Journal of Wildlife and Biodiversity 3(2): 18-28 (2019)

(http://jwb.araku.ac.ir/)

DOI: 10.22120/jwb.2019.99032.1044

Research Article

Phylogeny, genetic diversity and population structure of Brandt`s hedgehog *Paraechinus hypomelas*, inferred from the mitochondrial evidences

Ehsan Kashani[1], Hamid Reza Rezaei[2*], Nematolah Khorasani[3] , Morteza Naderi[4, 5]

[1]Department of Environmental Sciences, Faculty of Natural Resources and Environment Islamic Azad University, Science and Research Branch, Tehran, Iran

[2*]Department of Environment, Faculty of Fisheries and Environmental Science, Gorgan University of Agricultural Sciences and Natural Resources, Iran

[3]Department of Fisheries and Environmental Sciences, University of Tehran, Tehran, Iran,

[4]Department of Environmental Science, Faculty of Agriculture and Natural Resources, Arak University, Arak , Iran

[5]Koç University College of Sciences, Rumeli Feneri Yolu 34450 Sarıyer, İstanbul, Türkiye

*email: rezaei@gau.ac.ir

Received: 29 January 2019 / Revised: 28 February 2019 / Accepted: 15 March 2019 / Published online: 27 April 2019. Ministry of Sciences, Research and Technology, Arak University, Iran.

Abstract

The Brandt`s hedgehog *Paraechinus hypomelas* (Brandt 1836), is a relatively widespread species which range from Arabian Peninsula and Iran, through southern areas of Central Asia to western South Asia.The phylogenetic position of the species tat is little known in Iran, although it has been studied in different parts of its distributional range. To this aim, during 2017-2018, the species was sampled in a non-invasive method (n= 34) from the southeast of Iran. Genetic variation and polymorphic sites were determined from cyt*b* (1120bp). Totally 22 haplotypes and haplotype diversity ranging from 0.859 to 1.099 were detected from cyt*b*. The average value of the nucleotide differences among the cyt*b* sequences was calculated as 4.68. The Tamija's *D* test (-1.88) and Fu`s *FS* test (-15.73) revealed negative value, indicating significantly deviations from neutrality which both indicate of recent population expansion. Investigation on pairwise differences, mismatch distributions, indicated of past expansions (SSD= 0.0033, P value = 0.38). The Iran south east population constitute a phylogenetic clade which is completely distinct from other known lineages in the species distributional range. Relatively high amount of the haplotype and nucleotide diversity can be related to the high effective population of the species, the rate of gene flow among the populations and also the sudden expansion in the past.

Keywords: Brandt`s hedgehog, genetic structure, population expansion, phylogeny.

Introduction

Genetic diversity, as a basis of evolutionary potenitial, plays crucial role species adaptive variations (Pablos *et al*. 2015). Lowering the genetic diversity will lead to a decrease in species adaptive variations potential and fitness which in turn will affect its long-term survival (Desalle and Amato 2004, Frankham *et al*. 2004). One of the key data in the species conservation programs and setting priorities refer to this parameter (Avise 2000, Harley *et al*. 2005). Hedgehogs are small-sized, insectivorous and nocturnal mammalian species which distributed Africa and Asia as an indigenous species (He *et al*. 2012). The investigation on the spicies phytogeographic status shed light on the paths of occupying new habitats after the last glacial maximum (LGM) era (Hewitt 2000). Occupying new habitats with considerable difference in the features has

led to diverse gradient through the species molecular and even morphological peculiarities (Berggern *et al*. 2005). Although the phylology of the spicies has been studied well (Seddon *et al*. 2001, Berggren *et al*. 2005), there is no published data on its variations and structure especially throughout Iran.

Four species and three genera of the hedgehogs have been identified in Iran including desert hedgehog (*Hemiechinus aethiopicus*), the long-eared hedgehog (*Hemiechinus auritus*) and the Brandt`s hedgehog (*Paraechinus hypomelas*) which all are completely widespread in the southeastern Iran (Etemad 1985).

The Brandt`s hedgehog is well distributed in almost across the country steppe regions, particularly Baluchistan regions (Etemad 1985). In the current study we aimed to investigate the genetic diversity of the Brandt hedgehog using mitochondrial cyt*b*, and determining the phylogenic position of the species among other documented lineages.

Material and methods

Study area

During 2016 - 2017 the sampling was carried out in different parts of the southeastern Iran (Fig. 1). During the field work, we collected 34 tissue samples through a non-invasive methods like collecting samples from animals killed in vehicles collision and sampling from ear pad after live trapping based on DoE`s permission number of 95.27706 (2016). The dominant climate of the study area is arid and semi-arid and the majority of precipitation occurs in winter and fall.

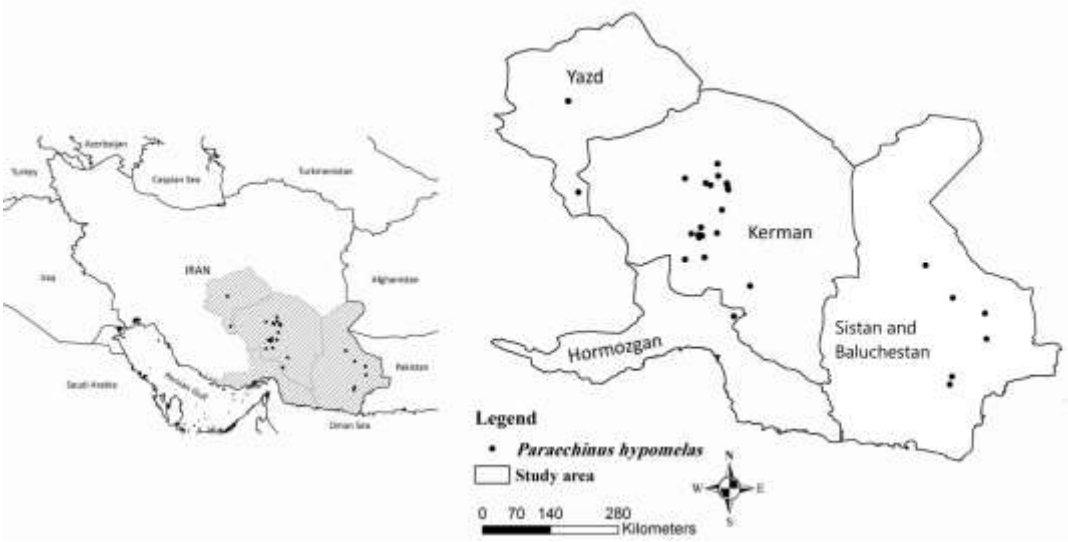

Figure 1. Study area (hatched) and sampling locations for the Iranian hedgehog

DNA extraction and amplification

Phenol–chloroform method was adopted in DNA extraction and the quality and quantity of the extracted DNA was tested by both gel electrophoresis and spectrophotometry methods. Amplification of 1120 bp segments of the marker (mtDNA cyt*b*) was done using primers L-hemi (TCGTGACCCATGATATG AAAAATC) and H-eri (GAATATCAACTT TGGGTGTTGATGGTA) (Banikova *et al*. 2014).

Polymerase chain reaction (PCR) was done using 20µl of the Ampliqon 2X Taq Permix consitituted of 10 µl of Taq polymerase enzyme, 1 µl DNA, 0.4 micro molars mixture of initiator at the concentration of 5 pico molars, 200 micro molars dNTP, 1 µl MgCl2. Twice distilled water was added to the mixture to reach a volume of 20µl. The thermal cycle of the amplification process was regulated as 5 minutes 95°C, followed by 35 cycles which included 45 seconds at 94°C, 45 seconds at 58°C, one minute at 72°C, and the final expansion at 73°C for 10 minutes. The amplified fragments were documented by 2% agarose gel electrophoresis. Finally PCR

products were sequenced by ABI 3730 xl, SeqGen company, USA.

Data analysis

The obtained sequences were first checked using SeqScape 2.6 (Applied Biosystems) and errors were corrected manually. Thereafter, sequences were aligned using BioEdit 7.2.5 and the Clustal w algorithm (Thompson *et al*. 1994). Determining the number of haplotypes (H), polymorphic sites (S), haplotype diversity (Hd) (Nei, 1987), and nucleotide diversity (π) (Lynch and Crease, 1990) was done by DnaSP 5 (Librado and Rozas 2009) long-eared hedgehogs (*Hemiechinus auritus*) was used as an outgroup in the phylogenetic tree (Table 1).

Table 1. Sequence information for sequences obtained from the gene bank, used to draw haplotype network and phylogeny tree

Reference	Species	Accession
He *et al*. 2011	*Paraechinus aethiopicus*	HQ857538
He *et al*. 2011	*Paraechinus aethiopicus*	HQ857537
Bannikova *et al*. 2014	*Paraechinus aethiopicus*	KF783142
He *et al*. 2011	*Mesechinus hughi*	HQ857531
He *et al*. 2011	*Mesechinus hughi*	HQ857530
He *et al*. 2011	*Mesechinus dauuricus*	HQ857529
He *et al*. 2011	*M. dauuricus*	HQ857528
He *et al*. 2011	*H. auritus*	HQ857522
Bannikova *et al*. 2014	*H. auritus*	KF783136
Nikaido *et al*. 2003	*Hemiechinus auritus*	NC_005033
Bannikova *et al*. 2014	*Erinaceus roumanicus*	KF783124
Bannikova *et al*. 2014	*Erinaceus concolor*	KF783121
	H. auritus	-
	H. auritus	-

Analysis of past demographic changes of the species populations (bottleneck and population expansion) was done by Tajima's neutrality test (D) (Tajima 1989) and Fu's Fs index (Fu 1997) in DnaSP 5.10. The significance of the D and Fs indices was investigated with 1000 permutations. The negative value of these indices indicates on the expansion of the species populations in the past. Mismatch distribution test in Arlequin 3.0 (Schneider and Excoffier 1999) was used to investigation on the demographic history of the species and its population expansion in the past. Goodness of fit between the expected and the observed nucleotide differences in the obtained sequences was evaluated by a bootstrap test (1000 bootstrap samples) and sum of the square differences (SSD).

Genetic Structure

In order to investigate the genetic structure of the species in the southeastern Iran, clustering algorithm employed used based on a Bayesian model in STRUCTURE (Pritchard *et al.* 2000). This model evaluate divers loci based on differences in their allele frequency and the lengths. It is assumed that the intended samples formed K genetic groups and depending on sample genotype each one may belong to one or several groups. The membership of the samples to any population can be determined by Qi index which is the average of posterior probability of the intended sample, which reflects belonging to a common ancestor with the Kth population. Because of the complexity which is usually occurs among the various populations, a Bayesian clustering algorithm was used by exploiting the admixture model. To that end, by placing the number of probable populations (K) from 1 to 6 (MAXPOPS = 1) and thereafter calculating the maximum likelihood of data (Ln P (D)), the suitable number of distinctive genetic structures was calculated in the sampled individuals. In this study, the intended threshold assumed to place

one sample in the intended group (Qi), was 0.8 (Randi and Lucchini 2002, Khosravi *et al*. 2018). By exploiting the admixture model and correlate allele frequency, Pritchard algorithm was used with five repletion for each population (K) and 100000 simulations of Markov chain. The Pritchard model calculates maximum likelihood of data (Ln P (D)) for any performance. Therefore, the average maximum likelihood for the data was calculated in STRUCTUTRE first. As a result, the highest value of Ln P (D) was considered as the most probable number of genetic structures. In the second method, ΔK was calculated in STRUCTURE Harvester. ΔK was calculated as formulated by Evanno *et al*. (2005) and the highest value for this index was assumed as the real extent of the genetic structure among the Iranian hedgehogs.

Haplotype network and phylogenic tree

A neighbor-joining network was created by uncorrected patristic distance and bootstrap analysis with the 1000 repetition in SPLITSTREE v4.6 (Huson and Bryant 2006). Also, a haplotype network was drawn using the statistical parsimony method in TCS v1.21 (Clement *et al*. 2000). To select the best nucleotide substitution model, JModelTest 2.1.5 was used (Darriba and Posada 2014). Subsequently, based on the chosen model, MrBayes 3.2.2 was used to draw the Bayesian tree (Huelsenbeck and Ronquist 2001). Bayesian analyses were done on the basis of Markov Chain Monte Carlo with 10^{12} repetitions and a sample frequency of 500. Finally, trees were drawn by Consensus Majority-Rule putting *E.roumanicus* as an out group.

Results

Haplotype and nucleotide diversity

Totally 34 samples were successfully sequenced which yielded 1121 bp of cyt*b* gene, mostly belong to Kerman (n=22), Yazd (n=5), Sistan and Baluchistan (n=6) and one sample from Hormozgan. Our analysis resulted in finding 24 haplotypes and 39 polymorphic sites, haplotype diversity (Hd) was 0.979 ±012 and nucleotide diversity (Pi) calculated as 0.004 ± 0.009. The average genetic distance between the sequences was calculated as 0.0115 (SE= 0.010). Past population distribution analyses based on the samples showed Tajima's *D* to be significant (P value < 0.05). The average nucleotide difference among paired sequences was 4.68. Mismatch distribution test indicates of population expansion in the past (SSD=0.0033, P value =0.38, and r=0.021, P value =0.44). Further, for *P. hypomelas* samples, the observed and expected frequency curve of nucleotides among pairs of sequences was unimodal which confirms past population expansion based on the sudden expansion model (Fig. 2). The index of population expansion (Tau value) and raggedness index were respectively calculated as 3.74 and 0.02. The fact that this index is not significant (P-value > 0.01) points to the rejection of the null hypothesis and hence confirms again the past population expansion.

Evaluating genetic structure

In figure 3 the maximum probability with five independent petitions per K is calculated. Larger value of LnP(D) for a K indicates the superiority of that K in determining the size of a population in terms of better precision and higher probability. As shown in figure 3a, as K increases, the maximum probability decreases after K=2 whitch indicate that analyzed group can be divided to two subgroups. Estimating population size using ΔK revealed the highest value ΔK to occur with two populations (K=2) (Fig. 3b). Also, the extent of group membership for each sample, was calculated based on two and three populations (Fig. 3c). Despite the prediction of two genetic groups, consideration of the samples geographical locations does not indicate distinctive genetic structuring based on the geographical location of the samples, which can be seen in the corresponding figures. It should be noted that the differences between LnP(D) for K=1 and K=2 is very small. Consequently, we could conclude that for

southeast samples, we couldn't find distinct genetic structure.

Haplotype network and phylogeny tree

Phylogenetic trees revealed a significant divergence among the analyzed samples (Fig. 4). Sequences of long-eared hedgehogs (*H. auritus*) were placed in the haplotype network very closed to the desert hedgehogs (*P. aethiopicus*). The long-eared hedgehog lives in parts of northern and southeastern Iran. In some areas, the long-eared hedgehogs and Iranian hedgehogs have overlapping habitats. The desert hedgehog has only been reported to live on Tonb Island in Iran (Karami *et al*. 2002). Also, the extracted sequences of European hedgehogs (*Erinaceus concolor*), present in parts of northeastern Iran, are placed with a short distance from northern white-breasted hedgehogs (*Erinaceus roumanicus*).

In the haplotype network (Fig.5 and 6), although some of the samples from Yazd province are placed in a separate group, however no specific order can be drown among the known haplotypes. This facts shows the insignificant differences in nucleotide sequences of field samples, in addition to demonstrating that there is no clear genetic structuring among populations of Brandt`s hedgehogs in the region. The Bayesian phylogenetic tree shown in figure 7. All the identified haplotypes from the southeastern Iran were placed in one clade. All long-eared hedgehogs were placed next to *Mesechinus hughi* and *Mesechinus dauuricus* with short genetic distance.

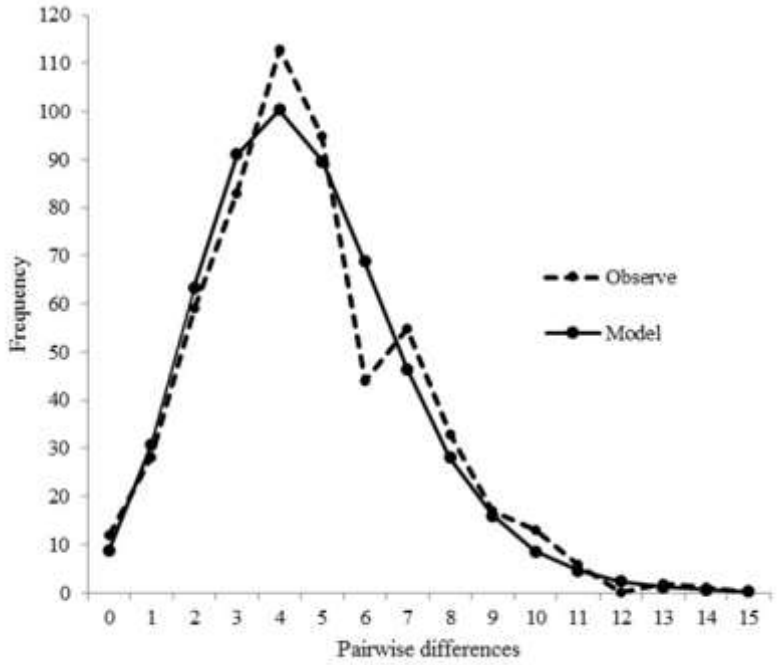

Figure 2. observed and expected values for differences in nucleotides between sequence, pairs in the samples, based on cytochrome *b* sequencing and the mismatch distribution, used to evaluate the probability of population expansion

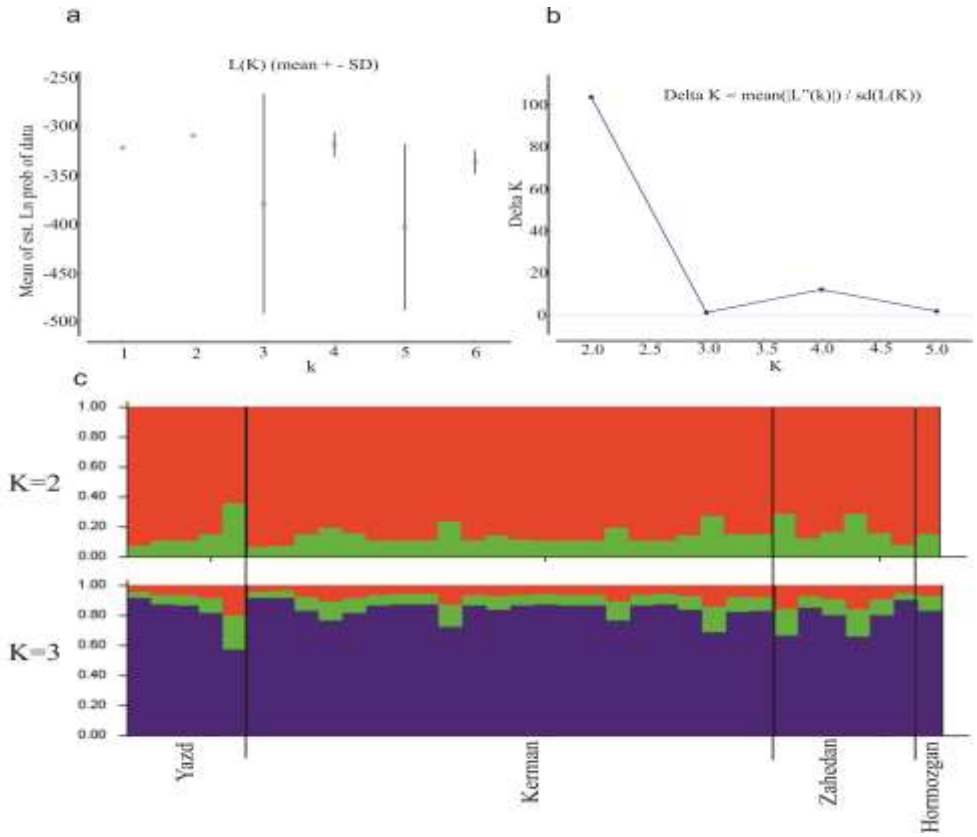

Figure 3. Changes in maximum LnP(D) with the increase in number of populations (a); changes in ΔK with the increase in number of populations (b); group membership for each sample with K=2 and K=3 (c)

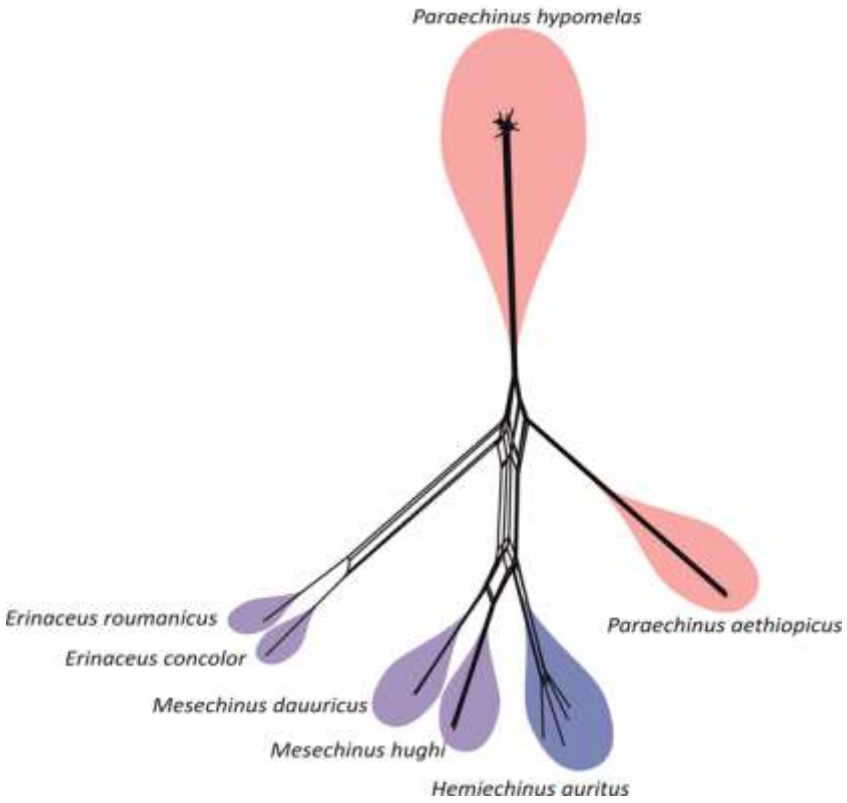

Figure 4. Haplotype network for Iranian hedgehog samples and sequences from other hedgehog species, using neighbor joining and uncorrected genetic distance in SPLITSTREE

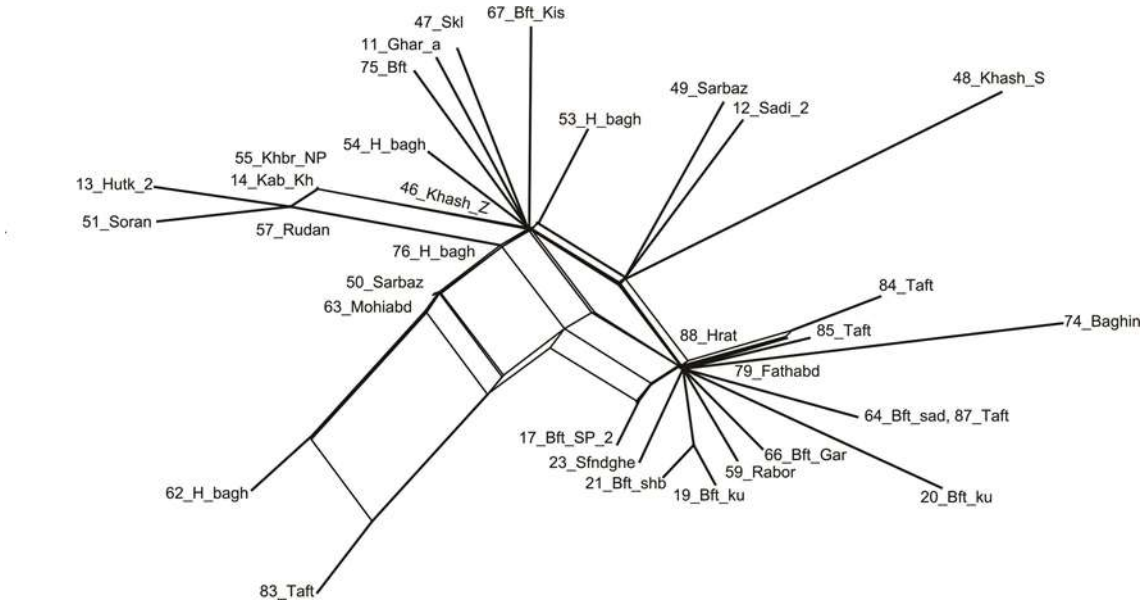

Figure 5. Haplotype network for Iranian hedgehog samples using neighbor joining and uncorrected genetic distance in SPLITSTREE

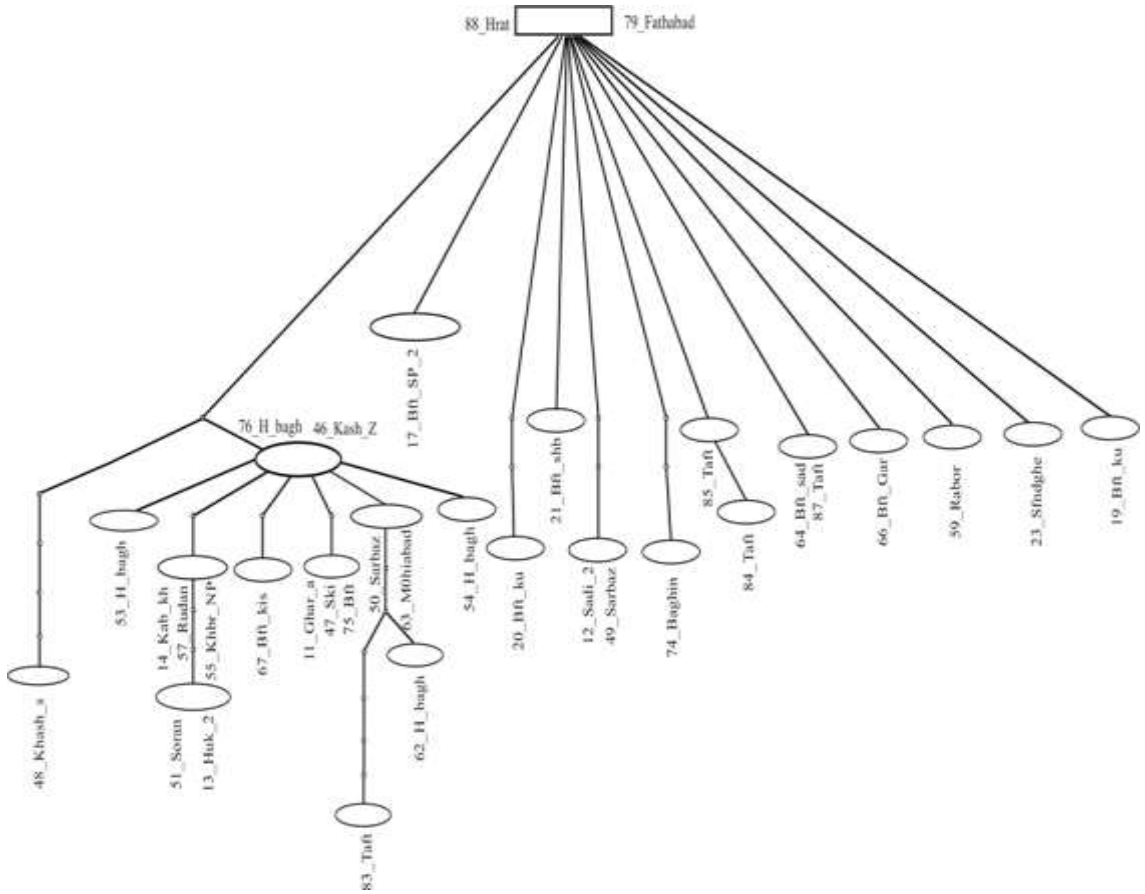

Figure 6. Haplotype network for Iranian hedgehog samples using statistical parsimony in TCS

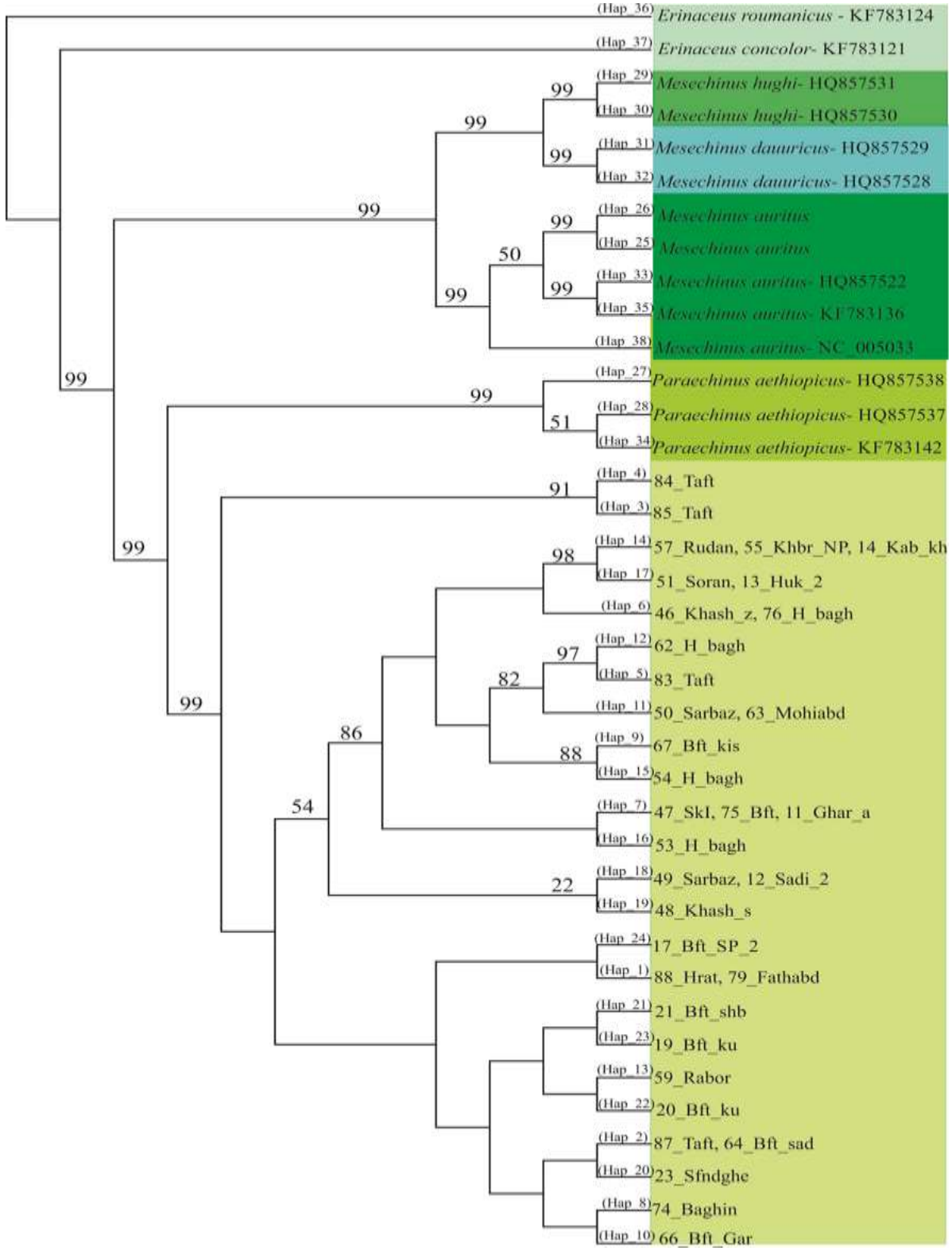

Figure 7. Phylogeny tree of hedgehog species using Bayesian inference. The numbers on branches show posterior probability.

Discussion

Steppes, deserts, and mountains cover the most parts of the southeastern Iran where host different species of hedgehogs especially *P. hypomelas*. Genetic variation and phylogeographic status of such animals and *P. hypomelas* in particular are still poorly known especially in Iran.

As indicated from our data, the genetic diversity (haplotype and nucleotide) of the Iranian hedgehog populations is relatively higher than

other areas like Serbia, Montenegro, Bosnia and Herzegovina and Macedonia (Djan *et al*. 2017). Stefanović *et al.* (2016) investigation on northern white-breasted hedgehogs in Siberia focusing mtDNA resulted in 12 identified haplotypes and nucleotide diversity of 0.827 however these values for Czech and Slovakia was reported as 0.289 for Hd and 0.077 for nucleotide diversity respectively (Bolfikova and Hulva 2012).

Demographic analyses also indicated of population expansion in the past which was already reported for northern white-breasted hedgehog (Bolfikova and Hulva 2012). Relatively higher genetic diversity of the Iranian hedgehogs can be related to their probable slower expansion in the past (Bolfikova and Hulva 2012).

We found no distinctive genetic structure among the analyzed samples however Pritchard algorithm suggested two genetic clusters which shows future need to supplementary further analysis be done in the future using better markers line SNP and microsatellites (Fig. 3).

Conclusion

Grouping all analyzed samples in one cluster confirmed the phylogenetic position of Iran's Brandt`s hedgehogs which was reported already (He *et al*. 2012, Bolfikova and Hulva 2012). As a pioneer research, this study yielded some findings on little known desert animals of the country and can be completed focusing on more effective molecular markers.

References

Avise J. 2000. Phylogeography: The History and Formation of Species. Cambridge, Massachusetts and London: Harvard University Press.

Bannikova A.A., Lebedev V.S., Abramov A.V., Rozhnov, V.V. 2014. Contrasting evolutionary history of hedgehogs and gymnures (Mammalia: Erinaceomorpha) as inferred from a multigene study. Biological Journal of the Linnean Society 112(3): 499-519.

Berggren K. T., Ellegren H., Hewitt G. M., Seddon J. M. 2005. Understanding the phylogeographic patterns of European hedgehogs, *Erinaceus concolor* and *E. europaeus* using the MHC. Heredity 95(1): 84-90.

Bolfíková B., Hulva P. 2012. Microevolution of sympatry: landscape genetics of hedgehogs *Erinaceus europaeus* and *E. roumanicus* in Central Europe. Heredity 108(3): 248-255.

Clement M., Posada D., Crandall K.A. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1659.

Darriba D., Taboada G. L., Doallo R., Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature methods 9(8): 772-780.

Derouiche L., Bouhadad R., Fernandes C. 2016. Mitochondrial DNA and morphological analysis of hedgehogs (Eulipotyphla: Erinaceidae) in Algeria. Biochemical Systematics and Ecology (64): 57-64.

DeSalle R., Amato G. 2004. The expansion of conservation genetics. Nature Reviews Genetics 5(9): 702.

Djan M., Stefanović M., Veličković N., Ćosić N., Ćirović D. 2017. Mitochondrial genetic diversity and structuring of northern white-breasted hedgehogs from the Central Balkans. Turkish Journal of Zoology 41: 774-782.

Etemad E. 1985. Mammals of Iran. Tehran, National council of Nature conservation.

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

Frankham R., Ballou J. D., Briscoe D. A. 2004. A primer of conservation genetics: Cambridge University Press.

Fu Y.X. 1997. Statistical tests of neutrality of mutations against population growth,

hitchhiking and background selection. Genetics 147(2): 915-925.

Habib M., Lakra W.S., Mohindra V., Khare P., Barman A.S., Sing A. 2011. Evaluation of cytochrome b mtDNA sequences in genetic diversity studies of *Channa marulius* (Channidae: Perciformes). Molecular Biology Research 38: 841–846.

Harley E. H., Baumgarten I., Cunningham J., O'ryan C. 2005. Genetic variation and population structure in remnant populations of black rhinoceros, *Diceros bicornis*, in Africa. Molecular Ecology 14(10): 2981-2990.

He K., Chen J.H., Gould G.C., Yamaguchi N., Ai H.S., Wang Y.X., Zhang Y.P., Jiang X.L. 2012. An estimation of Erinaceidae phylogeny: a combined analysis approach. PloS One 7: e39304.

Hewitt G.M. 2000. The genetic legacy of the Quaternary ice ages. Nature 405: 907–913.

Huelsenbeck J. P., Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17(8): 754-755.

Huson D.H., Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Molecular Biology Evolution 23: 254–267.

Karami M., Huttere R., Siahsarvie R., Benda P. 2002. Annotated check-list of the mammals of Iran.

Khosravi R., Hemami M.R., Malekian M., Silva T.L., Rezaei H. R., Brito J. C. 2018. Effect of landscape features on genetic structure of the goitered gazelle (*Gazella subgutturosa*) in Central Iran. Conservation Genetics 19(2): 323-336.

Lerp H., Wronski T., Pfenninger M., Plath M. 2011. A phylogeographic framework for the conservation of Saharan and Arabian Dorcas gazelles (Artiodactyla: Bovidae). Organisms Diversity & Evolution 11(4): 317-321.

Librado P., Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25(11): 1451-1452.

Lynch M., Crease T. 1990. The analysis of population survey data on DNA sequence variation. Molecular biology and evolution 7(4): 377-394.

McVeigh H. P., Davidson W. S. 1991. A salmon id phylogeny inferred from mitochondrial cytochrome b gene sequences. Journal of Fish Biology 39(sA): 277-282.

Nei M. 1987. Molecular evolutionary genetics: Columbia university press.

Pablos M. T. P., Saragusty J., Santiago-Moreno J., Stagegaard J., Göritz F., Hildebrandt T. B., Hermes R. 2015. Cryopreservation of onager (*Equus hemionus onager*) epididymal spermatozoa. Journal of Zoo and Wildlife Medicine 46(3): 517-525.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Randi E., Lucchini V. 2002. Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) population by Bayesian admixture analyses of microsatellite variation. Conservation Genetics 3: 31-45.

Schneider S., Excoffier L. 1999. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152(3): 1079-1089.

Seddon J.M., Santucci F., Reeve N.J., Hewitt G.M. 2001. DNA footprints of European hedgehogs, *Erinaceus europaeus* and *E. concolor*: Pleistocene refugia, postglacial expansion and colonization routes. Molecular Ecology 10: 2187–2198.

Silva T. L., Godinho R., Castro D., Abáigar T., Brito J. C., Alves P. C. 2015. Genetic

identification of endangered North African ungulates using noninvasive sampling. Molecular Ecology Resources 15(3): 652-661.

Spielman D., Brook B.W., Frankham R. 2004. Most species are not driven to extinction before genetic factors impact them. Population Biology 101: 15261-15264.

Stefanović M., Veličković N., Ćosić N., Ćirović D., Đan M. 2016. Genetic variability of northern white-breasted hedgehogs (*erinaceus roumanicus*) from serbia. Balkan Journal of Wildlife Research 3 (1): 1-6.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3): 585-595.

Thirstrup J.P., Pertoldi C., Loeschcke V. 2008. Genetic aalysis, breed assignment and conservation priorities of three native Danish horse breeds. Animal Genetics 39: 496- 505.

Thompson J.D., Higgins D.G., Gibson T.J. 1994. improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Research 22(22): 4673-4680.