



ORIGINAL ARTICLE

Mitochondrial DNA genetic variations among four horse populations in Egypt



Othman E. Othman *, Karima F. Mahrous, Heba I. Shafey

Cell Biology Department, National Research Centre, Dokki, Egypt

Received 9 January 2017; revised 7 May 2017; accepted 10 June 2017

Available online 6 July 2017

KEYWORDS

MtDNA;
Genetic variations;
Horse populations;
Arabian;
Thoroughbred;
Baladi;
Sports

Abstract Horses are one of the early domesticated animals in the world that changed societies and civilizations on a continent-wide scale. Due to the rare information about the genetic characterization of different horse populations in Egypt, this study aimed to identify the genetic biodiversity and relationships between four horse populations reared in Egypt. Genomic DNA was extracted and mtDNA region was amplified using polymerase chain reaction (PCR). The alignment of 384-bp amplified fragments showed the presence of 41 polymorphic sites resulting in 29 haplotypes which their sequences were submitted to GenBank under the accession numbers: **KX909898-KX909926**.

The phylogeny tree for tested horses declared the presence of mixing maternal lineages between the four tested populations but still there are some separated lineages especially for Arabian and Thoroughbred horses. The sequences of 72 tested sequences were aligned with 13 published sequences as references, 11 of them for different *Equus caballus* whereas the other two reference sequences for *Equus burchellii* and *Equus asinus*. The results showed that all tested horses from the four populations are grouped with reference sequences of *Equus caballus* and separated from the other two reference sequences of *Equus burchellii* and *Equus asinus*.

It is concluded that sequence analysis of mtDNA control region is still the most informative tool for the identification of genetic biodiversity and phylogeny of different horse breeds and populations. The horse populations reared in Egypt possess low genetic diversity and all of them are belonged to *Equus caballus* breed.

© 2017 Production and hosting by Elsevier B.V. on behalf of Academy of Scientific Research & Technology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Animal and plant domestication played an important role in the appearance of modern civilizations. The identification of

origin and phylogenetic relationships between animals is a key for understanding of cultural changes in these civilizations and consequently has public and scientific attractions [1,20]. Horses are considered one of the early domesticated animals in the world [21] that changed societies and civilizations on a continent-wide scale.

The genetic history was identified using three different genetic sources nuclear genome, Y chromosome and mitochondrial genome [26]. Mitochondrial DNA is maternal

* Corresponding author at: Cell Biology Department, National Research Centre, Dokki, Giza, Egypt. Fax: +20 2 33370931.
E-mail address: othmanmah@yahoo.com (O.E. Othman).
Peer review under responsibility of National Research Center, Egypt.

inherited, so all individuals within a maternal lineage are shared together in mtDNA genome [14]. Recently, mtDNA sequencing is used at a wide research scale concerning with molecular biodiversity, phylogeny and the identification of modern species origin [9,29].

MtDNA D-loop region that is the most variable part of mtDNA [15], so it is considered the most informative genetic materials for phylogenetic relationship studies within and between different species [28]. MtDNA sequences were used to identify the genetic biodiversity and relationships of horses within breeds [13,24], between breeds [18,28] and the origins of different wild and domestic horse breeds [8,10].

There are different horse breeds and populations around the world used for many purposes such as sports; for eventing, jump humping and racing competitions. Also, horses attained a prominent role as animals of transport and warfare in some societies [5]. Many reports discussed the fact that the Arabian horse breed is considered the oldest horse breeds and distributed all over the world [8,17]. The Arabian horses are used for genetical improvement of other horse breeds for enhancing their morphology and movement characteristics [3,32]. The information about the genetic characterization and variation between different horse populations in Egypt is still rare, so this study aimed to identify the genetic biodiversity and relationships between four horse populations reared in Egypt using mtDNA control region sequences.

2. Materials and methods

2.1. Blood samples and genomic DNA extraction

Blood samples were collected from 72 animals belonging to four horse populations reared in Egypt: 18 Arabian, 16 Baladi, 19 Sports and 19 Thoroughbred horses. Genomic DNA was extracted from the whole blood according to the method described by Miller et al. [27] with minor modifications. Briefly, blood samples were mixed with cold 2× sucrose-triton and centrifuged at 5000 rpm for 15 min at 4 °C. The nuclear pellet was suspended in lysis buffer, sodium dodecyl sulfate and proteinase K and incubated overnight in a shaking water bath at 37 °C. Nucleic acids were extracted with saturated NaCl solution. The DNA was picked up and washed in 70% ethanol. The DNA was dissolved in 1×TE buffer. The DNA concentration was determined, using Nano Drop1000 thermo scientific spectrophotometer and then diluted to the working concentration of 50 ng/μl.

2.2. Polymerase chain reaction (PCR)

The PCR amplifications were conducted in a 50 μL volume containing 5 μL of 10× reaction buffer, 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.2 μM each primer, 1.5U *Taq* DNA polymerase (Fermentas, Germany) and approximately 100 ng genomic DNA. The reaction was cycled for 1 min at 95 °C for 5 min, followed by 30 cycles: each consisting of 1 min at 94 °C for denaturation, 1 min at 55 °C for annealing, 2 min at 72 °C for extension and then a final extension for 10 min at 72 °C. The PCR products were electrophoresed on 2% agarose gel stained with ethidium bromide to test the amplification success. The PCR products were purified with a DNA purification kit (ExoSap-IT, USB Corporation). Sequencing was

performed by MacroGen Company (Seoul, South Korea). The primer sequences used in this study were designed according to published horse mtDNA [31]

Forward: 5'- CGC ACA TTA CCC TGG TCT TG -3'

Reverse: 5'- GAA CCA GAT GCC AGG TAT AG -3'.

2.3. Data analysis

MtDNA D-loop sequences were aligned using the BioEdit software [12] in order to identify and trace individual haplotype mutations. Haplotype structure, sequence variation, average number of nucleotide differences (D) and average number of nucleotide substitutions (Dxy) per site between populations were calculated using DnaSP 5.00 software [22]. Neighbor-joining (NJ) tree for all samples was constructed using Mega version 5.0 software [30].

3. Results and discussion

In Egypt, there are different horse populations used for many purposes, these populations include Arabian, Sport, Thoroughbred and Baladi horses. Sport horses are characterized by specific qualities in the conformation and movement making them suitable for eventing and jump humping. Thoroughbred horses are commonly used as sport horses and also for racing competitions. They are crossbred between England mares with Arabian and Turkoman stallions. The Baladi or Egyptian horse originally hailed from Egypt and due to the cross proximity of the Arab horses, there were significant genetic changes in Baladi horses. These animals are used for pleasure riding, racing and showing [6,8,25].

In spite of there are some studies dealing with the mtDNA genetic variations between horses from different breeds, countries and geographic regions [19], still there is some missing data about this issue concerning with the variations among some horse populations belonging to the same breed and present in the same geographical basins [23]. Due to this fact, the present work aimed to identify mtDNA biodiversity and genetic relationships between four different horse breeds reared in Egypt for helping in their genetic conservation.

The fragments at 509-bp from horse mtDNA were amplified using specific primers for horse mtDNA (Fig. 1). The PCR products were purified, sequenced and the 384-bp fragments represented the control region of mtDNA were aligned. The alignment of 72 sequences from Arabian, Baladi,

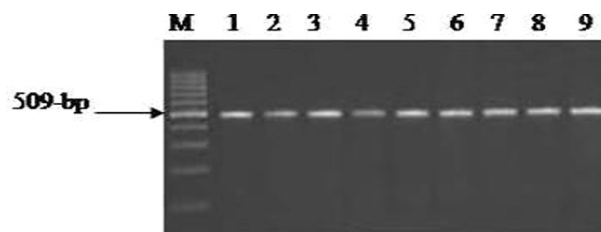


Figure 1 Ethidium bromide-stained gel of PCR products representing the amplified horse mtDNA. Lane M: 100-bp ladder marker. Lanes 1–9: 509-bp PCR products amplified from horse mtDNA.

Thoroughbred and Sports horses was done using BioEdit software. DnaSP 5.00 software was used to identify the sequence variation and polymorphic sites in the aligned sequences.

The result showed that the presence of 41 variation or polymorphic sites and the formation of 29 haplotypes (Fig. 2). The nucleotide sequences of these haplotypes were submitted to GenBank under the accession numbers: **KX909898-KX909926**. The most common haplotype appears in 8 sequences, all of them from Arabian horses. Also, there is one haplotype appears in 7 sequences (6 from Thoroughbred and one from Sport horses) and one haplotype appears in 6 sequences (3 from Thoroughbred, 2 from Arabian and one from Sports horses). There are 15 haplotypes; each of them appears in one sequence, 6 haplotypes in 2 sequences, one haplotype in 4 sequences and 4 haplotypes in 5 sequences. Regarding to the specific haplotypes, there are 4 haplotypes for each of Arabian and Thoroughbred horses, 7 haplotypes for Sport horses and 6 haplotypes for Baladi in addition to 8 haplotypes in which two or more different populations shared in them.

The statistical analysis of genetic diversity among 4 tested horse populations (Table 1) showed that the highest number

of haplotypes (14 haplotypes) was found in Sport horses where its 19 animals possessed 32 polymorphic sites. Ten haplotypes were present in both Baladi (16 animals with 28 polymorphic sites) and Thoroughbred horses (19 animals with 23 polymorphic sites). The lowest haplotype number (7 haplotypes) was appeared in Arabian horse population where its 18 animals possessed 23 polymorphic sites.

The haplotype diversity in four tested horse populations ranged from 0.7843 in Arabian horses (with average number of pairwise differences K: 7.01961) to 0.9650 in Sport horses (with K: 8.84211). Whereas the haplotypes diversity was 0.9333 in Baladi horses (with K: 8.57500) and 0.8830 in Thoroughbred horses (with K: 8.35088). Within all tested populations, the haplotype diversity and average number of pairwise differences were 0.9546 and 8.52543, respectively.

The result showed that the Arabian horses possessed lowest nucleotide diversity (0.01828) followed by nucleotide diversity of 0.02175 in Thoroughbred and 0.02233 in Baladi horses whereas the highest nucleotide diversity was 0.02303 in Sport horses with total nucleotide diversity of 0.02220 for all 4 tested populations.

```

111111111111111111222222233333333333
22267771222333345688989134557900003344455
26860467789345698712199851252823798912389
Hap_1: CTACACTGAAATTTGCACAATGAATCACACCTCATCAACAG
Hap_2: .....G...T.....G.C...T...GC.G...A
Hap_3: .....T...G...T.T.G.A..C.....C..C.....A
Hap_4: .CGT...A....C.T..G.....C.....A
Hap_5: .CGT...A....CAT..G.....C.....A
Hap_6: .....C..T..G.A.....C..C...GA
Hap_7: .....A.....A
Hap_8: .....C..T.....GT...C..C.....
Hap_9: .....A....C..T..G.A.....C..C..G.GA
Hap_10: .....G.....AT.....C.....GC...T.A
Hap_11: .....AT.....C...G...GC...T.A
Hap_12: .....G...T..G.....C..C...GA
Hap_13: .....T.A..G...T.T.G.A..C.....C..C.....A
Hap_14: .....C.....T.T.G..G.....T.CT...A
Hap_15: .....A....C..T..G.A.....C..C..G..A
Hap_16: .CGT.....C.T..G.....C.....A
Hap_17: T.....C..T..G.A.....C..CT...GA
Hap_18: .....C.....
Hap_19: ...G..A.....T..G...T.....C...GA
Hap_20: .....A....C..TG..G.A.....C..C..G.GA
Hap_21: .....A..C..T.....
Hap_22: .CGT...A....CAT..G.....C.....
Hap_23: ...G...G...T.....T.....C...GA
Hap_24: .....G..AT.....G.C.....GC.G...A
Hap_25: .....G...T..G.....TC..C...GA
Hap_26: .....C..T..G.A.....C..C.....A
Hap_27: .CGT...A....CC.T..G.....C.....A
Hap_28: .....C.....C.....
Hap_29: .....T.....T.T.G.A..C.....C..C.....A

```

Figure 2 Nucleotide sequences of 29 haplotypes. The variable sites' numbers are in red.

Table 1 The genetic diversity data.

Population	No. of sequences (N)	No. of polymorphic sites (S)	No. of haplotypes (H)	Haplotype diversity (HD)	Average number of pairwise differences (K)	Nucleotide diversity (π)
Arabian	18	23	7	0.7843	7.01961	0.01828
Baladi	16	28	10	0.9333	8.57500	0.02233
Sport	19	32	14	0.9650	8.84211	0.02303
Thoroughbred	19	23	10	0.8830	8.35088	0.02175
Total	72	41	29	0.9546	8.52543	0.02220

Table 2 Average pairwise differences between populations. Average number of nucleotide difference between populations D (below). Average number of nucleotide substitution per site between populations, Dxy (above).

Population	Arabian	Baladi	Sport	Thoroughbred
Arabian	–	0.02271	0.02210	0.02207
Baladi	8.722	–	0.02226	0.02244
Sport	8.485	8.549	–	0.02321
Thoroughbred	8.474	8.615	8.914	–

The genetic distances between four populations were estimated by average number of nucleotide difference between populations (D) and the average number of pairwise differences (Dxy). The lowest distance was observed between Arabian and Thoroughbred horses (D: 8.474 and Dxy: 0.02207) followed by distance between Arabian and Sport horses (D: 8.485 and Dxy: 0.02210) then distance between Baladi and Sport horses (D: 8.549 and Dxy: 0.02226) and distance between Baladi and Thoroughbred horses (D: 8.615 and Dxy: 0.02244), while the highest distance was observed between Sport and Thoroughbred horses (D: 8.914 and Dxy: 0.02321).

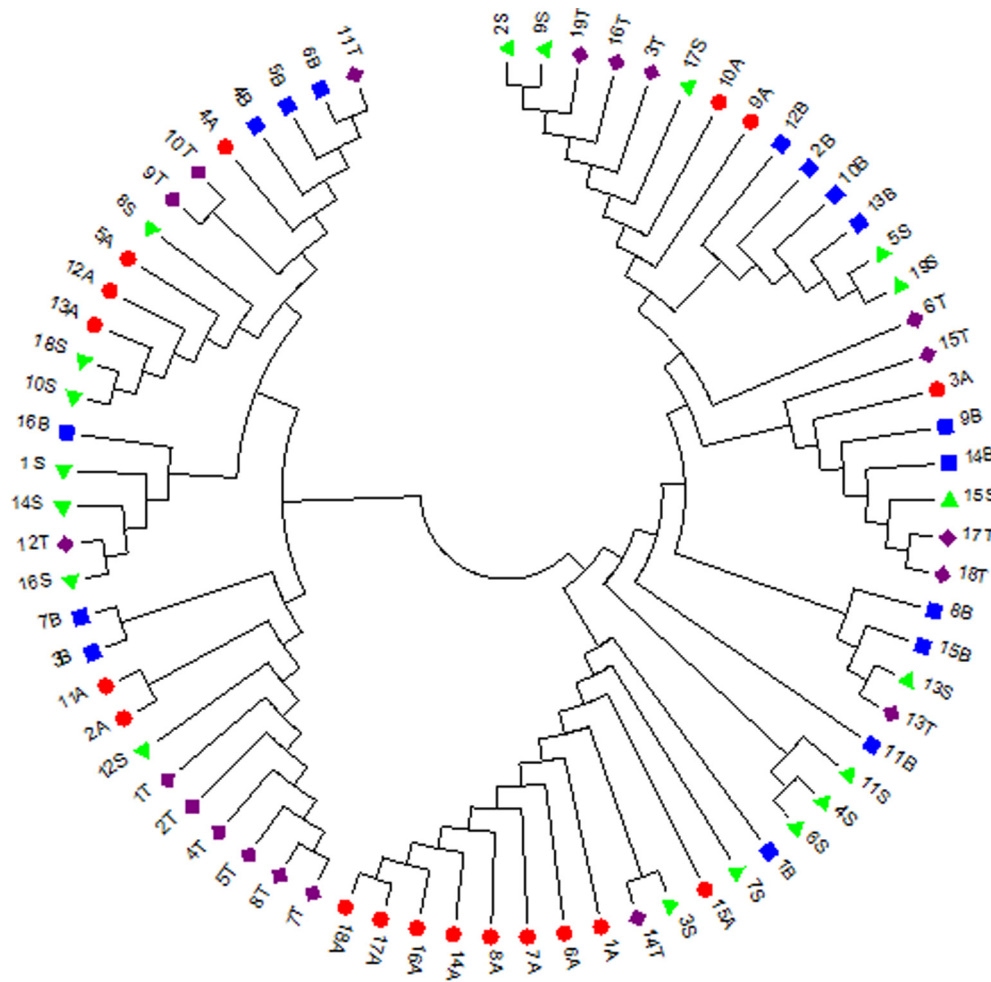


Figure 3 Neighbor-joining (NJ) tree of the tested horse animals. Arabian is in red, Baladi in blue, Sport in green and Thoroughbred in violet.

0.02321) then distance between Arabian and Baladi (D: 8.722 and Dxy: 0.02271 (Table 2).

Neighbor-joining (Phylogeny) tree for 72 tested sequences was constructed using the Mega 5.0 software. The tree declared the presence of mixing maternal lineages between the four tested populations but still there are some separated lineages especially for Arabian and Thoroughbred horses in contrast to Sport and Baladi lineages which mixed with each other and with other two populations (Fig. 3).

The sequences of 72 sequences were aligned with 13 published sequences as references, 11 of them were from different *Equus caballus* horse isolates around the world: X79547, KT211112, KT211090, KT211202, KT211171, KT211074, AY246174, AY575139, AF132568, D14991 and AF072995 whereas the other two reference sequences are JX312729 (*Equus burchellii*) and X97337 (*Equus asinus*) (Fig. 4). The results showed that all 72 animals belonging to the four tested horse populations are grouped with 11 reference sequences of *Equus caballus* and separated from the other two reference sequences of *Equus burchellii* and *Equus asinus*. This result declared the four tested horse populations are belonging to *Equus caballus* breed.

Jansen et al. [16] sequenced mtDNA D-loop from oriental and European horse breeds to answer of the question about the number of horse domestication origin. The phylogenetic tree showed that most tested mtDNA sequences are belonged to 17 distinct clusters, several of them related to the breeds or regions. To discover the relations between ancient and modern horses, Cieslak et al. [5] analyzed D-loop sequences from ancient and modern horses distributed over a wide range of regions and times. They showed that the ancient and domestic horses shared in 56 haplotypes of 87 haplotypes which are found in ancient horses. It is concluded that large mtDNA lineages mtDNA diversity represents ancestral variability and does not a result of horse breeding. Mitochondrial genomes of 83 modern horses from different continents were analyzed [2]. They revealed the presence of 18 major haplogroups, all of them were detected in modern horses from Asia. A wide range of lineages from the extinct *Equus ferus* was transmitted to modern *Equus caballus* as a result of domestication in the Eurasian steppes.

MtDNA D-loop sequence variation was used to discover the genetic relationships among different horse breeds and population around the world. Czerneková et al. [7] analyzed

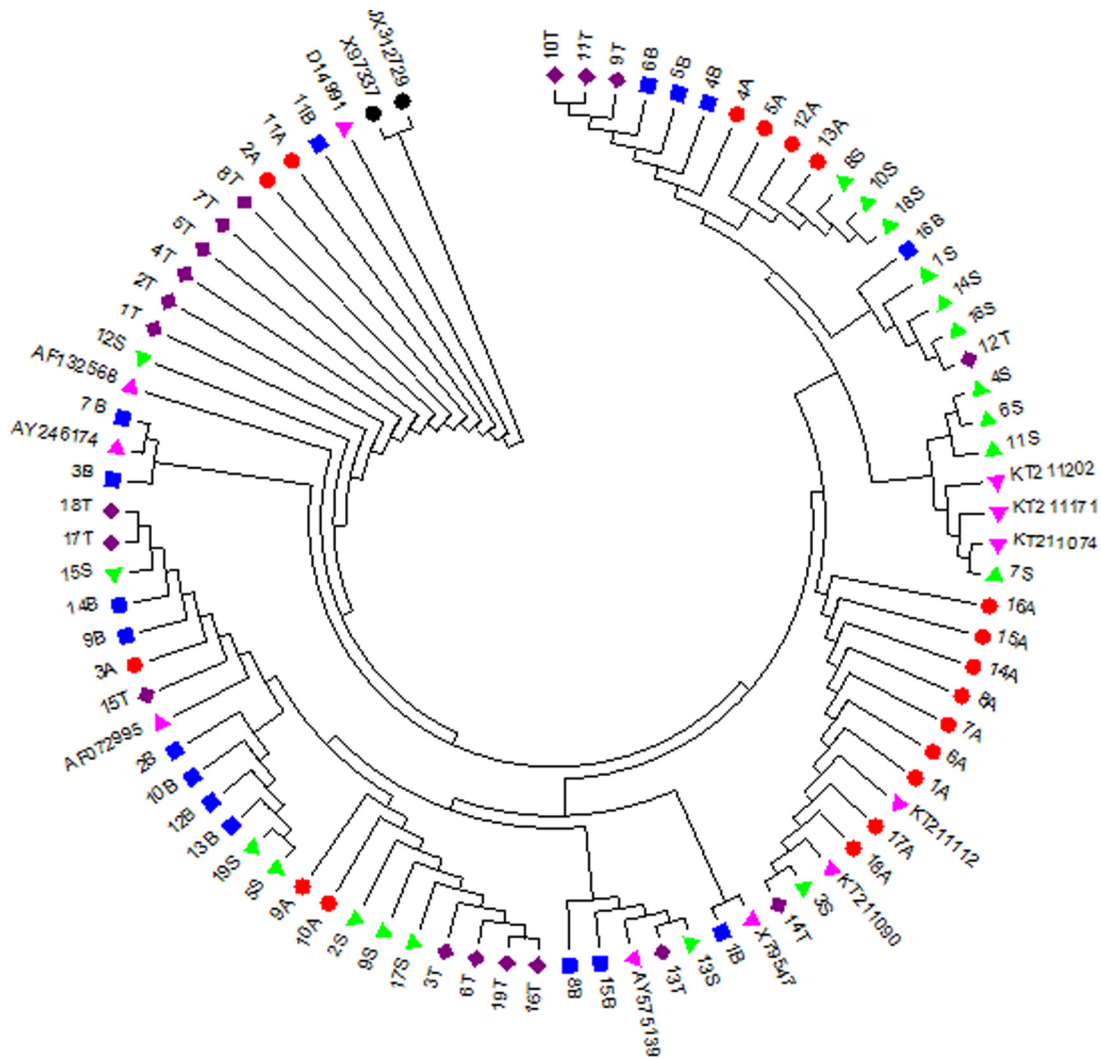


Figure 4 Phylogeny tree of 72 tested horses with reference sequences. Arabian is in red, Baladi in blue, Sport in green and Thoroughbred in violet. *Equus caballus* sequence references are in pink, *Equus burchellii* and *Equus asinus* reference sequences are in black.

165 Hucul horses mtDNA and the result revealed 38 polymorphic sites resulting in 14 haplotypes belonging to six haplogroups. This genetic information is the most important information used for breed conservations. The sequences of the same mtDNA region from maternal family lines of Zemaitukai, Lithuanian and European horse breeds were analyzed by Gus Cothran et al. [11]. The difference in nucleotides between haplotypes ranged from 6 to 11 nucleotides in Zemaitukai horse breed whereas the nucleotide difference between Lithuanian horse breed and reference sequences was 20 nucleotides without any clear pattern of genetic relationship between breeds.

Maternal genetic variation among native Arabian horse populations from the Middle East was tested by Khanshour and Cothran [17]. They sequenced mtDNA D-loop from 251 Arabian horses from different maternal family lines and the D-loop variability revealed differences among the haplotypes which have identical sequences in hypervariable region I. American-Arabian horses showed low diversity whereas Syrian population was the most variable haplogroup. No differentiation pattern was observed among all tested horse population and the results showed that several individuals from different strains shared a single haplotype.

Cardinali et al. [4] provided an idea about genetic variability and phylogeny in Italian horses where they analyzed the sequences of mtDNA control-region from ten Italian riding horse breeds and from Arabian horses. The authors reported the presence of previously described mtDNA haplogroups in Italian horse breeds with high haplotype diversity and the significant values of genetic differentiation were detected only in the geographically isolated contexts. This geographic effect was confirmed where the Italian pool stands in an intermediate position together with other Mediterranean stocks.

4. Conclusion

It is concluded that mtDNA control region sequence analysis is still the most informative tool for the identification of genetic biodiversity and phylogeny of different horse breeds and populations. The horse populations in Egypt possess low genetic diversity and shared in many haplotypes in spite of there is some separated lineages especially for Arabian and Thoroughbred horses and all of tested horse populations are belonged to *Equus caballus* breed.

References

- [1] D.W. Anthony, P. Bogucki, E. Coms, M. Gimbutas, B. Jovanovic, *Curr. Anthropol.* 27 (4) (1986) 291–313.
- [2] A. Achilli, A. Olivieri, P. Soares, H. Lancioni, B.H. Kashani, U. A. Perego, S.G. Nergadze, V. Carossa, M. Santagostino, S. Capomaccio, *Proc. Natl. Acad. Sci. USA* 109 (7) (2012) 2449–2454.
- [3] A. Bowling, A. Ruvinsky, *Genetic Aspects of Domestication. The Genetic of the Horse*, CAB International, UK, 2000.
- [4] I. Cardinali, H. Lancioni, A. Giontella, M.R. Capodiferro, S. Capomaccio, L. Buttazzoni, G.P. Biggio, R. Cherchi, E. Albertini, A. Olivieri, K. Cappelli, A. Achilli, M. Silvestrelli, *PLoS ONE* 11 (4) (2016) e0153004.
- [5] M. Cieslak, M. Pruvost, N. Benecke, M. Hofreiter, A. Morales, M. Reissmann, A. Ludwig, *PLoS ONE* 5 (12) (2010) e15311.
- [6] E.P. Cunningham, J.J. Dooley, R.K. Splan, D.G. Bradley, *Anim. Genet.* 32 (6) (2011) 360–364.
- [7] V. Czerneková, T. Kott, I. Majzlík, *Czech J. Anim. Sci.* 58 (10) (2013) 437–442.
- [8] I. Głazewska, *Livest. Sci.* 129 (1–3) (2010) 49–55.
- [9] M.W. Gray, G. Burger, B.F. Lang, *Science* 283 (1999) 1476–1481.
- [10] A.M. Guastella, A. Zuccaro, A. Criscione, D. Marletta, S. Bordonaro, *J. Heredity* 102 (6) (2011) 753–758.
- [11] E. Gus Cothran, R. Juras, V. Maciejauskiene, *Genet. Mol. Biol.* 28 (4) (2005) 677–681.
- [12] T.A. Hall, *Nucl. Acids Symp. Ser.* 41 (1999) 95–98.
- [13] E.W. Hill, D.G. Bradley, M. Al-Barody, O. Ertugrul, R.K. Splan, I. Zakharov, E.P. Cunningham, *Anim. Genet.* 33 (4) (2002) 287–294.
- [14] C.A. Hutchison, J.E. Newbold, S.S. Potter, M.H. Edgell, *Nature* 251 (5475) (1974) 536–538.
- [15] N. Ishida, T. Hasegawa, K. Takeda, M. Sakagami, A. Onishi, S. Inumaru, M. Kamtsu, H. Mukoyama, *Anim. Genet.* 25 (1994) 215–221.
- [16] T. Jansen, P. Forster, M.A. Levine, H. Oelke, M. Hurles, C. Renfrew, J. Weber, K. Olek, *Proc. Nat. Acad. Sci.* 99 (2002) 10905–10910.
- [17] A.M. Khanshour, E. Gus Cothran, *BMC Genet.* 14 (2013) 83–94.
- [18] K.I. Kim, Y.H. Yang, S.S. Lee, C. Park, R. Ma, J.L. Bouzat, H. A. Lewin, *Anim. Genet.* 30 (2) (1999) 102–108.
- [19] C.Z. Lei, R. Su, M.A. Bower, C.J. Edwards, X.B. Wang, *Anim. Genet.* 40 (6) (2009) 933–944.
- [20] M.A. Levine, *Late Prehistoric Exploitation of the Eurasian Steppe*. Cambridge McDonald Institute, 1999.
- [21] M.A. Levine, *The Domestication of horse: the origins, development and management of its behavior*. Cambridge University Press, 2005.
- [22] P. Libardo, J. Rozas, *Bioinformatics* 25 (11) (2009) 1451–1452.
- [23] J. Lira, A. Linderholm, C. Olaria, M. Brandstro, M.T.P. Gilbert, *Mol. Ecol.* 19 (2010) 64–78.
- [24] C. Luis, C. Bastos-Silveira, E.G. Cothran, M.M. Oom, *Genet. Mol. Biol.* 25 (2002) 309–311.
- [25] K.F. Mahrous, M. Hassanane, M. Abdel Mordy, H. Shafey, N. Hassan, *Genet. Eng. Biotech.* 9 (2) (2011) 103–109.
- [26] J.R. Meadows, O. Hanotte, C. Drogemuller, *Anim. Genet.* 37 (2006) 444–453.
- [27] S.A. Miller, D.D. Dykes, H.F. Polesky, *Nucl. Acids Res.* 16 (1988) 1215.
- [28] P.M. Mirol, P. Peral Garcia, J.L. Vega-Pla, F.N. Dulout, *Anim. Genet.* 33 (2002) 356–363.
- [29] P. Sutovsky, G. Schatten, *Int. Rev. Cytol.* 195 (2000) 1–65.
- [30] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, *Mol. Biol. Evol.* 28 (2011) 2731–2739.
- [31] X. Xu, U. Arnason, *Gene* 148 (1994) 657–662.
- [32] P. Zechner, J. Solkner, I. Bodo, T. Druml, R. Baumung, R. Achmann, E. Marti, F. Habe, G. Brem, *Livest. Prod. Sci.* 77 (2–3) (2002) 137–146.