Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Detection of Whole Genome Selection Signatures of Pakistani Teddy Goat

Rashid Saif [1,2], Jan Henkel[3], Tania Mahmood[1], Aniqa Ejaz[1], Fraz Ahmed[2], Saeeda Zia[4]

[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: rashid.saif37@gmail.com

Abstract

Whole genome pooled sequence data of 12 Pakistani Teddy goats is analyzed for positive selection signatures as their breed defining characteristics. Selection imprints left in the Teddy genome are unveiled by genomic differentiation after the successful paired-end alignment of 635,357,043 reads with (ARS1) reference genome assembly. Pooled-heterozygosity (Hp) and Tajima's D (TD) are applied for validation and getting better hits of selection signals, while pairwise F_{ST} statistics is conducted on Teddy vs. Bezoar (wild goat ancestor) for genomic differentiation. Annotation of regions under positive selection reveals 59 genes underlying production and adaptive traits. Hp score $-ZHp \geq 5$ detected six windows having highest scores on Chr. 29, 9, 25, 15 and 14 that harbor *HRASLS5, LACE1* and *AXIN1* genes which are candidate for embryonic development, lactation and body height. Secondly, TD value of ≤ -2.2 showed 4 windows with very strong hits on Chr.5 & 9 harbor *STIM1* and *ADM* genes related to body mass and weight. Lastly, F_{ST} analysis generated three strong signals with threshold ≤ 0.42 on Chr.12 & 5 harbor *ITGB1* gene associated with milk production & lactation traits. Other significant selection signatures encompass genes associated with wool production, prolificacy, immunity and coat colors. In brief, this study identified the genes under selection in this Pakistani goat breed that will be helpful to refining future breeding policies and converging required productive traits within and across other goat breeds and to explore full genetic potential of this valued livestock species.

Keywords Whole genome pooled-seq, Pakistani Teddy goat, Genomic selection signatures

© 2020 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Introduction

Capra hircus are small ruminants whose domestication started dates back to ~10,000 years ago (Naderi et al. 2008). Goats are mainly reared for meat, milk and wool production (Harris 1962). Classical domestication and breeding practices allowed geneticists and animal breeders to explore the inheritance of special economical traits in this species. Meanwhile, industrial revolution and commercial needs triggered the masses to develop genomic technologies along with refining husbandry techniques to get maximum outputs by adaptation of this species to diverse environments and make them specialized for valued products. Estimated goat population in Pakistan is 76.1 million which produces ~940 thousand tonnes of milk and ~344,000 tonnes of meat annually (S. Ejaz Wasti and Mr. Shujaat Malik Awan 2018-2019). Spontaneous phenotypic mutants have been studied and selected by artificial selection methods which are hallmarks of today's goat breeds like coat colors, meat and milk production, fecundity and adaptation traits. Teddy goat in Punjab, Pakistan is characterized by its short stature, weighing approximately 23-34 Kg, quality meat production and weather tolerant characters (Tahir et al. 1995). In our previous attempts for searching selective sweeps, we identified coat colour structural CNVs in 20 domesticated goat breeds (Henkel et al. 2019), and also analyzed 8 Pakistani breeds for putative variants responsible for large body size traits (Saif et al. 2020).

Aim of the current study is to report complete breed defining hotspots with reduced heterozygosity called "selective sweeps" in this Teddy goat breed from Pakistan, which is primarily raised for meat purposes in this regions. Ultimate goal of this endeavor is to aid in selecting particular characteristics of this breed through artificial selection which could also be helpful to explore the full genetic potential of this breed in its successive generations, to change the genetic merit and to conserve the genetic resources of this species.

Materials and methods

Sample collection and whole genome pooled sequencing

Whole blood samples of Teddy goat (n = 12) from home tract of this breed and Bezoar wild ancestor (n = 8) were collected from Punjab/Pakistan and Switzerland respectively. Genomic DNA of Pakistani breeds were extracted through standard protocol using TIANGEN biotech (Beijing) CO.,LTD, while Bezour DNA extraction was performed at Institute of Genetics, University of Bern. Both populations DNA were mixed into a single pool in equimolar ratios. High-throughput sequencing was conducted using Illumina HiSeq3000 platform which generated 150bp paired-end ~300mio reads. These were further submitted to European Nucleotide Archive (ENA) under Project ID: PRJEB23815 and sample accession number ERS2037817 for Teddy and ERS2037806 for Bezoar. Characteristics & representative

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

animals of Teddy (breed), San Clemente (Reference assembly) and Bezoar (wild ancestor) are shown (Table. 1; Fig. 1)

Table 1 Phenotypic traits of the Teddy (breed), Bezoar (wild ancestor) and San Clemente (Reference genome)

Breed	**Specific trait**	**Origin**	**Animals per pool**
Teddy (*Capra hircus*)	Meat producing, Tender meat quality, small to medium size.	Pakistan	12
Bezoar (*Capra aegagrus*)	Wild ancestor	Switzerland	08
San Clemente (ARS1 reference goat genome)	Meat producing, relatively small, mostly red and tan with black markings	USA	-

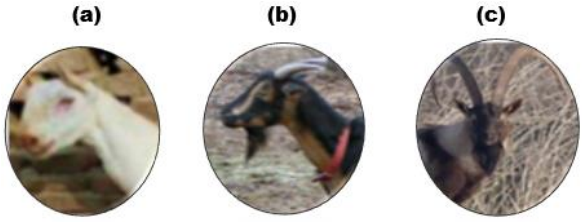

Fig. 1 True representatives of (a) Teddy, (b) San Clemente (ARS1-reference) and (c) wild ancestor Bezoar.

Mapping and SNVs calling

Quality checks were performed on both pools Fastq files using FastQC (v0.11.8) software. While Trimmomatic (v0.36) was applied for base quality filtration using SLIDINGWINDOW:4:20 MINLEN:2 parameters (Bolger et al. 2014). Filtered reads were aligned with ARS1 reference goat genome assembly using BWA-MEM algorithm v0.7.17 (Li et al. 2009). SAM files were converted to BAM files using samtools view and picard tools. BAM files were sorted on coordinate basis and adapter sequences were marked duplicate using Picard-SortSam and MarkDuplicates features respectively. Single Nucleotide Variants (SNVs) were detected using samtools mpileup which yielded combined mpileup file of Teddy and Bezoar as well as separate pileup file of Teddy for *Hp* and Tajima's D analysis (Rubin et al. 2012). Popoolation2 v1.201 tool scripts, mpileup2sync.jar with parameters --fastq-type sanger, –min-qual 20 and snp-frequency-diff.pl was applied on mpileup and pileup files, which generated synchronized (sync) combined mpileup and separate sync pileup files (Kofler et al. 2011).

Genome wide selection scanning

Three statistical tests were applied for the detection of genomics selection imprints left in this Teddy breed genome and generated the SNVs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Detecting selective sweeps using Pooled-heterozygosity (Hp)

First of all we calculated Hp score for both pools by using an in-house Ruby script which applies $Hp = 2\Sigma n_{MAJ}\Sigma n_{MIN}/(\Sigma n_{MAJ}+\Sigma n_{MIN})^2$, where ($n_{MAJ}$)) and ($n_{MIN}$) are major and minor allele counts with window-size of 150kb. The resulting (Hp) scores were Z-transformed by applying $-ZHp = (Hp\text{-}\mu Hp/\sigma Hp)$ such that if any $-ZHp$ value ≥ 5 as a best hit for considering that window under positive selection.

Detecting selective sweeps using Tajima's D (TD) statistics

Neutrality statistics as classical Tajima's D was computed which implies $D_{b.pool} = d_{b.pool}/\sqrt{Var(d_{b.pool})}$ in popoolation v1.2.2 tool script variance-sliding.pl (Korneliussen et al. 2013) that was run on separate pileup files with --min-count 1 --min-coverage 3 --max-coverage 50 and --fastq-type sanger.

Detecting genomics differentiation using Fixation index (F_{ST}) analysis

Thirdly, we ran fst-sliding.pl script of popoolation2 v1.201 (Wang et al. 2016) based on $F_{ST} = s^2/\bar{p}\,(1-\bar{p}) + s^2/r$ (Guo et al. 2018) with 50% overlapping window for each SNV value in earlier generated combined sync file with settings --min-count 2 –min-coverage 4 –max-coverage 50 –suppress-non informative and pool-size 12:8.

Goat reference genome

ARS1 goat reference genome accession number GCF_001704415.1 was obtained from NCBI and used for annotations.

SNP data visualization

R software was used for the construction of SNP density graph of $-ZHp$, TD and F_{ST} scores using CMplot package (Zhou et al. 2019). For $-ZHp$, TD and F_{ST} scores, manhattan plots were constructed by qqman package on R. Horizontal threshold lines were drawn on manhattan plots which signifies the chosen cutoff values ($-ZHp \geq 5$, TD $\leq$ -2.2 and $F_{ST} \leq 0.42$).

Quantile-Quantile plots and histogram

The $-ZHp$, TD and F_{ST} values were plotted against expected values as normal Q-Q plots using the function qqnorm on R software (Shaffer et al. 2008). Standard normal distribution diagonal line which represents the expected values was drawn with qqline function on R. To check the distribution of $-ZHp$, TD and F_{ST} values across all autosomes, histograms were constructed using the function hist.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Results

Quality checks and SNPs calling

Total 248,890,548 variants were called from the Teddy genome using aforementioned three applied statistics after quality checks, trimming and mapping steps using ARS1 reference genome assembly. A total of 635,357,043 (95.83%) reads passed the quality threshold for onward genome wide positive selection signature scanning (Fig. S1). Hp analysis called 26,115,502 SNPs and after applying $-ZHp \geq 5$ threshold, 33,324 SNPs were obtained. TD analysis called 66,775,895 SNPs and after applying TD ≤ -2.2 thresholds 45,512 SNPs were obtained. Similarly, F_{ST} analysis called 155,999,151 SNPs and after setting threshold of $F_{ST} \leq 0.42$, 162,218 SNPs were obtained for downstream analysis.

Selective sweeps and harbor genes

The windows under positive selection obtained after setting thresholds on the basis of previously published data and the rationale observation of our own data, are further fine-mapped and annotated which revealed that body weight/mass, reproduction, milk production, litter size, wool production, coat color and immune system related genes harbor in these selective sweeps (Table 2).

Table 2. List of genes under positive selection in Teddy breed and its associated traits.

Genes under selection	Selection traits	References
ITGB1, LRRIQ3, CCDC152, *NBEA*, *ADM*, *NUP98*, *GPR21*, *DPH6*, *GNB1*, *IPO9*, *RELN*, *ASH1L*, *STRBP, STIM1, TGFBR3, MYCBP2, MAB21L1, KREMEN1*, *SH2B3*, *FAM149B1*, *XRCC4*, *FGGY, TRPC7, HMBOX1, SEC63, PRRC2C, SSU72, ZNRF3, VWDE*, *FREM3*, *SP8, CEP57L1, NR2E1,* and *ATL3*	Meat production, quality and tenderness, body mass, body weight at birth, skeleton development	(Raza et al. 2020; Tao et al. 2020; Zonaed Siddiki et al. 2020)
*BIRC6, RTEL1**, *IFT88, TADA2A, SYNRG, HRASLS5, PRR12* and *SPEF2*	Reproductive performance, embryonic development, fecundity rate, number of teats, litter size	(El-Halawany et al. 2016; Yurchenko et al. 2019)
*FOXO3, LACE1, TTC27, RTEL1**, *TDRD3* and *VPS13B*	Milk production and lactation persistency	(Nayeri et al. 2019; Zhang et al. 2019)
SCAPER, AXIN1, STRBP, TCDC152, DLG4, DCBLD1, PDE5A and *GATAD2B*	Body height, Short stature	(Carty et al. 2012)
PKIA	Hair fleece development, wool production	(Xu and Li 2017)
KIT	Coat color	(Henkel et al. 2019)
MUC6	Innate immune response	(Zheng et al. 2020)

*Denotes genes involved in more than one function.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

The distribution of number of SNPs within 10,000 MB window calculated byHp, TD and F_{ST} analysis are shown (Fig. S2).

Selection footprints by pooled-heterozygosity analysis

Hp is applied using sliding window approach. Based on the generated scores (Fig. 2), six windows have highest $-ZHp$ values of 11.91, 11.45, 11.20, 11.34, 10.24 and 9.18 harbor Chr.29:42,075-42,225 kb, next two regions on Chr.9:28,875-29,100 kb, Chr.25 in a window from 150-300 kb, Chr.15:81,750-81,900 kb and Chr.14:16,050-16,200 kb region having 284, 337, 423, 94, 345 and 409 SNPs respectively. Total 82 windows exhibits strong signals ($-ZHp \geq 5$) which harbor forty genes related to body mass/weight, body height, milk/lactation, coat color, hair fleece development involved in wool production, number of teats and litter size, for embryonic development and reproduction. Twenty windows comprise LOCs (genes not having orthologues), while 6 windows harbor no genes Supplementary Table S1.

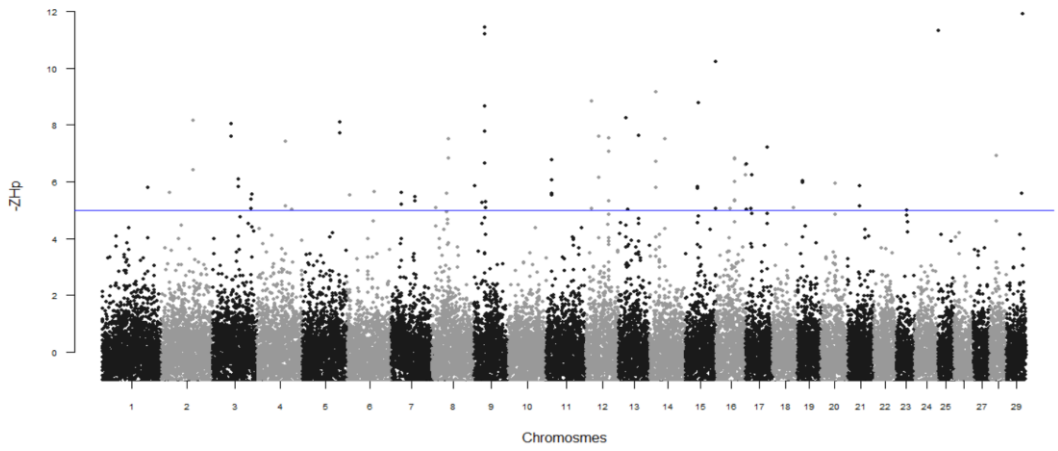

Fig. 2 Manhattan plot demonstrating $-ZHp$ values from Teddy goat breed. The blue horizontal line directs the suggested significant cutoff threshold of $-ZHp \geq 5$ for better hits. The scores were calculated using 150Kb window with sliding step size of 75 kb so that each dot signifies 150 kb windows.

In addition, theoretical distribution of millions of SNPs on x-axis are mapped with observed $-ZHp$ scores on y-axis (Fig. S3a), where black line shows deviated SNPs from the tail at both ends as compared to red line of theoretical distribution. Likewise, histogram of $-ZHp$ scores across all autosomes on x-axis vs. its frequencies on y-axis reveal that only handful of SNPs have either very low or very high $-ZHp$ values (Fig. S3b).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Genetic hitchhiking by Tajima's D statistics

By applying TD statistics, 41 top hit windows of 150 kb with the cutoff threshold of ≤ -2.2 harbor strong selective sweep (Fig. 3).

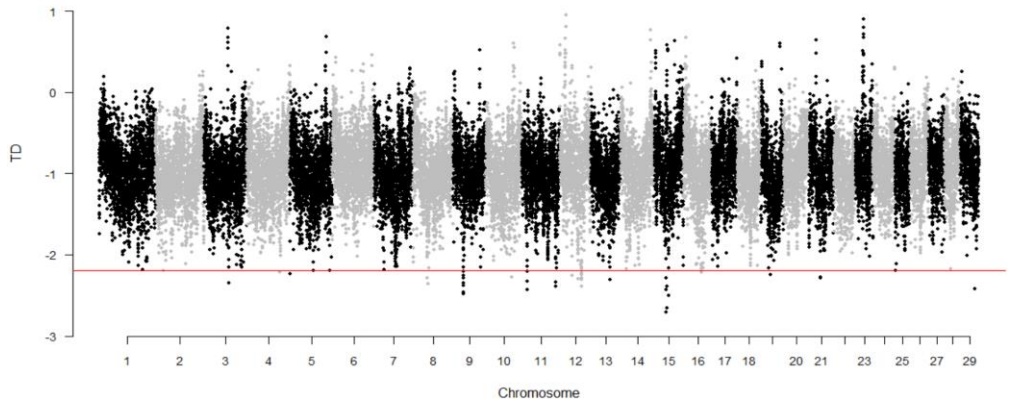

Fig. 3 Results of genome wide selective sweep scan by Tajima's D approach which were visualized as manhattan plot. Horizontal red bar is drawn to indicate the significant threshold of TD ≤ -2.2. The values calculated were based on sliding window approach with 150 kb window size and 75 kb step size so each dot represents a 150 kb window.

The 4 putative windows containing outliers are positioned on Chr.15:32,100-32,250 kb, 32,175-32,325 kb, 33,975-34,125 kb and 39,900-40,050 kb region with TD values -2.705, -2.71, -2.66 and -2.506 including 857, 836, 810 and 1024 number of SNPs respectively. Other significant signatures in the genome are located on Chr.29: 39525000-39675000 kb region, Chr.15:32,025,000-32,175,000 kb, Chr.11:15,075,000-15,225,000 kb, and on Chr.9:29,250,000-29,400,000 kb containing 1345, 969, 1089 and 1084 SNPs in total respectively. Further, the identified twenty genes under selection regions are related to body mass/weight, body height, lactation and reproduction while rest of 3 regions lack genes and 3 windows appeared with LOC genes Supplementary Table S2.

The distribution of TD values across all autosomes are observed as shown in black line against the expected standard normal distribution in red line (Fig. S4a) along with frequency distribution graph of TD values (Fig. S4b). A plethora of observed polymorphisms at the tail of the q-q plot deviates from the bulk of empirical distribution which deemed statistically significant and illustrates SNPs that are predominantly linked with particular traits of Teddy.

Genomic differentiation by Fixation Index

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Highly differentiated regions of Teddy vs. Bezoar (Fig. 4) are selected by setting the permissive threshold of $F_{ST} \leq 0.42$ that comprise 40 windows.

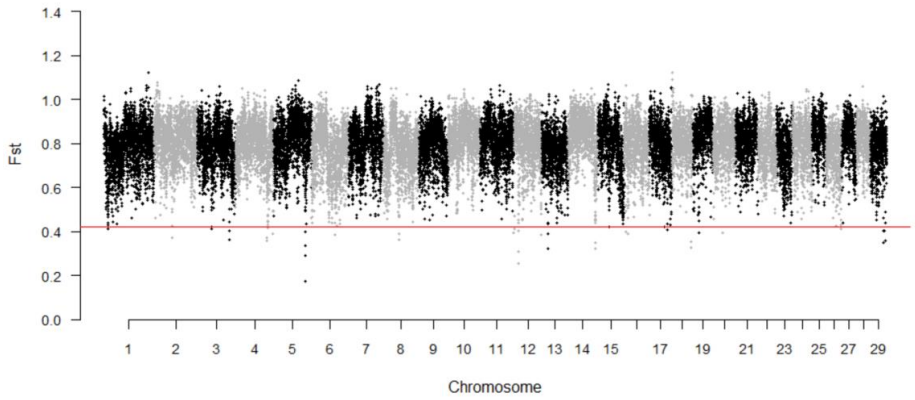

Fig. 4 Manhattan plot illustrating the distribution of selection signals across the whole genome in Pakistani Teddy goat. Horizontal red line points the preferred significant threshold for absolute values of F_{ST} (≤ 0.42). These values were calculated by sliding window approach setting 150 kb as window size with 75 kb step size considering all autosomes.

On Chr.5, two of the regions are under positive selection as they appeared very far from our set threshold having F_{ST} value of 0.174 containing 294 SNPs at 99,000-99,150 kb region, while the second signal were found in the genomic region of 98.925-99,075 kb comprising of 1236 SNPs with F_{ST} value of 0.291. Simultaneously, Chr.12:13,725-13,875 kb also span a selective sweep region with $F_{ST} = 0.253$ having 1486 SNPs. Other windows with strong signals are located on Chr.5:100,725-100,875 kb, Chr.12:13,800-13,950 kb, Chr.13:19,275-19,425 kb and on Chr.14:84,825-84,975 kb region containing 3498, 3145, 4708 and 5890 SNPs respectively. Fine mapping of under selection windows harbor seven genes related to body weight/ mass lactation and immunity while fourteen windows have LOCs and 13 regions are devoid of genes Supplementary Table S3.

To evaluate the significance of F_{ST} scores, the Q-Q plot is also constructed which shows that the observed (sample) quantiles of the F_{ST} values detected against the expected (theoretical) quantiles has some outliers at tail (Fig. S5a). These deviated SNPs whose values are near to zero are expected to be responsible for the genetic differentiation between the two populations. Moreover, the frequencies of fixation index profiles are also observed by computing histogram (Fig. S5b).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Common selective sweeps observed by *Hp*, *TD* and F_{ST} statistics

Several significant windows were found common among at least two of the three applied statistics and proposed as the best selection hits. Total 10 common positive signatures harbor genes associated with meat, milk and reproduction traits (Table 3).

Table 3 Common selective sweeps observed by more than one statistical approach

Chr.	Common selective sweeps	Statistical approaches	Gene	Selected traits	Significant values
3	68,850-69,000 kb	*Hp*, TD	*TGFBR3*	Skeleton development	$-ZHp$ = 5.481, TD = -2.345
4	90,900-91,050 kb	*Hp*, TD	*SP8*	Body weight	$-ZHp$ = 5.04, TD = -2.217
9	28,800-28,950 kb	*Hp*, TD	*FOXO3*	Milk production	$-ZHp$ = 7.022, TD = -2.462
9	28,875-29,025 kb 28,950-29,100 kb	*Hp*, TD	*LACE1*	Milk production	$-ZHp$ = 10.334, 10.11, TD = -2.4, -2.384
11	15,075-15,225 kb	*Hp*, TD	*TTC27*	Milk production	$-ZHp$ = 6.111, TD = -2.429
11	15,000-15,150 kb	*Hp*, TD	*BIRC6*	Embryonic development	$-ZHp$ = 6.05, TD = -2.327
12	32,100-32,250 kb 32,175-32,325 kb	*Hp*, TD	*STIM1*	Body weight/mass	$-ZHp$ = 5.231, 5.181, TD = -2.705, -2.71
13	53,325-53,475 kb	*Hp*, TD	*RTEL1*	Milk production, reproduction	-*ZHp* = 6.886, TD = -2.305
13	19,275-19,425 kb	*Hp*, F_{ST}	*ITGB1*	Milk production	$-ZHp$ = 7.425, F_{ST} = 0.39, 0.32
21	31,350-31,500 kb	*Hp*, TD	*SCAPER*	Body height	$-ZHp$ = 5.282, TD = -2.28

Discussion

Selective sweep regions in the Teddy breed that are either under natural or artificial selections are detected by three statistics e.g. *Hp*, *TD* and F_{ST}. The main trait of short statured Teddy goat breed is the quality meat production whose consumption in Pakistan is preferred among other mutton breeds/species due to its tender and leaner properties and its significance on various religious occasions. Also improved reproductive performance in goats and better fertility rate are other valued attribute of this goat breed. Recently, this meat producing Teddy breed is artificially selected for milk yield, as we find six genes affecting lactation. Similarly, strong selection for wool production, coat color and immunity in Teddy goats due to economic interest might have led to the detected selective sweeps.

Our study find 59 promising genes in Teddy goat from which 34 genes influence meat production. Examples include *LRRIQ3, TGFBR3* and *FGGY* that controls the number and diameter of muscle fiber which affects meat quality and tenderness, fat deposition and muscle growth in swine and average daily gain in Nellore cattle (Nonneman et al.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

2013; Santana et al. 2014; Jeong et al. 2015; Zhang et al. 2020). Eight genes putatively associated with embryonic development, high fecundity rate that additionally control litter size are also observed e.g. *BIRC6*, *RTEL1, SYNRG* etc. (El-Halawany et al. 2016; Xu and Li 2017). Similarly, we identified *FOXO3, LACE1, TTC27, RTEL1, TDRD3* and *VPS13B* related to lactation (Gao et al. 2017; Nayeri et al. 2019; Zhang et al. 2019). The genetic architecture underlying body height trait in Teddy presents 8 candidate genes e.g. *SCAPER* and *AXIN1* associated with body height trait in humans (Carty et al. 2012). Other selection hits include genes *PKIA, KIT* and *MUC6* which functions for hair fleece development enhancing wool production, define coat color phenotypes and trigger host innate immune responses respectively (Xu and Li 2017; Henkel et al. 2019; Zheng et al. 2020).

In conclusion, we find several signals with the strong ones around distinguished genomic regions especially harboring Chr. 5, 9, 12, 15, 14, 25 and 29. Congruent significant windows selected by more than one of the aforementioned methods are present on Chr. 3, 4, 9, 11, 12, 13, 21. Further, fine-mapping is still needed for more comprehensible understanding of selective sweeps discussed in this study. This research provides genome wide maps of selection footprints in Pakistani Teddy goat that will help in better understanding the genomic architecture effected by various artificial selection initiatives.

References

Beaumont RN, Warrington NM, Cavadino A, Tyrrell J, Nodzenski M, Horikoshi M, Geller F, Myhre R, Richmond RC, Paternoster L (2018) Genome-wide association study of offspring birth weight in 86 577 women identifies five novel loci and highlights maternal genetic effects that are independent of fetal genetics. Human molecular genetics 27:742-756

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120

Carty CL, Johnson NA, Hutter CM, Reiner AP, Peters U, Tang H, Kooperberg C (2012) Genome-wide association study of body height in African Americans: The women's health initiative SNP health association resource (share). Human molecular genetics 21:711-720

El-Halawany N, Zhou X, Al-Tohamy AF, El-Sayd YA, Shawky A, Michal JJ, Jiang Z (2016) Genome-wide screening of candidate genes for improving fertility in Egyptian native Rahmani sheep. Anim. Genet 47:10.1111

Gao Y, Jiang J, Yang S, Hou Y, Liu GE, Zhang S, Zhang Q, Sun D (2017) CNV discovery for milk composition traits in dairy cattle using whole genome resequencing. Bmc Genomics 18:265

Guo J, Tao H, Li P, Li L, Zhong T, Wang L, Ma J, Chen X, Song T, Zhang H (2018) Whole-genome sequencing reveals selection signatures associated with important traits in six goat breeds. Scientific reports 8:1-11

Harris D (1962) The distribution and ancestry of the domestic goat. In: Proceedings of the Linnean Society of London. Wiley Online Library, pp 79-91

Henkel J, Saif R, Jagannathan V, Schmocker C, Zeindler F, Bangerter E, Herren U, Posantzis D, Bulut Z, Ammann P (2019) Selection signatures in goats reveal copy number variants underlying breed-defining coat color phenotypes. PLoS genetics 15:e1008536

Jeong H, Song K-D, Seo M, Caetano-Anollés K, Kim J, Kwak W, Oh J-d, Kim E, Jeong DK, Cho S (2015) Exploring evidence of positive selection reveals genetic basis of meat quality traits in Berkshire pigs through whole genome sequencing. BMC genetics 16:1-9

Kofler R, Pandey RV, Schlötterer C (2011) PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27:3435-3436

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Korneliussen TS, Moltke I, Albrechtsen A, Nielsen R (2013) Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC bioinformatics 14:1-14

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079

Naderi S, Rezaei H-R, Pompanon F, Blum MG, Negrini R, Naghash H-R, Balkız Ö, Mashkour M, Gaggiotti OE, Ajmone-Marsan P (2008) The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. Proceedings of the National Academy of Sciences 105:17659-17664

Nayeri S, Schenkel F, Fleming A, Kroezen V, Sargolzaei M, Baes C, Cánovas A, Squires J, Miglior F (2019) Genome-wide association analysis for β-hydroxybutyrate concentration in Milk in Holstein dairy cattle. BMC genetics 20:58

Nonneman D, Shackelford S, King D, Wheeler T, Wiedmann R, Snelling W, Rohrer G (2013) Genome-wide association of meat quality traits and tenderness in swine. Journal of Animal Science 91:4043-4050

Raza SHA, Khan R, Gui L, Schreurs NM, Wang X, Mei C, Yang X, Gong C, Zan L (2020) Bioinformatics analysis and genetic polymorphisms in genomic region of the bovine SH2B2 gene and their associations with molecular breeding for body size traits in qinchuan beef cattle. Bioscience Reports 40

Rubin C-J, Megens H-J, Barrio AM, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB (2012) Strong signatures of selection in the domestic pig genome. Proceedings of the National Academy of Sciences 109:19529-19536

S. Ejaz Wasti MSH, Muhammad Asif, Muhammad Shoaib Malik, Mr. Zaila Husnain, Ms. Samina Khatoon, Mr. Attaullah Shah, Mr. Asadullah Qureshi, Mr. Omer Farooq, Ms. Sidra Saddiq, Ms. Nargis Mazhar,, Mr. Shujaat Malik Awan MAH, Muhammad Arslan, Hafiz Syed Muhammad Azeem, Mr. Faheem Anwar, Muhammad Faisal Shamim (2018-2019) Pakistan Economic Survey (ed. by Economic Adviser's Wing FD). In: Economic Adviser's Wing FD (ed)

Saif R, Henkel J, Jagannathan V, Drögemüller C, Flury C, Leeb T (2020) The LCORL Locus is under Selection in Large-Sized Pakistani Goat Breeds. Genes 11:168

Santana M, Utsunomiya Y, Neves H, Gomes R, Garcia J, Fukumasu H, Silva S, Leme P, Coutinho L, Eler J (2014) Genome-wide association study for feedlot average daily gain in Nellore cattle (Bos indicus). Journal of Animal Breeding Genetics 131:210-216

Shaffer L, Young T, Guess F, Bensmail H, León R (2008) Using R software for reliability data analysis. International Journal of Reliability

Application 91:53-70

Tahir M, Younas M, Raza S, Lateef M, Iqbal A, Raza P (1995) A study on estimation of heritability of birth weight and weaning weight of Teddy goats kept under Pakistani conditions. Asian-Australasian Journal of Animal Sciences 8:595-597

Tao L, He X, Pan L, Wang J, Gan S, Chu M (2020) Genome-wide association study of body weight and conformation traits in neonatal sheep. Animal Genetics 51:336-340

Tetens J, Widmann P, Kühn C, Thaller G (2013) A genome-wide association study indicates LCORL/NCAPG as a candidate locus for withers height in G erman W armblood horses. Animal genetics 44:467-471

Wang X, Liu J, Zhou G, Guo J, Yan H, Niu Y, Li Y, Yuan C, Geng R, Lan X (2016) Whole-genome sequencing of eight goat populations for the detection of selection signatures underlying production and adaptive traits. Scientific reports 6:38932

Xu S-S, Li M-H (2017) Recent advances in understanding genetic variants associated with economically important traits in sheep (Ovis aries) revealed by high-throughput screening technologies. Frontiers of Agricultural Science and Engineering 4:279-288

Yurchenko AA, Deniskova TE, Yudin NS, Dotsev AV, Khamiruev TN, Selionova MI, Egorov SV, Reyer H, Wimmers K, Brem G (2019) High-density genotyping reveals signatures of selection related to acclimation and economically important traits in 15 local sheep breeds from Russia. BMC genomics 20:294

Zhang L, Guo Y, Wang L, Liu X, Yan H, Gao H, Hou X, Zhang Y, Guo H, Yue J (2020) Genomic variants associated with the number and diameter of muscle fibers in pigs as revealed by a genome-wide association study. animal 14:475-481

Zhang S, Zhou C, Shen D, Li C, Cai W, Liu S, Yin HW, Shi S, Cao M (2019) Comparative transcriptomic and proteomic analyses identify key genes associated with milk fat traits in Chinese Holstein cows. Frontiers in genetics 10:672

Zheng Z, Wang X, Li M, Li Y, Yang Z, Wang X, Pan X, Gong M, Zhang Y, Guo Y (2020) The origin of domestication genes in goats. Science Advances 6:eaaz5216

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 December 2020 doi:10.20944/preprints202012.0421.v1

Zhou C, Li C, Cai W, Liu S, Yin H, Shi S, Zhang Q, Zhang S (2019) Genome-wide association study for milk protein composition traits in a Chinese Holstein population using a single-step approach. Frontiers in genetics 10:72

Zonaed Siddiki A, Miah G, Islam M, Kumkum M, Rumi MH, Baten A, Hossain MA (2020) Goat Genomic Resources: The Search for Genes Associated with Its Economic Traits. International Journal of Genomics 2020