Pakistan J. Zool., vol. 51(5), pp 1-4, 2019. **DOI**: http://dx.doi.org/10.17582/journal.pjz/2019.51.5.sc6

Short Communication

Genetic Diversity and Population Structure of Seven Tibet Yak Ecotype Populations using Microsatellite markers

Yan-Bin Zhu[1], Wang-Dui Basang[1], Zhan-Dui Pingcuo[1], Yang-Ji Cidan[1], Sang Luo[1], Dun-Zhu Luosang[1], Yang-La Dawa[1] and Guang-Xin E[2,3,*]

[1]Institute of Animal Husbandryand Veterinary Medicine, Tibet Academy of Agriculture and Animal Husandry Science, Lhasa, 850009, China
[2]State Key Laboratory of Barley and Yak Germplasm Resources and Genetic Improvement (Tibet Academy of Agricultural and Animal Husbandry Sciences (TAAAS)), Lhasa, 850002, China
[3]Chongqing Key Laboratory of Forage & Herbivore, Chongqing Engineering Research Centre for Herbivores Resource Protection and Utilization, College of Animal Science and Technology, Southwest University, Chongqing, China

Yan-Bin Zhu and Wang-Dui Basang contributed equally to this article.

Article Information
Received 04 February 2019
Revised 11 April 2019
Accepted 14 May 2019
Available online 27 June 2019

Authors' Contributions
W-DB, Y-BZ and G-XE conceived and designed the experiments. Y-BZ, Z-DP, Y-JC, D-ZL and SL performed the lab work. Y-LD, Y-BZ and G-XE analyzed the data and wrote the paper.

Key words
Domestic Yak, Diversity, Microsatellite, Population Structure, Tibet

ABSTRACT

The aim of the present study was to perform genetic diversity and population structure estimation on 382 individuals from seven Chinese Tibetan ecotype yak population using twenty-one microsatellites. The results revealed that the HO ranged from 0.4854±0.0194 in NNV to 0.6086±0.0267 in YRY, and the NA ranged from 3.86±1.98 in XMY to 6.05±3.37 in NNV. The least number of markers which deviated from Hardy-Weinberg equilibrium within the population was W, and largest number of deviations was in the NNV population. Consistent of phylogenetic relationships between the seven populations was identified by Phylogenetic N-J network (Reynold's genetic distance), FST and Principal components factor (PCA) analysis. These analyses inferred that the population cluster data was not only consistent with the populations' geographic habits but could also be influenced by artificial selection and feeding style. Lastly, two credible genetic backgrounds were identified from the yak populations in this study using STRUCTURE software, which corresponded to previous knowledge about different molecular genetic markers. Therefore, unexpectedly, our study indicated that the diversity of some of the populations was decreased, leaving us to improve and refine our conversional strategy. In addition, this resulted in a greater understanding of human yak phylogenetic differentiation as well as providing data support for understanding the evolution and migration of yak population in future studies.

Yak (Bos grunniens) are found on the Qinghai-Tibetan Plateau (Altitude > 3500m) and have more than 5,000 years of domesticated history (Zhang *et al*., 2014). They are an important farm animal species of Qinghai-Tibet Plateau both economically and culturally. Recently, a number of studies have estimated the diversity of yak populations using different genetic markers, such as the Y-Chromosome marker (Ma *et al*., 2015, 2018), mitochondrial DNA (Song *et al*., 2015; Basang *et al*., 2018), microsatellite (Zhang *et al*., 2008; Pei *et al*., 2018), and wide-genome sequences (Qiu *et al*, 2015; Lan *et al*., 2018).

Naqu is a city under the jurisdiction of the Tibet Autonomous Region. It is located in the north of Tibet, between the Tanggula Mountains, the Nyainqentanglha Mountains and the Gangdise Mountains. It is about 1156 kilometers long from east to west and 760 kilometers wide from north to south. It is connected to Changdu City in the east, Ali in the west, and Lhasa and Linzhi in the south. The city is adjacent to the north, bordering Xinjiang Uygur Autonomous Region and Qinghai Province. The total area is 369,674 square kilometers, with a total population of 501,300. At the same time, the existing yak population of Nagqu is about 3 million, which is the main producing area of yak in Qinghai-Tibet Plateau.

Here, to promote the development of local domestic Yak breeding in Naqu region, it is urgent that the genetic diversity be evaluated. In this study, we aim to estimate the genetic diversity of seven local yak population (six from

* Corresponding author: eguangxin@126.com
0030-9923/2019/0005-0001 $ 9.00/0

Table I.- Original distribution and sample information for the seven Tibetan yak populations.

Name	Code	SZ	N	E	Native Location
Xiama Population	XMY	20	30°47′51.61″	92°40′36.28″	Xiama No.3 villege, Jiali, Naqu
Jiali Population	JLY	111	30°38′36.76″	93°13′51.74″	Jiali Farm, Naqu
Neirong Farm Population	NRF	61	32°06′36.73″	92°18′7.68″	Neirong Farm, Neirong, Naqu
Neirong nine village Population	NNV	47	31°44′44.43″	92°04′4.36″	Nima No.9 Villege, Neirong, Naqu
Neirong eleven village Population	NEV	59	31°44′46.32″	92°04′2.03″	Nima No.11 Villege, Neirong, Naqu
Yare Population	YRY	23	31°29′42.03″	82°19′59.05″	Yare, Geji, Arli
Dangxiong Population	DXY	61	30°31′39.23″	91°17′20.96″	Longren, Dangxiong, Naqu

SZ, Sample size; N, North latitude; E, East longitude and Code, short name of breed.

Naqu, and one from Ali region) using 21 microsatellite markers, and the evolution of their genetic divergence and population structure.

Material and methods

The experimental conditions used in this study were approved by the Committee on the Ethics of Animal Experiments of Southwest University (No. [2007] 3) and accordant with the Animal Protection Law of China. DNA of 382 individuals from seven Tibet yak populations (Table I) was extracted by a standard phenol-chloroform protocol from vein blood samples.

All individuals were genotyped at the 21 microsatellite markers (Supplementary Table I) as well as with genotyping being performed on a Genetic Analyzer 3130xl (Applied Biosystems, US), a detailed description of the genotyping can be found in E *et al*. (2018). Conventional genetic diversity parameters, including observed (*Ho*) heterozygosity, expected heterozygosity (H_E), mean number of alleles (N_A), and polymorphism information content (*PIC*), were estimated with the Microsatellite Toolkit (Park, 2008). Deviations of markers from Hardy-Weinberg equilibrium (*HWE*) within populations were identified using GENEPOP 3.4 software (Raymond and Rousset, 1995). The inbreeding coefficient (F_{IS}) was estimated, and Fisher's exact test with Bonferroni correction was performed using FSTAT 2.9.3.2 (Goudet, 1995). Pairwise differences between populations (F_{ST}) and loci under selection were identified using Arlequin software version 3.5.1.3 (Excoffier and Lischer, 2010).

Principal components factor analysis (PCA) performed with R-Script. Bayesian clustering algorithm was implemented in STRUCTURE 2.3.3 (Pritchard *et al*., 2000; Falush *et al*., 2003) with 50000 burn, and 100000 Markov Chain Monte Carlo (MCMC) in 100 iterations from K=1 to K=7. To estimate the most optimal K reference as deta K = m|L"(K)|/s|L(K)| by Structure Harvester (Earl and vonHoldt, 2012).

Results and discussion

In total, 193 alleles were identified in 7 local yak populations across 21 microsatellites. Across populations, an average of 9.2 alleles of each marker were observed, ranging from 2 in BGR3001 to 21 in TGLA53. The H_O and H_E of each marker within all individuals was 0.5239 (from 0.0304 (BGR3012) to 0.735 (CSRD247)) and 0.555 (from 0.034 (BGR3012) to 0.7986 (SPS115)), respectively. The mean *PIC* across markers was 0.4969 and ranged from 0.0323 (BGR3012) to 0.8091 (TGLA73) across populations (Supplementary Table II). Our analysis not only indicated that most of the microsatellites which were used to estimate the diversity of yak populations would show high levels of polymorphism, but also indicated that the markers used in this study are qualified to represent the genetic diversity of these yak populations.

Across markers, the H_E within a breed ranged from 0.5333±0.0646 in XMY to 0.5832±0.0524 in YRY. The H_O ranged from 0.4854±0.0194 in NNV to 0.6086±0.0267 in YRY, and the N_A ranged from 3.86±1.98 in XMY to 6.05±3.37 in NNV (Table II). This indicated that those yak ecotype populations were still existing with a high level of diversity within the population.

Based on the Hardy-Weinberg equilibrium (HWE) analysis, nearly all markers in YRY populations were at *HWE*, expected BM2113. Only two markers deviated from HWE (*dHWE*) in B (BGR3004, ILSTS008), and DXY (BM2113, AGLA293) and other four populations exhibited 6 to 9 markers of *dHWE*. The range of F_{IS} within populations was from 0.153 in NNV to -0.059 in YRY population. Three of these populations (JLY, NEV, NNV) had P-value for F_{IS} considered significant at the adjusted nominal level (5%, $P < 0.00034$). According to the *HWE* and F_{IS} results, inbreeding did not occur in DWY, YRY and XMY population. This indicates that the other four populations are under potential risk of becoming endangered. This should be a reminder to the government and management units involved that they should pay more attention to the protection of the genetic diversity in these populations and improve their current protection and conservation methods.

In addition, highest average number of pairwise differences within populations was observed in XMY

Table II.- Genetic diversity and pairwise differences of seven Tibet ecotype yak populations.

	Population genetic diversity						Pair-wise differences analysis						
Population	H_O (±SD)	H_E (±SD)	N_A (±SD)	F_{IS}	P-Value	dHWE	XMY	JLY	NRF	NNV	NEV	YRY	DXY
XMY	0.51±0.03	0.53±0.07	3.86±1.98	0.04	0.17	2	1.49	1.46*	1.57*	1.43*	1.60*	1.59*	1.57*
JLY	0.51±0.01	0.54±0.06	5.76±3.63	0.06	0.0003#	8	0.09*	1.23	1.32*	1.22	1.34*	1.32*	1.31*
NRF	0.51±0.02	0.55±0.06	4.95±2.87	0.07	0.001	8	0.16*	0.03*	1.34	1.30*	1.36	1.35	1.32
NNV	0.49±0.02	0.57±0.06	6.05±3.37	0.15	0.0003#	9	0.09*	0.006	0.03*	1.21	1.31*	1.29*	1.29*
NEV	0.49±0.02	0.56±0.06	5.19±3.14	0.11	0.0003#	6	0.17*	0.04*	0.004	0.01*	1.37	1.36	1.34
YRY	0.61±0.03	0.58±0.05	4.24±2.21	-0.06	0.78	1	0.19*	0.04*	0.01	0.03*	0.02	1.32	1.33
DXY	0.54±0.02	0.55±0.06	4.95±2.94	0.01	0.30	2	0.17*	0.04*	-0.006	0.03*	0.002	0.02	1.31

Pa, number of private alleles; *d*HWE, number of populations that deviated ($P < 0.01$) from Hardy-Weinberg equilibrium; #, indicates the adjusted nominal level (5%) for one table is 0.0006 based on 1680 randomizations of *P*-values for F_{IS}.

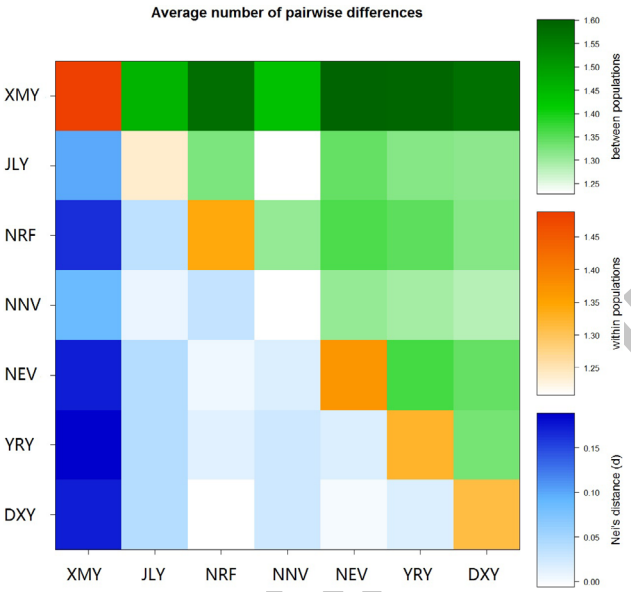

Fig. 1. Pairwise differences analysis between seven yak populations. About pairwise differences analysis as (1) above diagonal: average number of pairwise differences between populations (πXY); (2) Diagonal elements: average number of pairwise differences within populations (πX); and (3) Below diagonal: corrected average pairwise differences (πXY-(πX+πY).

(πX=1.48846) and the lowest was observed in NNV (πX=1.20773). The results of genetic divergence between populations (F_{ST}) based on corrected average pairwise differences (πXY-(πX+πY) and average number of pairwise differences between populations (πXY) indicated that the XMY population was carrying a higher divergence than the others (Table II, Fig. 1).

The genetic pattern of those populations from the phylogenetic Network tree can be seen in Figure 2A. The phylogenetic network revealed that XMY, and JLY population separated into Cluster I, which was easily understood being that the habitat of these three populations were around Jiali county. However, four populations, NRF, DXY, YRY, NEV population were separated into Cluster II, which was inconsistent with their geographic location, the YRY being from Ali region and others from the Naqu region of Tibet. In addition, the smaller sample size of this population (23, YRY) was not enough to present its correct population structure.

Regarding the genetic investigation of PCA (Fig. 2B) the investigation indicated that there could be a popular exchange of gene flow among those populations due to nomadic behavior, regular activity and human migration. Strangely, there are three populations NRF, NEV and NNV that had large divergences between each other. This could easily be caused by local artificial or cross breeding. What's more, due to the high occurrence of irregular breedable livestock selection human intervention was a possible reason leading to inbreeding within this population which was revealed from F_{IS} and *dHWD* in this study, this could also contribute to unreasonable population phylogenetic relationships within the populations.

Furthermore, we used STRUCTURE software to cluster the 7 Tibet ecotype populations into $2 \leq K \leq 7$ (Fig. 2C). The result revealed that the most credible *K* value was 2 by ΔK = m| L" (*K*)|/s|L(*K*)|, which was consistent with current knowledge that there are two genetic backgrounds within Chinese yak with different molecular markers, such as mitochondrial DNA (e.g., Basang *et al*., 2018), and Y –chromosome (e.g., Ma *et al*., 2018).

Conclusion

Seven Chinese Tibetan ecotype populations were genotyped using 21 microsatellites. The results indicated that the current genetic diversity in some of these populations is decreasing and conservation strategies should be improved What's more, the gene flow exchanged among those populations could be due to nomadic behavior

and normal activity etc. Last but not least two divergence genetic backgrounds were identified in this study which further corroborates the current knowledge of human yak phylogenetic differentiation as well as provides data support for understanding the evolution and migration of yak populations in future studies.

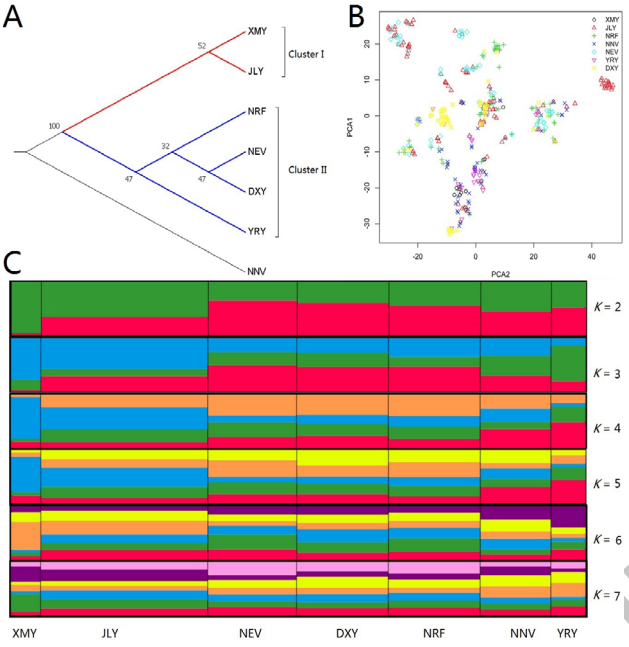

Fig. 2. Phylogenetic population structure of seven Tibet local yak populations. A, Phylogenetic network of 7 yak populations by Reynold's genetic distance; B, PCA pattern of 382 Tibet yak individuals from 7 populations; C, Cluster diagrams of 7 yak populations obtained using STRUCTURE.

Acknowledgements

This work was supported by National Technical System for Beef and Yak Industry (CARS-38), and the Open Project Program of State Key Laboratory of Barley and Yak Gemplasm Resources and Genetics Improvement (Tibet Academy of Agricultural and Animal Husbandry Sciences (TAAAS)), Lhasa Tibet 850002, China.

Conflict of interest

Authors declare no conflict of interest.

Supplementary material

There is supplementary material associated with this article. Access the material online at: http://dx.doi.org/10.17582/journal.pjz/2019.51.5.sc6

References

Basang, W.D., An, T.W., Danjiu, L.B., Zhu, Y.B., He, S.C., Luo, X.L., Ni, W.W., Wang, X., Cheng, S.Z., Wang, J. and E, G.X., 2018. *Pakistan J. Zool.*, **50**: 2051-2057. https://doi.org/10.17582/journal.pjz/2018.50.6.2051.2057

E, G.X., Zhao, Y.J., Chen, L.P., Ma, Y.H., Chu, M.X., Li, X.L., Hong, Q.H., Li, L.H., Guo, J.J., Zhu, L., Han, Y.G., Gao, H.J., Zhang, J.H., Jiang, H.Z., Jiang, C.D., Wang, G.F., Ren, H.X., Jin, M.L., Sun, Y.Z., Zhou, P. and Huang, Y.F., 2018. *Ecol Evol.,* **8**: 5111-5123.

Earl, D.A. and vonHoldt, B.M., 2012. *Conserv. Genet. Resour.*, **4:** 359-361.

Excoffier, L. and Lischer, H.E., 2010. *Mol. Ecol. Resour.*, **10**: 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Falush, D., Stephens, M. and Pritchard, J.K., 2003. *Genetics*, **164**: 1567–1587.Goudet, J., 1995. *J. Hered.*, **86**: 485-486.

Felsenstein, J., 2005. *PHYLIP* (Phylogeny Inference Package) version 3.6 Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Kovach, W.L., and Services, K.C., 2004. *MVSP- A multivariate statistical package for Windows*. ver 3.13m. http://www.kovcomp.co.uk/mvsp/index.html

Lan, D., Xiong, X., Mipam, T.D., Fu, C., Li, Q., Ai, Y., Hou, D., Chai, Z., Zhong, J. and Li, J., 2018. *G3 (Bethesda)*, **28**:945-952. https://doi.org/10.1534/g3.118.300572

Ma, Z.J., Xia, X.T., Chen, S.M., Zhao, X.C., Zeng, L.L., Xie, Y.L., Chao, S.Y., Xu, J.T., Sun, Y.G., Li, R.Z., Guanque, Z.X., Han, J.L. and Lei, C.Z., 2018. *Anim. Genet.*, **49**:618-622.

Ma, Z.J., Chen, S.M., Sun, Y.G., Xi, Y.L., Li, R.Z., Xu, J.T. and Lei, C.Z., 2015. *Genet. mol. Res.*, **14**:18859-62.

Pei, J., Bao, P., Chu, M., Liang, C., Ding, X., Wang, H., Wu, X., Guo, X. and Yan, P., 2018. *Peer J.*, **6**: e5946. https://doi.org/10.7717/peerj.5946

Pritchard, J.K., Stephens, M. and Donnely, P., 2000. *Genetics,* **155**: 945–959.

Qiu, Q., Wang, L., Wang, K., Yang, Y., Ma, T., Wang, Z., Zhang, X., Ni, Z., Hou, F., Long, R., Abbott, R., Lenstra, J. and Liu, J., 2015. *Nat. Commun.*, **22**:10283.

Raymond, M. and Rousset, F., 1995. *Heredity*, **86**: 248-249. https://doi.org/10.1093/oxfordjournals.jhered.a111573

Song, Q.Q., Chai, Z.X., Xin, J.W., Zhao, S.J., Ji, Q.M., Zhang, C.F., Ma, Z.J. and Zhong, J.C., 2015. *Genet. Mol. Res.*, **14**:1763-70. https://doi.org/10.4238/2015.March.13.3

Zhang, G., Chen, W., Xue, M., Wang, Z., Chang, H., Han, X., Liao, X. and Wang, D.J., 2008. *J. Genet. Genom.*, **35**: 233-238.

Pakistan J. Zool., vol. 51(5), pp 1-4, 2019. **DOI**: http://dx.doi.org/10.17582/journal.pjz/2019.51.5.sc6

Supplementary Material

Genetic Diversity and Population Structure of Seven Tibet Yak Ecotype Populations using Microsatellite markers

Yan-Bin Zhu[1], Wang-Dui Basang[1], Zhan-Dui Pingcuo[1], Yang-Ji Cidan[1], Sang Luo[1], Dun-Zhu Luosang[1], Yang-La Dawa[1] and Guang-Xin E[2,3,*]

[1]Institute of Animal Husbandryand Veterinary Medicine, Tibet Academy of Agriculture and Animal Husandry Science, Lhasa, 850009, China
[2]State Key Laboratory of Barley and Yak Germplasm Resources and Genetic Improvement (Tibet Academy of Agricultural and Animal Husbandry Sciences (TAAAS)), Lhasa, 850002, China
[3]Chongqing Key Laboratory of Forage & Herbivore, Chongqing Engineering Research Centre for Herbivores Resource Protection and Utilization, College of Animal Science and Technology, Southwest University, Chongqing, China

Yan-Bin Zhu & Wang-Dui Basang contributed equally to this article.

* Corresponding author: eguangxin@126.com
0030-9923/2019/0005-0001 $ 9.00/0

Supplementary Table I.- Primer information for the 21 genome microsatellites in Yak

Marker	Primer sequence (5′–3′)	Tm (°C)	FL(bp)	LM
BM2113	F: GCTGCCTTCTACCAAATACCC R: CTTAGACAACAGGGGTTTGG	57	123-151	FAM
SPS113	F: AAAGTGACACAACAGCTTCTCCAG R: AACGAGTGTCCTAGTTTGGCTGTG	55	123-151	ROX
TGLA122	F: CCCTCCTCCAGGTAAATCAGC R: AATCACATGGCAAATAAGTACATAC	54	143-175	FAM
TGLA126	F: CTAATTTAGAATGAGAGAGGCTTCT R: TTGGTCTCTATTCTCTGAATATTCC	55	107-121	FAM
ETH152	F: TACTCGTAGGGCAGGCTGCCTG R: GAGACCTCAGGGTTGGTGATCAG	56	194-212	ROX
ILSTS008	F: GAATCATGGATTTCTGGGG R: TAGCAGTGAGTGAGGTTGGC	58	175-187	FAM
ILSTS013	F: CTTGATCCTTATAGAACTGG R: ACACAAAATCAGATCAGTGG	58	121-135	FAM
ILSTSO50	F: AAATCAGACACCCAGTTTCC R: GTTTTTCTACACGAGTTGGC	55	161-183	ROX
BM1824	F: GAGCAAGGTGTTTTTCCAATC R: CATTCTCCAACTGCTTCCTTG	57	180-194	FAM
ETH225	F: GATCACCTTGCCACTATTTCCT R: ACATGACAGCCAGCTGCTACT	64	144-162	FAM

MGTG7	F: TTCATTGCAGCAACTATTTACAATAG R: TAAGTTCCCTGTATCATTTTTGAA	55	278-312	FAM
SPS115	F: AAAGTGACACAACAGCTTCTCCAG R: AACGAGTGTCCTAGTTTGGCTGTG	64	234-254	ROX
BGR3001	F: CTGGAGAAGGAAATGGCAAC R: AAGAGAAGCAGATGGCAGC	54	200-300	FAM
BGR3004	F: GGTTTGGGTTTACATCACTGGT R: CACTGGAGAAGGAAATGGCAAC	57	200-300	ROX
BGR3005	F: CGACTTCACTTTCACTTTTC R: TCCTCAATCTTTTCTCCTCT	52	200-300	FAM
BGR3011	F: TGTTTTGGTTCTTTGGTCTC R: CTTCACTTTCGCTTTTCACT	54	200-300	ROX
BGR3012	F: GACAGGAAGTAGGATGATGGTT R: TGGGAGATGGTGAAGGACAGGG	52	200-300	FAM
BGR3013	F: TCACTTTTCACTTTCCTGC R: GTAAGGGGAAAAAAGATGTG	53	200-300	ROX
BGR3019	F: CAGTTTGAAGGAGAAGCATG R: GTGCCAACTTCTCTTTTTCT	55	200-300	FAM
BGR3020	F: GACTTCCCTTTCACTTTTCAC R: AGAGGAACGCAACAGGGGAAT	51	200-300	ROX
TGLA53	F: GCTTTCAGAAATAGTTTGCATTCA R: ATCTTCACATGATATTACAGCAGA	55	143-191	FAM
ILSTS013	F: ACACAAAATCAGATCAGTGG R: CTTGATCCTTATAGAACTGG	58	121-135	FAM
TGLA73	F: GAGAATCACCTAGAGAGAGGCA R: CTTTCTCTTTAAATTCTATATGGT	55	111-143	ROX
AGLA293	F: GAAACTCAACCCAAGACAACTCAAG R: ATGACTTTATTCTCCACCTAGCAGA	55	210-240	FAM

F, forward primer; R, reverse primer; Tm, annealing temperature; FL, fragment length and LM, means label marker.

Supplementary Table II.- Genetic diversity statistics for loci across the even yak populations

Locus	H_O	H_E	PIC	Na
BGR3012	0.0304	0.034	0.0323	4
BM2113	0.6605	0.4273	0.3393	9
SPS115	0.729	0.7986	0.7542	18
BGR3001	0.1429	0.0952	0.0536	2
ETH152	0.7021	0.7943	0.7387	7
TGLA122	0.6754	0.7511	0.7072	11
BGR3004	0.555	0.6235	0.5365	10
ETH225	0.7043	0.7102	0.6512	9
MGTG7	0.5856	0.8441	0.8008	16
TGLA73	0.8152	0.8411	0.8091	16
AGLA293	0.2204	0.5423	0.4738	14
BGR3020	0.5091	0.4848	0.3753	5
ILSTSO50	0.6664	0.6827	0.6261	11
TGLA53	0.7505	0.7556	0.7112	21

BGR3005	0.0173	0.0172	0.0164	5
BGR3013	0.5015	0.5317	0.414	4
TGLA126	0.6854	0.6927	0.629	7
BGR3011	0.2656	0.3005	0.2481	3
BM1824	0.6983	0.704	0.6456	8
BGR3019	0.3983	0.46	0.4021	5
ILSTS008	0.6887	0.5745	0.4708	8
Mean	0.5239	0.5555	0.4969	9.2

Na, the total number alleles of each marker in all populations.