. H. Z., Saini, K. S., Alsulaimany, F. A., Hanafy, A., & Sabir, J. S. M. (2019). SNP mapping and phylogenetic analysis of Saudi Arabian horse breeds based on mitochondrial genome sequencing. *Indian Journal of Experimental Biology*, 58, 225–230.
- Sheikh, A. (2023). Mitochondrial DNA sequencing of Kehilan and Hamdani horses from Saudi Arabia. *Saudi Journal of Biological Sciences*, 30(9), 103741. <http://nopr.niscpr.res.in/handle/123456789/51265>
- Steinberg, S., & Cedergren, R. (1994). Structural compensation in atypical mitochondrial tRNAs. *Nature Structural Biology*, 1(8), 507–510. <https://doi.org/10.1038/nsb0894-507>
- Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7), 3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Vilstrup, J. T., Seguin-Orlando, A., Stiller, M., Ginolhac, A., Raghavan, M., Nielsen, S. C., & Orlando, L. (2013). Mitochondrial phylogenomics of modern and ancient equids. *PLOS ONE*, 8(2), 55950. <https://doi.org/10.1371/journal.pone.0055950>
- Wada, K., Okumura, K., Nishibori, M., Kikkawa, Y., & Yokohama, M. (2010). The complete mitochondrial genome of the domestic red deer (*Cervus elaphus*) of New Zealand and its phylogenetic position within the family Cervidae. *Animal Science Journal*, 81(5), 551–557. <https://doi.org/10.1111/j.1740-0929.2010.00762.x>
- Xiufeng, X., & Árnason, Ú. (1994). The complete mitochondrial DNA sequence of the horse, *Equus caballus*: Extensive heteroplasmy of the control region. *Gene*, 148(2), 357–362. [https://doi.org/10.1016/0378-1119\(94\)90766-3](https://doi.org/10.1016/0378-1119(94)90766-3)
- Xu, X., Gullberg, A., & Arnason, U. (1996). The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. *Journal of Molecular Evolution*, 43, 438–446. <https://doi.org/10.1007/BF02337515>
- Yilmaz, O. (2013). Two new genetic sources specific to Ardahan-Iğdir-Kars Basin: Colored horse and Dikkulak dog. *Turkey-Kars Governorship International Sarikamis Symposium*. <https://www.sariktur.org.tr/>
- Yong, H. S., Song, S. L., Lim, P. E., Chan, K. G., Chow, W. L., & Eamsobhana, P. (2015). Complete mitochondrial genome of *Bactrocera arecae* (Insecta: Tephritidae) by next-generation sequencing and molecular phylogeny of Dacini tribe. *Scientific Reports*, 5(1), 15155. <https://doi.org/10.1038/srep15155>
- Zeder, M. A. (2008). Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proceedings of the National Academy of Sciences*, 105(33), 11597–11604. <https://doi.org/10.1073/pnas.0801317105>