https://doi.org/10.52973/rcfcv-e33262 Revista Científica, FCV-LUZ / Vol. XXXIII, rcfcv-e33262, 1 – 7

Phylogenetic and bottleneck analysis of the Turkish Arabian and Thoroughbred horse populations

Análisis filogenético y de cuello de botella de las poblaciones de caballos Turcos Árabes y de Pura sangre

Cevdet Yarali[1], Abdurrahman Köseman[2]*, Yusuf Özşensoy[3], İbrahim Şeker[4], Burhan Toprak[5], Kemal Zengin[1]

[1]*Veterinary Control Central Research Institute, Genetics Laboratory, Ankara, Türkiye.*
[2]*Malatya Turgut Ozal University, Akcadag Vocational School, Department of Crop and Animal Production. Malatya, Türkiye.*
[3] *Sivas Cumhuriyet University, Faculty of Veterinary Medicine, Department of Veterinary Genetics. Sivas, Türkiye.*
[4] *Firat University, Faculty of Veterinary Medicine, Department of Zootechny. Elazig, Türkiye.*
[5] *Yozgat Bozok University, Faculty of Veterinary Medicine, Department of Anatomy. Yozgat, Türkiye.*
Corresponding author: abdurrahman.koseman@ozal.edu.tr

ABSTRACT

This study aimed to determine the phylogenetic and bottleneck analyses of Turkish Thoroughbred and Arabian horse populations. In the study, genotyping was performed using a total of 17 microsatellite markers in the samples taken from 959 Thoroughbred and 813 Arabian horses. The average effective allele number in Arabian horses was 3.338 and the average number of alleles was 7.412 in Thorougbred horses. Genetic distance and genetic identity between Thoroughbred and Arabian breeds was 0.411 and 0.663, respectively. Also genetic identity in each Arabian and Thoroughbred horse populations was 1.000. The FCA table showed that the two breeds were completely separated from each other and were compatible. In conclusion; the bottlenecks of Thoroughbred and Arabian horse populations were in a normal L distribution and these horse breeds do not appear to have succumbed to introgression. Therefore, they are not yet at risk of extinction any time soon.

Key words: Arabian horse; bottleneck; microsatellite; phylogenetic analysis; thoroughbred horse

RESUMEN

Este estudio tuvo como objetivo determinar los análisis filogenéticos y de cuello de botella de las poblaciones de caballos árabes y de pura sangre turcos. En el estudio, el genotipado se realizó utilizando un total de 17 marcadores microsatélites en las muestras tomadas de 959 caballos pura sangre y 813 caballos árabes. El número medio de alelos efectivos en los caballos árabes fue de 3,338 y el número medio de alelos fue de 7,412 en los caballos pura sangre. La distancia genética y la identidad genética entre las razas Pura Sangre y Árabe fue de 0,411 y 0,663, respectivamente. También la identidad genética en cada población de caballos árabes y pura sangre fue de 1.000. La tabla FCA mostró que las dos razas estaban completamente separadas entre sí y eran compatibles. En conclusión; los cuellos de botella de las poblaciones de caballos pura sangre y árabes estaban en una distribución L normal y estas razas de caballos no parecen haber sucumbido a la introgresión. Por lo tanto, aún no están en riesgo de extinción en el corto plazo.

Palabras clave: Caballo árabe; cuello de botella; microsatellite; análisis filogenético; caballo pura sangre

Received: 25/04/2023 Accepted: 22/06/2023 Published: 22/07/2023

INTRODUCTION

Thoroughbred and Arabian horses constitute 4.2 and 15.1% of the racing and riding population, respectively. These two breeds account for 5.6 and 18.3% of parental lineage and 7.9 and 32.9% of founding lineage, respectively. On the species scale, both breeds account for altogether 26% of the founding origins [1].

The Thoroughbreds are one of the versatile horse breeds that influenced the development of many other breeds [2]. There are concerns about the loss of genetic diversity of this closed breed [3].

Genetic characterization studies have an important place in the determination of inter and intra breeds genetic diversity and identification of races. For this reason, these studies are carried out to protect the purity of breeds and to identify breeding strategies [4, 5]. Although the deoxyribonucleic acid (DNA) sequences lining the repeat regions in microsatellites are the same in individuals of a species, the number of repeats may vary between individuals and even between homologous chromosomes of individuals. Therefore, microsatellites are preferred in many molecular biology studies [5]. Microsatellite loci are more useful on the contrary classical genetic markers to determine the genetic relationships of close populations and to examine many polymorphic loci [6].

While the number of Thoroughbred horses running in the hippodromes in Türkiye in 2021 was 3734 (55%), the number of Arabian horses running was 3085 (45%). In the same year, the number of races organized exclusively for Thoroughbreds was 3246 (54%), while the number of races organized exclusively for Arabian horses was 2765 (46%) [7].

This research was aimed at studying phylogenetic and bottleneck analysis of the Turkish Arabian and Thoroughbred horse populations using microsatellite markers.

MATERIALS AND METHODS

Animals

All Arabian and Thoroughbred horses born in Turkiye or imported from any Country must be parentage tested by DNA STR analysis in an ISAG-approved laboratory for recording as a purebred horse to the Turkish Studbook.

This study's research material consisted of 959 Thoroughbred and 813 Arabian horses, which were applied to different Provincial or District Directorates of Agriculture and Forestry in Türkiye for their registration in the pedigree between 2006-2021.

Blood samples

The blood samples tested in the study were collected from 959 Thoroughbred and 813 Arabian horses, which were sent to Ankara Veterinary Control Central Research Institute Genetics Laboratory by different Provincial or District Directorates of Agriculture and Forestry in Türkiye between 2006-2021 for parentage verification test.

The blood samples used in the genetic analyzes were collected by the official veterinarians using anticoagulant tripotassium ethylenediaminetetraacetic acid (K_3EDTA) tubes and delivered to the laboratory under the cold chain. During the blood collection process, the horses' name, chip number, age, gender, breeder and address information and photographs were recorded in the database created on the computer. These processes were performed on IPhone 11 phone camera (Apple, USA) and Asus Notebook (X542UR-GQ436T Intel® Core i5-8250U, Asus, Taipei). All necessary transactions were carried out according to the provisions of the related Regulations of the Turkish Ministry of Agriculture and Forestry.

DNA Extraction

In the extraction of DNA from blood and some hair samples, both manual and automatic isolation methods were performed. DNA isolation from some of the blood samples was performed by Chelex 100 and Proteinase K using the manual extraction method. DNA was extracted from the remaining samples using automatic Roche Magna Pure LC, Qiagen Biorobot M48 and QIAsymphony workstations and kits according to the Roche Magna Pure LC kit, Qiagen M48 mini kit (Catalog No: 953336) and QIAsymphony DSP DNA Mini Kit (REF: 937236) protocols.

Microsatellite markers genotyping

In this study, 17 microsatellites for parentage testing were selected in Short tandem repeat (STR) analysis. Polymerase chain reaction (PCR) amplification of isolated genomic DNA samples was performed by the ABI StockMarks 17-plex Horse Genotyping kit (Cat. no. 4336405) according to manufacturer protocols. Capillary electrophoresis of samples was performed by ABI 3130 Genetic Analyzer (Applied Biosystems ®). STR alleles were detected by using the GeneMapper software (Version 4.0).

Statistical data analysis

General population parameters including allele numbers (Na), effective allele numbers (Ne) and Pairwise Population Matrix of Nei Genetic Distance and Identity were calculated using the GenAlEx6.501 program [8]. Factorial Correspondence Analysis (FCA) was performed using the GENETIX 4.05 package program [9] and the figure was drawn, which revealed the relationship between each subject with others in terms of family. Dendrograms (Neighbor-Joining Tree) were drawn to determine inter-breed genetic distance using the Populations 1.2.32 statistical program [10]. Moreover, it was tested if the population was at risk of extinction using the Bottleneck 1.2.02 program [11]. All these parameters were analyzed using the genotypic data recommended by Raymond and Rousset [12].

The official permission required to carry out the current study was taken from the General Directorate of Food and Control, The Ministry of Agriculture and Forestry (Date: 15/08/2022 and Issue: 6621366), and ethical approval was obtained from the Local Ethics Committee of Veterinary Control Central Research Institute (Date: 21/07/2022 and Issue: 2022/18).

RESULTS AND DISCUSSION

Allele numbers and frequencies

Allele numbers detected in Arabian and Thoroughbred horses were shown in TABLE I. A total of 145 different alleles (Na) were detected (mean 8.53), the highest Na were determined in Arabian horses (13 alleles) and Thoroughbreds (11 alleles) for ASB2. However, the lowest Na were determined as HTG6 and HTG7 (4 alleles) in Arabian horses; and were AHT4, HMS1, HMS6, HTG4, and HTG7 (5 alleles) in Thoroughbred horses.

In the study, the mean effective number of alleles (Ne) was 3.338 ranging from 1.477 (HTG7) to 4.790 (HMS3) in Arabian horses, and was 3.534 ranging from 1.836 (HMS2) to 6.456 (ASB2) in Thoroughbred horses. The mean Ne was smaller than the mean Na in Arabian horses (7.412) and Thoroughbred horses (6.529) (TABLE I).

TABLE I
Sample size (N), number of alleles (Na) and number of effective alleles (Ne) over all loci for Arabian and Thoroughbred horse populations

	Arabian horses			Thoroughbred horses			Total
Locus	**N**	**Na**	**Ne**	**N**	**Na**	**Ne**	**Na**
AHT4	813	9	3.615	959	5	3.475	9
AHT5	813	7	3.063	959	7	2.905	9
ASB2	812	13	3.176	958	11	6.456	14
ASB17	813	8	3.220	957	7	4.063	10
ASB23	813	6	3.753	957	8	4.353	9
CA425	813	6	3.908	959	8	2.197	8
HMS1	813	11	2.718	959	5	2.810	11
HMS2	813	7	2.630	959	6	1.836	9
HMS3	813	8	4.790	673	6	3.104	8
HMS6	813	6	3.083	959	5	2.583	6
HMS7	813	6	3.654	958	6	4.761	6
HTG4	813	5	3.323	959	5	2.144	5
HTG6	813	4	2.161	959	6	2.664	7
HTG7	813	4	1.477	959	5	2.712	6
HTG10	808	8	4.280	956	8	5.358	9
LEX3	428	8	4.107	555	7	4.713	8
VHL20	813	10	3.793	959	6	3.939	11
Mean	790	7.412	3.338	917.882	6.529	3.534	8.53

Almarzook *et al.* [13] detected a total of 251 polymorphic Na in Arabian horses ranging from 3 (HTG4 and HTG7) to 11 (ASB17 and LEX3), and also overall mean of Ne was 4.141. Machmoum *et al.* [14] concluded that the mean Na per locus was 6.52, 6.35, and 7 in the Desert breed, Straight Egyptian, and Polish Arabian, respectively. Cosenza *et al.* [15] reported that Na had 118 alleles (mean 7.375) per locus ranging from 5 (HMS1) to 10 (HMS2) and also, Ne had 3.385 alleles per locus ranging from 1.723 (HMS2) to 5.308 (ASB2) in the Thoroughbred horses.

The Na and Ne are one of the main indicators of genetic diversity, also Ne indicates a higher probability of extinction. Even if the values we determined in our work differed significantly from those studies because of population structure, horse numbers, and/or the selected microsatellite panel; all of the markers in the current study have a credible polymorphism to assess genetic variation in Arabian and Thoroughbred horse populations.

Phylogenetic analysis

Genetic distance and genetic identity matrix in Arabian and Thoroughbred horse populations was shown in TABLE II [16], and the genetic relationship of the two breeds to each other was shown in FIG 1. In addition, Neighborind Joining Tree (NJT) of Arabian and Thorougbred bred horses was drawn separetely (FIG 2 and FIG 3). Grouping values that ensured reliability with the bootstrap test (1,000 repetitions) were shown on the NJT. It is seen that the Arabian and Thoroughbred horse populations are divided into numerous groups by phylogenetic trees.

Genetic distance and genetic identity between Thoroughbred and Arabian breeds was found 0.411 and 0.663, respectively, in the pairwise population matrix. Also genetic identity in each Arabian and Thoroughbred horse populations was found 1.000 (TABLE II). Genetic identity and genetic distance are opposite each other. A high genetic identity creates a low genetic distance. Accordingly, it can be said that there is a moderate genetic identity and genetic distance between the Arabian and Thoroughbred horse populations (between breeds). A high genetic identity creates a low genetic distance. Accordingly, it can be said that there is a moderate genetic identity and genetic distance between Arabian and Thoroughbred horse populations. However, there is no genetic distance in Arabian and also Thoroughbred horses themselves.

TABLE II
Genetic distance and genetic identity matrix in Arabian and Thoroughbred horse populations

Breed	**Arabian horses**		**Thoroughbred horses**	
	Genetic distance	**Genetic identity**	**Genetic distance**	**Genetic identity**
Arabian horses	0.000	1.000	–	–
Thoroughbred horses	0.411	0.663	0.000	1.000

In the FCA for Arabian and Thoroughbred horses observed that one group was concentrated in the middle, and the others were located apart around its consistent with the NJT chart. It was seen from the FCA chart the two breeds were completely separated from each other and compatible. According to the results, these horse breeds or populations seem not submitted to introgression.

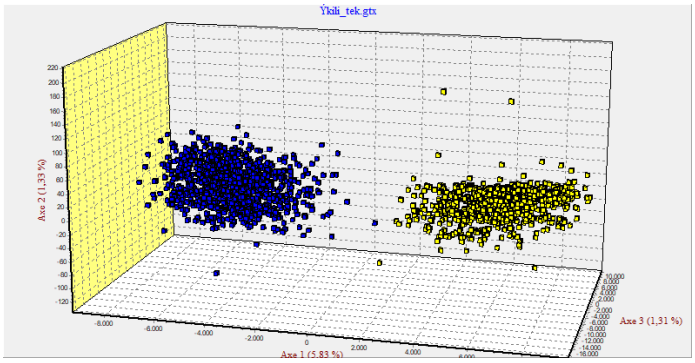

FIGURE 1. FCA chart of the Arabian and Thoroughbred horses (Yellow: Arabian horses, Blue: Thoroughbred horses)

Mahrous *et al.* [17] reported that the highest genetic distance was found to be 0.4405 in Thoroughbred and Native breeds, and the lowest genetic distance was found to be 0.2586 in Native and Arabian horses. The values determined in our research are lower than these values.

According to the NJT analysis, Jung *et al.* [18] determined that Halla horses were between Thoroughbred and Jeju horses and were different the Thoroughbred and Jeju horses. Almarzook *et al.* [13] reported that the three breeds were distributed all over the adjacent conjugation tree without clear distinction and also in the median–

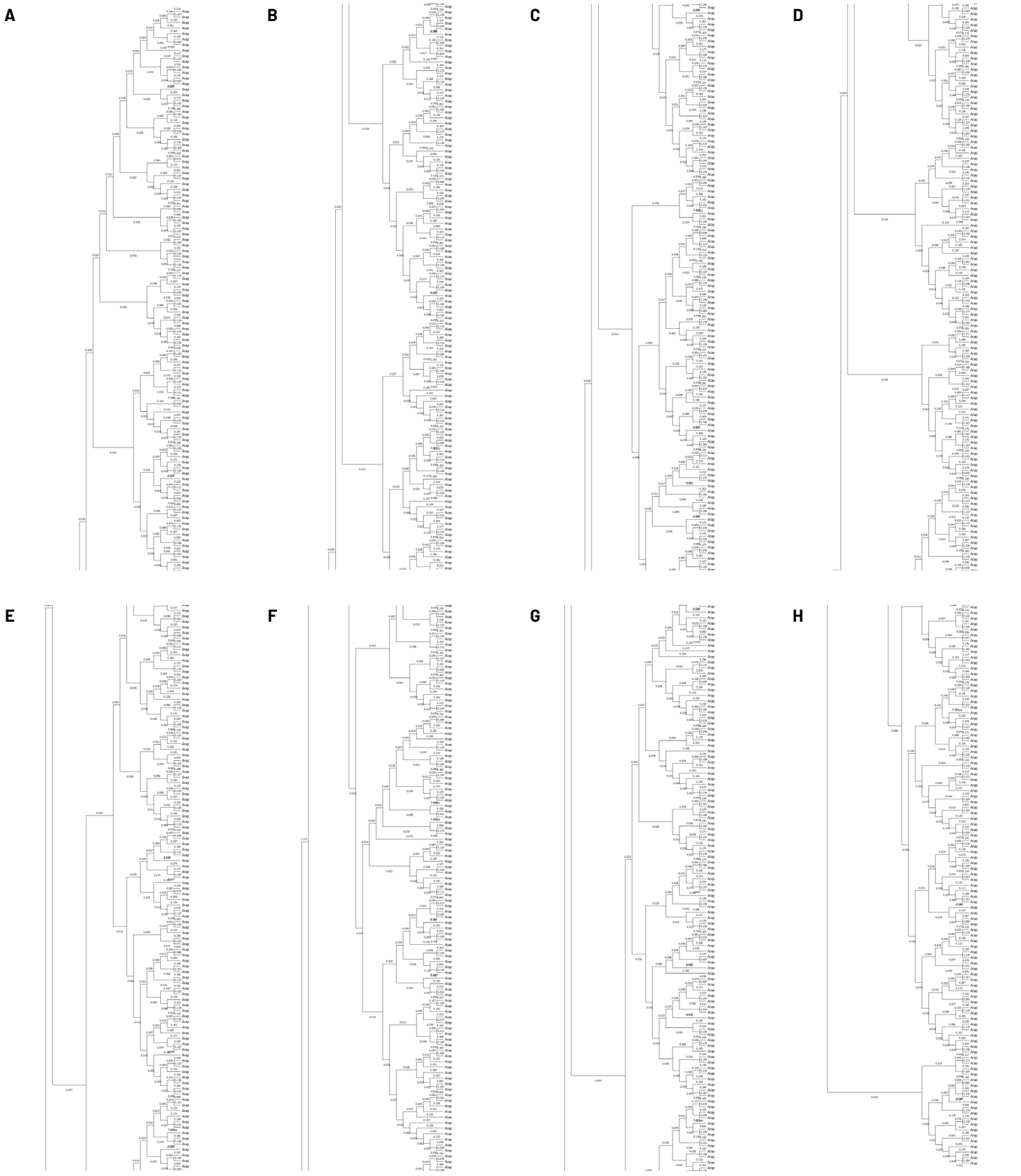

FIGURE 2. (A-H) NJT tree generated with STRs to show bootstrap values of Arabian horse genotype (n=813)

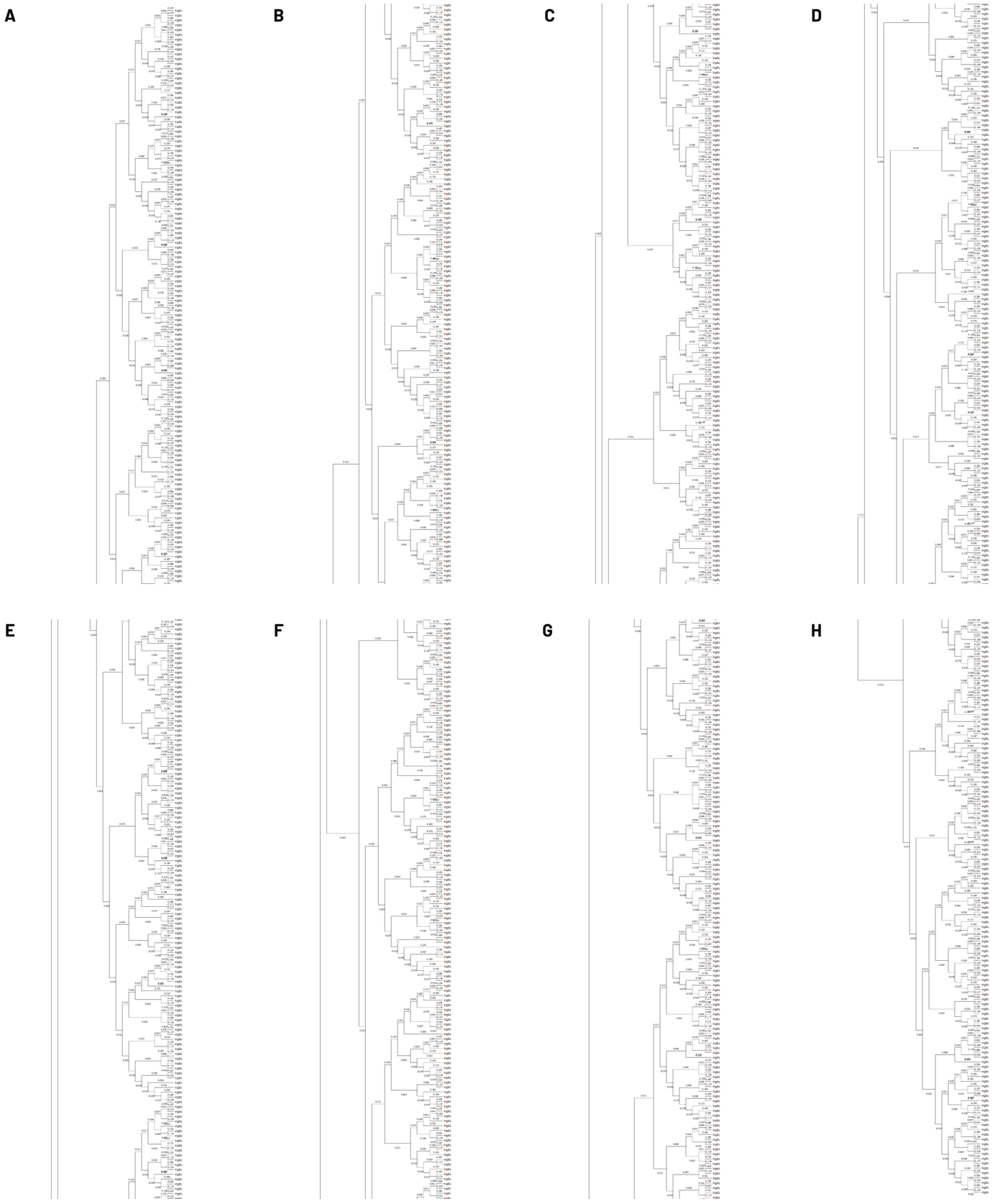

FIGURE 3. (A-H) NJT tree generated with STRs to show bootstrap values of Thoroughbred horse genotype (n=959)

joining network, the Syrian horses were grouped into seven major haplogroups (FIG 2 and FIG 3).

Cosgrove *et al.* [19] stated that there was a high degree of genetic variation and complex ancestry in Arabian horses from the Middle East region. Moreover, in this work it was concluded that powerful verification the interbreeding of Thoroughbreds with Arabians but opposite to popular creed, no significant genomic bequest of the Arabian breed to the Thoroughbred horses.

Limiting inbreeding in the horse population will reduce the effect of undesirable hereditary traits.

Bottleneck analysis

The analyses of bottleneck risk in Arabian and Thoroughbred horses are given in FIGS. 4 and 5. The bottleneck of the Arabian and Thoroughbred horse populations was at a normal L distribution. Duru [20] also reported that the Arabian horse breed might not be at risk. In a study that analyzed the relationship between inbreeding and racing performance, it was found that selective breeding did not effectively alleviate the genetic load of the Australian Thoroughbred population [21].

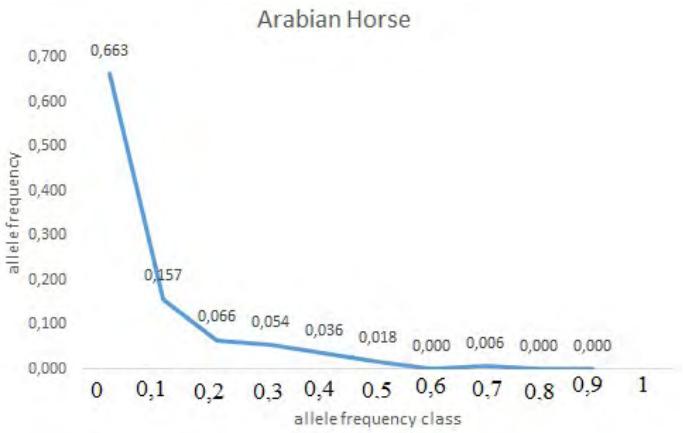

FIGURE 4. Bottleneck chart of the Arabian horse population

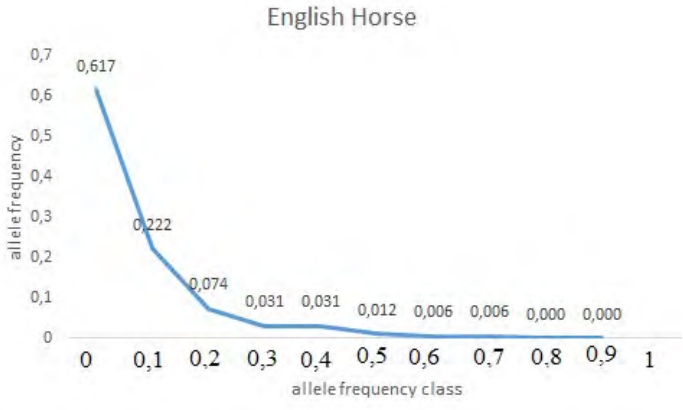

FIGURE 5. Bottleneck chart of the Thoroughbred horse population

In a study of Małopolski horses (Anglo–Arabian x oriental horse breeds) using 18 microsatellites, it was determined that there was no danger of inbreeding, and Thoroughbreds were the largest contributor to the gene pool of the current Małopolski population [22]. But, in a study on sport horses it was concluded that if the population was under bottleneck effect, the genetic variation reduction was observed in the non–founder generations [23]. Moreover, Koseman *et al.* [24] stated that the bottleneck assessment of the colored horse population was at a normal L distribution and there was no risk of extinction in the population.

The findings show that, despite the similarities and contrasts in the literature, there is no bottleneck risk yet in the Arabian and Thoroughbred populations we have studied

CONCLUSION

The results provided decent data about the sophisticated genetic structure to understand the efficacy of phylogenetic and bottleneck. Moreover, it underlined the availability and great utilility of the Arabian and Thoroughbred horse gene stock at least at Türkiye scale. For this reason the results might give disclosure and suggestion concerning Arabian and Thoroughbred horse breeds upkeep. In this context, the following results are highlighted:

- Arabian and Thoroughbred horses have an important genetic potential in Türkiye.
- There is very high genetic variability in Arabian and Thoroughbred horse populations.
- Both genera showing a normal "L distribution" in the phylogenetic tree are divided into many groups within themselves.
- Neither horse breed population is at risk of extinction any time soon.

ACKNOWLEDGMENTS

We would like to thank the Republic of Türkiye Ministry of Agriculture and Forestry for its cooperation in providing data.

Conflict of interest statement

The authors declare there is no conflict of interest.

BIBLIOGRAPHIC REFERENCES

[1] Pirault P, Danvy S, Verrier E, Leroy G. Genetic structure and gene flows within horses: a genealogical study at the french population scale. PLoS One. 2013; 22(4):e61544.

[2] Thiruvenkadan AK, Kandasamy N, Panneerselvam S. Inheritance of racing performance of Thoroughbred horses. Livest. Sci. 2009; 121:308–326.

[3] Cunningham EP, Dooley JJ, Splan RK, Bradley DG. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to Thoroughbred horses. Anim. Genet. 2001; 32(6):360–4.

[4] Buróczıová M, Řıha J. Horse breed discrimination using machine learning methods. J. Appl. Genet. 2009; 50(4):375–377.

[5] Özşensoy Y, Kurar E. Markör sistemleri ve genetik karakterizasyon çalışmalarında kullanımları. J. Cell Mol. Biol. 2012; 10(2):11–19.

[6] Takezaki N, Nei M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genet. 1996; 144(1):389–99.

[7] General Statistics [Internet]. Istanbul: Türkiye Jockey Club. 2022 [Accessed 19 March 2023]. Available in: https://bit.ly/3rfOk2D.

[8] Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioin App Note. 2012; 28(19):2537–2539.

[9] Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations [Internet]. Montpellier: Université Montpellier II; 2004 [Accessed 19 March 2023]. Available in: https://bit.ly/3pCQzMQ.

[10] Langella O. Populations 1.2.32 population genetic software [Internet]. 2011; [Accessed 19 March 2023]. Available in: https://bit.ly/3XJqIoG.

[11] Piry S, Luikard G, Cornuet JM. Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. J. Hered. 1999; 90(4):502–503.

[12] Raymond M, Rousset F. Genepop (version 1.2): population–genetics software for exact tests and ecumenicism. 1995; J. Hered. 86:248–9.

[13] Almarzook S, Abdel–Shafy H, Ahmed AS, Reissmann M, Brockmann GA. Genetic diversity of Arabian horses using microsatellite markers. Egyptian J. Anim. Prod. 2022; 59(1):19–27.

[14] Machmoum M, Boujenane I, Azelhak R, Badaoui B, Petit D, Piro M. Genetic diversity and population structure of Arabian horse populations using microsatellite markers. J. Equine Vet. Sci. 2020; 93:103200.

[15] Cosenza M, La Rosa V, Rosati R, Chiofalo V. Genetic diversity of the Italian Thoroughbred horse population. Ital. J. Anim. Sci. 2019; 18(1):538–545.

[16] Nei M, Kumar S. Molecular evolution and phylogenetics. New York: Oxford University Press. 2000; 88 p.

[17] Mahrous KF, Hassanane M, Abdel–Mordy M, Heba I, Shafey HI, Hassan N. Genetic variations in horse using microsatellite markers. JGEB. 2011; 9(2):103–109.

[18] Jung JS, Seong J, Lee GH, Kim Y, An JH, Yun JH, Kong HS. Genetic diversity and relationship of Halla horse based on polymorphisms in microsatellites. J. Anim. Sci. Biotechnol. 2021; 36:76–81.

[19] Cosgrove EJ, Sadeghi R, Schlamp F, Holl HM, Moradi–Shahrbabak M, Miraei–Ashtiani SR, Abdalla S, Shykiind B, Troedsson M, Stefaniuk–Szmukier M, Prabhu A, Bucca S, Bugno–Poniewierska M, Wallner B, Malek J, Miller DC, Clark AG, Antczak DF, Brooks SA. Genome diversity and the origin of the Arabian horse. Sci Rep. 2020; 10(1):9702.

[20] Duru, S. Pedigree analysis of the Turkish Arab horse population: structure, inbreeding and genetic variability. Anim. 2017; 11(9):1449–1456.

[21] Todd ET, Ho SY, Thomson PC, Ang RA, Velie BD, Hamilton Na. Founder–specific inbreeding depression affects racing performance in Thoroughbred horses. Sci. Rep. 2018; 8(1):6167.

[22] Zabek T, Zyga A, Radko A, Słota E. Analysis of genetic variation in Małopolski horses using molecular and pedigree data. Nat. Res. Instit. Anim. Prod. 2006; 6(1):13–27.

[23] Próchniak T, Kasperek K, Knaga S, Rozempolska–Rucińska I, Batkowska J, Drabik K, Zięba G. Pedigree analysis of Warmblood horses participating in competitions for young horses. Front. Genet. 2021; 12:658403.

[24] Koseman A, Ozsensoy Y, Erdogan M, Yarali Y, Toprak B, Zengin K, Seker I. Investigation of genetic variations using microsatellite markers in Colored horses in Turkey. Russ. J. Genet. 2020; 56 (5):592–602.