DNA Analysis of Brown Bear fecal samples from Kintrishi, Machakhela, Mtirala, Algeti, Vashlovani and BKNP (partially) protected areas

Introduction

The brown bear (*Ursus arctos*, Linnaeus) is one of the most widespread large carnivores in the world. The species inhabits Europe, Asia and North America. According to the International Union for Conservation of Nature (IUCN), as of 2017, the brown bear was assessed as a species of Least Concern with a stable population trend. This assessment indicates that, despite facing various threats, the overall population of brown bears is not currently at risk of significant decline. However, habitat of brown bears in Europe is fragmented which can lead to challenges in genetic diversity, movement and long-term survival of the species in certain areas (Swenson et al. 2000).

Brown bear was abundant and widespread throughout the Caucasus in the 18th and 19th centuries (Dinnik 1914). Since the late 19th century it has declined in numbers, primarily due to habitat loss and unregulated hunting. Georgian populations of the species survives in wooded and rugged habitats and protected areas. The country hosts three populations inhabiting in Great Caucasus, Lesser Caucasus, and most Eastern part of the country, Vashlovani national park and its surrounding areas (Lortkipanidze 2010). Brown bear is endangered in the country (Georgian Red List (doc # 2356. 2014). In recent decades, various management and protection approaches have been suggested, including human-bear conflict management plans, habitat conservation, and management plans for national parks with a special focus on large carnivores. In 2004, the first genetic monitoring efforts were implemented in Borjom-Kharagauli National Park by NACRES (Murtskhvaladze & Tarkhnishvili 2006).

In the decision-making process concerning effective conservation strategies for Georgia's brown bear population, assessing genetic diversity, structure, demographic dynamics and migration patterns holds paramount importance for the species' long-term survival. These assessments play a key role in determining susceptibility to stochastic mortality factors, and ultimately the risk of population extinction.

Non-invasive genetic methods (based on scat samples) are commonly used in management of rare species, especially in carnivore populations. Large mammals, particularly those, like the brown bear, living in mountainous and forested areas, are always difficult to observe: their crepuscular habit, low density, large home ranges, and elusive behaviour make it difficult and expensive to use traditional field surveys. The non-invasive genetic method does not require live trapping of animals and target species can be studied without physical capturing and thereby behavioural responses and risks trap-related injures are reduced.

Genetic characterisation of target populations can be performed using various genetic markers. Microsatellite loci have been widely adopted in non-invasive genetic monitoring. Microsatellites, also known as short tandem repeats (STRs) or simple sequence repeats (SSRs), are repeating sequences of 2-6 base pairs of DNA (Ellegren 2004). These repeats are scattered throughout the genome and are highly polymorphic, meaning they vary greatly in length between individuals. This high level of variability makes microsatellites particularly useful as genetic markers for a variety of applications (Barlow 2023). The method is cost effective and can be used for paternity studies, kinship analysis, for estimation population size, genetic variability, population dynamics. Microsatellites are suitable for examining populations with low genetic diversity and can be successfully amplified from low-quality DNA, which is frequently obtained from non-invasive samples (Murphy et al. 2002). However, analysing a large number of loci is costly and raises the possibility of genotyping errors. Consequently, researchers often choose a limited number of loci. It is essential to carefully select these markers to ensure precise and consistent estimates of population structure, genetic diversity, and individual assignment.

The aim of the study was to obtain the following information:

(a) Identify unique bear individuals including their sex from 403 scat samples;
(b) Evaluate relatedness between bear individuals within and between national parks;
(c) Estimate the genotyping success rate within national parks and determine if the sample age (freshness) affects genotyping success.
(d) Compare results of the current study, in particular, how many bears were recaptured from the study conducted by NACRES in 2014.

Materials and methods

Sample collection

403 brown bear scat samples were provided by NACRES in 20023. Samples were stored in alcohol and preserved in cold temperature prior to lab work. Along with the samples, a database containing sample IDs, collection dates, and collector names was sent by NACRES (file Data Final_Bear.xlsx).

DNA extraction

DNA was extracted from the samples using a Ox-GEn scat extraction kit according manufacturers instruction (appendix 1). To check for contamination and pipetting error, negative control was used per each extraction procedure. Extracted DNA was eluted in 60 ul AE buffer.

Genotyping - Nuclear microsatellite markers

In our previous work (Murtskhvaladze et al 2010), we defined a set of loci that best identifies unique genotypes (i.e. individual bears). Sex identification was performed using specific primers that co-amplify a bear specific Y marker and a bear specific X marker. If the samples is a male bear heterozygote alleles appears and homozygote allele appears in the case of a female bear. A set of eight loci: Cxx20, G10B, G10C, G10L, G10P, G10J, Mu50, Mu59 (Waits et al. 2000; Jackson et al. 2008) was sufficient to distinguish between full siblings at P (IDsib) < 0.0001.

Polymerase chain reaction (PCR) and loading multiplexes were developed to reduce the time and cost of genetic analyses. Nine loci were amplified in two multiplex PCR reactions using Qiagen multiplex kit. Thermal cycling was performed at 95 C for 15 min, at a touchdown of 15 cycles of 94 C for 30 s, 57.3 C for 90 s (with a stepwise decrease of 0.4 C at each cycle), 72 C for 1 min; then 22 amplification cycles of 94 C for 1 min, 52.5 C for 90 s, and 72C for 1 min, followed by a final extension at 60 C for 30 min.

Amplified DNA was run on SeqStudio Thermofisher, using deionized Formamide and Genescan size standard Liz 500 (Applied Biosystems Inc., Foster City, CA, USA) at Genome Laboratory, R. Lugar Center, National Centre for Disease Control and Public Health (NCDC).

Genotypes were screened using Genemapper v 5.0 software package (Perkin Elmer, Waltham, MA, USA). All these eight loci were amplified and screened at least three times, genotyping errors (allelic dropouts: ADO and false alleles: FA and genotype reliability: RCI) were calculated (Waits et al. 2001; Miller et al. 2002, RelioType, GIMLET). RCI threshold value was set as 95%. In final data set, low quality and putatively mixed samples were excluded from further analyses. Genotyping was repeated for the loci of samples which shows discrepancy between the first two PCR assays.

Data analysis

Computer softwares: ML-Relate (Kalinowski 2006) and PAST v. 2.17c (Hammer et al 2001) were used for calculations of relatedness and genetic distances. To analyse genetic variability across spatial populations (according to sampling data), we performed analysis of molecular variance (AMOVA), using ARLEQUIN v3.1 (Excoffier et al. 2005). Each subpopulation was tested for heterozygosity excess to detect recent population bottlenecks, using the program BOTTLENECK under the two-phase model of microsatellite evolution with 10%of the infinite allele model and 90% of the stepwise mutation model (Luikart & Cornuet 1998). To detect genetic structure in our sample without the prior identification of subpopulations, we used a Bayesian clustering approach implemented in the software STRUCTURE (Pritchard et al. 2000). This software helps infer and delineate the true number of genetically homogenous groups of sampled individuals, K from their genotypes at multiple loci, and assign the individuals to the inferred groups. We used the admixture model, and the option of correlated allele frequencies between populations, setting the burn-in

period and the length of MCMC to 100 000 each. This model considers that individuals can have an admixed ancestry such that each individual can inherit a fraction of their genome from ancestors in population k. We tested the range of Ks from 1 to 10, To quantify the statistics of the posterior probability of population structure for a given K, we performed 10 independent runs for each K. The true number of K was selected using (i) plot of the smallest value of –2 log Pr(X/K) for each of the runs (Pritchard et al. 2000) and (ii) the rate of change in the log probability of the data (DK) between successive values (Evanno et al. 2005).

Results

All markers were amplified with an overall genotyping success rate 99.8%. All 8 loci were polymorphic, no markers were deviated from Hardy-Weinberg equation.

Summary of screened samples is given in table 1. We successfully screened 347 samples, each for more than 5 loci. It is noteworthy that the field survey team collected and stored fecal samples according to a standardised protocol, ensuring high-quality data collection. Only a few of these samples were difficult to score. However, an exception was found in the samples from BKNP, where 32% of the samples contained fragmented or insufficient DNA, making them challenging to analyse.

We compiled a comprehensive database, based on file Data Final_Bear.xlsx. that includes the following information for each unique bear individual: field ID of the sample, origin, and the names of the field officers. This database can be viewed in details from the file Ua_24.xls. All unsuccessfully screened samples are listed as unidentified in the file Ua_24.xls. Question marks (?) indicates name of unknown filed officer.

Table 1. Result summary

Sample location	Samples provided	Successfully Screened	Fragmented/no DNA	Number of Bear Individuals	Females	Males
ANP	8	8	0	4	0	4
VNP	36	31	5	16	3	13
KNP*	56	54	2	33	10	23
MaNP*	83	75	9	51	13	38
MtNP*	116	108	9	68	23	45
BKNP	104	71	31	51	18	33
Possible migrants*	-	-	-	5	1	4
Total	403	347	56	228	68	160
*	MaNP, MtNP; KNP share bear individuals (Possible migrants) see discussion for details					

228 unique bear individuals were identified from 403 fecal samples. Among them 160 were males and 68 females (table 1). The relationship between each pair of individuals that has the highest likelihood among the four following relationship: Unrelated (U), half sibs (HS), full subs (FS), parent/offspring (PO) is given in file Relationship_Matrix.xlx.

All samples were grouped according their origin, cluster analysis showed that there were four clusters Fig.1 The mean LnP(K) reaches its peak at k=4, suggesting that this value has the highest likelihood.

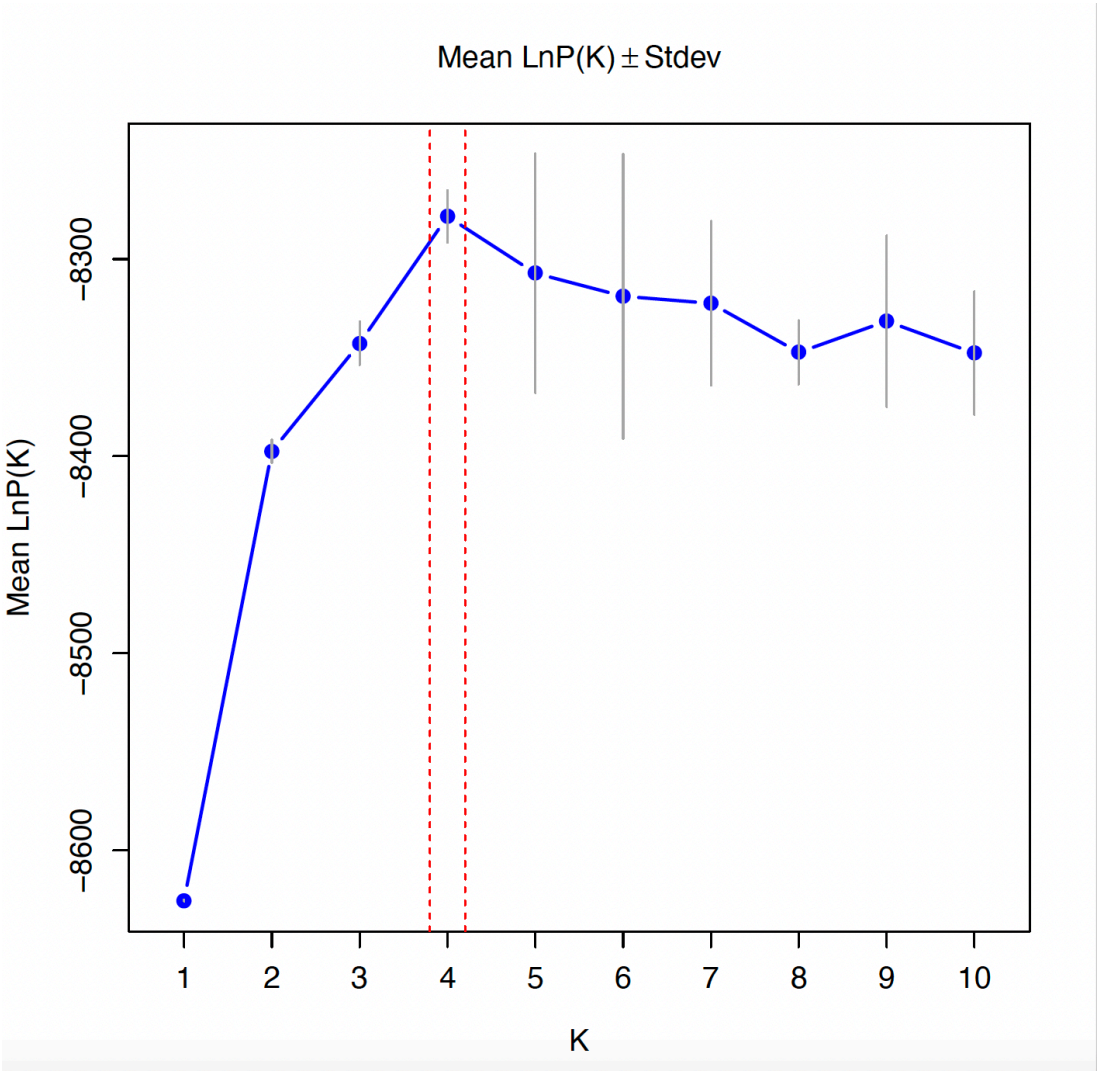

Figure 1. Evaluation of optimal number of K. Y-axis mean of the natural logarithm of *P*(*K*), model-based clustering for evaluation the likelihood of different models. X-axis- values of *K* the number of possible clusters (1-10) in a given data. The blue points show the mean LnP(K) for each value of *K*. The vertical error bars represent the standard deviation (Stdev) around these means. The vertical red dashed lines highlight a specific value of *K*, which appears to be around 4 in this plot.

Genetic structure of studied individuals cal be visual viewed on fig 2. Each colour bar represents each bear individual, each colour showed likelihood of membership of each cluster.

Analysis of population structure showed that the the studied population were divided into four groups. However, this result should be interpreted with caution because of the relatively small sample size from ANP and VNP. Relationship between bear individuals are shown in file Bear_tree.jpg.

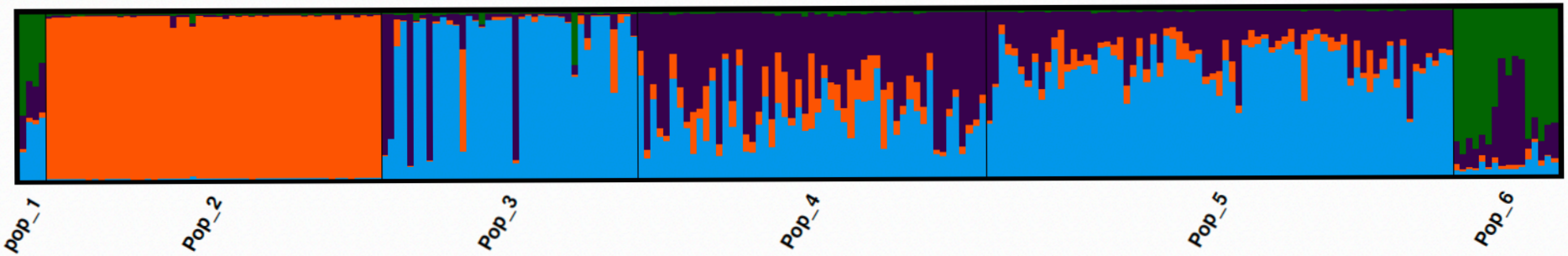

Figure 2. **Population Structure Analysis**: Each individual is represented by a vertical bar that is divided by K colored segments representing the likelihood of a membership to each cluster. The population codes are as follows: Pop_1 consists of individuals from ANP, pop_2 from BKNP, pop_3 from KNP, pop_4 from MaNP, pop_5 from MtNP, and pop_6 from VNP, respectively.

Average F-statistics for studied population is Fst = 0.04; Fit = 0.04; Fis = 0.161, Fit= 0.2. See discussion section for details.

We could not recaptured bears identified in 2014 in the sample set collected in 2023.

Population pairwise Fst values are given in table 2. * indicates P values > 0.05

Table 2. Population pairwise Fst values, * P> 0.5

Population	ANP	BKNP	KNP	MaNP	MtNP	VNP
ANP	0					
BKNP	0.093*	0				
KNP	0.056	0.078*	0			
MaNP	0.029	0.063*	0.021*	0		
MtNP	0.038	0.061*	0.007*	0.011*	0	
VNP	0.067*	0.07*	0.063*	0.034*	0.05*	0

Discussion

Bear077 and Bear106 were identified from KNP and MaNP; while Bear080 and Bear085 from KNP and MtNP; Bear086 from KNP, MaNP and MtNP. We cannot confirm whether these individuals migrate across the mentioned national parks, or there was mislabeling/contamination during sampling and storage. It is highly recommended to conduct additional research to answer this question.

Comparison of sample “freshness” and success rate was impossible to done due to lack of data, not all samples have included this type of information. We are unable to estimate how fresh were scat samples in the moment of collection, after field survey the samples were stored in alcohol which caused any indication of the freshness to be lost.

The sex ratio was biased towards males, among screened individuals 70% were males and 30% females. This disparity can be explained by several factors: a) Male mobility: Male bears tend to have larger home ranges and are more likely to travel greater distances in search of food, mates, or new territories. This increased mobility makes them more likely to encounter and be captured by screening efforts. b) Seasonal Factors: Seasonal behaviours can also contribute to this difference. During certain times of the year, male bears may be more active and visible, e.g. during mating season when they are actively searching for females. In contrast, females may be less active, especially if they are caring for cubs or if it is a period when they are less likely to move around extensively. c) Female Vulnerability: Female bears may be more vulnerable to various threats that

reduce their visibility or likelihood of being screened. They might be more cautious and avoid areas where they could be detected by humans, especially if they have cubs to protect. Additionally, females with cubs might choose denser, more secluded habitats, further reducing their chances of being screened.

Average F-statistics indicates low genetic differentiation between studied sub-populations, in our case Fst = 0.04; Populations with this value are relatively similar genetically. There is no significant isolation between them. This could be due to gene flow between populations, or large population sizes reducing the effect of genetic drift.

Fis measures inbreeding within subpopulations. Fis value 0.161 indicates a moderate inbreeding. Moderate overall inbreeding and reduced genetic diversity make the population more vulnerable to threats such as low birth rates, developmental problems, low survival rates, and increased susceptibility to infections and parasites. However, this statement should be interpreted with caution due to potential sampling bias. Field officers may sometimes follow movement paths and collect samples from closely related groups of bears, which can skew the results and overestimate the degree of inbreeding. To obtain more accurate assessments, regular monitoring, such as every five years, and random sampling should be implemented to clarify this conclusion.

The parameter Fit shows genetic differentiation within a population relative to the total population. It's calculated as the correlation between alleles within individuals relative to the entire population. The value Fit 0.2 suggests that there's some departure from panmixia within the entire population. This can be attributed to the small sample sizes from ANP and VPA.

High P >0.5 for Fst among geographicaly distant populations in particularly VPA and the others (table 2), can be explained with relatively small sample size (16 bears from VPA) and/or/ by continuity ob populations in the past.

Recommendations

In conclusion we can say that Georgian brown bear population is genetically diverse. It is recommended to conduct additional research to estimate sex ration especially in areas where a small number of individuals were recorded.

Conservation efforts need to consider within-subpopulation and between-subpopulation characteristics, seasonal migrations, factors to improve genetic health. It is necessary to protect migration corridors, and regular monitoring (every five to seven years) is recommended to ensure the long-term maintenance of a genetically healthy and stable bear population in Georgia.

References:

Barlow 2023. Fundamentals of Conservation Biology.

Dinnik NI (1914) Zveri Kavkaza, Khishchnye. Khishnie, 2, 342–401. Tiflis: Tipografiya K. P. Kozlovskogo [Animals of the Caucasus. Carnivores, 2, 342–401. Tbilisi: Typography of K.P. Kozlovskiy]. (in Russian).

Ellegren H. (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5, 435–445.

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.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online, 1, 47–50.

Lortkipanidze, B. (2010). Brown bear distribution and status in the South Caucasus. *Ursus*, *21*(1), 97-103.

Luikart G, Cornuet J (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology, 12, 228–237.

Murphy MA, Waits LP, Kendall K. (2000) Quantitative evaluation of faecal drying methods for brown bear DNA analysis. Wildlife Society Bulletin, 28, 951–957.

Murphy M, Waits LP, Kendall K, Wasser S, Higbee J, Bogden R (2002) An evaluation of long-term preservation methods for brown bear (Ursus arctos) faecal DNA samples. Conservation Genetics, 3, 435–440.

Murtskhvaladze M, Gavashelishvili A, Tarkhnishvili D (2010) Geographic and genetic boundaries of brown bear (Ursus arctos) population in the Caucasus. Molecular Ecology 19, 1829–1841.

Murtskhvaladze M, Tarkhishvili D. (2004) Estimation brown bear abundance and population structure in Borjom_Kharagauli National Park with molecular-genetic method. Proc. Georgian Acad. Sci. Biol. See B. Vol 4, No 4 2006.

Pritchard KJ, Matthew S, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945–959.

Swenson JE, Gerstl N, Dahle B, Zedrosser A (2000) Action Plan for the conservation of the Brown Bear (Ursus arctos) in Europe. Nature and Environment, Council of Europe Publishing. 114, 15–18.

Waits LP, Taberlet P, Swenson J, Sandegren A, Franzen R (2000) Genetic diversity and gene flow in the Scandinavian brown bear. Molecular Ecology, 9, 421–431.

Waits LP, Luikart G, Taberlet P (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular Ecology, 10, 249–256.

Fecal DNA Purification protocol

Step 1:

- Weigh ≤ 20-60mg of fecal sample and place in a 2 ml in a microcentrifuge tube, illed with OxGEn Glass Beads;
- Add 500 µl of **Solution A1, 100 Solution A2** the supernatant. Mix thoroughly or vortex 10s, then add 100µl Solution A3, and vortex for 30s.
- Add 20µl of Proteinase K, and mix thoroughly by bead beating machine/ vortexing 15 min 5000 rpm.
- Centrifuge the tube for 5-10s 8000rpm remove drops from the lid.
- Incubate for 15 min at 56^0 C (overnight)

Step 2 :

- Freeze the sample for 20 min at -20^0 C
- Cool down to room temperature (RT), centrifuge at 13000 rpm for 5 min
- Transfer 200 µl of supernatant into a new 2ml microcentrifuge tube, add 200 µl Solution B and 200 µl 95% Ethanol. Vortex 2min Centrifuge the tube for 5-10s 8000rpm remove drops from the lid.
- Transfer 400 µl of lysate to a G-spin-column, centrifuge at 8000 rpm for 1. **Discard the flow trough;**
- Repeat step 9 with the remaining lysate until the entire lysate has passed through the G-spin/column. **Change the collection tube**
- Add 600 µl of Solution W1 and centrifuge at 13000 rpm for 2 min **Change the collection tube;**
- Add 600 µl of Solution W2 and centrifuge at 13000 rpm for 2 min **Discard the collection tube;**
- Remove residual buffer by centrifuging at centrifuge at 13000 rpm for 2 min **Discard the collection tube;**

Step 3 :

- Transfer the G-spin/column into a new 1.5 ml microcentrifuge tube;
- Add 60 µl of preheated (56^0 C) Solution E to the G-spin/column.
- Incubate for 5 min at room temperature (RT).
- Elute the DNA by centrifuging at 13000 for 2 minutes.
- Store extracted DNA **-20^0 C**