GENETIC SIGNATURE OF STRONG RECENT POSITIVE SELECTION AT FSHβ GENE IN GOATS

Nabeel Ijaz, Guiqiong Liu, Xunping Jiang[*], Birhanu Tesema, Dangwei Wang，Fangzheng Xu, Qintao Zhao and Huarong Huang

Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education, College of Animal Science and Technology,Huazhong Agricultural University, Wuhan 430070, People's Republic of China.
[*]Corresponding author's e-mail: xpjiang@mail.hzau.edu.cn

Follicle stimulating hormone (FSH) is a glycoprotein hormone, secreted by gonadotrophs of the anterior pituitary gland. It regulates the development and functions of gonads and is essential for normal reproductive functions in mammals. Determination of selection is arising as one of the most significant steps in the investigation of multilocus population genetic. This study investigated outlier loci detection using a F_{ST} based technique and testing of single nucleotide polymorphisms (SNPs) genotyped in individuals of four goat breeds of China. The utilization of outlier detection approach using SNPs resulted in the identification of the FSHβ to be under selection. Nucleotide sequence of 10 species were also included to determine the evolutionary selection of FSHβ gene and the results showed that this gene previously was under purified selection but recently it is under positive selection. The nucleotide sequence of FSHβ gene of goats showed closer coding sequence similarity to FSHβ gene of ovine (55%) and equine (55%) than to that of cattle (54%), rabbit (49%) and human (48%). The outlier loci recognized in this study may be useful to elucidate adaptive variability in the studied species attributed by wide climate variation in their native areas and vigorous import/export which suggest their ability to adapt to a new environment. In conclusion, the FSHβ gene which was previously under purified evolutionary selection is now remarkably under strong positive selection.
Keywords: Genetic variation, FSHβ, positive selection, adaptation, SNP

INTRODUCTION

China is blessed with 58 native domestic goat breeds, scattered in diverse environmental areas including the Qinghai-Tibet plateau region, north and south agricultural region, the north pastoral region and the mixed pastoral-agricultural region (Wei *et al*., 2014). These breeds have been displaying great genetic potential due to which these breeds have been attaining immense attention, making conservation of domestic breeds crucial to meet the future breeding requirements.

Hormones are habitually essential for the conversion of genotype to phenotype, adopting key steps in progress and integrating the expression of suites of functionally vital traits (Nijhout, 2003; Adkins-Regan, 2005). Fitness is estimated thoroughly by the hormones, and they regularly interact with the components of life mechanisms by endurance and reproduction (Adkins-Regan, 2005; Bonier *et al*., 2009). In spite of anticipated association between fitness and hormones, there is little knowledge regarding the effect of selection on individual's deviation of hormone level and the traits regulated by the hormones in goats (Adkins-Regan, 2005). Hormones have been associated with conception rates in a number of species (Alatalo *et al*., 1996; Mills *et al*., 2007) and previous studies have revealed the relationship between corticosterone, a glucocorticoid stress hormone, and

endurance or reproductive success (Breuner *et al*., 2008; MacDougall- Shackleton *et al*., 2009).

Function of the hormones during the evolution of natural populations are mostly possessed by the "phenotypic engineering" studies, in which altered levels of hormones are used experimentally to monitor their effects on the phenotypic variants and fitness (Adkins-Regan, 2005; Reed *et al*., 2006). However, these experiments present just restricted information regarding evolutionary processes, since the process of hormonal exploitation alters the phenotypic variation on which selection works. Due to this reason, dose dependent controlled experiments should be complemented by studies focusing on the causes and outcomes of individual hormonal deviation (Williams, 2008; Ketterson *et al*., 2009).

Follicle-stimulating hormone (FSH) is a pituitary gonadotrophin, belonging to the glycoprotein family of hormones. This hormone consists of 2 subunits, α−subunit is common in all the gonadotropins, and is noncovalently connected to a specific β−subunit which determines the biological activities of FSH. This hormone is responsible for the ovarian follicular development, maturation, priming of the LH receptors in granulosa cells and it also promotes the progesterone and steroid hormones synthesis. It also stimulates aromatase enzyme for conversion of testosterone to estrogen, and enhances the production of the plaminogen

activators (Ward, 1991). FSHβ is a candidate gene for the evaluation of signatures of selection among genetic markers of productive performance of goats that could be useful in explaining the historical drives and activities that shaped the history of a population.

The importance of molecular markers for genetic investigation studies of different livestock breeds has been demonstrated by several researchers (Mukesh *et al.*, 2004). Due to indiscriminate distribution within the genome and high degree of polymorphism and probability of mechanized scoring of genotypes, microsatellite markers have been found to be the major dominant apparatus for determining genetic variability and evaluating genetic distances between strongly linked populations of ruminant species (Buchanan *et al.*, 1994; Ellegren *et al.*, 1997).

In this experiment, we worked on the genetic variability to reveal the molecular evolution of the FSHβ gene of goat and to present constructive data for understanding the positive selection and association of single nucleotide polymorphism in interspecific FSHβ gene.

MATERIALS AND METHODS

Ethics statement: All the experimental protocols of this study were approved by the Law of Animal Husbandry in People's Republic of China (Dec 29, 2005). Blood samples were collected after permission from owners of animals. The procedures for collection of blood samples of experimental goats were reviewed and permitted by the Biological Studies Animal Care and Use Committee of National Animal Husbandry Service, Hubei, PR China. All efforts were made to minimize any discomfort to animals during blood collection.

Selection of goats: A total of 526 animals belonging to 4 indigenous goat breeds including Nanjiang yellow goat from householders in Youzhou, Chongqing (n=30), Youzhou black goat from Youzhou Black Goat Farm, Youzhou Chongqing (n=140), Macheng black goat from Hubei Jinyang Stock Farm, Macheng, Hubei (n=163) and Tongshan black goats from Hubei Tongshan black-boned goat corporation, Tongshan, Hubei, China(n=193) were included in this study. All these goats were above the age of one year and were reared under their traditional system. Only less than 6 animals were sampled per herd to enhance the breed representativeness and to minimize the relationship among animals being sampled.

Collection of blood samples and extraction of DNA: Blood samples were collected through jugular vein puncture method in vials containing EDTA. The genomic DNA was extracted from these blood samples by the phenol-chloroform method (Zumbo, 2014), dissolved in TE buffer (10 mmol/L Tris-HCl and 1 mmol/L EDTA, pH 8.0), and kept at -20°C.

Genotyping and sequencing: The genomic DNA was used for the extension of 13 polymorphic loci. Exon sequences were screened for detection of polymorphism by Tsinke company. Sequence analysis of the PCR products was performed using Seqman software (Swindell et al., 1997). The aligned tested sequence represented genotypes including both homozygous and heterozygous.

Detection of selection: To find out the selection effects, the approach adopted previously by Beaumont and Nichols (1996) was applied using Lositan software. Fixation index (F_{ST}) values were calculated using the allele frequencies for each locus conditional on heterozygosity (*He*) and P-values for each locus were calculated. Simulations included 13 loci, 526 individuals, 4 populations and an expected F_{ST} of 0.102. This method provided support for divergent selection by finding outliers with higher F_{ST} values than expected, controlling for heterozygosity. Using 100000 simulations on real data population datasets were built. For provisional mutual allocation of F_{ST} against mean heterozygosity, quantiles were assumed keeping the confidence interval of 95%. Loci showing a typical differentiation behavior (i.e. F_{ST}) and lying outside the simulated neutral distribution were taken as outliers (Antao et al., 2008).

Sequence analysis: Coding sequences of FSHβ gene of 10 species including sheep, goat, cattle, horse, rabbit, rat, mouse, human, frog and chicken were obtained from GenBank (www.ncbi.nlm.nih.gov/genbank). MEGA 4.0 program was used for alignment of concluded sequences of proteins (Tamura *et al.*, 2007) with default alignment parameters, followed by manual adjustment. The coding sequences were consequently aligned in accordance with the protein alignment. By using the maximum likelihood model method positive selection and amino acid sites under selection were identified. In this experiment, four models (M1a, M2a, M7 and M8) in Codeml programe of PAML package were used, involving two steps procedure. Firstly, positive selection was determined using the likelihood ratio test (LRT) and that is, for the existence of sites with ω>1. Comparison with a null model was used to achieve this which did not allow sites with ω>1 and a more-general model that did.

For determining the subsequent LRT, the null model M7, which assumed a beta distribution B (*p*, *q*) for ω (in the interval 0<ω<1, 0 indicated complete constraint and 1 was the expectation under no selection pressure) was compared with alternative model M8, in which an extra class of sites with ω estimated was added. In this way, in model M8 a proportion of sites (p0) came from the beta distribution B (*p, q*) and the remaining sites *(p1 = 1-p0)* had a ω ratio estimated from data that could be >1. If the LRT was significant (*P*<0.05), the model M8 fit the sequences better than M7, and therefore, the positive selection was inferred.

Phylogenetic analysis: For analysis of evolutionary relationship of FSHβ gene of goat with other species, it was aligned with the nucleotide sequences of the FSHβ gene of

sheep, cattle, horse, human, mouse, fish and rabbit. Their nucleotide sequences were obtained from the NCBI (National Centre Biotechnology Information, Rockville Pike Bethesda, USA). A phylogenetic tree was constructed using MEGA version 4 with the neighbor joining method (Tamura *et al.*, 2007). Nucleotide sequences of FSHβ obtained from NCBI were introduced to ClustalW for similarity with reported sequences. The results of analysis of FSHβ of goat were compared with 10 species and the genetic similarity results were recorded.

RESULTS

The sequence characterization among the animals of four goat breeds revealed 13 SNPs in 7364 bp sequence, giving an average density of single SNP for every 566 bases. Genotyping was performed on 13 SNPs and these SNP polymorphisms resulted in the discovery of outlier loci. F_{ST} based method identified FSHβ gene loci as under selection sweep in the examined breeds.

Positive selection of FSHβ gene by FDIST analysis: Using the outliers approach, FSHβ gene lied outside the 95% confidence interval of the provisional combined allocation of *Fst* and mean heterozygosity by Lositan FDIST analysis (Fig. 1). FSHβ gene was significantly in the area corresponding to positive selection (p<0.05) (Table 1).

Table 1. Locus, heterozygosity (He) and fixation index (F_{ST}) for each of 13 genotyped SNPs

Locus	*Het*	*Fst*	*P*
PGR	0.494	0.123	0.632
GnRHR	0.172	0.035	0.172
FSHβ	0.248	0.427	0.998*
INH	0.442	0.185	0.855
LHPA	0.295	0.018	0.048
LHPB	0.229	0.015	0.025
Mc1rA	0.126	0.082	0.541
Mc1rB	0.331	0.065	0.314
AgoutiB	0.498	0.094	0.480
Myf5	0.330	0.172	0.834
AANAT	0.367	0.134	0.688
IGF	0.459	0.186	0.869
NPYY1	0.295	0.108	0.573

Table 2. Results of positive selection tests for FSHβ gene.

Model	Parameter Estimates	*LnL*	*LRTs*	Positive selection sites
Model1	p_0=0.811 p_1=0.189 ω_0=0.078 ω_1=1.000	-1300.478	0	Not Allowed
Model2	p_1= 0.811 $p_2$0.123 p_3= 0.066 ω_1= 0.077 ω_2= 1.000 ω_3= 1.000	-1300.478		15A
Model7	p= 0.632 q= 4.189	-1284.935	0.476	Not Allowed
Model8	p_0= 0.989 p= 0.672 q= 4.935 p_1=0.010 ω =1.787	-1284.697		15A

(*: $P>95\%$; **: $P>99\%$), (ω) was the ratio of nonsynonymous-to-synonymous substitutions, Purified selection (p_0), neutral selection (p_2), positive selection (p_3), substitution ratio for all sites, p and q were the β distribution parameters, LnL log likelihood; LRT likelihood ratio
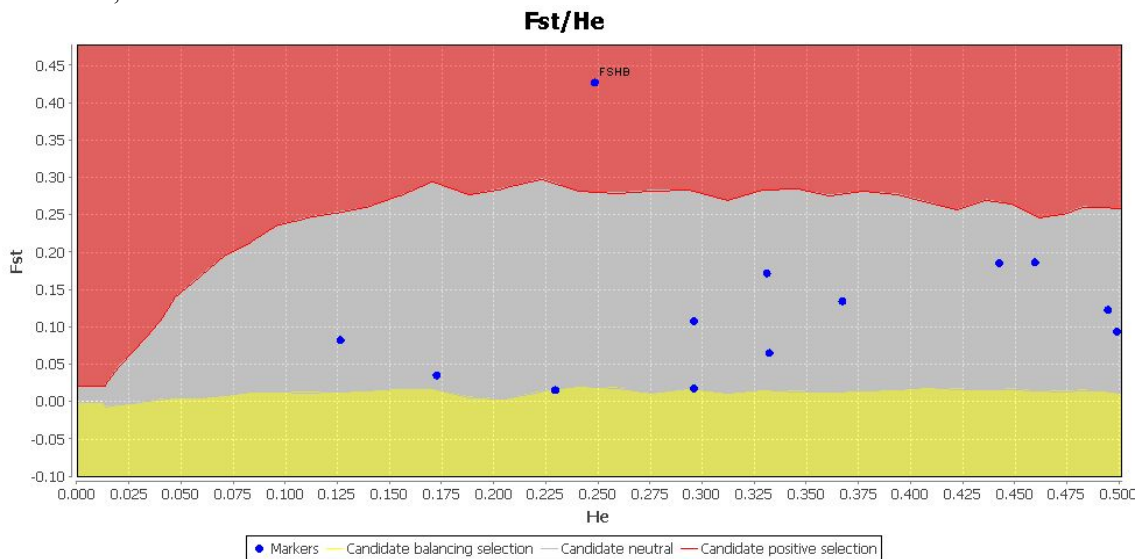

Figure 1. Candidate gene FSHβ under positive selection keeping the 95% confidence interval. F_{st} (Fixation index), He (Heterozygosity)

Heterozygosity (*He*) and Fixation Index (*F_{ST}*) *P* (Simulated *Fst*<sample *Fst*

***Evolutionary analysis of positive selection*:** Exuberance of non-synonymous by synonyms replacements was a molecular proof for positive genetic selection. Database was investigated using two pairs of models (M1–M2; M7–M8). LRT was determined by using the likelihood logs; the effort was to determine if the substitute models (M2, M8) adjusted the series superior to null models (M1, M7). Model M1 with the assessment of $\omega 0 = 0.0778$ gave an LnL value of -1300.48 for the first LRT and with the assessment of $\omega_2 = 1.000$ model M2 gave an LnL value of -1300.48 (Table 2). The test statistic was $2\Delta LnL = 0$, with df = 2 and P<0.001. The model M1was rejected and model M2 was acknowledged because it did not grant evolution of positive selection sites. Model M2 demonstrated that purifying selection was found in 81% sites, neutral evolution was

found in 12.3% and positive selection covered only 6.6% sites. Values of ωs were =0.077, 1.000 and 1.000, respectively. Next LRT correlated additional parameter-rich models M7 and M8, two times the log-likelihood difference was 0.476. Correlated with the significance of $\chi 2$ distribution (df = 2), it rejected the M7 model and accepted the M8 model by keeping level of significance P<0.001. Therefore, M8 model appeared to be considerably more competent for the sequences than model M7. M8 model demonstrated, ωs of 98.9% sites suited the beta distribution β (0.672, 4.935), and only 1% sites accepted a $\omega 1 = 1.787$ and were occupied by positive selection. The substitute models M8 and M2 were together accepted. Consequently, evolution analysis of 10 species revealed that the FSHβ gene was under purified selection and there was no evidence of positive selection found previously among 10 species.

***Phylogenetic relationship of FSHβ gene among species*:**

Table 3. Coding sequences similarity analysis and values of pair wise comparison of FSHβ gene among different species.

Species	Sheep	Goat	Human	Cattle	Mouse	Rat	Rabbit	Horse	Chicken	Frog
Sheep		55	48	54	48	47	49	56	39	28
Goat	55		48	54	48	47	49	55	39	30
Human	48	48		38	39	42	40	41	54	29
Cattle	54	54	38		47	46	49	50	40	28
Mouse	48	48	39	47		54	49	48	39	27
Rat	47	47	42	46	54		54	48	39	29
Rabbit	49	49	40	49	49	54		54	34	38
Horse	56	55	41	50	48	48	54		51	28
Chicken	39	39	54	40	39	39	34	51		35
Frog	28	30	29	28	27	29	38	28	35	

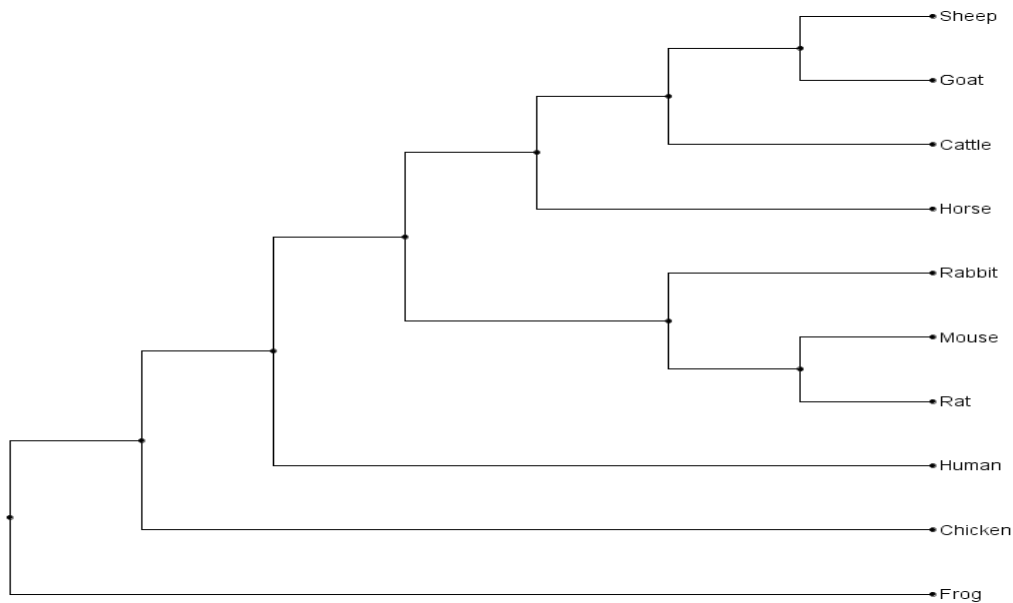Similar values in the under triangle (%),.

Figure 2. Phylogenetic relationship of FSHβ gene in different species.

From GenBank public database, 10 species complete coding sequence was collected and kept goat as reference. All collected sequences and their corresponding amino acids were aligned through MEGA4.0 (Tamura et al., 2007). Before using the software, manual editing of collected sequences was performed for phylogenetic analysis. The Neighbour Joining tree (Fig. 2) shows that the caprine FSHβ was closely related to FSHβ of sheep, rat, mouse, cattle, human, horse and rabbit (BP = 100%). The nucleotide sequence of FSHβ gene of goats showed much closer genetic similarity to the FSHβ gene of sheep (55%) and horse (55%) than to that of cattle (54%), rabbit (49%) and human (48%) (Table 3).

DISCUSSION

The ongoing progress of significant catalogs of genetic deviation has encouraged new interests in discovering targets of positive selection, which ultimately, will help to clarify the role of selection and drift evolution. Moreover, genomic regions important for genome functioning are restricted by the signatures of positive selection. Thus, recognition of this type of genomic regions will help the recognition of genetic deviation that affects phenotype assortment and helps elucidating genome function. Furthermore, each target of positive selection describes a lot of information about the chronological forces and proceedings which shaped the history of a population.

Identification of the evidence for selection is difficult, as the effect of selection on the distribution of genetic diversity can be suppressed by population analysis history (Akey *et al.*, 2002). The SNPs genetic markers can provide the prospect to recognize a genome-wide scan for signature of selection (Sunyaev *et al.*, 2000; Fay *et al.*, 2001). In the outlier methods, the genes under selection are recognized in the extreme tails of the sample distributions. These methods have developed as an extensively used approach for identification of selection signatures in genome wide scans (Akey *et al.*, 2002; Storz *et al.*, 2004). While studying multiple populations, the inconsistency of F_{ST} gene among the loci is lowered which favors the detection of outlier loci (Beaumont and Nichols, 1996).

By manipulating F_{ST} and mean heterozygosity to determine genetic discrimination for each locus, we have identified FSHβ gene which has been potentially marked for positive selection following outlier method to elucidate the distribution of genetic variability. Goat, a domestic animal reared under natural environmental conditions faces two different types of pressures including humans and environment that proposes its probable selection for FSHβ gene. The goat breeds used in this experiment belonged to four different regions of China having different types of temperate areas which may be a reason for the genetic variation among the goats and lead to selection.

FSHβ gene was under evolution and showed variation from neutrality in all populations in study performed by Grigorova *et al.* (2007), suggesting a possible effect of balancing selection but in our study this gene was under positive selection. The possible reason of the selection of this gene might be evolution of this gene from balancing to positive selection due to adaptation to the new environment.

By manipulating F_{ST} and mean heterozygosities as a determinant of genetic seclection for each locus, we recognized FSHβ gene in goat populations, which has been found as potential target of selection according to the distribution of its genetic diversity. This gene is mainly associated with reproductive system and can be selected by adaptive environmental pressure. This gene also showed evidence of balancing selection (Grigorova *et al.*, 2007) previously, while now it seems to be under positive selection. Enriched genes encompassing targets of positive selection are obtained from the simple outlier technique employed in this study. In fact, all outliers are not a depiction of target gene, and the extreme values observed in a distribution comprise of both misleading and actual positives. The power and specificity of genetic selection is estimated by its magnitude. Hence, a strong selection is expected to identify about 50% of the genes under consideration, while a moderately intense selection has almost zero power to detect so many genes. The simulations, recorded here signify intrinsic hurdles to differentiate between confusing impact of genetic drift and natural selection. Our simulations were obtained by simplifying the genome and variability in the frequency of mutation, recombination and selection coefficient across loci, while the population disturbances were not considered. Thus, this study highlights need of further theoretical knowledge and appropriate techniques to harvest maximum information from the studies on genetic variation and selection.

***Conclusion*:** Among the 10 species we did not find any previous positive selection using the likelihood models but recent result of 13SNPs revealed that FSHβ loci was identified as outliers and positively selected. FSHβ gene identified in this study can be useful for breeding purposes in selected goat species.

REFERENCES

Adkins-Regan, E. 2005. Hormones and Animal Social Behavior. Princeton Univ. Press, USA.

Akey, J.M., G. Zhang, K. Zhang, L. Jin and M.D. Shriver. 2002. Interrogating a high-density SNP map for

signatures of natural selection. Genome Res. 12:1805-1814.

Alatalo, R.V., J.H. Glund, A. Lundberg, P.T. Rintamaki and B. Silverin. 1996. Testosterone and male mating success on the black grouse leks**. Proc. R. Soc. Lond. B: Biol. Sci. 263:1697-1702.

Antao, T., A. Lopes, R.J. Lopes, A. Beja-Pereira and G. Luikart. 2008. LOSITAN: a workbench to detect molecular adaptation based on a *Fst*-outlier method. BMC Bioinformatics 9:323.

Beaumont, M.A. and R.A. Nichols. 1996. Evaluating loci for use in the genetic analysis of population structure. Proc. R. Soc. Lond. B: Biol. Sci. 263:1619-1626.

Bonier, F., P.R. Martin, I.T. Moore and J.C. Wingfield. 2009. Do baseline glucocorticoids predict fitness? Trends. Ecol. Evol. 24:634-642.

Breuner, C.W., S.H. Patterson and T.P. Hahn. 2008. In search of relationships between the acute adrenocortical response and fitness. Gen. Comp. Endocrinol. 157:288-295.

Buchanan, F.C., L.J. Adams, R.P. Littlejohn, J.F. Maddox and A.M. Crawford. 1994. Determination of evolutionary relationships among sheep breeds using microsatellites. Genomics 22:397-403.

Ellegren, H., S. Moore, N. Robinson, K. Byrne, W. Ward and B.C. Sheldon. 1997. Microsatellite evolution– a reciprocal study of repeat lengths at homologous loci in cattle and sheep. Mol. Biol. Evol. 14:854-860.

Fay, J.C., G.J. Wyckoff and C.I. Wu. 2001. Positive and negative selection on the human genome. Genetics 158:1227-1234.

Grigorova, M., K. Rull and M. Laan. 2007. Haplotype structure of *FSHB,* the beta-subunit gene for fertility-associated follicle-stimulating hormone: Possible influence of balancing selection. Ann. Hum. Genet. 71:18-28.

Ketterson, E.D., J.W. Atwell and J.W. McGlothlin. 2009. Phenotypic integration and independence: hormones, performance, and response to environmental change. Integr. Comp. Biol. 49:365-379.

Mills, S.C., A. Grapputo, E. Koskela and T. Mappes. 2007. Quantitative measure of sexual selection with respect to the operational sex ratio: a comparison of selection indices. Proc. R. Soc. Lond. B: Biol. Sci. 274:143-150.

Mukesh, M., M. Sodhi, S. Bhatia and B.P. Mishra. 2004. Genetic diversity of Indian native cattle breeds as analysed with 20 microsatellite loci. J. Anim. Breed. Genet. 121: 416-424.

Nijhout, H.F. 2003. Development and evolution of adaptive polyphenisms. Evol. Dev. 5: 9-18.

Reed, W.L., M.E. Clark, P.G. Parker, S.A. Raouf, N. Arguedas, D.S. Monk, E. Snajdr, V. Nolan and E.D. Ketterson. 2006. Physiological effects on demography: a long-term experimental study of testosterone's effects on fitness. Am. Nat. 167:667-683.

Storz, J.F., B.A. Payseur and M.W. Nachman, 2004. Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. Mol. Biol. Evol. 21:1800-1811.

Sunyaev, S.R., W.C. Lathe Iii, V.E. Ramensky and P. Bork. 2000. SNP frequencies in human genes: an excess of rare alleles and differing modes of selection. Trends Genet. 16:335-337.

Swindell, S.R and T.N. Plasterer. 1997. SEQMAN. In Sequence data analysis guidebook. Springer, New York, USA. pp.75-89.

Tamura, K., J. Dudley, M. Nei and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596-1599.

Ward, D.N. 1991. Gonadotropins. In: P.T. Cupps (ed.), Reproduction in Domestic Animals, 4th Ed. San Diego, USA, Academic Press. pp.25-80.

Wei, C., J. Lu, L. Xu, G. Liu, Z. Wang, F. Zhao and C. Liu. 2014. Genetic structure of Chinese indigenous goats and the special geographical structure in the southwest China as a geographic barrier driving the fragmentation of a large population. PLOS one. Doi: 10.1371/journal.pone.0094435

Williams, T.D. 2007. Individual variation in endocrine systems: moving beyond the "tyranny of the Golden Mean. Phil. Trans. R. Soc. B. 363:1687-1698.

Zumbo, P. 2014. Phenol-chloroform Extraction. Weill Cornell Medical College. Available online at http://physiology.med.cornell.edu/faculty/mason/lab/zumbo/files/PHENOL-CHLOROFORM.pdf