IRSTI 34.23.35 https://doi.org/10.26577/bb108320268

R.E. Amzeyev[1*], M.Sh. Yermekova[2], K. Kirikci[3], K.Zh. Dossybayev[2], T.T. Nurkenov[1]

[1]Abai Kazakh National Pedagogical University, Almaty, Kazakhstan
[2]Al-Farabi Kazakh National University, Almaty, Kazakhstan
[3]Ahi Evran University, Kırshehir, Turkiye
[*]e-mail: rauanbiolog@gmail.com

PHYLOGENETIC ANALYSIS OF THE KAZAKH FAT-TAILED COARSE-WOOL SHEEP BREED

In this study, a phylogenetic analysis was conducted to determine the genetic diversity of the Kazakh fat-tailed coarse-wool sheep breed raised in Kazakhstan using SNP polymorphisms of the mitochondrial DNA (mtDNA) d-loop region consisting of 840 base pairs (bp). A total of 80 Kazakh fat-tailed coarse-wool sheep were initially sampled from two breeding farms in Kazakhstan. Following DNA extraction, PCR amplification, and sequencing, 10 samples were successfully sequenced for the mitochondrial DNA (mtDNA) D-loop region. For comparative phylogenetic analysis, these 10 newly generated sequences were combined with 15 reference sequences retrieved from the NCBI GenBank database, resulting in a final dataset of 25 sequences. DNA molecules were extracted from peripheral blood samples, and polymerase chain reaction (PCR) amplification was performed. Sequencing analysis of the amplified products revealed nucleotide compositions of 31.46% adenine, 36.19% thymine, 16.42% guanine, and 15.93% cytosine. The obtained sequences were compared with available nucleotide sequences of domestic sheep and wild mouflon breeds from the NCBI database. Phylogenetic tree analysis conducted using the MEGA X software demonstrated that the Kazakh fat-tailed coarse-wool sheep formed a distinct cluster with a 98–100% genetic similarity, indicating a unique genetic diversity and confirming its status as an autochthonous sheep breed. Furthermore, the Kazakh fat-tailed coarse-wool sheep showed a close genetic relationship with Mongolian, Tibetan, and Gissar sheep breeds. In contrast, European sheep breeds were positioned at a greater genetic distance in the phylogenetic tree, suggesting that they belong to different evolutionary lineages.

Keywords: phylogenetics, sequencing, genetic diversity, D-loop, DNA markers, SNP50, PCR, genetic differentiation.

Р.Е. Амзеев[1*], М.Ш. Ермекова[2], К. Қырыкчы[3], К.Ж.Досыбаев[2], Т.Т. Нуркенов[1]

[1]Абай атындағы Қазақ ұлттық университеті, Алматы, Қазақстан
[2]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан
[3]Ахи Евран Университеті, Қыршехир, Түркия
*e-mail: rauanbiolog@gmail.com

Қазақтың құйрықты қылшық жүнді қой тұқымына жүргізілген филогенетикалық талдау

Бұл зерттеуде Қазақстанда өсірілетін қазақтың құйрықты қылшық жүнді қой тұқымының генетикалық әртүрлілігін анықтау мақсатында ұзындығы 840 жұп нуклеотидтен (ж.н.) тұратын митохондриялық ДНҚ-ның (мтДНҚ) D-ілмек аймағындағы SNP-полиморфизмдер негізінде филогенетикалық талдау жүргізілді. Зерттеу үлгілері ретінде популяциядан барлығы 80 қой кездейсоқ таңдап алынды. Перифериялық қан үлгілерінен ДНҚ бөлініп алынып, полимеразды тізбекті реакция (ПТР) әдісімен амплификация жүргізілді. Амплификацияланған өнімдерді секвенирлеу нәтижесінде нуклеотидтік құрамы 31,46% аденин, 36,19% тимин, 16,42% гуанин және 15,93% цитозинді құрады. Алынған нуклеотидтік тізбектер NCBI деректер базасында қолжетімді үй қойлары мен жабайы муфлондардың референстік тізбектерімен салыстырылды. MEGA X бағдарламасы арқылы жүргізілген филогенетикалық талдау қазақтың құйрықты қылшық жүнді қой тұқымының тізбектері 98–100% жоғары bootstrap статистикалық қолдауымен жеке кластер түзетінін көрсетті, бұл талдауға енгізілген референстік қой популяцияларынан айқын генетикалық дифференциациясын көрсетеді. Сонымен қатар қазақтың құйрықты қылшық жүнді қой тұқымы моңғол, тибет және гиссар қой тұқымдарымен жақын генетикалық туыстықты көрсетті. Ал еуропалық қой тұқымдары филогенетикалық ағашта генетикалық қашықтықтың жоғары болуымен ерекшеленді, бұл олардың әртүрлі эволюциялық линияларға жататынын көрсетеді.

Түйін сөздер: филогенетика, секвенирлеу, генетикалық әртүрлілік, D-ілмегі, ДНҚ маркерлер, SNP50 , ПТР.

 © 2026 Al-Farabi Kazakh National University

Р.Е. Амзеев[1*], М.Ш. Ермекова[2], К. Қырықчы[3],
К.Ж. Досыбаев[2], Т.Т. Нуркенов[1]

[1]Казахский национальный педагогический университет имени Абая, Алматы, Казахстан
[2]Казахский национальный университет имени аль-Фараби, Алматы, Казахстан
[3]Университет Ахи Еврана, Кыршехир, Турция
*e-mail: rauanbiolog@gmail.com

Филогенетический анализ казахской грубошёрстной овцы

В данном исследовании был проведен филогенетический анализ для определения генетического разнообразия казахской курдючной грубошерстной породы овец, разводимой в Казахстане, с использованием SNP-полиморфизмов D-петли митохондриальной ДНК (мтДНК) длиной 840 пар оснований (п.н.). Всего из популяции были случайным образом отобраны 80 овец. Из периферической крови выделяли ДНК, после чего проводили амплификацию методом полимеразной цепной реакции (ПЦР). В результате секвенирования амплифицированных продуктов был определен нуклеотидный состав: 31,46% аденина, 36,19% тимина, 16,42% гуанина и 15,93% цитозина. Полученные последовательности сравнивали с доступными в базе данных NCBI нуклеотидными последовательностями домашних овец и диких муфлонов. Филогенетический анализ, проведенный с использованием MEGA X, показал, что последовательности казахской курдючной грубошерстной породы овец образуют отдельный кластер с высокой статистической поддержкой bootstrap 98–100%, что свидетельствует о четкой генетической дифференциации от референсных популяций овец, включенных в анализ. Кроме того, казахская курдючная грубошерстная порода овец продемонстрировала близкое генетическое родство с монгольской, тибетской и гиссарской породами овец. В отличие от них, европейские породы овец располагались на большем генетическом расстоянии на филогенетическом дереве, что указывает на их принадлежность к различным эволюционным линиям.

Ключевые слова: филогенетика, секвенирование, генетическое разнообразие, D-петля, ДНК-маркеры, SNP50, ПЦР.

Introduction

Animal husbandry is a traditional and strategically important branch of agricultural production, deeply rooted in the natural and climatic conditions of Kazakhstan and based on centuries of livestock-breeding culture and experience. In particular, sheep breeding is closely intertwined with the lifestyle of the nomadic population and represents an integral part of the national culture and tradition. Among the local livestock breeds, the Kazakh fat-tailed coarse-wool sheep occupies a special place. Its origin dates back to ancient times, having been developed through folk selection in the steppe and semi-desert regions of Kazakhstan. This breed is characterized by high adaptability, resilience, and productivity in both meat and fat yield, representing a unique component of the country's genetic resources. Between 1930 and 1950, the Kazakh fat-tailed coarse-wool sheep was described based on specialized zootechnical studies and officially recognized as an independent breed according to its ethnological and productive characteristics. Today, this breed is widely distributed across Kazakhstan, particularly in the southern, western, central, and eastern regions of the country (Kanapin, 2009).

Scientific investigation of the genetic diversity and biological characteristics of native livestock breeds has become one of the priority directions in enhancing agricultural productivity, ensuring food security, and strengthening export potential. The preservation and efficient use of genetic resources play a decisive role in maintaining the nation's food security and guiding breeding and selection programs (Wood & Phua, 1996; Ozerov et al., 2006; Pozharskiy et al., 2020). Morphologically, the Kazakh fat-tailed coarse-wool sheep is large, compact, and characterized by a fat-storing tail and coarse-type fleece. The average live weight of adult rams ranges from 90 to 120 kg, while ewes typically weigh between 55 and 75 kg. Lambs can reach 20–30 kg by the age of four months, and these values can further increase under favorable feeding and grazing conditions. The tail fat alone may reach 8–14 kg, and under special care, up to 20 kg. At present, studying the genetic diversity of this breed is a pressing issue. Intensive crossbreeding, large-scale selection programs aimed at maximizing production, and the potential loss of regional genetic resources pose serious threats to the genetic uniqueness and breed integrity of local sheep populations (Hiendleder et al., 1998; Kumar et al., 2024; Osipov, 2007).

These factors necessitate an in-depth investigation of the intra-breed genetic structure using an integrated approach that combines morphological and molecular-genetic methods (Erokhin et al., 2014). Since the Kazakh coarse-wool sheep is distributed across Central Asia, its mitochondrial

DNA (mtDNA) haplotypes are likely to belong to haplogroups A or C, which distinguishes it genetically from European sheep breeds (Deniskova et al., 2019; Meadows et al., 2005). However, the level of genetic polymorphism in fat-tailed coarse-wool sheep remains insufficiently studied. To analyze the genetic structure of sheep populations, efficient and accessible molecular-genetic markers such as short tandem repeats (STR loci) and the non-coding D-loop region of mitochondrial DNA (mtDNA) are widely used. Because mtDNA is maternally inherited, it exhibits high intraspecific variability and serves as a reliable marker for studying genetic diversity and phylogenetic relationships (Marzanov et al., 2004; Ozerov, 2004; Tajima, 1989). Sequencing of mtDNA is recognized as one of the most effective methods for identifying mtDNA haplotypes in sheep (Hiendleder et al., 2002; Pedrosa et al., 2005). In the study conducted by I.E. Mukhametzharova, the level of genetic diversity among six sheep breeds reared in Kazakhstan-Kazakh semi-fine-wool meat-wool, Edilbay, South Kazakhstan merino, Kazakh fat-tailed semi-coarse wool, and Akzhaik fine-wool meat-wool-was investigated based on mtDNA. The mtDNA sequences from samples of each breed were analyzed and comparatively characterized according to their phylogenetic structure and haplotypic composition. The results showed that among 64 samples representing six local breeds, 43.75% (28/64) belonged to haplogroup A, 39.06% (25/64) to haplogroup B, and 17.19% (11/64) to haplogroup C (Hiendleder et al., 2002). Currently, international databases such as GenBank and EMBL contain extensive microsatellite data on domestic animals. For sheep alone, more than 300 microsatellite markers have been characterized. Due to their high polymorphism, these markers are widely employed to assess the genetic characteristics of livestock species and breeds (Álvarez et al., 2013; Mukhametzharova, 2022).

Materials and methods

For this study, peripheral blood samples were collected from 80 Kazakh fat-tailed coarse-wool sheep from two breeding farms in Kazakhstan: 40 animals from the “Kundyzdy” breeding farm and 40 animals from the “Aksunkar” breeding farm. Animals were randomly selected from the available breeding population at each farm, without intentional selection based on specific production or phenotypic traits. The 80 animals represented the initial study population. Following DNA extraction, PCR amplification of the mitochondrial DNA D-loop region, and assessment of PCR product quality, 10 samples were successfully sequenced and retained for downstream sequence analysis. The collection of peripheral blood samples was carried out by qualified veterinary specialists. Blood samples were drawn using sterile needles and collected into tubes containing EDTA reagent to prevent coagulation. The samples were then transported to the laboratory at a temperature of +4 °C. Genomic DNA was extracted from the peripheral blood using the QIAamp DNA Blood Mini Kit (250) (QIAGEN, Germany). The concentration and purity of the isolated DNA were assessed using a NanoDrop One spectrophotometer (Thermo Fisher Scientific, USA). Primers specific to the mitochondrial DNA D-loop region were synthesized by Macrogen Inc. (Seoul, South Korea). PCR amplification of the D-loop region was performed using a Mastercycler thermal cycler (Eppendorf, Germany). The polymerase chain reaction (PCR) conditions were as follows: initial denaturation at 94 °C for 6 min; followed by 34 cycles of denaturation at 96 °C for 46 s, annealing at 54 °C for 32 s, and extension at 74 °C for 1 min; with a final extension step at 74 °C for 1 min. For comparative phylogenetic analysis, 15 reference mtDNA D-loop sequences representing domestic and wild sheep breeds were retrieved from the NCBI GenBank database. Phylogenetic relationships were reconstructed using the Neighbor-Joining (NJ) method implemented in MEGA X. The analysis was based on the aligned mitochondrial DNA D-loop sequences of the 10 newly generated KFT-CWS samples and 15 reference sequences from GenBank. Bootstrap support for the inferred tree topology was evaluated using 1,000 replicates. The resulting phylogenetic tree was visualized and annotated using MEGA X. These reference sequences were selected to represent geographically and genetically diverse sheep populations relevant to the phylogenetic analysis. Thus, the final dataset used for phylogenetic reconstruction consisted of 25 sequences, including 10 newly generated KFT-CWS sequences and 15 sequences obtained from GenBank.

For comparative phylogenetic analysis, 15 reference mitochondrial DNA (mtDNA) D-loop sequences representing domestic and wild sheep breeds were retrieved from the NCBI GenBank database. The final dataset consisted of 25 sequences, including 10 newly generated KFT-CWS sequences and 15 reference sequences from GenBank. Multiple sequence alignment was performed using the ClustalW algorithm implemented in MEGA X. The aligned sequences were manually inspected and trimmed to ensure comparable nucleotide positions

across all sequences. Phylogenetic relationships were reconstructed using the Neighbor-Joining (NJ) method implemented in MEGA X, and the reliability of the inferred tree topology was assessed using 1,000 bootstrap replicates. Gaps and missing data were treated according to the default settings of MEGA X. Nucleotide substitution probabilities and the corresponding substitution matrix were calculated using MEGA 12 under the Tamura–Nei (1993) model. The resulting phylogenetic tree was visualized and annotated using MEGA X.

Results and discussion

Polymerase chain reaction (PCR) analysis was performed using specific primers corresponding to the mitochondrial DNA (mtDNA) D-loop region (Figure 1). The amplified PCR products were subsequently analyzed by electrophoresis in a 2% agarose gel. The molecular weight of the synthesized amplicons was evaluated using a DNA ladder ranging from 100 to 1000 base pairs (bp).

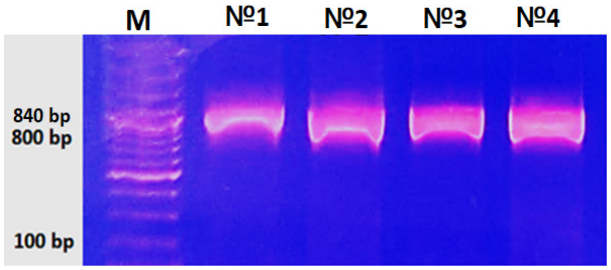
Table 1
Representative DNA concentration and purity measurements for selected Kazakh fat-tailed coarse-wool sheep samples

№	ng/µl	A260/280	A260/230
1	10.6	1.07	0.14
2	9.8	2.44	0.80
3	12.7	1.55	0.67
4	13.8	2.74	0.92
5	10.7	2.62	0.92
6	13.5	2.44	0.62
7	12.6	1.38	0.86

Table 1 presents DNA concentration and purity measurements for seven representative samples from the initial collection of 80 animals. These seven samples were selected to illustrate the range of DNA quality observed in the extracted material and were not intended to represent the complete set of 80 animals. DNA quality was assessed using the A260/280 and A260/230 absorbance ratios together with DNA concentration measurements. Samples considered suitable for PCR amplification were those that produced a sufficient amount of DNA and yielded successful amplification of the expected mtDNA D-loop fragment. Following successful PCR amplification, samples producing clear amplicons of the expected size and adequate sequence quality were retained for sequencing. Ultimately, 10 samples generated sequences of sufficient quality for downstream analysis.

To assess the quantitative and qualitative integrity of the amplified PCR products, horizontal gel electrophoresis and spectrophotometric analysis were employed. The residual nucleotides and unbound primers in the samples were removed using the ExoSAP-IT reagent to ensure high-purity DNA templates suitable for sequencing.

Figure 1
Electrophoretic profile of the mitochondrial DNA (mtDNA) D-loop region of the Kazakh fat-tailed coarse-wool sheep breed. M – DNA marker; each fragment corresponds to 100 base pairs (bp)



Electrophoretic profile of PCR-amplified products of the mitochondrial DNA (mtDNA) D-loop region of the Kazakh fat-tailed coarse-wool sheep breed on a 2% agarose gel

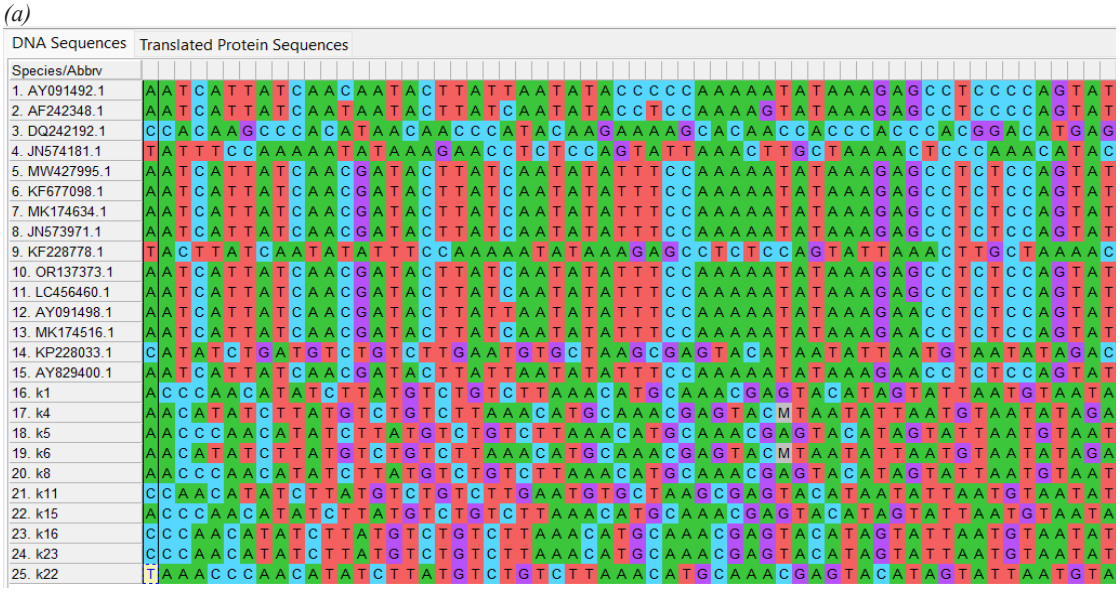
Lanes 1–4 represent the analyzed samples. M – DNA marker, where each fragment corresponds to 100 base pairs (bp). The PCR amplification of the D-loop region was performed using the following primer pairs: D-1: forward 5′-AACGGCAATAGCCCTTGA-3′ and reverse 5′-TGTACGTGTACGTGGGTGTA-3′;D-2: forward 5′-CCCGTCGCAGTCAAATCAA-3′ and reverse 5′-GTTGTATGCCGATTCAAAGGC-3′. The total size of the amplified fragment was approximately 840 bp. Sequencing analysis was conducted to determine the nucleotide composition of the amplified mtDNA D-loop region. The obtained nucleotide composition consisted of 31.46% adenine (A), 36.19% thymine (T), 16.42% guanine (G), and 15.93% cytosine (C). The 10 newly generated KFT-CWS mtDNA D-loop sequences were compared with 15 reference sequences retrieved from the NCBI GenBank database. The final comparative dataset therefore comprised 25 sequences: 10 sequences generated in the present study and 15 publicly available reference sequences representing domestic and wild sheep breeds. Multiple sequence

alignment was performed using the ClustalW algorithm implemented in MEGA X.

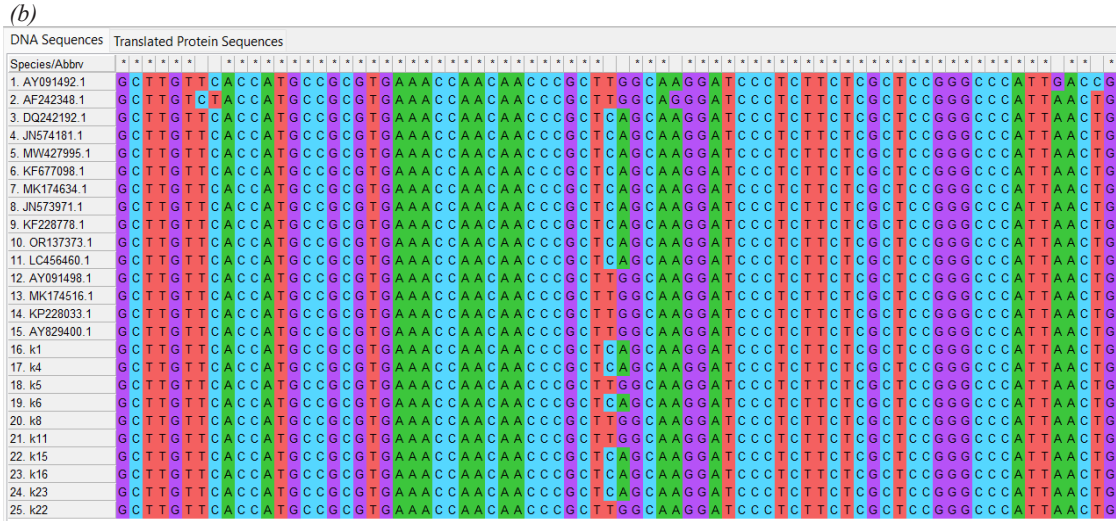
The reference sequences included Ovis ammon collium (AY091492.1), Ovis ammon darwini (AF242348.1), and domestic sheep accessions DQ242192.1, JN574181.1, MW427995.1, KF677098.1, MK174634.1, JN573971.1, KF228778.1, OR137373.1, LC456460.1, AY091498.1, MK174516.1, KP228033.1, AY829400.1, and 10 newly generated Kazakh fat-tailed coarse-wool sheep (KFT-CWS) sequences (K1–K10). Comparative sequence alignment revealed several single nucleotide polymorphisms (SNPs) among tandem repeats in the mtDNA D-loop region (Figure 2).

Figure 2
Alignment of nucleotide sequences of the mitochondrial DNA (mtDNA) D-loop region of the Kazakh fat-tailed coarse-wool sheep breed before (a) and after (b) processing using the ClustalW algorithm in the MEGA X software

(a)



(b)

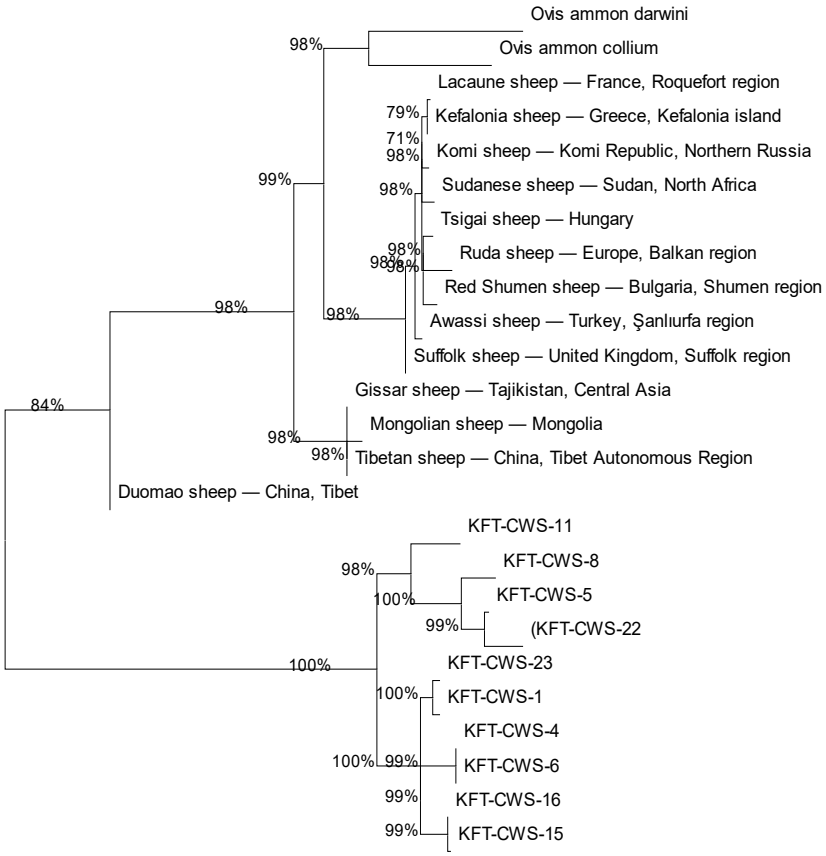


Reference Sequences Used for Phylogenetic Analysis *1. AY091492.1 – Ovis ammon collium (Wild sheep, Central Asia) 2. AF242348.1 – Ovis ammon darwini (Wild sheep, Central Asia) 3. DQ242192.1 – Komi sheep (Komi Republic, Northern Russia) 4. JN574181.1 – Ruda sheep (Europe, Balkan region) 5. MW427995.1 – Tsigai sheep (Hungary) 6. KF677098.1 – Awassi sheep (Turkey, Şanlıurfa region) 7. MK174634.1 – Suffolk sheep (United Kingdom, Suffolk region) 8. JN573971.1 – Kefalonia sheep (Greece, Kefalonia region) 9. KF228778.1 – Lacaune sheep (France, Roquefort region) 10. OR137373.1 – Red Shumen sheep (Bulgaria, Shumen region) 11. LC456460.1 – Sudanese sheep (Sudan, North Africa) 12. AY091498.1 – Gissar sheep (Tajikistan, Central Asia) 13. MK174516.1 – Tibetan sheep (China, Tibet Autonomous Region) 14. KP228033.1 – Duomao sheep (China, Tibet) 15. AY829400.1 – Mongolian sheep (Mongolia) 16–25. (k1–r22) – Kazakh fat-tailed coarse-wool sheep (Kazakhstan)*

The phylogenetic tree was constructed using the Neighbor-Joining (NJ) method implemented in the MEGA X software based on mitochondrial DNA (mtDNA) D-loop nucleotide sequences. The reliability of the inferred tree topology was assessed using bootstrap analysis with 1,000 replicates. According to the tree topology, the domestic and wild sheep sequences were divided into four major clades, with distinct sub-clusters representing individual sheep breeds (Figure 3).

Figure 3

Phylogenetic tree illustrating the evolutionary relationships among 25 sheep sequences, including 10 newly generated Kazakh fat-tailed coarse-wool sheep (KFT-CWS) sequences and 15 reference sequences retrieved from the NCBI GenBank database. The tree was constructed using the Neighbor-Joining method based on mitochondrial DNA (mtDNA) D-loop sequences. Bootstrap values calculated from 1,000 replicates are shown at the corresponding nodes



The first clade consisted of wild argali (*Ovis ammon darwini* and *Ovis ammon collium*), which were placed distantly from the domestic sheep lineages. These taxa served as the outgroup, representing the ancestral lineage of domestic sheep. This positioning also helped determine the root of the tree, showing a relatively high genetic distance (0.025–0.030) from the other clades. The second clade included European, African, and Western Asian sheep breeds–such as Kefalonia, Lacaune, Ruda, Sudanese, Tsigai, Red Shumen, Komi, Awassi, and Suffolk sheep–which were grouped into a single major branch with 100% bootstrap support, indicating a shared ancestral origin and close genetic relatedness among these populations. The third clade comprised the Central Asian breeds–Tibetan (China), Gissar (Tajikistan), and Mongolian (Mongolia) sheep–forming a cluster with 100% bootstrap support. The Duomao sheep (China, Tibet) were closely associated with this clade, with 85% bootstrap support, although they showed noticeable genetic differentiation from the other sequences within the clade. The fourth clade represented the 10 newly generated Kazakh fat-tailed coarse-wool sheep (KFT-CWS) sequences (K1–K10), which formed a distinct monophyletic cluster with strong bootstrap support of 98–100%.

This high bootstrap support indicates strong statistical support for the clustering of the KFT-CWS sequences. This clustering pattern strongly supports the breed's unique genetic diversity and confirms its status as an autochthonous (indigenous) sheep breed. Overall, the phylogenetic analysis demonstrated that the Kazakh fat-tailed coarse-wool sheep possess a distinct genetic differentiation while maintaining a close phylogenetic affinity with other Central Asian breeds such as Mongolian, Tibetan, and Gissar sheep. In contrast, they are genetically distant from European and African domestic sheep breeds, reflecting a long-term isolated evolutionary development within the Central Asian region.

Table 2

Results of the assessment of natural selection effects and demographic changes in sheep populations based on the mitochondrial DNA (mtDNA) D-loop region, evaluated using Tajima's Neutrality Test implemented in the MEGA X software

m	*S*	p_s	Θ	π	*D*
25	233	0.234406	0.062079	0.083974	1.408268

According to the results of Tajima's Neutrality Test, the calculated Tajima's D value was 1.408268, indicating a positive deviation from neutrality. This suggests the possible presence of balancing selection or population substructure (subdivision) within the studied population. Based on 25 sequences included in the comparative dataset of the mitochondrial D-loop region, a total of 233 polymorphic sites (S) were identified. Furthermore, the nucleotide diversity ($\pi = 0.083974$) was found to be higher than the theta ($\theta = 0.062079$) value, demonstrating a high level of genetic variation within the population (Table 2). These findings imply a deviation from the neutral mutation model, indicating that evolutionary forces such as selection or demographic expansion may have influenced the genetic structure of the studied sheep population.

Table 3

Maximum Likelihood matrix of nucleotide substitution probabilities calculated using the Tamura–Nei (1993) model in the MEGA 12 software

	A	T/U	C	G
A	-	*2.39*	*1.66*	12.55
T/U	*2.60*	-	19.00	*1.29*
C	*2.60*	27.24	-	*1.29*
G	25.35	*2.39*	*1.66*	-

Based on the Tamura–Nei (1993) evolutionary model, nucleotide substitution analysis of the mitochondrial D-loop region among the examined sheep breeds revealed an uneven distribution of transition and transversion probabilities. The highest substitution probabilities were observed between C→T/U (27.24) and G→A (25.35) transitions, demonstrating that purine–pyrimidine transitions occur much more frequently than transversions. The lowest substitution probabilities were recorded for G→C (1.66), T/U→G (1.29), and C→G (1.29) pairs. These results indicate that nucleotide substitutions contributing to sequence polymorphism occur at a higher rate, and the applied evolutionary model effectively accounts for such mutational biases (Table 3).

Table 4 presents the pairwise genetic distance matrix (Genetic Distance Matrix) among the studied sheep breeds, reflecting the degree of genetic similarity and divergence between populations. The analysis revealed varying levels of genetic differentiation among breeds, allowing an estimation of their evolutionary relationships. Comparatively, the Kazakh fat-tailed coarse-wool sheep (KFT-CWS) showed closer genetic similarity with certain Asian breeds, particularly Tibetan, Gissar, and Mongolian sheep. The genetic distance between KFT-CWS1 and Mongolian sheep was 0.1454, suggesting a shared ancestral origin. The Duomao sheep (China, Tibet) exhibited a very close relationship with Tibetan sheep (0.0008), indicating that they likely originated from a common selective lineage or population branch. The genetic distances between KFT-CWS and Duomao (0.0385), KFT-CWS and Gissar (0.0333), and KFT-CWS and Tibetan sheep (0.0321) were relatively small, supporting a close evolutionary relationship. Although the Duomao breed is genetically related to the Kazakh fat-tailed sheep, it remains slightly more distant compared to Gissar and Tibetan breeds. In contrast, European breeds such as Komi, Lacaune, and Suffolk showed considerably higher genetic distances, reflecting distinct evolutionary lineages and long-term isolation from Asian sheep populations.

Table 4

Pairwise genetic distance values among sheep breeds based on the mitochondrial DNA (mtDNA) D-loop region, calculated using the Pairwise Distances method in the MEGA X software.

	Ovis ammon collium	Ovis ammon darwini	Komi sheep	Ruda sheep	Tsigai sheep	Awassi sheep	Suffolk sheep	Kefalonia sheep	Lacaune sheep	Red Shumen sheep	Sudanese sheep	Gissar sheep	Tibetan sheep	Duomao sheep	Mongolian sheep	KFTCWS Kazakhstan
Ovis ammon collium	0															
Ovis ammon darwini	0,0520															
Komi sheep	0,0627	0,0719														
Ruda sheep	0,0638	0,0667	0,0083													
Tsigai sheep	0,0604	0,0632	0,0041	0,0069												
Awassi sheep	0,0576	0,0623	0,0027	0,0096	0,0059											
Suffolk sheep	0,0567	0,0595	0,0041	0,0087	0,0051	0,0042										
Kefalonia sheep	0,0567	0,0613	0,0027	0,0069	0,0033	0,0042	0,0033									
Lacaune sheep	0,0634	0,0730	0,0027	0,0088	0,0037	0,0037	0,0037	0,0000								
Red Shumen sheep	0,0576	0,0623	0,0055	0,0078	0,0042	0,0068	0,0059	0,0042	0,0049							
Sudanese sheep	0,0576	0,0604	0,0041	0,0078	0,0042	0,0068	0,0059	0,0042	0,0037	0,0051						
Gissar sheep	0,0558	0,0539	0,0474	0,0374	0,0356	0,0330	0,0303	0,0338	0,0445	0,0347	0,0329					
Tibetan sheep	0,0567	0,0530	0,0474	0,0374	0,0347	0,0321	0,0294	0,0330	0,0432	0,0338	0,0321	0,0008				
Duomao sheep	0,0553	0,0532	0,0448	0,0380	0,0348	0,0306	0,0286	0,0327	0,0447	0,0359	0,0317	0,0040	0,0040			
Mongolian sheep	0,0586	0,0567	0,0519	0,0402	0,0383	0,0356	0,0330	0,0365	0,0485	0,0374	0,0356	0,0025	0,0033	0,0070		
KFTCWS Kazakhstan	0,2469	0,2639	0,2218	0,2169	0,2169	0,2154	0,2169	0,2169	0,2904	0,2169	0,2169	0,2488	0,2488	0,1454	0,2472	0

Sheep Breeds Included in the Pairwise Genetic Distance Analysis: *Ovis ammon collium – Wild argali (Central Asia), Ovis ammon darwini – Wild argali (Central Asia), Komi sheep – Komi Republic, Northern Russia, Ruda sheep – Europe, Balkan region, Tsigai sheep – Hungary, Awassi sheep – Turkey, Şanlıurfa region, Suffolk sheep – United Kingdom, Suffolk region, Kefalonia sheep – Greece, Kefalonia island, Lacaune sheep – France, Roquefort region, Red Shumen sheep – Bulgaria, Shumen region, Sudanese sheep – Sudan, North Africa, Gissar sheep – Tajikistan, Central Asia, Tibetan sheep – China, Tibet Autonomous Region, Duomao sheep – China, Tibet, Mongolian sheep – Mongolia, Kazakh fat-tailed coarse-wool sheep (KFT-CWS) – Kazakhstan*

Conclusion

Mitochondrial DNA (mtDNA) is one of the most informative molecular markers for phylogenetic studies in sheep, as it allows for an accurate assessment of genetic diversity within populations and the elucidation of their evolutionary origins. In this study, the control region (D-loop, 840 bp) of mtDNA from the Kazakh fat-tailed coarse-wool sheep was analyzed to investigate its genetic diversity and phylogenetic relationships. Sequencing analysis revealed the nucleotide composition of the D-loop region as 31.46% adenine (A), 36.19% thymine (T), 16.42% guanine (G), and 15.93% cytosine (C). Comparative alignment of the obtained sequences with reference sheep breeds available in the NCBI database identified several single nucleotide polymorphisms (SNPs) among tandem repeats. The phylogenetic tree constructed based on the D-loop sequences clearly demonstrated the genetic diversity, geographic distribution, and evolutionary relatedness among sheep breeds, grouping them into four major clades. The first clade included wild argali (Ovis ammon darwini and Ovis ammon collium), which served as the outgroup with a high genetic distance (0.025–0.030) from domestic breeds. The second clade comprised European sheep breeds (Kefalonia, Lacaune, Ruda, Sudanese, Tsigai, Red Shumen, Komi, Awassi, Suffolk), indicating a close genetic relationship and a shared ancestry. The third clade included Central Asian breeds such as Tibetan (China), Gissar (Tajikistan), and Mongolian sheep, grouped with 100% bootstrap support. The Duomao sheep (China, Tibet) showed an 85% genetic affinity to this clade, suggesting partial divergence. The fourth clade consisted exclusively of the 10 newly generated KFT-CWS sequences, which formed a distinct cluster, indicating their genetic differentiation from the reference sheep populations included in the analysis. This clustering pattern indicates a clear genetic differentiation of the analyzed Kazakh fat-tailed coarse-wool sheep sequences from the other reference populations included in the dataset. Overall, the phylogenetic analysis demonstrated that the Kazakh fat-tailed coarse-wool sheep shares a close genetic relationship with Mongolian, Tibetan, and Gissar breeds, while remaining genetically distinct from European lineages. These findings highlight the independent evolutionary development of the Kazakh breed and its importance as a valuable genetic resource within Central Asia.

Author Contributions

Amzeyev R.E.: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft; Dossybayev K.Zh.: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft, Supervision; Kirikci K.: Methodology, Data curation; Yermekova M.Sh. – Investigation, Writing – review & editing; Nurkenov T.T.: Visualization.

References

Álvarez, I., Capote, J., Traoré, A., Fonseca, N., Pérez, K., Cuervo, M., Fernández, I., & Goyache, F. (2013). Mitochondrial analysis sheds light on the origin of hair sheep. Animal Genetics, 44(3), 344–347. https://doi.org/10.1111/j.1365-2052.2012.02398.x

Deniskova, T., Dotsev, A., Lushihina, E., Shakhin, A., Kunz, E., Medugorac, I., Reyer, H., Wimmers, K., Khayatzadeh, N., Sölkner, J., Sermyagin, A., Zhunushev, A., Brem, G., & Zinovieva, N. (2019). Population structure and genetic diversity of sheep breeds in Kyrgyzstan. Frontiers in Genetics, 10, 1311. https://doi.org/10.3389/fgene.2019.01311

Erokhin, A. I., Kotarev, V. I., & Erokhin, S. A. (2014). Ovcevodstvo [Sheep breeding]. Voronezh State Agrarian University.

Hiendleder, S., Kaupe, B., Wassmuth, R., & Janke, A. (2002). Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. Proceedings of the Royal Society B: Biological Sciences, 269, 893–904. https://doi.org/10.1098/rspb.2002.1975

Hiendleder, S., Lewalski, H., Wassmuth, R., & Janke, A. (1998). The complete mitochondrial DNA sequence of the domestic sheep (Ovis aries) and comparison with other major ovine haplotypes. Journal of Molecular Evolution, 47, 441–448. https://doi.org/10.1007/PL00006401

Kanapin, K. K. (2009). *Edilbaevskaya ovtsa* [Edilbay sheep]. Bastau.

Kumar, S., Stecher, G., Suleski, M., Sanderford, M., Sharma, S., & Tamura, K. (2024). Molecular evolutionary genetics analysis version 12 for adaptive and green computing. Molecular Biology and Evolution, 41, 1–9.

Marzanov, H. S., Nasibov, M. G., Zhiryakov, A. M., Klenovitskyi, P. M., Kaplinskaya, L. I., Marzanova, L. K., Ozerov, M. Y., & Marzanov, Y. S. (2004). Metodicheskie rekomendatsii po ispol'zovaniyu geneticheskikh markerov v razvedenii ovets [Methodological recommendations for the use of genetic markers in sheep breeding]. VGNII Zhivotnovodstva.

Meadows, J. R. S., Hiendleder, S., & Kijas, J. W. (2005). Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. Heredity, 94, 623–631. https://doi.org/10.1038/sj.hdy.6800658

Mukhametzharova, I. E. (2022). Materialy mezhdunarodnoy nauchno-prakticheskoy konferentsii "Seyfullinskie chteniya–18(2): Nauka XXI veka – epokha transformatsii" [Materials of the International Scientific and Practical Conference "Seifullin Readings–18(2): Science of the 21st Century – the Era of Transformation"]. Vol. 1, Part 2, pp. 156–158.

Osipov, Y. S. (Ed.). (2007). Bol'shaya Rossiyskaya entsiklopediya [Great Russian Encyclopedia] (Vol. 8). Bol'shaya Rossiyskaya Entsiklopediya.

Ozerov, M. Y. (2004). Kharakteristika allelofonda u razlichnykh porod ovets po mikrosatellitam [Characterization of the gene pool of different sheep breeds using microsatellites] [Candidate of Biological Sciences dissertation, All-Russian Research Institute of Animal Husbandry].

Ozerov, M. Y., Tapio, M., Marzanov, N. S., Nasibov, M. G., Shaidullin, I. N., & Kantanen, J. (2006). Mikrosatellitnyy analiz evolyutsionno-geneticheskikh svyazey u razlichnykh porod ovets [Microsatellite analysis of evolutionary-genetic relationships in different sheep breeds]. Doklady Rossiyskoy Akademii Sel'skokhozyaystvennykh Nauk, (2), 30–33.

Pedrosa, S., Uzun, M., Arranz, J. J., Gutiérrez-Gil, B., San Primitivo, F., & Bayón, Y. (2005). Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. Proceedings of the Royal Society B: Biological Sciences, 272(1577), 2211–2217.

Pozharskiy, A., Khamzina, A., Gritsenko, D., Khamzina, Z., Kassymbekova, S., Karimov, N., Karymsakov, T., & Tlevlesov, N. (2020). SNP genotyping and population analysis of five indigenous Kazakh sheep breeds. Livestock Science, 241, 104252. https://doi.org/10.1016/j.livsci.2020.104252

Tajima, F. (1989). Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics, 123, 585–595. https://doi.org/10.1093/genetics/123.3.585

Wood, N. J., & Phua, S. H. (1996). Variation in the control region sequence of the sheep mitochondrial genome. Animal Genetics, 27(1), 25–33. https://doi.org/10.1111/j.1365-2052.1996.tb01173.x

Information about the authors:

Amzeev Rauan Esengeldievich – Lecturer, State-Owned Enterprise "Gymnasium No. 159 named after Y. Altynsarin" (Almaty, Kazakhstan, e-mail: rauanbiolog@gmail.com).

Yermekova Marina Shamshadinovna – Master of Natural Sciences, Junior Researcher, Laboratory of Molecular Genetic Expertise, Limited Liability Partnership "Kazakh Research Institute of Livestock and Fodder Production", PhD student (2nd year) at Al-Farabi Kazakh National University, (Almaty, Kazakhstan, e-mail: nurii_90@mail.ru).

Koray Kırıkçı – Doctor of Agricultural Sciences, Associate Professor, Kırşehir Ahi Evran University, Faculty of Agriculture, Department of Animal Science, Animal Breeding and Improvement (Kırşehir, Türkiye, e-mail: koray.kirikci@ahievran.edu.tr).

Dossybayev Kairat Zhumagaliyevich – PhD, Senior Researcher, Head of the Laboratory of Molecular Genetic Expertise, Limited Liability Partnership "Kazakh Research Institute of Livestock and Fodder Production", Senior Lecturer at Al-Farabi Kazakh National University, (Almaty, Kazakhstan, e-mail: kairat1987_11@mail.ru).

Nurkenov Tulendy Tuleshevich – PhD, Senior Lecturer, Abai Kazakh National Pedagogical University, (Almaty, Kazakhstan, e-mail: t.nurkenov@abaiuniversity.edu.kz).

Авторлар туралы мәлімет:

Амзеев Рауан Есенгелдиевич – оқытушы, ШЖҚ "Ы.Алтынсарин атындағы №159 гимназия, (Алматы, Қазақстан, e-mail: rauanbiolog@gmail.com).

Ермекова Марина Шамшадиновна – жаратылыстану ғылымдарының магистрі, «Қазақ мал шаруашылығы және жемшөп өндірісі ғылыми-зерттеу институты» ЖШС, Молекулярлық-генетикалық сараптама зертханасының кіші ғылыми қызметкері, Әл-Фараби атындағы Қазақ ұлттық университетінің 2-курс докторанты (Алматы, Қазақстан, e-mail: nurii_90@mail.ru).

Корай Қырыкчи – ауыл шаруашылығы ғылымдарының докторы, доцент, Қыршехир Ахи Евран университеті, Ауыл шаруашылығы факультеті, Зоотехния кафедрасы, Жануарларды өсіру және селекциялау бағыты (Қыршехир, Түркия, e-mail: koray.kirikci@ahievran.edu.tr).

Досыбаев Қайрат Жумагалиевич – PhD аға ғылыми қызметкер, «Қазақ мал шаруашылығы және жемшөп өндірісі ғылыми-зерттеу институты» ЖШС, Молекулярлық-генетикалық сараптама зертханасының меңгерушісі, Әл-Фараби атындағы Қазақ ұлттық университетінің аға оқытушысы, (Алматы, Қазақстан, e-mail: kairat1987_11@mail.ru).

Нуркенов Туленды Тулешевич – PhD, аға оқытушы, Абай атындағы Қазақ ұлттық педагогикалық университеті (Алматы, Қазақстан, e-mail: t.nurkenov@abaiuniversity.edu.kz).

Сведения об авторах:

Амзеев Рауан Есенгелдиевич – преподаватель, ГКП «Гимназия № 159 имени Ы. Алтынсарина» (Алматы, Казахстан, e-mail: rauanbiolog@gmail.com).

Ермекова Марина Шамшадиновна – магистр естественных наук, младший научный сотрудник лаборатории молекулярно-генетической экспертизы ТОО «Казахский научно-исследовательский институт животноводства и кормопроизводства», докторант 2 курса Казахского национального университета имени аль-Фараби (Алматы, Казахстан, e-mail: nurii_90@mail.ru.)

Корай Қырыкчи – доктор сельскохозяйственных наук, доцент, Университет Кыршехир Ахи Евран, факультет сельского хозяйства, кафедра зоотехнии, направление «Разведение и селекция животных» (Кыршехир, Турция, e-mail: koray.kirikci@ahievran.edu.tr).

Досыбаев Қайрат Жумагалиевич – PhD, старший научный сотрудник, заведующий лабораторией молекулярно-генетической экспертизы ТОО «Казахский научно-исследовательский институт животноводства и кормопроизводства», старший преподаватель Казахского национального университета имени аль-Фараби (Алматы, Казахстан, e-mail: kairat1987_11@mail.ru).

Нуркенов Туленды Тулешевич – PhD, старший преподаватель Казахского национального педагогического университета имени Абая (Алматы, Казахстан, e-mail: t.nurkenov@abaiuniversity.edu.kz).

Received January 16, 2025
Re-uploaded August 30, 2026
Accepted September 15, 2026