

Whole Genome Selective Sweeps Analysis in Pakistani Kamori Goat

Rashid Saif (✉ rashid.saif37@gmail.com)

Gulab Devi Education Complex, Decode Genomics <https://orcid.org/0000-0001-8915-242X>

Jan Henkel

Institute of Genetics, University of Bern, Switzerland

Tania Mahmood

Decode Genomics

Aniqa Ejaz

Decode Genomics

Saeeda Zia

National University of Computer and Emerging Sciences - Lahore Campus

Original Article

Keywords: Kamori goat, Whole genome pooled sequencing, selection signatures, Bezoar wild ancestor, SNVs

Posted Date: February 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-182850/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Gene Reports on November 1st, 2021. See the published version at <https://doi.org/10.1016/j.genrep.2021.101429>.

Whole Genome Selective Sweeps Analysis in Pakistani Kamori Goat

Rashid Saif^{1,2*}, Jan Henkel³, Tania Mahmood², Aniq Ejaz², Saeeda Zia⁴

¹ Institute of Biotechnology, Gulab Devi Educational Complex, Lahore, Pakistan

² Decode Genomics, Punjab University Employees Housing Scheme (II), Lahore, Pakistan

³ Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland

⁴ Department of Sciences and Humanities, National University of Computer and Emerging Sciences, Lahore, Pakistan

*Corresponding Author:

Dr. Rashid Saif

Associate Professor

Gulab Devi Educational Complex

Gulab Devi Chest Hospital, Ferozepur Road, Lahore, Pakistan

rashid.saif37@gmail.com

Abstract

Natural and artificial selection fix certain genomic regions of reduce heterozygosity which is an initial process in breed development. Primary goal of the current study is to identify these genomic selection signatures under positive selection and harbor genes in Pakistani Kamori goat breed. High throughput whole genome pooled-seq of Kamori (n = 12) and Bezoar (n = 8) was carried out. Raw fastq files were undergone quality checks, trimming and mapping process against ARS1 reference followed by calling variant allele frequencies. Selection sweeps were identified by applying pooled heterozygosity (*Hp*) and Tajima's D (TD) on Kamori while regions under divergent selection between Kamori & Bezoar were observed by Fixation Index (F_{ST}) analysis. Genome sequencing yielded 619,031,812 reads of which, 616,624,284 were successfully mapped. Total 98,574 autosomal selection signals were detected; 32,838 from *Hp* and 32,868 from each F_{ST} & TD statistics. Annotation of the regions with threshold ($-ZH_p \geq 5$, $TD \leq -2.72$ & $F_{ST} \leq 0.09$) detected 60 candidate genes. The top hits harbor Chr.1, 6, 8 & 21 having genes associated with body weight (*GLIS3*, *ASTE1*), coat color (*DOCK8*, *MIPOLI*) & body height (*SLC25A21*). Other significant windows harbor milk production, wool production, immunity, adaptation and reproduction trait related genes. Current finding highlighted the under-selection genomic regions of Kamori breed and likely to be associated with its vested traits and further useful in breed improvement, and may be also propagated to other undefined goat breeds by adopting targeted breeding policies to improve the genetic potential of this valued species.

Keywords: Kamori goat, Whole genome pooled sequencing, selection signatures, Bezoar wild ancestor, SNVs

1. Introduction

Domestic goat breeds have been under well documented and strong selection for several years through natural and artificial methods for various quantitative traits especially meat, milk and reproduction. The selection strategies impose selection pressure on particular regions of the genome that control such traits. For instance, intensive crossbreeding strategy was applied on Pakistani breeds, Pak-Angora and Beetal with Hairy goat, for bulk quality production of Mohair [1]. Similarly, hybrid Kamori goat (Kamori x Patairee) are popular in Pakistan for beautiful coat color, appearance and being less expensive than its pure breed [2]. So ascertaining the genes and genomic regions

under positive selection affected due to selection force is vital for understanding the phenotypic diversity among breeds due to genetic variations [3]. Whole genome analysis based on population genetic statistics have been developed that significantly assesses the selective sweeps in all livestock species without known phenotypes. These include; F_{ST} for genomic differentiation [4], cross-population composite likelihood ratio (XP-CLR) that uses variability in allele frequencies at linked loci only between two populations, Hp uses variability estimator in allele counts [5], extended haplotype homozygosity (EHH) approach using SNP data [6], TD statistics to calculate possible distortions in distribution of allele frequencies [7] and others have been used to reveal selection footprints in various livestock species.

In this study, we used three statistical methods (i) Hp (ii) TD and (iii) F_{ST} to explore genomic selective sweeps in Kamori goat vs. the wild ancestor Bezoar from Switzerland. Kamori is famous for its long ears, neck and unique coat colors usually dark brown with small dark or coffee colored patches over its entire body. They are medium to large size milch breed mainly found in Nawabshah, Dadu and Larkana districts in Sindh province of Pakistan. On average a Kamori doe gives ~1.5 liters milk [1]. The primary goal of this research was to identify genomic regions and underlying candidate genes affected by selection in Kamori goat.

2. Materials and methods

2.1 Animals and high-depth WG pooled sequencing

Genomic DNA was extracted for high throughput sequencing, from Kamori ($n = 12$) using TIANGEN biotech (Beijing) CO., LTD and extraction from Bezoar goat ($n = 8$) was carried out at Institute of Genetics, University of Bern, Switzerland from whole blood samples. Same amount of extracted DNA per individual was mixed into a single pool. Whole genome sequencing was performed using Illumina HiSeq 2000 platform that generated 150bp paired-end ~300mio reads. Sequencing files are available at European Nucleotide Archive (ENA) under Project ID: PRJEB23815. Summary of the characteristics (Table 1) along with the figures of all three breeds used in this study are shown in Fig. (1).

Table 1 Features summarizing Kamori (breed), Bezoar (wild ancestor) and San Clemente (reference goat) used in this study.

Breed	Specific trait	Origin	Animals per pool
Kamori (<i>Capra hircus</i>)	Milk and meat producing breed, unique coat colors, long ears and neck, medium to large size breed, brown or coffee colored patches over its entire body.	Pakistan	12
Bezoar (<i>Capra aegagrus</i>)	Wild ancestor	Switzerland	08
San Clemente (reference goat)	Meat producing, relatively small, mostly red and tan with black markings	USA	-



Fig. (1). Representative animals. (a): Kamori (subject goat). (b): San Clemente (reference goat). (c): Bezoar goat (wild ancestor).

2.2 Calling SNVs from whole genome pooled-seq data

Filtration of both pools fastq files for base quality was done by Trimmomatic (v0.36) using SLIDINGWINDOW:4:20 MINLEN:2 parameters. Filtered reads were aligned to ARS1 reference genome assembly using BWA-MEM algorithm v0.7.17 [8] followed by SAM to BAM conversion using its samtools view feature. Sorting on coordinate basis and tagging adapter sequences was accomplished using SortSam and MarkDuplicates features of Picard tools [9]. Quality of reads were checked using FastQC (v0.11.8) software. Samtools mpileup was run on both pools bam files together and on single Kamori bam file to call SNVs. Resulting mpileup and pileup files were subjected to Popoolation2 v1.201 tool [10], mpileup2sync.jar and snp-frequency-diff.pl scripts, that generated synchronized (sync) combined mpileup and separate sync pileup file.

2.3 Selective sweep detection using H_p , TD and F_{ST} statistics

An in-house Ruby script with window-size of 150kb was applied to calculate $-ZH_p$ scores by first applying $H_p = 2\sum n_{MAJ}\sum n_{MIN}/(\sum n_{MAJ} + \sum n_{MIN})^2$ and then Z-transforming it by using $-ZH_p = -(H_p - \mu_{H_p})/\sigma_{H_p}$ to search for positive selection hits [5]. On separate Kamori pileup file, Popoolation v1.2.2 tool's script, variance-sliding.pl was run [7] to compute classical Tajima's D applying $D_{b,pool} = d_{b,pool}/\sqrt{\text{Var}(d_{b,pool})}$. For genomic differentiation between the two goat breeds, Popoolation2 v1.201 tool's script, fst-sliding.pl based on $F_{ST} = s^2/\bar{p}(1-\bar{p}) + s^2/t$ [4] was applied on each SNV value obtained in combined sync file with 50% overlapping window.

2.4 Data visualization

Manhattan plots of $-ZH_p$, TD and F_{ST} values were constructed using manhattan function of qqman package while SNP density graph of these scores was constructed using CMplot package on R software [11]. In addition, the $-ZH_p$, TD and F_{ST} scores were constructed against the theoretical quantiles as Q-Q plots using qqnorm function while the distribution of these values across the no. of windows was plotted as histogram using hist function on R [12].

3. Results

3.1 Quality control and pooled-seq data processing

Genome sequencing yielded 619,031,812 total reads which are further trimmed. Almost 616,624,284 reads with a coverage of 99.61% are mapped against the ARS1 reference genome assembly which are quality checked for further onward selection signature study Fig. (S1). SNV calling generated 98,574 autosomal selection signals including 32,838 from H_p and 32,868 variants each from F_{ST} and TD statistics. The distribution of number of SNPs within 100

MB window size by each of the method is shown in Fig. (S2). Likewise, the Q-Q plot of $-ZH_p$, TD and F_{ST} values across all autosomes are observed in black line against the expected standard normal distribution in red line along with the frequency distribution graphs of $-ZH_p$, TD and F_{ST} values in Fig. (S3; S4 and S5).

3.2 Candidate selection signals and harbor genes

The windows under positive selection in Kamori goat obtained after setting thresholds ($-ZH_p \geq 5$, $TD \leq -2.72$ and $F_{ST} \leq 0.09$) on the basis of previously published data and rationale observation of our own data, are further fine-mapped and annotated which revealed genes related to body mass, coat color, wool type, milk production, immunity, adaptation to environment, body height and reproduction (Table 2).

Table 2 Genes under selection in Kamori goat revealed by H_p , F_{ST} , TD and its associated traits.

Under selection genes	Related traits	References
<i>USP33, HMBOX1, GLIS3, TJP2, FAM189A2, APBA1, MYCBP2, KLHL1, RAPH1, STIM1, ASTE1, ATP8B3, RAPGEF6, SDHAF4, B3GNT6, FAM168A, ACSF3</i>	Meat production, body mass and body weight	[13-15]
<i>ANKRD11, RERE, PRP6</i>	Reproduction and embryonic development	[16]
<i>DMRT2, DMRT1, ZFPM1, ATP2C1</i>	Milk protein composition and production	[17, 18]
<i>SAXO1, ADAMTSL1, DUSP15, EIF2S2, PIEZO1, CDH15, CPNE7, TDP2, SPIRE2, RIC1, TPX2, STAG1, CCDC91, ZNF609, MGA, ATAD2B, SIK3, SLC255A21, TET1, ADAMTS6, NCAPG</i>	Body height	[9, 19, 20]
<i>FH, WDPCP, TRMT11</i>	Environmental adaptation	[21]
<i>SLC38A8</i>	Silky hairs/wool type	[22]
<i>KIT, KANK, DEF8, DOCK8, MLANA, RALY, MIPOL1, IRF4</i>	Coat color	[10, 23]
<i>CBFA2T3, RFX3, CTLA4</i>	Immunity	[24]

3.3 Signatures of positive selection by H_p analysis

The H_p analysis of Kamori breed returned 62 windows under selection with $-ZH_p \geq 5$ Fig. (2). The top hits appear on Chr.8:40,500-40,650 kb, Chr.6:70,800-70,950 kb and Chr.8:40,425-40,575 kb regions having $-ZH_p$ scores 5.96, 5.91 and 5.90 containing 321, 280 and 267 number of SNPs respectively. The 62 putative regions under selection harbor 28 known genes related to body weight, body height, coat color, reproduction and milk protein composition while 10 windows have LOC genes and 11 regions have no genes (Table S1).

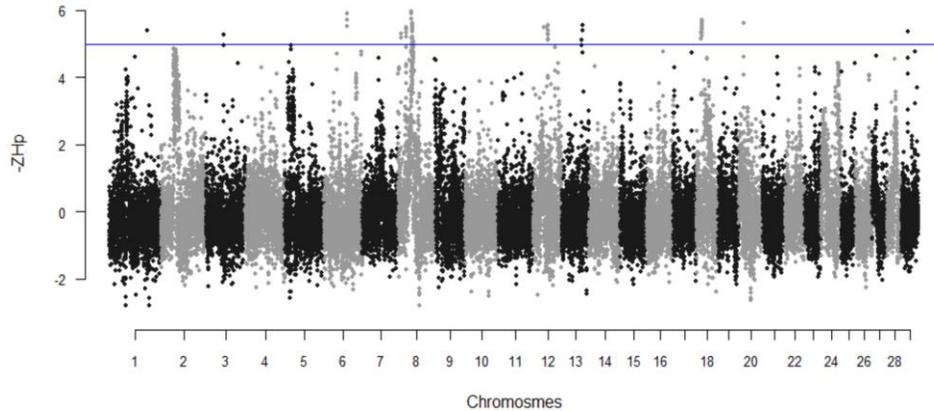


Fig. (2). Manhattan plot of $-ZHp$ analysis on Kamori goat breed. The blue horizontal line directs the suggested significant cutoff threshold of $-ZHp \geq 5$ for better hits. The sliding window approach was used with window size of 150 kb and 75 kb step size considering all autosomes.

3.4 Regions under purifying selection by Tajima's D statistics

Based on the calculations of TD, 63 significant windows (having $TD \leq -2.72$) corresponding to the regions that have undergone a recent bottleneck or a selective sweep are revealed Fig. (3). The top most significant windows with TD values -2.81 and -2.79 are located on Chr.8:4,025-4,175 kb and 40,500-40,650 kb region possessing 2,218 and 2,188 SNPs respectively. Twenty three genes are identified lying on these putative selective sweeps that are associated with bone length/bone height, body weight/meat production, milk protein, immunity, coat color and cold-induced thermogenesis for adaptation to the environment. Twelve of the regions have LOC genes while 07 putative windows are devoid of genes (Table S2).

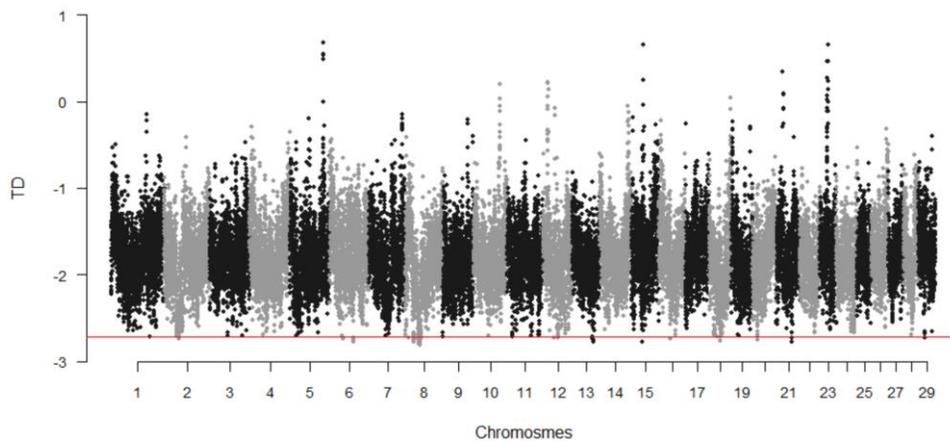


Fig. (3). Manhattan plot showing TD scores from Kamori goat breed. The red horizontal line indicates the stringent threshold of $TD \leq -2.72$ that is computed using 50% overlapping window size such that each dot represents a 150kb window.

3.5 Regions under divergent selection by F_{ST} analysis

F_{ST} methodology is applied on Kamori and Bezoar to focus on the genomic regions that differ between the two. Thirty four candidate regions are divulged after adjusting the F_{ST} threshold to ≤ 0.09 . Of these significant windows, highly differentiated regions are positioned on Chr.21:46,875-47,025 kb, 46,800- 46,950 kb and on Chr.1:138,450-138,600 kb region having 0.66, 0.73 and 0.72 F_{ST} values comprising 3,995, 4,081 and 4,919 SNPs respectively. Fine mapping of regions under selection revealed 23 genes related to milk protein and fat, body height and weight, immunity, reproduction, coat color, adaptation to high altitude and in cold environment and silky hair, 7 LOC genes and only 1 window has no gene (Table S3).

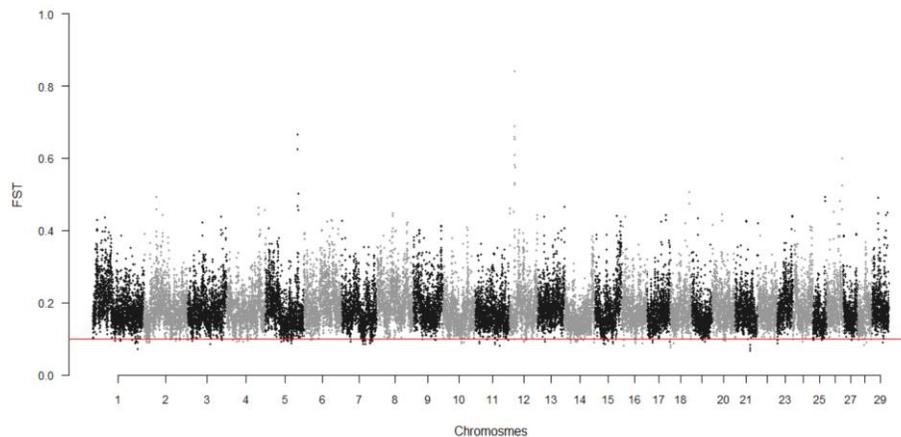


Fig. (4). Manhattan plot based on F_{ST} values distributed across all autosomes of Kamori goat. The preferred threshold of $F_{ST} \leq 0.09$ is illustrated by horizontal red line. By using sliding window approach, window size was set to 150 kb and step size 75 kb.

3.6 Congruent Selection signatures identified by –ZHp, TD and F_{ST} approaches

Table 3 list the putative selection signature and underlying genes that are observed above the preferred thresholds by at least two of the three methods. Interestingly, 23 windows that harbor 15 genes are commonly selected by –ZHp and TD while none of the signals by both methods corresponds to the signals generated by F_{ST} .

Table 3 Common selective sweeps and genes identified by three approaches– ZHp, TD and F_{ST}.

Chr.	Common selective sweeps	Statistical methods	Gene	Selected traits	Significant values
6	70,725-70,875 kb	Hp, TD	<i>KIT</i>	Coat color	–ZHp = 5.532, TD = -2.734
6	70,800-70,950 kb 70,875-71,025 kb	Hp, TD	<i>LOC102174549</i>	-	–ZHp = 5.91, 5.72, TD = -2.75, -2.77
8	9,450-9,600 kb 9,525-9,675 kb	Hp, TD	<i>HMBOX1</i>	Body weight, meat production	–ZHp = 5.30, 5.18, TD = -2.72, -2.73
8	25,200-25,350 kb 25,275-25,425 kb 25,350-25,500 kb 25,425-25,575 kb	Hp, TD	<i>ADAMTSL1</i>	Body height	–ZHp = 5.22, 5.44, 5.33, 5.48 TD = -2.74, -2.75, -2.74, -2.77
8	38,850-39,000 kb	Hp, TD	<i>RIC1</i>	Body height	–ZHp = 5.18, TD = -2.76
8	40,350-40,500 kb	Hp, TD	<i>TRNAG-UCC</i>	tRNA gene	–ZHp = 5.73, TD = -2.78
8	40,425-40,575 kb 40,500-40,650 kb	Hp, TD	<i>GLIS3</i>	Body weight	–ZHp = 5.90, 5.96, TD = -2.78, -2.79
8	43,200-43,350 kb 43,275-43,425 kb	Hp, TD	<i>DMRT2</i>	Milk protein	–ZHp = 5.44, 5.54 TD = -2.76, -2.77
8	43,350-43,500 kb 43,425-43,575 kb	Hp, TD	<i>DMRT1</i>	Milk protein	–ZHp = 5.53, 5.18 TD = -2.73, -2.76
8	43,575-43725 kb	Hp, TD	<i>LOC102183789</i>	-	–ZHp = 5.04, TD = -2.74
8	43,650-43,800 kb	Hp, TD	<i>KANK1</i>	Coat color	–ZHp = 5.21, TD = -2.77
8	45,450-45,600 kb	Hp, TD	<i>APBA1</i>	Body mass	–ZHp = 5.05, 5.08, TD = -2.73, -2.73
12	33,750-33,900 kb	Hp, TD	<i>MYCBP2</i>	Body weight	–ZHp = 5.47, TD = -2.72
12	44,250-44,400 kb	Hp, TD	<i>LOC102173756</i>	-	–ZHp = 5.38, TD = -2.72
20	14,250-14,400	Hp, TD	<i>ADAMTS6</i>	Body height	–ZHp = 5.63, TD = -2.74

4. Discussion

This genome wide selection scan presents the results of regions under positive selection using complementary statistical tests, –ZHp and TD on Kamori breed and also of genomic diversity analysis applied on Kamori vs. Bezoar using F_{ST} approach.

By assessing the SNVs in Kamori breed using –ZHp and TD, we observed selection footprints that are under the influence of both natural and artificial selection. Integration of these two approaches highlighted several concordant regions that are under selection (Table 3) which possibly include true positive selection signals with more confidence. These regions harbor genes related to body weight and height, coat color and milk production. Several regions harbor genes controlling body weight, immunity and body height traits are under genetic hitchhiking selection in Kamori such as on Chr. 2, 8, 13, 15 and 16. Among them, *RALY* gene has also been identified in Iranian Markhoz goat responsible for its black and brown coat color [23]. Similarly, *RAPH1* and *STIM1* is associated with body weight while *TPX2*, *NCAPG* etc. are related to body height in humans and in mice [9, 20]. Moreover, regions solely under old selection are also observed by Hp test that are located on Chr.3, 8, 13 and 18. Some examples include *ZFPM1* that is under selection in African sheep for milk yield [18]. The *ACSF3* gene is reported to be under selection in Korean cattle for meat production [19]. The *PIEZO1*, *CDH15*, *TDP2*, *EIF2S2*, *DUSP15* genes have previously been found under selection in humans for body height trait [20]. All the positive selection signatures found in Kamori adds support to the particular characteristics of this breed.

Our work also revealed clear genomic differences between the two goats under study which can be attributed to the regions varying majorly in body height, coat colors and milk protein among others. It has been shown that *STAG1*, *SIK3*, *SLC25A21*, *CCDC91*, *ZNF609*, *MGA*, *ATAD2B* and *TET1* are associated with body height in humans and in various livestock species [20] which differentiates the goats included here. For the *MIPOLI* and *IRF4* genes, studies have suggested their potential role in eumelanin pigmentation in chicken plumage and in human hair color respectively [20, 25] thus, they may explain the dark brown coat color of Kamori. The *TRMTIL* gene associated with chicken adaptation and survival in hot conditions [21] appeared under selection in Kamori which is likely due to its habitat in hot environment of Sindh province. One selective sweep harboring *ATP2CI*, a candidate gene for milk protein composition in Holstein bulls [17], appeared on a differentiated genomic region in this milch breed. None of the selection signals observed by F_{ST} above the threshold are consistent with the signals generated by $-ZHp$ and TD, indicating that only divergent selection is acting on these regions.

5. Conclusion

With the aim to identify selection footprints in Pakistani Kamori goat breed, we applied three statistical tests; $-ZHp$, TD and F_{ST} . A total of 98,574 autosomal positive as well as divergent selection signals were found that are likely associated with body mass, coat color, wool type, milk production, immunity, environmental adaptation, body height and reproduction. Our findings call for further investigation of Pakistani goat genome with the outcome that can support for selection and breeding programs for improved production and adaptive traits.

Acknowledgements: Authors acknowledged Prof. Dr. Tosso Leeb's great supports to accomplish this research endeavor, and we also acknowledge Dr. Safdar Ali Fazlani for being helpful in the Kamori sample collection.

Funding: Sequencing of Pakistani goats were funded by Swiss National Science Foundation (SNSF) (31003A_172964), when RS was postdoc fellow at University of Bern, under Swiss Government Excellence Scholarship/Hans Sigrist Foundation and now analyzing this already generated publicly available data further.

Availability of data and material: Relevant data is available in the manuscript including supplementary files and supporting information.

Conflict of interests: There is no competing interests among the authors.

Code Availability: F_{ST} , TD scripts are available publicly, while in-house ruby script was used for *Hp* statistics, Manhattan and other supplementary plots were plotted using Bioconductor qqMan package using R.

Authors Contribution: Rashid Saif (RS) envisaged the idea, involved in critical thinking, analysis of data, editing, proofread and correspondence with journal. Jan Henkel (JH) was involved in software and Linux features understandings, analysis of data and provision of in-house *Hp* script. Tania Mahmood (TM) and Aniqah Ejaz (AE) helped in data analysis and initial write-up. Saeeda Zia (SZ) helped in understanding the statistical methods.

Ethics Approval: All Pakistani animal were sampled with the consent of their owners by obeying the local regulations.

Consent to participation: As this is not a human based study, but authors informed and took farmer's consent before collecting samples of their animals.

Consent for publication: Authors have taken permission from the PI of this SNSF funded project to further analyze this publicly available data and has no objection to publish this work.

Supplementary Material

Fig. (S1). Illustration of quality checks results.

Fig. (S2). SNP density plot.

Fig. (S3). Distribution of $-ZHp$ scores

Fig. (S3). Distribution of TD scores

Fig. (S3). Distribution of F_{ST} scores

Table S1. Genomic selection signatures $-ZHp \geq 5$

Table S2. Genomic selection signatures $TD \leq -2.72$

Table S3. Genomic selection signatures $F_{ST} \leq 0.09$

References

1. Country Report on State of Animal Genetic Resources of Pakistan; Prepared for Submission to FAO for First Report on the State of the World's Animal Genetic Resources (SoWAnGR) 2003. p. 1-48.
2. Nabeel DAFaDMS. Kamori – a famous goat breed of Pakistan. FARMER REFORMER: A monthly e-magazine about Agriculture and Livestock 2017.
3. Biswas S, Akey JM. Genomic insights into positive selection. *TRENDS in Genetics*. 2006;22(8):437-46.
4. Porto-Neto L, Lee S, Sonstegard T, Van Tassell C, Lee H, Gibson J, et al. Genome-wide detection of signatures of selection in Korean Hanwoo cattle. *Animal genetics*. 2014;45(2):180-90.
5. Rubin C-J, Megens H-J, Barrio AM, Maqbool K, Sayyab S, Schwochow D, et al. Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences*. 2012;109(48):19529-36.
6. Rothammer S, Seichter D, Förster M, Medugorac I. A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC genomics*. 2013;14(1):908.
7. Tajima F. DNA polymorphism in a subdivided population: the expected number of segregating sites in the two-subpopulation model. *Genetics*. 1989;123(1):229-40.
8. Guo J, Tao H, Li P, Li L, Zhong T, Wang L, et al. Whole-genome sequencing reveals selection signatures associated with important traits in six goat breeds. *Scientific reports*. 2018;8(1):1-11.
9. Saif R, Henkel J, Jagannathan V, Drögemüller C, Flury C, Leeb T. The LCORL Locus is under Selection in Large-Sized Pakistani Goat Breeds. *Genes*. 2020;11(2):168.
10. Henkel J, Saif R, Jagannathan V, Schmocker C, Zeindler F, Bangerter E, et al. Selection signatures in goats reveal copy number variants underlying breed-defining coat color phenotypes. *PLoS genetics*. 2019;15(12):e1008536.
11. Turner S. Annotated Manhattan plots and QQ plots for GWAS using R, Revisited. *Nature Precedings*. 2011:1-.
12. Shaffer LB, Young TM, Guess FM, Bensmail H, León RV. Using R software for reliability data analysis. *International Journal of Reliability and Applications*. 2008;9(1):53-70.
13. de Lemos MVA, Peripolli E, Berton MP, Feitosa FLB, Olivieri BF, Stafuzza NB, et al. Association study between copy number variation and beef fatty acid profile of Nelore cattle. *Journal of applied genetics*. 2018;59(2):203-23.
14. SanCristobal M, Rohart F, Lascor C, Bouffaud M, Trouilh L, Martin PG, et al. Exploring transcriptomic diversity in muscle revealed that cellular signaling pathways mainly differentiate five Western porcine breeds. *BMC genomics*. 2015;16(1):1055.
15. Sheet S, Krishnamoorthy S, Cha J, Choi S, Choi B-H. Identification of Candidate Genes and Pathways Associated with Obesity-Related Traits in Canines via Gene-Set Enrichment and Pathway-Based GWAS Analysis. *Animals*. 2020;10(11):2071.
16. Edea Z, Dadi H, Dessie T, Kim K-S. Genomic signatures of high-altitude adaptation in Ethiopian sheep populations. *Genes & genomics*. 2019;41(8):973-81.
17. Lin S, Zhang H, Hou Y, Liu L, Li W, Jiang J, et al. Correction: SNV discovery and functional candidate gene identification for milk composition based on whole genome resequencing of Holstein bulls with extremely high and low breeding values. *PloS one*. 2019;14(11):e0225747.
18. Marina H, Reverter A, Gutiérrez-Gil B, Alexandre PA, Porto-Neto LR, Suárez-Vega A, et al. Gene Networks Driving Genetic Variation in Milk and Cheese-Making Traits of Spanish Assaf Sheep. *Genes*. 2020;11(7):715.
19. Edea Z, Jung KS, Shin S-S, Yoo S-W, Choi JW, Kim K-S. Signatures of positive selection underlying beef production traits in Korean cattle breeds. *Journal of Animal Science and Technology*. 2020;62(3):293-305.
20. GeneCards®: The Human Gene Database. <https://www.genecards.org/> Accessed.
21. Walugembe M, Bertolini F, Dematawewa CMB, Reis MP, Elbeltagy AR, Schmidt CJ, et al. Detection of selection signatures among Brazilian, Sri Lankan, and Egyptian chicken populations under different environmental conditions. *Frontiers in Genetics*. 2019;9:737.
22. Li D, Li Y, Li M, Che T, Tian S, Chen B, et al. Population genomics identifies patterns of genetic diversity and selection in chicken. *BMC genomics*. 2019;20(1):263.
23. Nazari-Ghadikolaei A, Mehrabani-Yeganeh H, Miarei-Aashtiani SR, Staiger EA, Rashidi A, Huson HJ. Genome-wide association studies identify candidate genes for coat color and mohair traits in the Iranian Markhoz goat. *Frontiers in genetics*. 2018;9:105.
24. Gautier M, Flori L, Riebler A, Jaffrézic F, Laloé D, Gut I, et al. A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. *BMC genomics*. 2009;10(1):550.
25. Yang L, Du X, Wei S, Gu L, Li N, Gong Y, et al. Genome-wide association analysis identifies potential regulatory genes for eumelanin pigmentation in chicken plumage. *Animal genetics*. 2017;48(5):611-4.

Figures



Figure 1

Representative animals. (a): Kamori (subject goat). (b): San Clemente (reference goat). (c): Bezoar goat (wild ancestor).

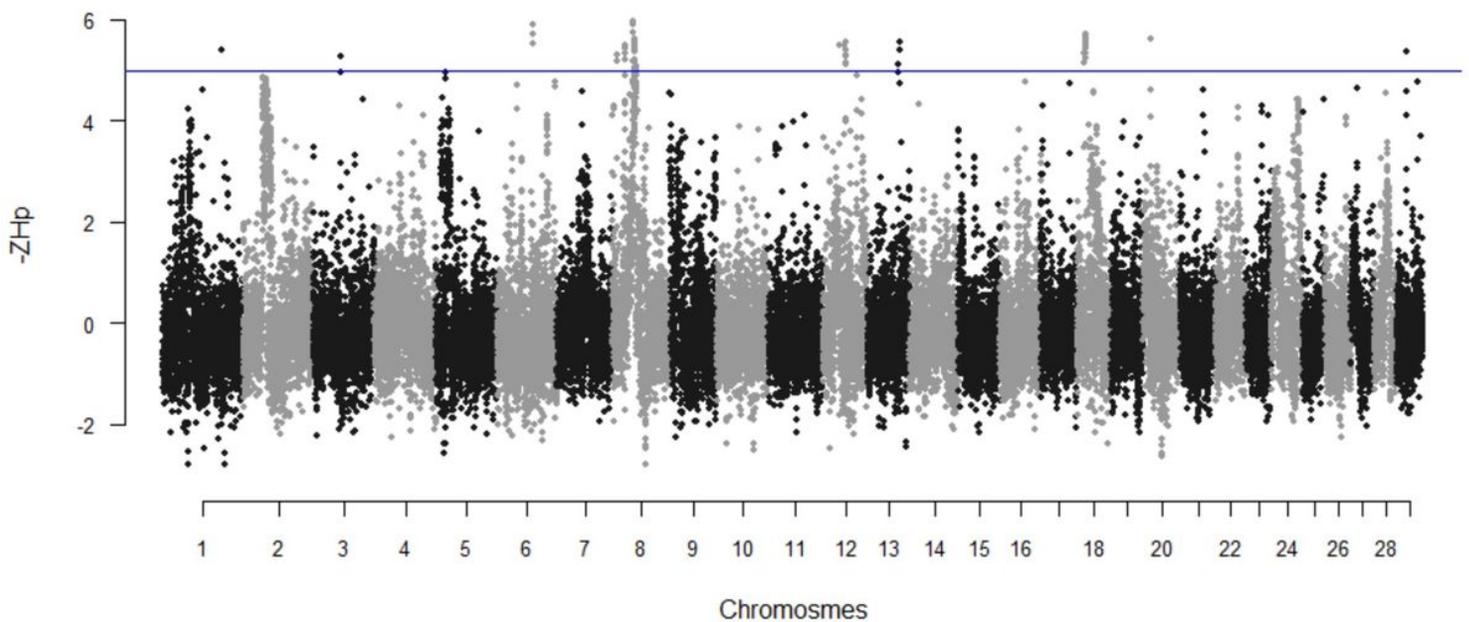


Figure 2

Manhattan plot of $-Z_{Hp}$ analysis on Kamori goat breed. The blue horizontal line directs the suggested significant cutoff threshold of $-Z_{Hp} \geq 5$ for better hits. The sliding window approach was used with window size of 150 kb and 75 kb step size considering all autosomes.

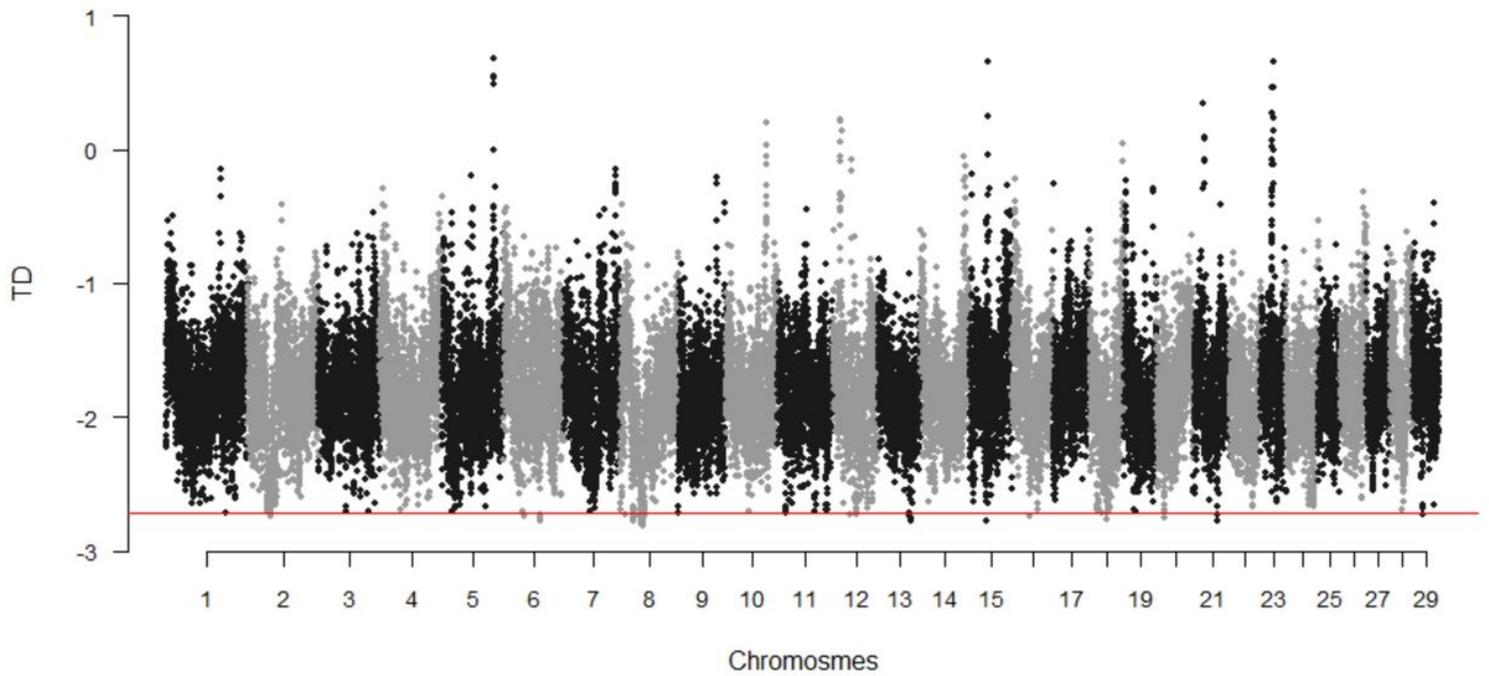


Figure 3

Manhattan plot showing TD scores from Kamori goat breed. The red horizontal line indicates the stringent threshold of $TD \leq -2.72$ that is computed using 50% overlapping window size such that each dot represents a 150kb window.

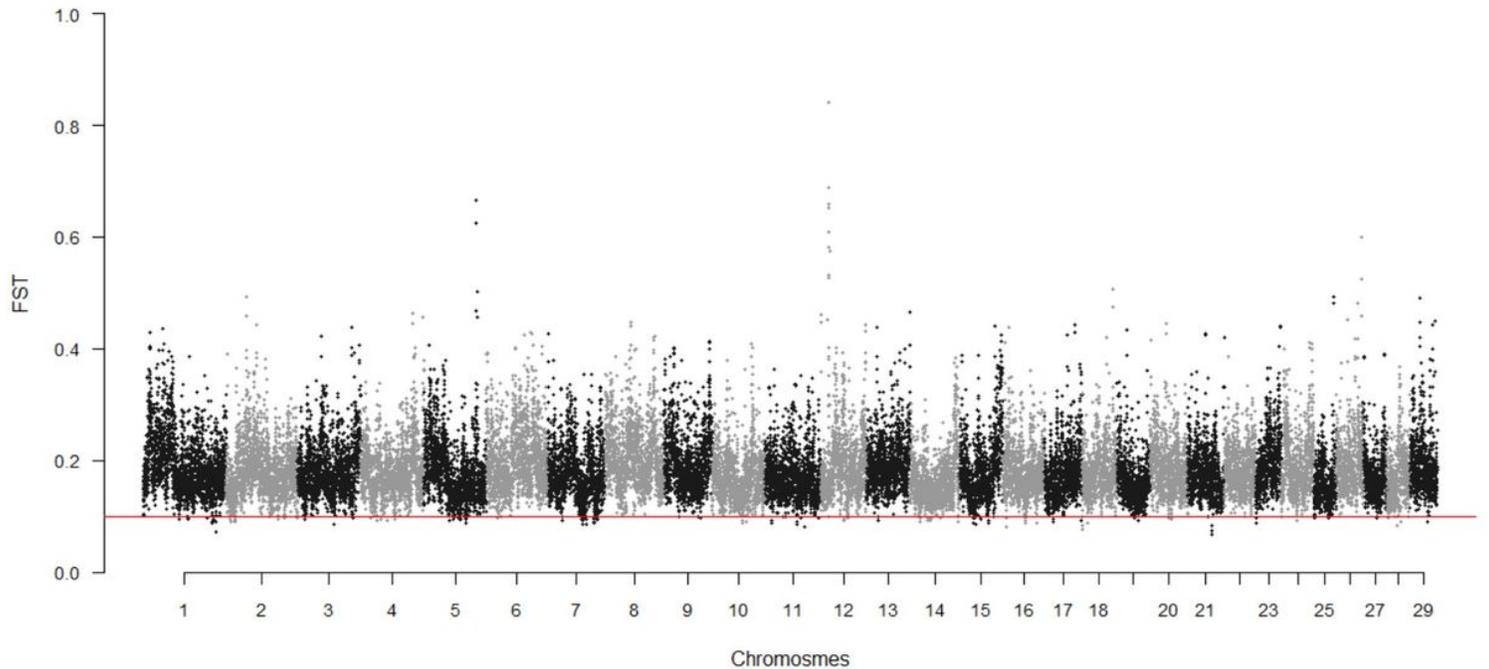


Figure 4

Manhattan plot based on F_{ST} values distributed across all autosomes of Kamori goat. The preferred threshold of $F_{ST} \leq 0.09$ is illustrated by horizontal red line. By using sliding window approach, window size was set to 150 kb and step size 75 kb.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFiguresKAM.pdf](#)
- [TableS1Hp.xlsx](#)
- [TableS2TD.xlsx](#)
- [TableS3FST.xlsx](#)