B-1371
[1-6]

RESEARCH ARTICLE Indian Journal of Animal Research

Re-sequencing of Whole Genome Jordanian Awassi Ewes using Hiseq Sequencing Technology: Major Step Towards Sheep Genomic Selection

Hussein Migdadi[1], Nizar Haddad[1], Ruba AlOmari[1], Mohammad Brake[2], Mustafa AlShdaifat[1], Sadeer Amasha[1], Monther sadder[3]

10.18805/IJAR.B-1371

ABSTRACT

Background: Jordanian Awassi sheep (*Ovis aries*) is the dominant fat tail sheep breed that appeals to customers because of its various production systems, including fiber, meat and milk. This report is the first whole ewe genome sequence (WGS) of *O. aries* submitted in the NCBI database from Jordan.

Methods: 64 Paired-end sequencing libraries were constructed and subjected to Illumina Hiseq 2500 sequencing system. High-quality reads were aligned against the reference sheep genome and detecting comprehensive sources (SNPs, InDels, SV, CNVs) of genetic variations. We have deposited data sequences at the NCBI under SRA (sequence reads archives) under the accession numbers SRR11128863, PRJNA574879.

Result: Genome resequencing of Jordanian Awassi ewe was carried out with approximately 93.88 Gb with a mapping rate and effective mapping depths were 99.28% and 36.32. Around 19 million SNPs, 3,6 million InDels, 35,180 Structure variation and 13,524 copy number variation among the Jordanian ewe genome were detected. This wide range of genetic variation provides a framework for further genetic studies that will help understand the molecular basis underlying phenotypic variation of economically important traits in sheep and improve intrinsic defects in domestic sheep breeds.

Key words: Ewes, Insertions/deletions, *Ovis aries,* Selection, The copy number of variations.

INTRODUCTION

Domestication of sheep (*Ovis aries*) in the Fertile Crescent is back to 11,000 years ago (Zeder, 2008). It spread west throughout Europe, south into North Africa and east into Asia (Chessa *et al.,* 2009) and became adapted to a wide range of environments and has been selected for varied production systems, including fiber, meat and milk. During domestication, natural and artificial selection play a role in remarkable changes occur in sheep behavior, physiological and morphological appearance and other essential traits (Wang *et al.,* 2019) and a great variety of sheep breeds have been developed (Larson *et al.* 2014). Before wool and milk, man first reared the sheep for access to meat (Chessa *et al.,* 2009), which played an essential role in human society, increased human subsistence stability and spread globally. The environmental heterogeneity and differences affect the phenotypic and genetic variation across sheep populations in climatic factors. Studying the impact of climate on selective signatures well help in the detection of the genetic basis of local adaptation and speciation in response to changing climates (Joost *et al*., 2007), which will provide insight into functionally important variants (Nielsen *et al.,* 2007; Sah and Dixit, 2021) and will help in developing appropriate breeding programs under scenarios of future climate changes (Howden *et al.,* 2007). Approximately 25% of the world's sheep breeds have belonged to fat-tailed sheep (Farid *et al*., 1983). These breeds can survive in harsh environments and satisfy human intake of dietary fat.

[1]National Agricultural Research Center, Baqa, 19381, Jordan.
[2]Department of Science, Jerash University, Jordan.
[3]Department of Horticulture and Crop Science, Faculty of Agriculture, University of Jordan Amman, Jordan.

Corresponding Author: Hussein Migdadi, National Agricultural Research Center, Baqa, 19381, Jordan.
Email: h.migdadi@gmail.com

How to cite this article: Migdadi, H., Haddad, N., AlOmari, R., Brake, M., AlShdaifat, M., Amasha, S. and Sadder, M. (2021). Resequencing of Whole Genome Jordanian Awassi Ewes using Hiseq Sequencing Technology: Major Step Towards Sheep Genomic Selection. Indian Journal of Animal Research. DOI: 10.18805/IJAR.B-1371.

Submitted: 09-04-2021 **Accepted:** 29-04-2021 **Online:** 28-07-2021

Jordan has only Awassi sheep of various ecotypes and strains that originated and were genetically developed for many past generations (Al-Atiyat *et al.,* 2012). However, more information and details of Jordanian Awassi sheep breeds, including morphological, phenotypical and production characteristics and genetic diversity and structure, were reported. Awassi sheep (*Ovis aries*) is the dominant fat tail sheep breed, which is also dominant in Mediterranean countries. They well characterized it for their carpet wool, body weight and morphological characteristics at different ages (Tabbaa, 2003). Based on mitochondrial DNA (mtDNA) and microsatellite-based evidence, Al-Atiyat *et al.* (2018) and Jawasreh *et al.* (2018) reported that

Jordanian Awassi has common haplotype groups (B and A) and group C with Saudi, Kuwaiti, Egypt and Yemen sheep. Progress toward genetic improvements of sheep using conventional crossbreeding methods for traits associated with meat, milk, wool and reproduction was slow.

National Agricultural Research center set up objectives to decipher molecular mechanisms of the response of locally adapted genotypes and breeds to environmental stress and use the information generated to select elite lines and breeds that cope with harsh environmental conditions.

The first whole-genome resequencing of Jordanian Awassi ram was published (Haddad *et al.,* 2020), which reported the availability of wide genome variability (e.g., SNPs, InDel, SV, CNV) that could be helpful in the signature's detection of selection to reveal new insight into the mechanism of advanced breeding, artificial selection and genes associated essential traits. Genetic studies on Jordanian Awassi revealed genetic variability among and within breeds (Al-Atiyat *et al.,* 2014; Jawasreh *et al.,* 2018), genetic parameters related to growth performance, carcass characteristics and meat quality (Jawasreh *et al.,* 2019). To increase our knowledge of Awassi sheep evolution and breeding, national research on sheep breeding strategy will focus on selection, including an advanced genotyping test for economic traits found in Awassi. To advance genome selection techniques, speed up Awassi sheep improvement, putative selection, understanding disease inheritance and get insights into the genetic control of milk, meat and wool production, we conducted whole-genome resequencing of Jordanian Awassi ewe.

MATERIAL AND METHODS

Animal Samples

All animals were handled according to the ethical approval of relevant national legislation (Approval Code No. (G/6) 2006): Instructions and conditions for the acquirer of test animals and testing them.

Sample collection and DNA extraction

Whole-blood samples (10 mL) via EDTA-coated vacutainer tubes were collected from 55 individuals of Awassi landraces from three main Awassi sheep-producing areas in Jordan to minimize potential bias because of overrepresentation of local effects (*e.g.*, inbreeding) and kept in the freezer until DNA extraction. Genomic DNA was extracted using the DNA purification kit (Promega, Wisconsin, USA) following the manufacturer's instructions.

Genome sequencing

For genome sequencing, 0.5 µg of genomic DNA from each sample fragmented and after electrophoresis, DNA fragments of the desired length are gel purified. 64 Paired-end sequencing libraries were constructed according to the manufacturer's instructions (Illumina Inc., San Diego, CA, USA) and subjected to Illumina Hiseq 2500 sequencing system (Illumina, San Diego, CA, USA).

Raw reads contain adapters, unknown or low-quality bases discarded. The parameters of SOAPnuke (Chen *et al.,* 2018) software for this project is "-n 0.05 -l 20-q 0.2 -G -Q 2." The filter steps included removing reads more than half of the bases' qualities are less than five and get clean reads. High-quality reads were aligned against the reference sheep genome assembly using Burrows-Wheeler Aligner (BWA; Li and Durbin 2009) software. BAM format files are used to store the alignment information. After fixing mate-pair information, they were adding read group information and marking duplicate reads caused by polymerase chain reaction, variant calling and detecting single-nucleotide polymorphisms (SNPs) and small Insertion/Deletions (InDels) by GATK software (https://www.broadinstitute.org/gatk/) (McKenna *et al.,* 2010). BreakDancer software (http://breakdancer.sourceforge.net/) (Chen *et al.,* 2009) used for the identification of Structure Variants (SVs) and Copy Number Variants (CNVs). At each stage of the analysis, quality control was applied for the alignment and the called variant (Fig 1). The genome raw data sequences of Jordanian Awassi sheep have been deposited with the NCBI under SRA (sequence reads archives) under the following accession number SRR11128863, PRJNA574879. SRA records accessible with the following link: https://www.ncbi.nlm.nih.gov/sra/PRJNA574879.

RESULTS AND DISCUSSION

Pools of DNA of fifty-five Jordanian Awassi ewe sheep (Fig 2), were sequenced at approximately 60× coverage and then analyzed jointly with publicly available the sheep reference genome covering with an average of 97.78% of the reference genome (Jiang *et al.,* 2014). 99.22 Gb of genome data were generated in the present study. For quality insurance, the raw data was modified by deleting the adapter pollution in reads and then the reads which contain over 50% low-quality bases (quality value <= 12) were removed and the stringent quality filtering yielded 93.88 Gb of genome data.

Results in Table (1) summarized the raw and cleaned reads of the whole genome sequence. The Q20 base rate of each lane is above 97%, so the data quality is very high (Fujimoto *et al*., 2010)

The genome size is 2,584,832,510 bp and the effective size is 2,560,768,911bp. Burrows-Wheeler Aligner (BWA) software used for sequence reads alignment to the reference genome. The mapping rate and the final effective mapping depth were 99.28% and 36.32. Based on the consensus sequence, the polymorphic loci between our breed and the reference filtered and a high-fidelity SNP data set generated (Table 2). 18,994,659 SNPs among Jordanian ewe genomes. The SNP results are listed as CDS, exon and genes. In this study, 152,684 synonymous and 102,531 non-synonymous SNPs were annotated among this sheep genome. Furthermore, the distribution of predicted large-effect SNPs is presented in Fig 3. Premature stop codons were 1016 and 1071 were disrupt splicing donor or acceptor sites.

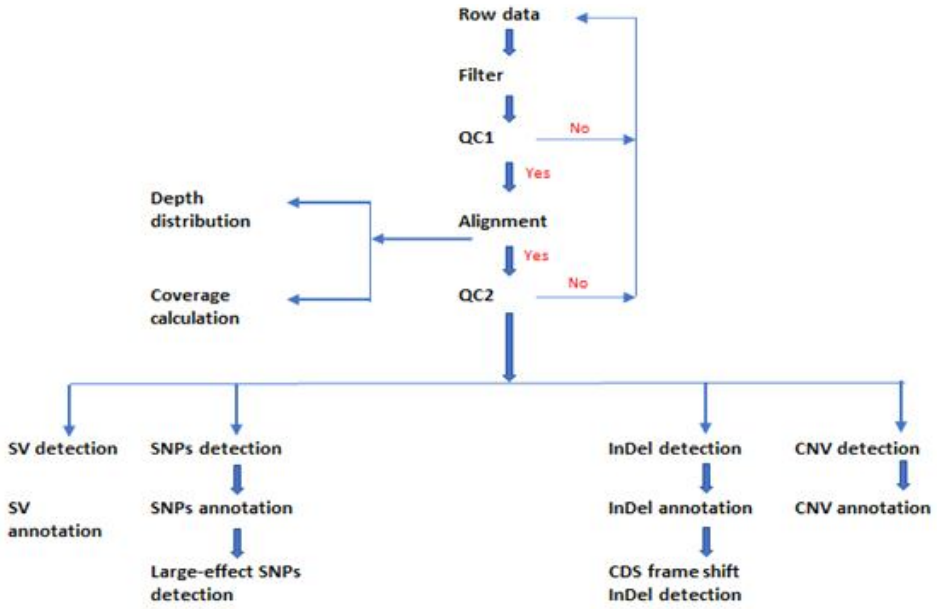

Fig 1: The pipeline of Standard Bioinformatics Analysis. The pipeline based on the family of SOAP software. QC; quality control.

Fig 2: A typical Awassi sheep (Ovis aries) is the dominant fat tail sheep breed in Jordan.

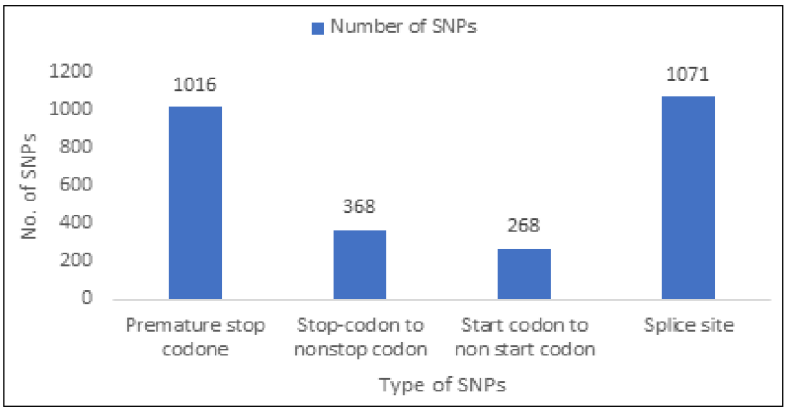

Fig 3: Number of SNPs generated and the distribution of predicted large-effect of SNPs.

Short InDel detection and annotation

The Genome Analysis Toolkit (GATK) was used to detect InDels. 3,626,673 InDels generated and distributed as 1,782,830 insertions and 1,843,843 were deletions (Table 3). Furthermore, the frameshift InDel based on the InDel annotation has resulted in 43,614.

Structure *and* copy number variation detection and annotation

The structural variation detected includes deletion, insertion, duplication, inversion and transposition. 43.14% were deletion type of SV. Over 189326 were in gene sites. The SV results are presented in Table (4).

Copy Number Variation (CNV) is a basic form of structural variation among individuals of the same species. Genome regions deleted or duplicated on some chromosomes between the sequencing individual and the reference have corresponded to CNVs. The CNV results listed in Table (5); Over 72% of the CNV were down-regulated. Recently, the whole-genome resequencing of Jordanian Awassi Ram (*Ovis aries*) using Hiseq sequencing technology was reported (Haddad *et al.,* 2020). The report considered this research as the first step towards sheep genomic selection of Awassi ram. Based on this report, the authors urged similar work to be performed on Awassi ewe's genome. They justified that whole-genome resequencing of Jordanian Awassi Ram was considered a preliminary stage requiring similar whole sequencing of Awassi ewe's genome to start Awassi sheep genomic research selection. Genome-wide studies used High-throughput sequencing and screening technologies on sheep are reported. They included single-nucleotide polymorphisms, induction variations, structural variation, copy number variants at the whole-genome level, transcriptome and DNA methylation sequences and comprehensive information on functional genes genetic variants associated with economically important traits. Studies to decipher the genetic basis of animal domestication are dramatically increased and considered an active research area. Genomic signatures related to domestication and improvement were found in response to demographic and selective differences between wild and domestic populations (Florian J. Alberto *et al.,* 2018). Li *et al.* (2020) stated that the availability of whole-genome sequences provides an opportunity to study domestication at the gene mutation level. They re-sequenced the whole genomes of 248 wild, landrace and improved sheep and detected genomic regions: targets of domestication, breeding and selection. Furthermore, they found non-synonymous mutations in candidate genes and significant differences in their allele frequency distributions across breeds, such as PDGFD (platelet-derived growth factor D) gene for fat deposition in the tails of sheep. Pan *et al.* (2018) performed whole-genome sequencing of 99 sheep and identified candidate genes revealed local adaptations of the sheep associated with sensory perception, muscle strength, eating habits, mating process and aggressive behavior.

Table 1: Summary of the sequencing and alignment results from raw and clean reads in the Jordanian Awassi ewe Sheep.

Sequence	Value (Raw Data)	Value (Clean Data)
GC rate%	43.66%	43.50%
Q20 rate%	97.92%	98.73%
Q30 rate%	91.83%	93.96%
Reads	992.22	938.81 Mb
Bases	99.22	93.88 Gb
Clean Data/Raw	—-	94.62%
Sample Coverage rate (%)		97.78
Map reads rate (%)		99.30
Map bases rate (%)		99.30
Uni hit reads rate (%)		85.81
Uni hit bases rate (%)		85.21

Mb and Gb; mega and giga base pairs.

Table 2: Summary of genome-wide SNPs and their annotation distribution in the Jordanian Awassi ewe Sheep.

SNP distribution	Value
SNP	18,994,659
Homo	2,566,429
Hete	16,428,230
Syn CDS	152,684
Nonsyn CDS	102,531
Exons	798,273
genes	7,749,608
mRNA	22,021,539
ncRNA	901,190
transcript	664328
tRNA	798

CDS: coding sequence; Homo: homozygous; Hete: heterozygous; ncRNA non-coding RNA Nonsyn: non-synonymous mutations; Syn: synonymous mutations.

Table 3: Insertion and deletion, annotation, distribution and a frameshift mutation in CDS in the genome of Jordanian Awassi ewe Sheep.

Insertion and Deletion	Value
In Del	3,626,673
Insertion	1,782,830
Deletion	1,843,843
CDS	50,544
C- gene segment	604
V- gene segment	4,282
cDNA match	41,320
Exon	160,756
Gene	1,488,817
mRNA	4,005,145
Frame-shift Mutation CDS	43,614
3X-shift Mutation CDS	6,930
3X-shift Mutation CDS Phase 0	2,284
3X-shift Mutation CDS PhaseNo0	4,646

*CDS: Coding Sequence

Table 4: Structure variation annotation, distribution in the genome of Jordanian Awassi ewe Sheep.

Structure variation	Value
SV	35,180
Insertion	29
Deletion	15,178
Others	19,973
CDS	3,500,033
C- gene segment	41
V- gene segment	432
cDNA match	227,951
Exon	3,982,793
Gene	189,326
mRNA	30,135

*CDS: Coding Sequence

Table 5: Copy number variation distribution and annotation distribution in the the genome of Jordanian Awassi ewe Sheep.

Copy number variation	Value
CNV	13,524
Up-regulation	3,757
Down-regulation	9,767
CDS	3,797
C- gene segment	2
V- gene segment	57
cDNA match	1000
exon	5,085
gene	4,847
mRNA	11,690

*CDS: Coding Sequence

Furthermore, Florian J. Alberto *et al.* (2018) determined certain genomic regions carrying genes involved in the nervous system, immunity and productivity traits differentiating domestic breeds of sheep and goats from wild populations. Chang *et al.* (2018), reported two novel miRNAs from the *Ovis aries* may play an important role in inducing embryonic development, sex regulation, production of sperm and ovulation. Moreover, Chang *et al.* (2020), identified few genes serve as potential biomarkers for monitoring disease progression and abortion risk during pregnancy in sheep. Adam Abied *et al.* (2020) used high-density SNP Chip data to characterize the auto-zygosity of five local Chinese sheep breeds belonging to different geographical locations identified by the genomic regions with high runs of homozygosity frequencies within individuals of each breed. These regions included candidate genes associated with disease resistance traits, the innate and adaptive immune response, digestion and metabolism, growth, body size and developments. Brake *et al.* (2021) demonstrated using the complete mitochondrial genome of the Awassi-Jo breed (*Ovis aries*) data to construct the phylogenetic tree. They showed that Awassi-Jo diverged earlier than related breeds (Turkey, Italy, Germany and Netherland) with a common ancestor in haplogroup HB. The results revealed the importance of mitochondrial data in studying sheep evolution and domestication.

CONCLUSION

Consequent whole-genome resequencing (WGR) has provided geneticists with the opportunity to investigate the genome in depth. These investigations have been recent and cut-edge technology applied to many world sheep breeds. WGR provided contributions to a better understanding of evolutionary forces' effects, population dynamics, the discovery of genes, selecting signatures and starting specific maps and genomes. The present work describes the whole-genome resequencing of Jordanian Awassi ewe, which experienced specific selection pressures over several generations. The wide range of genetic variation provides a framework for further genetic studies that will help understand the molecular basis underlying phenotypic variation of economically important traits in sheep and improve intrinsic defects in domestic sheep breeds.

REFERENCES

Abied, A., Xu, L., Sahlu, B. W., Xing, F., Ahbara, A., Pu, Y., Lin, J., Berihulay H., Islam R., He X., Mwacharo J., Zhao Q. and Ma, Y. (2020). Genome-Wide Analysis Revealed Homozygosity and Demographic History of Five Chinese Sheep Breeds Adapted to Different Environments. Genes. 11(12): 1480.

Al-Atiyat, R.M., Aljumaah, R.S., Alshaikh, M.A., Abudabos, A.M. (2018). Microsatellite-Based Genetic Structure and Diversity of Local Arabian Sheep Breeds. Frontiers Genetics. 9: 408.

Al-Atiyat, R.M., Tabbaa, M.J., Salameh, N., Tarawneh, K. (2012). Analysis of genetic variation of fat tailed-sheep in the southern region of Jordan. Asian Journal Animal Veterinary Advances. 7:378-389.

Al-Atiyat, R.M., Aljumaah, R.S. (2014). Genetic distances and phylogenetic trees of different Awassi sheep populations based on DNA sequencing. Genetic and Molecular Research. 13: 6557-6568.

Alberto, F.J., Boyer, F., Orozco-terWengel, P., Streeter, I., Servin, B., de Villemereuil, P., *et al.* (2018). Convergent genomic signatures of domestication in sheep and goats. Nature communications. 9(1):1-9.

Brake, M.H., Migdadi, Hussein M., Sadder M.T., Awabdeh, S., Jawasreh, Kh., Obeidat W., Al Omari, R. and Haddad, N.J. (2021). Complete mitochondrial genome sequence of Awassi-Jo sheep breed (*Ovis aries*) in Jordan. Mitochondrial DNA Part B. 6(3): 1263-1264.

Chang, W., Wang, J., Zhang, Y., and Wu, J. (2018). Discovery of two novel miRNAs from the Ovis aries by a combinatorial approach of experiments and bioinformatics. Indian Journal of Animal Research. 52(8): 1155-1161.

Chang, W., Cui, Z. and Wang, J. (2020). Identification of Potential Disease Biomarkers in the Ovaries of Dolang Sheep from Xinjiang using Transcriptomics and Bioinformatics Approaches. Indian Journal of Animal Research. DOI: 10.18805/ijar.B-1265.

Chen, K., Wallis, J., McLellan, M., Larson, D., Kalicki, J., Pohl, C., *et al.* (2009). Break Dancer: an algorithm for high-resolution mapping of genomic structural variation. Nature methods. 6: 677.

Chen, N., Cai, Y., Chen, Q., Li, R., Wang, K., Huang, Y., *et al.* (2018). Wholegenome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. Nature communications. 9: 2337.

Chessa, B., Pereira, F., Arnaud, F., Amorim, A., Goyache, F., Mainland, R., Pemberton, J.M., Beraldi D., Stear M.J., *et al.* (2009). Revealing the history of sheep domestication using retrovirus integrations. Science. 324(5926): 532-6.

Farid, A., Izedifard, J., Edris, M. and Makarechian, M. (1983). Meat from culled old ewes of two fat-tailed Iranian breeds, II-Meat, subcutaneous fat and bone in the wholesale cuts. Iran Agricultural Research. 2(2): 93-114.

Fujimoto, A., Nakagawa, H., Hosono, N., Nakano, K., Abe, T., Boroevich, K.A. *et al*. (2010). Whole-genome sequencing and comprehensive variant analysis of a Japanese individual using massively parallel sequencing. Nature genetics. 42(11): 931-938.

Haddad, N., Migdadi, H., Al-Atiyat, R., Jawasreh, K., Awabdeh, S., Obeidat, W., AlOmari, R., Aldamra, M., Ababneh, H., Tabbaa, M.J. and Brake, M. (2020). Whole Genome re sequencing of Jordanian Awassi Rams (*Ovis aries*) Using Hiseq Sequencing Technology: The First Step Towards Sheep Genomic Selection. Genetics Molecular Research. 19(2): 1-14.

Howden, S.M., Soussana, J.F., Tubiello, F.N., Chhetri, N., Dunlop, M., and Meinke, H. (2007). Adapting agriculture to climate change. Proceedings of the National Academy of Sciences. 104(50): 19691-19696.

Jawasreh, K., Amareen, A.A., Aad, P. (2019). Effect and Interaction of β-Lactoglobulin, Kappa Casein and Prolactin Genes on Milk Production and Composition of Awassi Sheep. Animals. 9(6): 382.

Jawasreh, K., Ismail, Z.B., Iya, F., Castañeda-Bustos, V.J., Valencia-Posadas, M. (2018). Genetic parameter estimation for pre-weaning growth traits in Jordan Awassi sheep. Veterinary World. 11(2): 254-258.

Jawasreh, K.I., Ababneh, M.M., Ismail, Z.B., Younes, A.M., Sukhni, I.A. (2018). Genetic diversity and population structure of local and exotic sheep breeds in Jordan using microsatellite markers. Veterinary World. 11(6): 778-781.

Jiang, Y., Xie, M., Chen, W., Talbot R., Maddox J.F., Faraut T., Wu C., et al. (2014). The sheep genome illuminates biology of the rumen and lipid metabolism. Science. 344: 6188.

Joost, S., Bonin, A., Bruford, M.W., Després, L., Conord, C., Erhardt, G. and Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Molecular ecology. 16(18): 3955-3969.

Larson, G. and Fuller, D.Q. (2014). The evolution of animal domestication. Annual Review of Ecology, Evolution and Systematics. 45: 115-136.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25(14): 1754-1760.

Li, X., Yang, J., Shen, M., Xie, X. L., Liu, G. J., Xu, Y. X. and Li, M. H. (2020). Whole-genome resequencing of wild and domestic sheep identify genes associated with morphological and agronomic traits. Nature communications. 11(1): 1-16.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., *et al.* (2010). The genome analysis toolkit: a map reduce framework for analyzing next-generation DNA sequencing data. Genome Research. 20: 1297-1303.

Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C. and Clark, A.G. (2007). Recent and ongoing selection in the human genome. Nature Reviews Genetics. 8(11): 857-868.

Pan, Z., Li, S., Liu, Q., Wang, Z., Zhou, Z., Di, R., Miao, B., Hu, W., Wang, X., Hu, X. and Xu, Z. (2018). Whole-genome sequences of 89 Chinese sheep suggest the role of RXFP2 in the development of unique horn phenotype as response to semi-feralization. Giga Science. 7(4): 1-15.

Sah, R., Dixit, S. (2021). Genetic Diversity Analysis among 27 Indian Goat Populations using Microsatellite Markers. Indian Journal of Animal Research. 55(3): 243-254.

Tabbaa, M.J. (2003). Some factors affecting different body and fat-tail dimensions of Awassi ewes in Jordan. Dirasat. 30: 280-284.

Wang, W., Zhang, X., Zhou, X., Zhang, Y., La, Y., Zhang, Y., *et al.* (2019). Deep genome resequencing reveals artificial and natural selection for visual deterioration, plateau adaptability and high prolificacy in Chinese domestic sheep. Frontiers in genetics. 10: 300.

Zeder, M.A. (2008). Domestication and early agriculture in the Mediterranean Basin: origins, diffusion and impact. Proceedings of the national Academy of Sciences. 105: 11597-604.