

Genetic relationship between Kangal, Akbash and other dog populations

Evren Koban^{a,*}, Çiğdem Gökçek Saraç^a, Sinan Can Açıan^a, Peter Savolainen^b, İnci Togan^a

^a Department of Biological Sciences, Middle East Technical University, 06531 Ankara, Turkey

^b Department of Biotechnology, Royal Institute of Technology (KTH), 10691 Stockholm, Sweden

ARTICLE INFO

Article history:

Received 5 December 2006

Received in revised form 11 June 2007

Accepted 11 June 2008

Available online 18 September 2008

Keywords:

Kangal dog

Akbash dog

Canine

Mitochondrial DNA

Genetic relationship

ABSTRACT

Kangal and Akbash dogs are the two well-known shepherd dog breeds in Turkey. In order to contribute to the understanding of the genetic relationship between Kangal dogs, Akbash dogs and the dogs from different regions of Eurasia, 585 base pair (bp) segment of mitochondrial DNA (mtDNA) control region was sequenced from Kangals and Akbashes. Sequences of the Kangal and Akbash dogs examined in the present study were comparatively examined with those of previous studies on dogs. Consensus neighbour-joining tree with bootstrapping, which is constructed based on pairwise F_{ST} values between populations, indicated that Kangal dogs and Akbash dogs are on different branches of the tree. Furthermore, the nodes of these branches were supported with high bootstrap values. In conclusion, the present study indicated that Kangal and Akbash dogs might have descended maternally from different origins along the evolutionary history of domestic dogs.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Kangal dogs are the most popular dogs of Central Anatolia due to their strength, intelligence, and loyalty. Kangal dogs are physically distinguished by having an invariably black mask on a massive head and short, dense hair ranging in colour from dun (yellow) to steel grey [10]. On the other hand, Akbash dogs are pure white dogs of Western Anatolia. Recently, there has been much controversy about the origin of Kangal dogs. One of the hypotheses concerning the origin of Kangal dogs is given by Karadağ [7]. He claims that Kangal dogs were brought from Central Asia by the nomadic Turks.

The most commonly used marker for determining the history of populations within a species is the sequence analysis of mitochondrial DNA (mtDNA) [13]. The reason for the use of mtDNA in phylogenetic studies is its particular properties such as high mutation rates and absence of recombination. However, because it is inherited maternally it can give information only about the maternal history of the populations. Canine mtDNA sequence comparisons [15,14,16] as well as Y chromosome [9] and microsatellite markers [5] have been used to determine the origins of the domestic dog.

In Savolainen et al. [14]'s study, the genetic variation in 582 base pair (bp) segment of mtDNA obtained from 654 domestic dogs from Europe, Asia, Africa, Arctic America and from 38 Eurasian wolves were analysed. All dogs were grouped into six haplogroups (A to F), i.e. discrete groups of individuals who at some point in time shared a common ancestor, with respect to their maternal inheritance. Analysis revealed that domestic dogs originated about 15,000 to 40,000 years before present (B.P.) from a common gene pool in East Asia, and have subsequently spread all over the world [14]. If Savolainen et al. [14]'s study was considered from the perspective of Turkish dogs, there were 10 Kangal, 11 Akbash, 2 Anatolian Shepherd and one

* Corresponding author. Tel.: +90 532 5268608; fax: +90 262 6412309.

E-mail addresses: evrenkoban@yahoo.com (E. Koban), gcigdem@metu.edu.tr (Ç. Gökçek Saraç), can.acan@gmail.com (S.C. Açıan), savo@biotech.kth.se (P. Savolainen), togan@metu.edu.tr (İ. Togan).

¹ Current address: TÜBİTAK Marmara Research Center, GEBI, 41470, Kocaeli, Turkey.

In the frame of evolutionary history, we assumed that the breeds of the region might have developed from the population that had been introduced/migrated to the region from elsewhere. Therefore, we decided to use the breeds as the representatives of the population. Based on this assumption, to determine the genetic similarity of Kangal and Akbash dogs to each other and to the dogs from different regions of Eurasia, nine populations were employed. The names of the populations are: Kangal dogs, Akbash dogs, East Asian (China, Japan, and Korea), Middle Eastern (Iran, Saudi Arabia, Dubai, Israel, and Syria), Southwest Asian (Afghanistan, Uzbekistan, Tadjikistan, and Kazakhstan), European (Britain, Continent), Scandinavian, Siberian and African dogs. The sequence data of listed dog populations were retrieved from the study of Savolainen et al. [14]. Since those retrieved data were collectively analysed with the sequence data of Kangal and Akbash dogs obtained in the present study, sample size of Kangal and Akbash dogs increased to 115 and 20, respectively.

Sewall Wright [17] developed the conceptual and mathematical framework to describe the distribution of genetic variation within a species. One of the measures is F_{ST} which describes genetic divergence among populations [6]. Definition of F_{ST} is given below:

$$F_{ST} = (H_T - H_S) / H_T$$

where H_T is the expected heterozygosity of the metapopulation and H_S is the mean expected heterozygosity across subpopulations.

Using the aforementioned nine populations, pairwise F_{ST} estimations [17] for haploid data type were computed utilizing the options of the software package Arlequin version 3.01 [3]. By calculating pairwise F_{ST} , one of the genetic measures (Wright's F -statistics) that indicate the degree of differentiation between the populations has been converted into distance estimate. The statistical significance of F_{ST} values was estimated by permuting haplotypes among populations, using 10 000 permutations.

The obtained matrix of pairwise F_{ST} estimations between nine populations was then used to construct neighbour-joining (NJ) tree [11] of populations. The effect of the sampling variance on the tree topology was analysed following a bootstrap technique. Because of the fact that PHYLIP package does not have features defining populations and allowing trees of populations, SEQBOOT program in PHYLIP package 3.62 [4], was modified to accept command line parameters and with the help of a bash script it was run to obtain one thousand pseudo-data sets automatically. Although PHYLIP package provides excellent algorithms for our analysis, the user-unfriendly interface and the cumbersome file input-output preferences that need user intervention makes it inappropriate for repeated batch operations. We provided a solution by modifying PHYLIP's SEQBOOT such that, the user needs to enter only the necessary parameters (number of replications and the random seed number) just once from the command prompt at the beginning. The modified SEQBOOT acquires the input file names from a file (list.txt) created by the user, and names the output files accordingly. The other parameters (sequence type, bootstrap, non-interleaved sequences) are kept constant and sent automatically to the modified SEQBOOT program. This method not only reduces runtime, but also avoids any user errors. Then, the output files were parsed to become input file format in Arlequin version 3.01 software [3] by utilizing a program developed by S.C. Aan. The output files of the modified SEQBOOT consists of n times bootstrapped sequences in m population files. However, to be able to compare the populations and drawing trees of populations, the output files have to be arranged such that, the first sequence of the first population, the first sequence of the second population, and the first sequence of the m th population are put together in a file and in a second file, the second sequences are arranged, and the procedure continues until n th sequences are written in the n th file. In the end, we have n files with m sequences readable by Arlequin version 3.01 [3] software. The whole procedure was summarized with the help of a flow chart in Fig. 2.

By another program written by S.C. Aan, output files, consisting of 1000 F_{ST} matrices, in Arlequin 3.01 software [3] were transformed into input file format of CONSENSE program in PHYLIP package 3.62 [4]. In this way, 1000 trees were constructed and the consensus of all resulting NJ trees was built with the CONSENSE program as implemented in PHYLIP package 3.62 [4].

As stated previously, along with other populations', some of the Kangal ($n_{\text{Kangal}} = 10$) and Akbash ($n_{\text{Akbash}} = 11$) sequence data were retrieved from the study of Savolainen et al. [14]. To see the significant contribution of sequences of newly collected Kangal ($n_{\text{Kangal}} = 105$) and Akbash ($n_{\text{Akbash}} = 9$) dog samples to the phylogenetic tree, NJ tree of nine populations was also constructed using only the samples from the database.

3. Results

By comparing sequences of Kangal and Akbash dogs in the present study to the reference sequence "A2" [14], a total of 26 different haplotypes were determined in 114 dogs ($n_{\text{Kangal}} = 105$, $n_{\text{Akbash}} = 9$), six of them being newly described (Fig. 1). These haplotypes belong to the four major haplogroups A, B, C and D based on their D-loop haplotype in accordance with Savolainen et al. [14]. Haplogroups A and B were found in both of the studied dogs, whereas haplogroups C and D were found only in Kangal dogs. As was previously observed in Kangals, haplotype D5 was seen in 13 Kangal dogs in the present study. D5 was 2-3 mutational steps away from South European dogs while 5 steps away from Scandinavian dogs [15]. Haplogroup D was not seen in Akbash dog samples (see Introduction) as previously observed in [14].

To observe the genetic similarity of Kangal and Akbash dogs to each other and to the dogs from different regions of Eurasia, nine populations were determined (see Materials and Methods), then pairwise F_{ST} estimations between nine populations

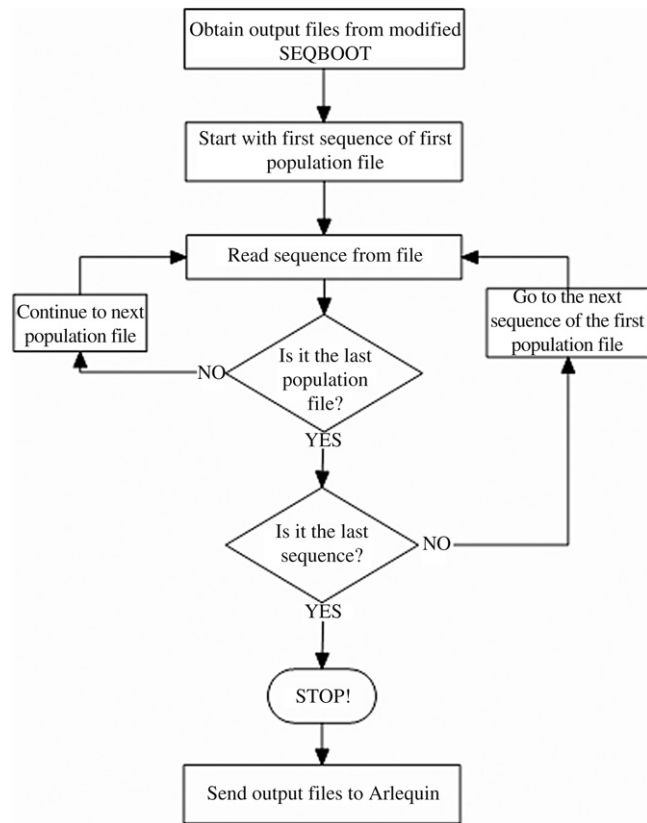


Fig. 2. Summarization of the procedure of converting modified SEQBOOT output files into the Arlequin files.

Table 1
Pairwise F_{ST} values of nine populations

	Africa	Akbash	East Asia	Europe	Kangal	Middle East	Scandinavia	Siberia	Southwest Asia
Africa	0								
Akbash	0.04651 ^{NS}	0							
East Asia	0.01859 ^{NS}	0.02430 ^{NS}	0						
Europe	0.02486 ^{NS}	0.01779 ^{NS}	0.01173 [*]	0					
Kangal	0.19094 [*]	0.20856 [*]	0.12679 [*]	0.14708 [*]	0				
Middle East	0.16982 [*]	0.12728 [*]	0.07198 [*]	0.08142 [*]	0.04721 [*]	0			
Scandinavia	0.25031 [*]	0.25874 [*]	0.19079 [*]	0.18756 [*]	0.05792 [*]	0.10474 [*]	0		
Siberia	0.01762 ^{NS}	0.04829 ^{NS}	0.00513 ^{NS}	0.00618 ^{NS}	0.12573 [*]	0.09801 [*]	0.16362 [*]	0	
Southwest Asia	0.21786 [*]	0.20357 [*]	0.10316 [*]	0.13402 [*]	0.01648 ^{NS}	-0.00467 ^{NS}	0.09088 [*]	0.12260 [*]	0

^{*} $p < 0.05$.
^{NS} not significant.

were obtained and shown in Table 1. In this matrix the negative F_{ST} value between Southwest Asia and Middle East was assumed as zero and negativity was assumed to be a computational artifact.

Furthermore, to determine population relationships in detail, neighbour-joining (NJ) tree of nine populations were constructed using the matrix of pairwise F_{ST} estimations between populations. The resulting NJ tree is shown in Fig. 3A. To verify the contribution of sequences of newly collected Kangal and Akbash dog samples to the phylogenetic tree, NJ tree of nine populations was also constructed using only the database samples (see Materials and Methods) and shown in Fig. 3B. From these trees, it can easily be seen that Kangal and Akbash dogs are on two separate clusters. The similarity of Kangal dogs to the Scandinavian and Southwest Asian (Afghanistan, Uzbekistan, Tadjhikistan, Kazakhstan) dogs was indicated by the first clusters in both trees (Fig. 3A, 3B) On the other hand, the highest similarity of Akbash dogs to African, European (Britain, Continent) and Siberian dogs was clearly depicted by Cluster II in Fig. 3A. Furthermore, in the first tree (Fig. 3A) robustness of these clusters was very high as can be seen from the bootstrap values (100% and 61%).

Consequently, although all the samples of Kangal and Akbash dogs were collected from Anatolia, genetic similarity between these two dogs was found to be low.

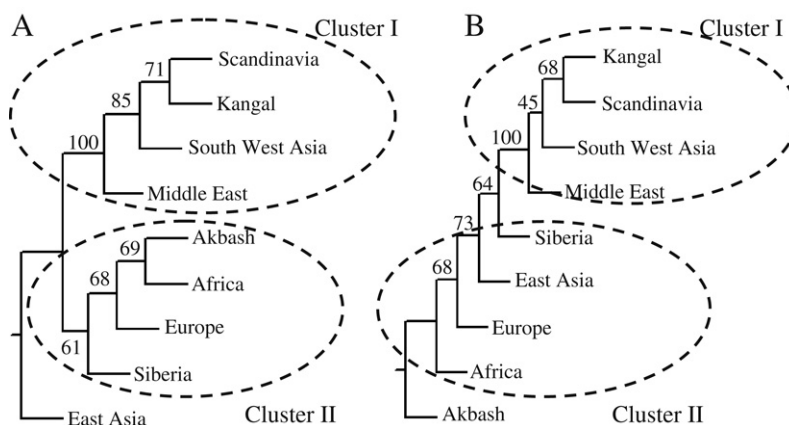


Fig. 3. Consensus neighbour-joining (NJ) trees of nine populations based on the pairwise F_{ST} estimates. Ellipses surround main clusters. Percent bootstrap supports in 1000 replicates are shown on the nodes. (A) NJ tree constructed from newly added Kangal and Akbash sequences to the database sequences, (B) NJ tree constructed from the former database sequences.

4. Discussion

Domestic animals have shared a joint history of evolution with humans for approximately the last 10,000 years, but domestic dogs, because they are believed to have been domesticated by the hunter gatherers, have perhaps shared a longer joint history with humans for about the last 15,000–40,000 years as suggested by Savolainen et al. [14]. If the origin of domestication of dogs was Eastern Asia as predicted [14], their migration to the west might have been through more than a single route. The two major clusters on the tree constructed might indicate two main routes of humans and dogs to the west from Eastern Asia. One is towards Africa and ancestors of the Akbash seem to have followed this route. This south route perhaps starting from Siberia but going to Africa to Anatolia in terms of Akbashes and Europe was clearly seen by Cluster II in Fig. 3A and it exhibited the significance of contribution of the new samples that are collected in the present study. On the other hand, during the latest glacial, around 18,000 years ago, Northern and Central Europe were covered with glaciers. With the retreat of glaciers, North Eurasian human migration might have brought the dogs towards Northern Europe. This migration might be represented by Cluster I in Fig. 3. The Kangal seems to be associated with this second “northern route”. During this migration which is represented by Cluster I in Fig. 3., different European wolves might have introgressed to the domestic dogs as was implied by the high degree of genetic differences between D1–D4 to others. This route might have left imprints in Eastern Central Anatolia in Kangal dogs by the migration branch through the Caucasus. In a recent study, the contribution of Turkic speaking Central Asian populations to the gene pool of Anatolian Turkish human population was analysed with admixture analysis [1]. In that study, based on the mtDNA analysis, it was determined that there was 22% female contribution from Central Asia to Turkey. Furthermore, based on the Y chromosome markers a significant ($p < 0.05$) and very heavy migration route from Central Asia through the south of the Caspian Sea was detected [1]. Migrating Turks were nomads and were heavily relying on livestock. It can be envisaged that during their migration(s), a part of the northern route dogs accompanied them and today, in Anatolia they are known as Kangal dogs. As a support to this hypothesis Southwest Asian (Afghanistan, Uzbekistan, Tadjikistan, and Kazakhstan) dogs were found to be the closest to the Kangal dogs in the present study. Furthermore, nomadic Turks initially settled in the Eastern and Central Eastern parts of Anatolia where the town of Kangal is located. This town is famous for its dogs, Kangal dogs.

As a conclusion, in this study a hypothesis which was proposed by Karadağ [7] and which was based on the linguistic and morphological evidence about the origin of Kangal dogs of Turkey was supported with the help of computational biology.

Acknowledgements

We are grateful to Assoc. Prof. Dr. Vahdettin Altunok, Vet. Tuğber Demirci, Sue Kocher, Margaret Mellor, Doğan Kartay, Onur Kanlı, İlker Ünlü, Akın Tülübaş and Banu Ergezen for helping us during the collection of the samples. The warm welcome by and invaluable help of Helen Angleby, Mattias C.R. Oskarsson and Lars-Göran Dahlgren will always be greatly appreciated. The experiments were carried out at KTH Biotechnology. Moreover, we would like to thank Havva Dinç and Ceran Şekeryapan for their valuable help. The project was supported by METU (Project Code: BAP-08-11-DPT.2002K120510-BTEK-16).

References

- [1] C. Berkman, Comparative analyses for the Central Asian contribution to Anatolian gene pool with reference to Balkans, Ph.D. Thesis, Middle East Technical University, Ankara, 2006.

- [2] H. Dinç, ALU Insertion Polymorphisms in Anatolian Turks, Ms Thesis, Middle East Technical University, Ankara, 2003.
- [3] L. Excoffier, G. Laval, S. Schneider, ARLEQUIN (version 3.0): An integrated software package for population genetics data analysis, *Evolutionary Bioinformatics Online* 1 (2005) 47–50.
- [4] J. Felsenstein, PHYLIP (Phylogeny Inference Package), version 3.6, Department of Genome Sciences, University of Washington, Seattle, 2004.
- [5] D.N. Irion, A.L. Schaffer, S. Grant, A.N. Wilton, N.C. Pedersen, Genetic variation analysis of the Bali street dog using microsatellites, *BMC Genetics* 6 (2005) 6.
- [6] M.A. Jobling, C. Tyler-Smith, M. Hurles, *Human Evolutionary Genetics: Origins, Peoples and Disease*, Garland Science, 2004.
- [7] H. Karadağ, An investigation on the opinions on the place name of Kangal, and the origin of Kangal dog, *Milli Folklor Dergisi*. 7 (2002) 24–45.
- [8] K.S. Kim, S.E. Lee, H.W. Jeong, J.H. Ha, The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome, *Molecular Phylogenetics and Evolution* 10 (1998) 210–220.
- [9] C. Natanaelsson, M.C.R. Oskarsson, H. Angleby, J. Lundeberg, E. Kirkness, P. Savolainen, Dog Y chromosomal DNA sequence: Identification, sequencing and SNP discovery, *BMC Genetics* 7 (2006) 45.
- [10] D.D. Nelson, A general classification of the native dogs of Turkey, in: *International Symposium on Turkish Shepherd Dogs*, 1996, pp. 19–96.
- [11] N. Saitou, M. Nei, The neighbor-joining method: A new method for reconstructing phylogenetic trees, *Molecular Biology and Evolution* 4 (1987) 406–425.
- [12] J. Sambrook, E.F. Fritsch, T. Maniatis, *Molecular Cloning: A laboratory Manual*, second ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989.
- [13] P. Savolainen, Mitochondrial DNA analysis of the control region in forensic and population genetic studies, Ph.D. Thesis, Royal Institute of Technology, Stockholm, 1999.
- [14] P. Savolainen, Y. Zhang, J. Luo, J. Lundeberg, T. Leitner, Genetic evidence for an East Asian origin of domestic dogs, *Science* 298 (2002) 1610–1613.
- [15] B. Van Asch, L. Pereria, F. Pereria, P. Santa-Rita, M. Lima, A. Amorim, MtDNA diversity among four Portuguese autochthonous dog breeds: A fine-scale characterisation, *BMC Genetics* 6 (2005) 37.
- [16] C. Vila, P. Savolainen, J.E. Maldonado, I.R. Amorim, J.E. Rice, R.L. Honeycutt, K.A. Crandall, J. Lundeberg, R.K. Wayne, Multiple and ancient origins of the domestic dog, *Science* 276 (1997) 1687–1689.
- [17] S. Wright, The genetical structure of populations, *Annals of Eugenics*. 15 (1951) 323–354.