J. Agr. Sci. Tech. (2023) Vol. 25(3): 581-595

Comparative Analysis of Differential Exon Usage between the Breeds of Sheep and Goats Using RNA-Seq Data

F. Nikookalam Azim[1], A. A. Masoudi[1]*, A. Ehsani[1], R. Vaez Torshizi[1], and P. Shariati Gazgazareh[1]

ABSTRACT

Alternative splicing, alternative transcript start site, and alternative transcript polyadenylation site are the main factors resulting in diversity of the transcripts of a gene. The main objectives of this study were to analyze the alternation process in breeds of sheep and goat, and to identify its role in differentiation of breeds of a species. RNA-seq data were prepared from ovarian tissue of two breeds of Shal and Sangsari sheep and two breeds of Tibetan and Jintang black goats. Reads were aligned to the reference genome and significant genes with respect to differential exon usage were identified. The statistical comparison revealed that 8,104 genes were significantly different in exon usage between the sheep breeds and 173 genes differed between the goat breeds. Out of the 121,861 studied exons, only 22.7% were preserved during future generations between the breeds, of which 99.3% did not display any alternatives. The high protection was probably due to the lack of involvement of the exons in alternative process. The genes with differential exon usage in goat had a higher percentage of alternatives than those in sheep. The interracial analysis showed that alternative splicing was the most influential type of alternatives in the breeds of sheep and goats. It seems that the conservation process of the exons is related to the contribution of these exons in alterative process in both sheep and goat breeds. The significant PI3K-Akt and alternative splicing pathways play a role in cell growth, development of ovaries, and mRNAs splicing.

Keywords: Alternative splicing, Interethnic analysis, Ovarian tissue, Transcription.

INTRODUCTION

After the discovery of genes, as inherited materials by Friedrich Miescher [22], investigators used various tools and methods to answer the question on how many genes are in the genome of animals [5, 33, 36, 39]. Plenty of researches in a few decades ago identified around 120,000 genes in the animal genomes [35], while the number of genes has been decreasing in various studies and has reached about 19000 genes in 2019 [24]. But, where did all this difference come from? Actually, it seems that the reason for this difference was the researchers' perception. At first, they assumed each gene expresses one type of protein and the number of genes is equal to the number of proteins [27]. However, the next studies led to the discovery of the alternative splicing process and showed that one gene could produce different proteins. Therefore, the move to discover the processes that lead to the production of different proteins from a single gene made it less biologically important to study the number of genes [16]. In many studies, the process type of these reactions has been used separately or simultaneously to detect a specific process because several diseases, including cancers, have been associated with dysregulation of alternative splicing [1, 9, 11, 29]. From another viewpoint, Differential Exon Usage (DEU) is a broader concept than the

[1] Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Islamic Republic of Iran.
*Corresponding author; e-mail: masoudia@modares.ac.ir

[DOI: 10.22034/jast.25.3.581]

alternative splicing that was developed in recent years. DEU includes alternative transcript start sites and the polyadenylation sites, which result in different uses of the exons at the 5′ and 3′ transcript boundaries [2]. RNA-seq data is a suitable method for studying alternative splicing and isoforms of gene expression [30, 4]. Over the past decades, RNA sequencing has become an essential tool for transcriptome-wide analysis to understand the expression of different genes [6, 7, 13, 23, 34, 37, 38]. However, a study examined the distribution of these alternatives among different tissues in 2018. The results indicated the alternative start site and alternative transcript polyadenylation sites had the greatest effect on the formation of different tissues. Additionally, from a genetics and animal breeding point of view, breeds are very important in various breeding production goals. A breed is a specific group of domestic animals that have specific phenotypes or characteristics. Breeds are formed through genetic isolation (natural adaptation to the environment), selective breeding, or a combination of the two [14].

The expression of eukaryotic genes is temporarily and multidimensionally controlled [32]. Only a relatively small set of the entire genome is expressed in each type of tissue, and the expression of genes depends on the stage of development [20]. Therefore, gene expression in eukaryotes is specific to each tissue [19]. Also, the amount of gene products made in the same tissue as well as in other tissues that make up that product, regulates the expression of that gene [18].

In this study, we aimed to examine the distribution process of the three processes of alternative splicing, alternative transcript start sites, and alternative polyadenylation sites between the ovarian tissue in two breeds of sheep (Iranian native Shall and Sangsari) and two breeds of goats (Chinese native Jintang black and Tibetan). Another objective was to answer the question of which types of alternatives play a role in creating different breeds of the same species, and in fact, which types of alternatives cause the main differences between the breeds. In addition, we planned to examine the differential usage of exons to determine the importance of alternative processes in creating breeds of one species.

MATERIALS AND METHODS

Data Preparation

In this study, total RNA sequences of ovarian tissue were prepared from 9 Iranian sheep including two breeds of Shall and Sangsari and 6 Chinese goats from Tibetan and Jintang black breeds by *in-vivo* or *in-silico* search in biological databases, respectively.

Preparation of Data by *in-Vivo*

Indigenous breeds of sheep, including 9 heads of two breeds of Shal and Sangsari ewes, were obtained from the Iranian research center for agricultural livestock and natural sources of Qazvin and Semnan provinces, respectively. Sheep were bred for 6 months in the livestock farms in the Faculty of Agriculture, Tarbiat Modares University, to uniform the animals' conditions. All ewes were grouped under normal temperature and lighting and they had free access to feed and water until slaughter. Animals were synchronized for estrus cycle using a standard protocol [12] and ovaries were taken for RNA extraction. Total RNAs were extracted by Gene JET TM RNA purification kit (Thermo Science, USA) according to the manufacturer's instructions. The quality and quantity of the extracted RNAs were tested by ultraviolet-visible spectrophotometry and electrophoresis on an agarose gel (Figure 1). Total RNA samples were frozen in liquid nitrogen immediately after extraction and stored at -70°C until the next analyses. Total RNAs were sequenced by the High-

[DOI: 10.22034/jast.25.3.581]

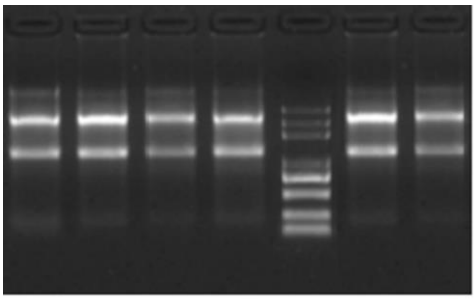

Figure 1. Electrophoresis of the ovarian RNAs of the samples on agarose gel (From left to right; Lane 5 is ladder).

throughput sequencing Illumina Hiseq 2000 system (Beijing Genomics Institute, China).

Preparation of Data by *in-Silico*

RNA sequences of the ovarian tissue of two species of Tibetan and Jintang black goats (3 of each breed) were taken from the Ensembl database with the accession numbers of GSM2342194, GSM2342195, GSM2342196, GSM2342197, GSM2342198, and GSM2342199 [40] using SRAtoolkit software v 2.3.2.

Identification of the Types of Alternatives in the Genome

All steps for analysis of the RNA-seq data in sheep and goat samples had a similar process as follow. First, the quality of all data was assessed using fastqc (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) under Linux system. In this study, the reference genomes of sheep species v3.1 and goat v2 reference genome obtained from the ensemble (https://asia.ensembl.org/info/data/ftp/index.html). Then, annotation files in GTF format were downloaded and GFF format was obtained from Python script, dexseq-prepare-annotation.py DEXSeq package. The alignment of reads was performed using STAR v2.5.3 software. SAM file sorting was performed by Samtools v0.11.3 software and reads were counted using another Python script in the DEXSeq package (dexseq-count.py).

Then, an expression of different isoforms of a gene and the expression of each exon in each gene was evaluated by General Linear Model (GLM) to count the reads. In fact, the special model we used in this study is as follow:

$K_{ijl} \simeq$ NB (Mean= $s_{j\mu ijl}$, Dispersion= α_{il})

NB (Negative binomial distribution).

In this model, K_{ijl} is the number of overlapping reads obtained from the negative binomial by DEXSeq package. We used the distribution obtained from the scattering and the mean, in which the scattering estimate is used by the McCarthy method and is calculated for each counting bin. Then, the calculation of the mean by a logarithm of the predictive linear model [2] is performed as follow:

$$Log\mu_{ijl} = B_i^G + B_{iL}^E + B_{i\rho j}^C + B_{i\rho jl}^{EC}$$

The linear predictor μ_{ijl} is decomposed into four factors as follows: B_i^G represents the baseline expression strength of gene i. B_{iL}^E is (up to an additive constant) the logarithm of the expected fraction of the reads mapped to gene i that overlap with counting bin l. $B_{i\rho j}^C$ is the logarithm of the fold change in the overall expression of gene i under condition ρj (the experimental condition of sample j). Finally, $B_{i\rho jl}^{EC}$ is the effect that condition ρj has on the fraction of reads falling into bin l [2]. As a result, the exons that show differential use in each gene are shown as a purple box (Figure 2).

Next, we used the HTML file extracted from the DEXSeq package to identify the types of alternative splicing in each gene. This file contained four types of charts: counts, expression, splicing, and transcripts for each gene, which show the normal number of reads for each exon, the expression of each exon, the amount of application of each exon, and the different forms of transcription of each gene, respectively. In this study, focusing on transcript diagrams, the types of alternatives created in genes like alternative transcript start sites (2C), alternative splicing (2B), and alternative transcript polyadenylation sites

[DOI: 10.22034/jast.25.3.581]

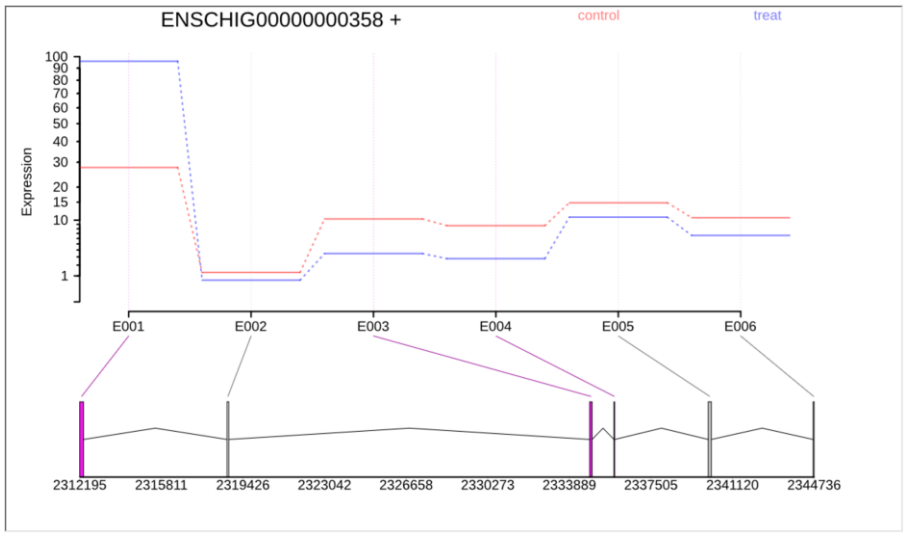

Figure 2. A diagram of differential exon usage (A sample: ENSCHIG00000000358 gene). Y-axis is RPKM measurement unit of expression. This figure shows the output of the transcript between Tibetan (red line; control) and Jintang black (blue line; treat) breeds, calculated according to the logarithm of the predictive linear model. The positive sign (+) on the accession number indicates the expression of the transcript in sense state. The purple boxes represent differential exon usage of the transcript.

(2D) with significant differences between the breeds were examined (Figure 3). To calculate the number of genes involved in the alternative process (Figure 4), we used the following equation:

$GWA= (NG_{ATSS}+NG_{AS}+NG_{ATPS})–[(NG_{ATSS\ \&\ AS}+ NG_{ATSS\ \&\ ATPS}+NG_{AS\ \&\ ATPS}) (NG_{ATSS\ \&\ AS\ \&\ ATPS})]$

There are several important factors for calculating genes with alternative splicing (GWA) as follows: Genes with Alternative Transcript Start Sites (GATSS), Genes involved in Alternative Splicing (GAS), Genes involved in Alternative Transcript Polyadenylation Sites (GATPS), genes involved in two types of simultaneous alternatives (alternative transcript start sites, and alternative splicing, alternative transcript start sites and alternative transcript polyadenylation, alternative splicing, and alternative transcript polyadenylation), and genes involved in three types of simultaneous alternatives (alternative transcript start sites, alternative splicing, and alternative transcript polyadenylation). Subsequently, we identified genes with differential exon usage that were simultaneously involved in some kinds of alternatives (Figure 5).

Gene Ontology

To investigate the structural and functional characteristics of differential exons that create different transcription isoforms between the breeds, the David v 6.8 (https://david.ncifcrf.gov/) database was used [1]. Function and metabolic pathways of the significant genes were determined using the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEEG) databases. The paths considered significance at $P< 0.01$.

RESULTS

Identification and Relationship between the Types of Alternatives in the Genome of Sheep and Goat

Following the procedures described previously, we identified the exons that show alternative usage in each gene. Figure 2 shows the output of a transcript between

[DOI: 10.22034/jast.25.3.581]

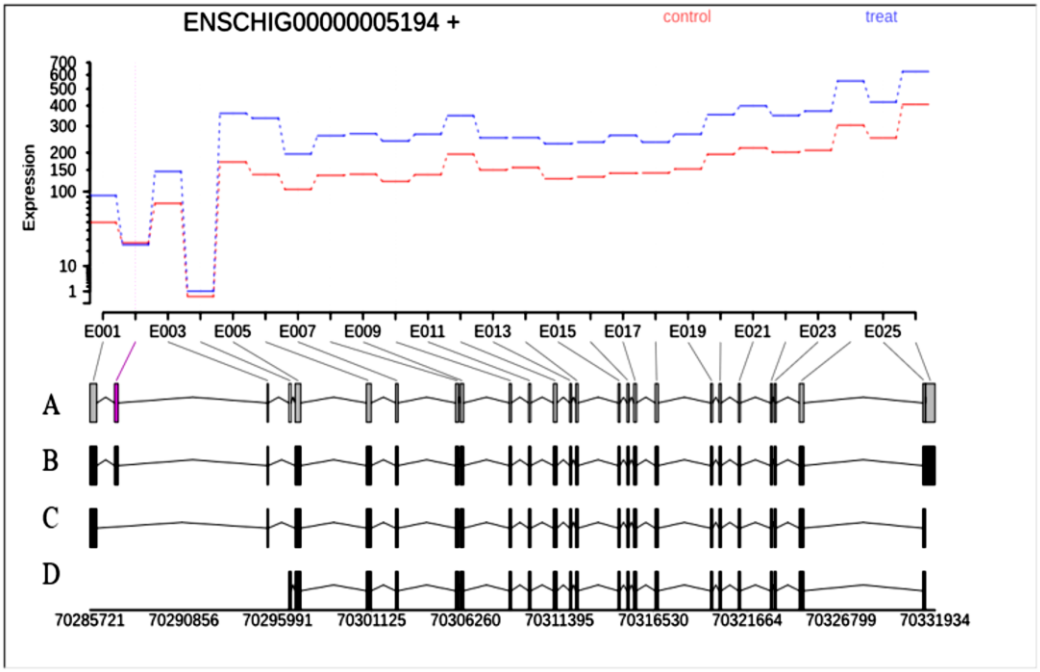

Figure 3. A diagram of the types of alternatives (A sample: ENSCHIG00000005194 gene). Y-axis is RPKM measurement unit of expression. **A**: Contribution of all the exons of the gene in transcription and no alternative occurs. **B**: Alternative splicing of the transcript by non-participation of exon 4. **C**: Alternative of polyadenylation sites of the transcript by non-expression of the last exon (26). **D**: Lack of expression of exons 1, 2, 3, and 26 resulting in alternative transcript start sites and alternative polyadenylation sites in the transcript, simultaneously. The positive sign (+) on the accession number indicates the expression of the transcript in sense state. Tibetan (red line; control) and Jintang black (blue line; treat) breeds.

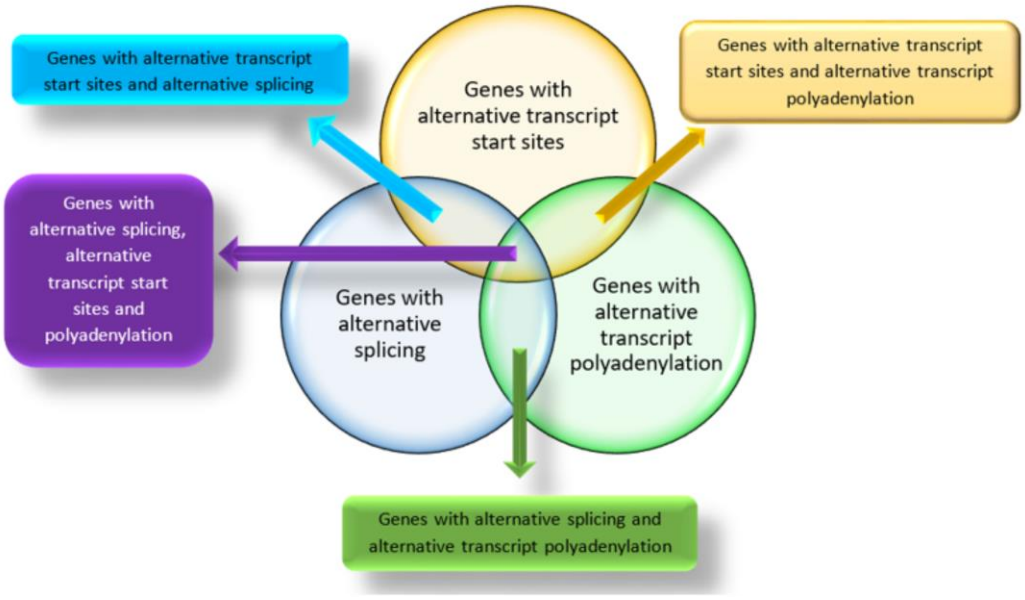

Figure 4. Diagram of the relationship between the types of alternatives in sheep and goat breeds.

[DOI: 10.22034/jast.25.3.581]

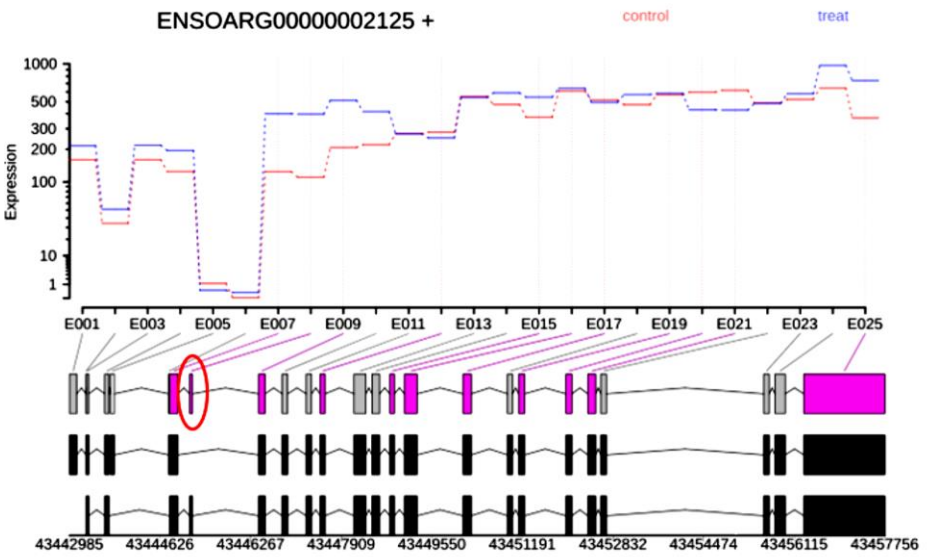

Figure 5. A diagram of the contribution of an exon in differential exon usage and alternative of transcripts (A sample: ENSOARG00000002125 gene). Y-axis is RPKM measurement unit of expression. The exon 8 (red circle) has both the differential use of an exon and generating multiple transcripts of a single gene by alternative splicing (DEUTA). The positive sign (+) on the accession number indicates the expression of the transcript in sense state. Tibetan (red line; control) and Jintang black (blue line; treat) breeds.

Tibetan (control) and Jintang black (treat) breeds, which is calculated according to the logarithm of the predictive linear model. The purple boxes (exons 1 and 3) represent differential exon usage of the transcript. In addition, by identification of the types of alternatives in genes, many kinds of alternatives in the genome were detected (Figure 3). Contribution of all exons of the gene in transcription and no alternative changes in figure 3-A and alternative splicing of the transcript by non-participation of the exon in are shown in figure 3-B. However alternative of polyadenylation sites of the transcript by non-expression of the exon (Figure 3-C) and lack of the expression of exons which resulted in alternative transcript start sites or alternative polyadenylation sites in the transcripts (Figure 3-D) were detected between the breeds. Diagram of the relationship between the types of alternatives in sheep and goat breeds are shown in Figure 4, which shows that several important factors for calculating the alternative splicing (GWA) of the Genes are Alternative Transcript Start Sites (GATSS), Alternative Splicing (GAS), Alternative Transcript Polyadenylation Sites (GATPS) of the genes, and a combination of them. A result of the contribution of an exon in differential exon usage and alternative of transcripts in goat breeds (A sample: ENSOARG00000002125 gene) is shown in Figure 5. The exon 8 (red circle) has both the differential use of an exon and generating multiple transcripts of a single gene by alternative splicing (DEUTA).

Alternative Splicing is the Most Effective Type of Alternative between the Two Breeds of Sheep and Goats

Throughout the interracial analysis by DEXSeq, 8,277 genes showed differential exon usage, of which 8,104 genes were related to differences between Shal and Sangsari sheep breeds, and 173 genes showed differences in exon usage between Tibetan and Jintang black goat breeds. Analyzing the HTML file obtained from the DEXSeq

[DOI: 10.22034/jast.25.3.581]

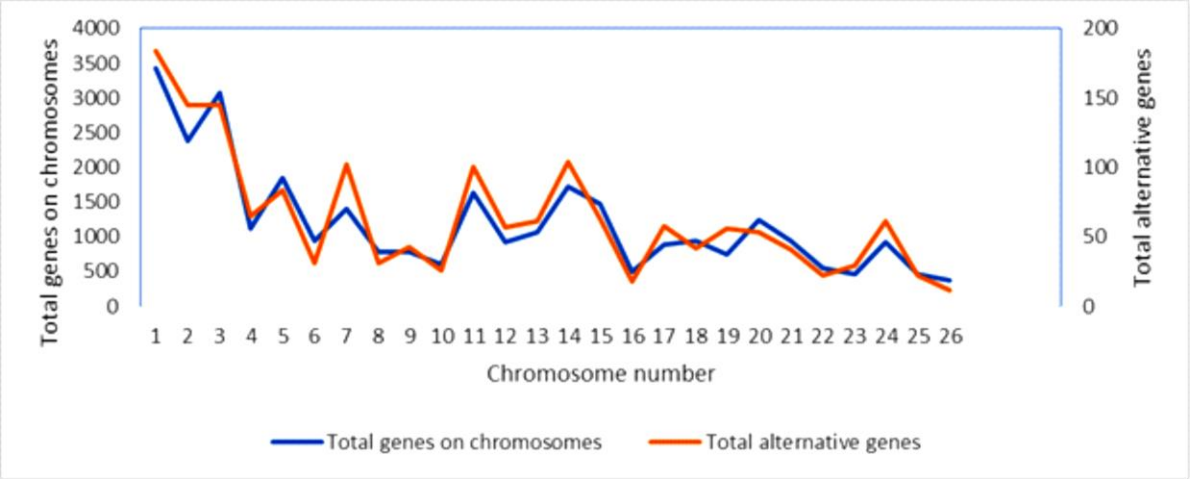

Figure 6. Relationship between the frequency of the alternative genes and the number of genes on each chromosome in sheep breeds.

Table 1. Types of alternatives and the number of genes associated with each alternative in interracial analysis of sheep (Shal and Sangsari) and goat breeds (Tibetan and Jintang black).

Alternation types[a]	Number of alternative genes	
	Sheep	Goat
ATSS	491	72
AS	790	109
ATPS	444	58
ATSS+AS	358	67
ATSS+ATPS	194	40
AS+ATPS	311	55
ATSS+AS+ATPS	125	40

[a] ATSS: Alternative Transcript Start Sites, AS: Alternative Splicing, ATPS: Alternative Transcript Polyadenylation Sites.

package, among the genes with differential exon usage (8,277), a total of 1,104 alternative genes were identified. Of these, 987 genes were related to interracial analysis of sheep and 117 genes were linked to the goats. Out of these genes, 139 and 16 DEUTA genes were identified in sheep and goats, respectively. In total, 121861 exons were studied, of which 27,747 exons are preserved during future generations between both breeds. Also, 178 exons involved in some kinds of alternatives and may be omitted in transcription. The remaining 27,569 exons were not involved in any kind of alternative process and participate in transcriptional activities. Towards identifying the number and type of alternatives in each gene, transcription diagrams were examined. The results showed that some genes had only one type of alternatives, while there were genes that displayed both, or even all, three types of alternatives, simultaneously (Table 1). What can be clearly stated is that interracial analysis identified 790 genes with alternative splicing between the two breeds of Shal and Sangsari, and 109 genes between Tibetan and Jintang black breeds. These are the most effective type of alternatives between both breeds. In this study, the frequency of the alternative transcript start sites was in the second place. Four hundred ninety-one genes in interracial analysis in sheep samples and 72 genes between goat breeds revealed alternative

[DOI: 10.22034/jast.25.3.581]

transcript start site isoforms. Interracial analyses of Shal and Sangsari sheep and Tibetan and Jintang black goats showed, respectively, 444 and 58 genes involved in alternative transcript polyadenylation sites. The number of genes in this process was less than the other forms of alternatives in these breeds of sheep and goat (Table 1).

Distribution of Alternating Genes on Chromosomes

The interracial analysis of the two breeds of Shal and Sangsari determined the distribution of a minimum of 2 genes with alternative transcript start sites on chromosome 16 and maximum of 50 genes on chromosome 1 (Figure 6). OAR26 with 6 genes and OAR1 with 90 genes were recognized by the lowest and the highest number of genes involved in the alternative splicing, respectively. Moreover, 2 genes on OAR26 and 44 genes on OAR1 were related to the process of alternative transcript polyadenylation sites in sheep breeds (Table 2). On the other hand, between the Tibetan and Jintang black, the minimum and maximum genes involved in the alternative transcript start sites were 0 on CHI12, CHI27 and 7 on CHI19, respectively (Figure 7). Besides, the process of alternative splicing did not occur on chromosomes 27 and 12, and CHI19 with 10 genes had the maximum alternative splicing. In goat breeds, seven chromosomes of 11, 12, 14, 20, 21, 27 and 28 with no gene and chromosomes 5, 6 and 7 only with four genes showed the highest number of genes involved in alternative transcript polyadenylation sites (Table 2).

Pathways of the Genes with Different Transcription Isoforms Derived from Differential Exons (DTISDEs)

Many functional and biological pathways of genes with differential exons that create different transcription isoforms between the breeds were significantly ($P< 0.01$) detected in sheep or goat species (Tables 3 and 4). The first pathway in sheep is insulin resistance, the integrated physiology of insulin resistance owes to defective insulin action at target cells. Two signaling pathways of PI3K-Akt and Ras were the main biological pathways in sheep breeds (Table 3). In goat specie, alternative splicing, which is the process of selecting different combinations of splice sites within a messenger RNA precursor to produce variably spliced mRNAs, was the significant

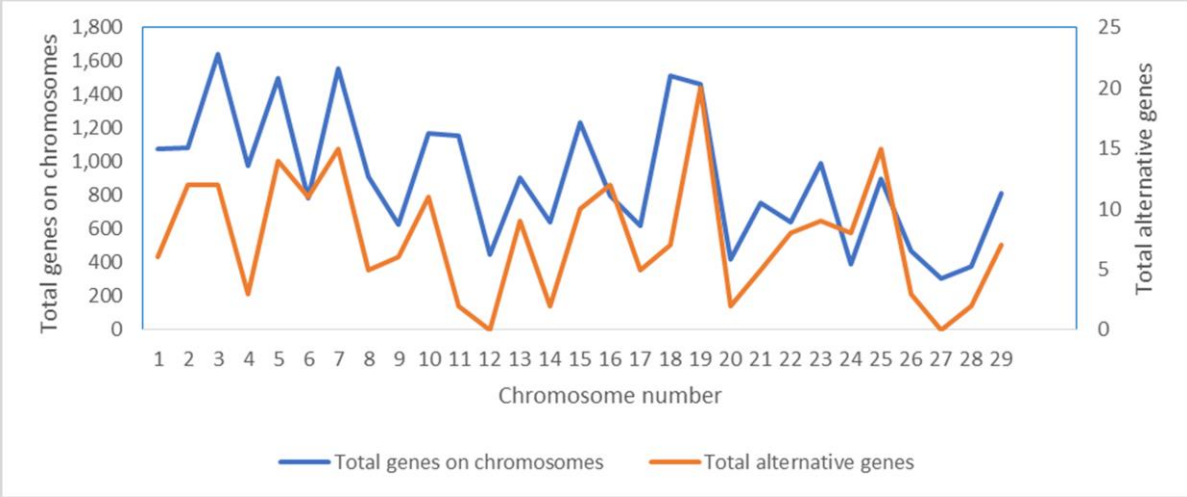

Figure 7. Relationship between the frequency of the alternative genes and the number of genes on each chromosome in goat breeds.

[DOI: 10.22034/jast.25.3.581]

Table 2. Distribution of the alternatives on chromosomes using interracial analysis in sheep and goat.

Chromosome number	Alternative start sites		Alternative splicing		Alternative polyadenylation sites	
	Sheep	Goat	Sheep	Goat	Sheep	Goat
1	50	1	90	3	44	2
2	49	2	57	5	39	5
3	44	5	65	6	36	1
4	18	1	25	1	22	1
5	25	4	38	6	21	4
6	12	3	15	4	4	4
7	31	5	41	6	30	4
8	13	1	12	2	6	2
9	8	2	24	2	11	2
10	5	3	11	5	10	3
11	22	1	55	1	23	0
12	16	0	28	0	13	0
13	15	2	30	4	16	3
14	26	1	51	1	27	0
15	14	3	29	4	20	3
16	2	5	10	4	6	3
17	17	1	26	1	15	3
18	9	1	22	6	11	2
19	14	7	25	10	17	3
20	16	1	23	1	14	0
21	13	2	20	3	8	0
22	6	3	13	4	3	1
23	12	4	8	4	9	1
24	13	2	30	4	18	2
25	8	4	9	8	5	3
26	4	1	6	1	2	1
27		0		0		0
28		1		1		0
29		3		3		1

Table 3. Pathways of the genes with DTISDEs in sheep specie.

Category	Term	P-value
KEGG_PATHWAY	oas04931: Insulin resistance	0.001573
KEGG_PATHWAY	oas04151: PI3K-Akt signaling pathway	0.002267
KEGG_PATHWAY	oas04520: Adherens junction	0.002633
KEGG_PATHWAY	oas04144: Endocytosis	0.006799
KEGG_PATHWAY	oas04014: Ras signaling pathway	0.009918
KEGG_PATHWAY	oas04510: Focal adhesion	0.013985
KEGG_PATHWAY	oas04141: Protein processing in endoplasmic reticulum	0.016069
KEGG_PATHWAY	oas03013: RNA transport	0.017448
KEGG_PATHWAY	oas04142: Lysosome	0.029424
KEGG_PATHWAY	oas03460: Fanconi anemia pathway	0.034811
KEGG_PATHWAY	oas03018: RNA degradation	0.046231
KEGG_PATHWAY	oas05100: Bacterial invasion of epithelial cells	0.051367
KEGG_PATHWAY	oas04015: Rap1 signaling pathway	0.055085
KEGG_PATHWAY	oas04722: Neurotrophin signaling pathway	0.061321
KEGG_PATHWAY	oas04146: Peroxisome	0.068797
KEGG_PATHWAY	oas03040: Spliceosome	0.07129
KEGG_PATHWAY	oas01212: Fatty acid metabolism	0.092245
KEGG_PATHWAY	oas04064: NF-kappa B signaling pathway	0.092965

[DOI: 10.22034/jast.25.3.581]

Table 4. Pathways of the genes with DTISDEs in goat specie.

Category	Term	p-value
UP_KEYWORDS	Alternative splicing	0.013526
KEGG_PATHWAY	hsa03022: Basal transcription factors	0.019499
GOTERM_BP_DIRECT	GO:0006367~transcription initiation from RNA polymerase II promoter	0.070176
KEGG_PATHWAY	hsa05168: Herpes simplex infection	0.077714
UP_SEQ_FEATURE	Mutagenesis site	0.086871

pathway (Table 4). Another important pathway between the goat breeds is hsa03022: Basal transcription factors. The basal transcription factors comprises many proteins that act in RNA transcription process.

DISCUSSION

Alternative splicing, alternative transcript start sites, and polyadenylation sites are three well-known factors that make transcriptional differences from a single gene. We determined three main points from the current study. First, alternative splicing is the most effective type of alternatives in creating transcriptional isoforms between the sheep and goat breeds. Second, distribution of alternating genes on chromosomes revealed a positive correlation between the number of genes on each chromosome and gene alternation. Third, goat genes are more affected by the alternative process than the sheep genes. Furthermore, we have identified DEUTAs that may not participate in the transcription and ultimately translation process due to their involvement in an alternative process. Recent studies in human tissues have shown that variation in the alternative start and termination sites of transcription drives the most transcript isoform differences across the human tissues [27]. However, to the best of our knowledge, the alternative process has not yet been determined among different breeds of animals. In a study by Belabdi *et al*. (2019), it was reported that breed is a natural or artificial choice of a group of individuals of the same specie that shows one or more phenotypes more clearly [3]. On the other hand, we found that alternative splicing is the most effective type of alternative between the two breeds of sheep and goats. It has been determined that alternative splicing usually affects the Untranslated Regions (UTRs) of the gene [17]. In contrast, another study revealed that these regions play a role in enhancing gene expression [10]. According to the previous findings and the findings of this study, the reason for the high frequency of alternative splicing between breeds in two species of sheep and goats can be the effect on UTRs and gene expression and isoforms, which results in differentiation of the breeds in one species. If this process does not occur, the individuals will only show normal expression of a gene and the uniformity of the phenotypes across the population. In a group of sheep, for example, alternative splicing may occur in areas that affect the UTRs that produce fat tail and, consequently, these sheep are categorized into a specific breed based on this particular trait. Distribution of the alternative process on chromosomes found an almost downward trend of spreading of alternatives, however, some chromosomes did not show this pattern. According to Figures 5 and 6, it can be said that the alternatives have a positive correlation with the number of genes on each chromosome, while the number of genes is not directly related to the number of chromosomes that validates the downward trend of gene localization across the genome [26]. We found that the genes with differential use of exons in goat species were less than in sheep species. Graphical data derived from the DEXSeq package show

[DOI: 10.22034/jast.25.3.581]

that this could be due to the involvement of most of these genes in the alternatives process, so that 67.63% of goat genes and only 12.17% of sheep genes were involved in the alternatives process. This can be explained by the small number of genes with differential use of exons in goat specie. As a result, the greater involvement of goat genes in these processes is likely compensated by their small number of genes. Reyes *et al*. (2013) demonstrated a direct relationship between exon differences and their persistence over later generations. Exons with significant differences are preserved in future generations [25]. Similarly, in the interbreeding analysis of two species of sheep and goats in this study, out of 121,861 exons, 27,747 exons showed special protection. Out of the 118,296 exons of sheep species genome and 3,565 exons of goat species genome, 27,537 exons in sheep and 210 exons in goat species were protected. Exons that have more protection over future generations not only seem to have significant differential application, but also have a very low level of participation in the alternatives process. Indeed, only 0.64% of exons were involved in a kind of alternatives. It seems that the conservation process of the exons is related to the contribution of these exons in alterative process.

The differential exon usage of different transcription isoforms exhibited eighteen significant ($P< 0.01$) functional and biological pathways for genes between sheep breeds (Table 3) and four biological pathways between goat breeds (Table 4). Insulin resistance is the first pathway significantly detected in sheep breeds. Insulin resistance pathway is the integrated physiology of insulin resistance to be obligated to defective insulin action at target cells. Insulin interaction with its receptor activates an intrinsic tyrosine protein kinase, which autophosphorylates the receptor as well as downstream substrates. In the main pathway, a family of proteins, known as insulin receptor substrates, activate a cascade of serine-protein kinases. Akt (protein kinase B), one of the significant pathways in this study, is a major branch point with numerous downstream substrates leading to a variety of physiological functions including the regulation of the homeostasis. The other signaling pathways that were identified in sheep breeds were of PI3K-Akt signaling pathway and Ras signaling pathway (Table 3). The phosphoinositide-3-kinase (PI3K, or Akt) plays a role in cell metabolism, growth, proliferation, and survival [15]. Ras signaling is an important intracellular signaling pathway that plays a role in cellular proliferation and differentiation, survival, and gene expression [21]. Similarly, ovarian gene expression patterns and their enrichment pathways in Hu sheep identified PI3K-Akt, estrogen metabolism and oogenesis pathways significantly related to ovarian different developmental stages [31]. Alternative splicing and hsa03022: Basal transcription factors were the main biological pathways in goat species (Table 4). Alternative splicing is a process by which exons can be either excluded or included in, or from, a pre-mRNA, resulting in multiple mRNA isoforms. Alternative splicing is an important mechanism in the developmental and cell-type specific control of gene expression, and as a mechanism for increasing the proteome diversity [8]. The basal transcription factors contains more than 33 proteins that act in RNA transcription process. Correct transcription of the genes are mainly related to the involvement of the basal transcription factors. Transcription factors are vital for the normal development of an organism, as well as for routine cellular functions and responses. During development of multicellular organisms, transcription factors are responsible for dictating the fate of individual cells [28].

The outcome of this study indicated that differential exon usage of genes in goat breeds were more than in sheep breeds. Moreover, alternative splicing is the most effective type of alternative between the two breeds of sheep and goats. It seems that the

[DOI: 10.22034/jast.25.3.581]

conservation process of the exons is related to the contribution of these exons in alternative process. In addition, eighteen biological pathways of genes between sheep breeds and four biological pathways between goat breeds were the most significant pathways. The PI3K-Akt pathway plays a role in cell growth and metabolism and development of ovaries.

ACKNOWLEDGEMENTS

The contributions of the Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, in collecting and processing the samples are gratefully acknowledged.

REFERENCES

1. Acland, A., Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C. and Church, D. M. 2014. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.,* **42(D1):** D7-D17.
2. Anders, S., Reyes, A. and Huber, W. 2012. Detecting Differential Usage of Exons from RNA-Seq Data. *Nat. Preced*., **22(10):** 2008-2017.
3. Belabdi, I., Ouhrouch, A., Lafri, M., Gaouar, S. B. S., Ciani, E., Benali, A. R. and Blanquet, V. 2019. Genetic Homogenization of Indigenous Sheep Breeds in Northwest Africa. *Sci. Rep*., **9(1)**: 1-13.
4. Blekhman, R., Marioni, J. C., Zumbo, P., Stephens, M., and Gilad, Y. 2010. SexSpecific and Lineage-Specific Alternative Splicing in Primates. *Genome Res.,* **20(2):** 180-189.
5. Brown, T. A. 2018. Genomes 4, 4th Edition, Published by Garland Science, 538 Pages.
6. Brooks, A. N., Yang, L, Duff, M. O., Hansen, K. D., Park, J. W., Dudoit, S. and Graveley, B. R. 2011. Conservation of an RNA Regulatory Map between Drosophila and Mammals. *Genome Res.,* **21(2):** 193-202.
7. Cline, M. S., Blume, J., Cawley, S., Clark, T. A., Hu, J. S., Lu, G. and Williams, A. 2005. ANOSVA: A Statistical Method for Detecting Splice Variation from Expression Data. *Bioinformatics*, **21(Suppl. 1):** i107i115.
8. Chen, K., Dai, X. and Wu, J. 2015. Alternative splicing: An important mechanism in stem cell biology, *World J.Stem Cells.,* **7(1):** 1–10.
9. Elmus, G. 2013. Beale Insulin Signaling and Insulin Resistance. J. *Investig. Med*., **61(1):** 11–14.
10. El-Shehawi, A. and Elseehy, M. 2017. Genome Size and Chromosome Number Relationship Contradicts the Principle of Darwinian Evolution from Common Ancestor. J. *Phylogenetics Evol. Biol*., **5:** 179.
11. Evfratov, S. A., Osterman, I. A., Komarova, E. S., Pogorelskaya, A. M., Rubtsova, M. P., Zatsepin, T. S. and Burnaev, E. 2017. Application of Sorting and Next Generation Sequencing to Study 5′-UTR Influence on Translation Efficiency in *Escherichia coli, Nucleic Acids Res*., **45(6):** 3487-3502.
12. Gottfredson, R. 2001. Hormonal Control of Ewe Reproduction. Departament of Animal Science, *University of Wisconsin-Madison*.
13. Griffith, M., Griffith, O. L., Mwenifumbo, J., Goya, R., Morrissy, A. S., Morin, R. D. and Marra, M. A. 2010. Alternative expression analysis by RNA sequencing. Nat. Methods, **7(10):** 843-847.
14. Grishkevich, V. and Yanai, I. 2014. Gene Length and Expression Level Shape Genomic Novelties. *Genome Res*., **24(9):** 1497-1503.
15. Hemmings, B. A. and Restuccia, D. F. 2012. PI3K-PKB/Akt Pathway. Published by Cold Spring Harbor Laboratory Press, *Cold Spring Harb Perspect in Biology*, **4:** a011189, PP. 1-3. 46. a011189.
16. Katz, Y., Wang, E. T., Airoldi, E. M. and Burge, C. B. 2010. Analysis and Design of RNA Sequencing Experiments for Identifying Isoform Regulation. *Nat. Methods***, 7(12):** 1009-1015.

[DOI: 10.22034/jast.25.3.581]

17. Mills, J. D. and Janitz, M. 2012. Alternative Splicing of mRNA in the Molecular Pathology of Neurodegenerative Diseases. *Neurobiol. Aging*., **33(5):** 1012. e11-1012. e24.
18. Mohammadabadi, M. 2020. Expression of Calpastatin Gene in Raini Cashmere Goat Using Real Time PCR. Agric. *Biotechnol*. J., **11(4):** 219-235.
19. Mohammadabadi, M., Bordbar, F., Jensen, J., Du, M. and Guo, W. 2021. Key Genes Regulating Skeletal Muscle Development and Growth in Farm Animals. *Animals***,** **11(3):** 835.
20. Mohammadabadi, M., Masoudzadeh, S. H., Khezri, A., Kalashnyk, O., Stavetska, R. V., Klopenko, N. I. and Tkachenko, S. V. 2021. Fennel (Foeniculum vulgare) Seed Powder Increases Delta-Like Non-Canonical Notch Ligand 1 Gene Expression in Testis, Liver, and Humeral Muscle Tissues of Growing Lambs. *Heliyon*., **7(12):** e08542.
21. Molina, J. R. and Adjei, A. A. 2021. The Ras/Raf/MAPK Pathway. *J. Thorac. Oncol.*, 1: 7–9.
22. Pertea, M. and Salzberg, S. L. 2010.Between a Chicken and a Grape: Estimating the Number of Human Genes, *Genome Biol*., **11(5):** 1-7.
23. Purdom, E., Simpson, K. M., Robinson, M. D., Conboy, J. G., Lapuk, A. V. and Speed, T. P. 2008. FIRMA: A Method for Detection of Alternative Splicing from Exon Array Data. *Bioinformatics*, **24(15):** 17071714.
24. Piovesan, A., Antonaros, F., Vitale, L., Strippoli, P., Pelleri, M. C. and Caracausi, M. 2019. Human Protein-Coding Genes and Gene Feature Statistics, *BMC Res*. Notes., **12:**1-5.
25. Reyes, A., Anders, S. and Huber, W. 2013. Inferring Differential Exon Usage in RNA-Seq Data with the DEXSeq Package. https://mirrors.nju.edu.cn/bioconductor/2.13/bioc/vignettes/DEXSeq/inst/doc/DEXSeq.pdf .
26. Reyes, A., Anders, S., Weatheritt, R. J., Gibson, T. J., Steinmetz, L.M. and Huber, W. 2013. Drift and Conservation of Differential Exon Usage across Tissues in Primate Species. *PNAS*, **110**: 15377-15382.
27. Reyes, A. and Huber, W. 2018. Alternative Start and Termination sites of Transcription Drive Most Transcript Isoform Differences across Human Tissues. *Nucleic Acids Res*., **46(2):** 582-592.
28. Reese, J. C. 2003. Basal Transcription Factors. *Current Opinion in Genetics & Developmen*., **13(2):** 114118.
29. Rischkowsky, B. and Pilling, D. 2007. The State of the World's Animal Genetic Resources for Food and Agriculture in Brief. *FAO*.
30. Stark, R., Grzelak. M. and Hadfield, J. 2019. RNA Sequencing: The Teenage Years. *Nat. Rev. Genet*., **20(11):** 631-656.
31. Shabbir, S., Boruah, P, Xie, L, Kulyar, M. F. E. A., Nawaz, M., Yousuf, S. and Miao, X. 2021. Genome-Wide Transcriptome Profiling Uncovers Differential miRNAs and lncRNAs in Ovaries of Hu Sheep at Different Developmental Stages. *Sci. Rep*., **11(1):** 1-12.
32. Shahsavari, M., Mohammadabadi, M., Khezri, A., Asadi Fozi, M., Babenko, O., Kalashnyk, O. and Tkachenko, S. 2021. Correlation between Insulin-Like Growth Factor 1 Gene Expression and Fennel (Foeniculum vulgare) Seed Powder Consumption in Muscle of Sheep. *Animal Biotechnology*., 1-11.
33. Tang, J. Y., Lee, J. C., Hou, M. F., Wang, C. L., Chen, C. C., Huang, H. W. and Chang, H.W. 2013. Alternative Splicing for Diseases, Cancers, Drugs, and Databases. *Science World* .
34. Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J. and Pachter, L. 2010. Transcript Assembly and Abundance Estimation from RNA-Seq Reveals Thousands of New Transcripts and Switching among Isoforms. *Nat. Biotechnol*., **28(5):** 511.
35. Vogel, F., A Preliminary Estimate of the Number of Human Genes. 1964. *Nature.*, **201(4921):** 847-847.
36. Wang, T., Birsoy, K., Hughes, N. W., Krupczak, K. M., Post, Y., Wei, J. J. and Sabatini, D. M. 2015. Identification and Characterization of Essential Genes in the Human Genome. *Science*, **350(6264):**1096-1101.

[DOI: 10.22034/jast.25.3.581]

37. Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C. and Burge, C. B. 2008. Alternative Isoform Regulation in Human Tissue Transcriptomes. *Nature*, **456(7221)**: 470-476.
38. Wu, J., Akerman, M., Sun, S., McCombie, W. R., Krainer, A. R., Zhang, M. Q. 2011. SpliceTrap: A Method to Quantify Alternative Splicing under Single Cellular Conditions. *Bioinform*., **27(21):** 3010-3016.
39. Yap, K. and Makeyev, E. V. 2013. Regulation of Gene Expression in Mammalian Nervous System through Alternative Pre-mRNA Splicing Coupled with RNA Quality Control Mechanisms. Mol. Cell. Neurosci., **56:** 420-428.
40. Zi, X. D., Lu, J. Y., Zhou, H., Ma, L., Xia, W., Xiong, X. R. and Wu, X. H. 2013. Comparative Analysis of Ovarian Transcriptomes between Prolific and Non-Prolific Goat Breeds via High-Throughput Sequencing. *Reprod. Domest. Anim*., **53(2):** 344-351.

بررسی تمایز افتراقی اگزونها بین نژادهای گوسفند و بز با استفاده از داده های RNA-seq

ف. نیکوکلام عظیم ، ع. ا. مسعودی، ع. احسانی، ر. واعظ ترشیزی، و پ. شریعتی

گزگزاره

چکیده

جفت وجور شدن اگزونها، محل شروع نسخه برداری رونوشتها و محل پلی آدنیلاسیون رونوشتها عوامل اصلی ایجاد رونوشت های متنوع یک ژن هستند. هدف اصلی این مطالعه بررسی روند تناوب ژنها در نژادهای گوسفند و بز و شناسایی نقش آن در تمایز نژادهای یک گونه بود. داده های RNA-seq از بافت تخمدان دو نژاد گوسفند شال و سنگسری و دو نژاد بز سیاه تبتی و سیاه جین تانگ تهیه شد. خوانشها با ژنوم مرجع همتراز شدند و ژن های معنی دار با توجه به استفاده افتراقی اگزونها شناسایی شدند. مقایسه آماری نشان داد که ۸۱۰۴ ژن در استفاده از اگزون بین نژادهای گوسفند و ۱۷۳ ژن بین نژادهای بز تفاوت معنی داری دارند. از ۱۲۱۸۶۱ اگزون بررسی شده، تنها ۲۲/۷ % در طول نسل های آینده بین نژادها حفظ می شود، که از اینها ۹۹/۳% هیچگونه جفت و جور شدن اگزونی را نشان نمی دهند که می تواند به علت حفاظت بالادر فرآیند جایگزینی باشد. ژن های با کاربرد افتراقی اگزون ها در گونه بز نسبت به گوسفند بیشتر بود. تجزیه و تحلیل بین نژادی نشان داد که جفت و جور شدن اگزونها تأثیرگذارترین نوع برای ایجاد رونوشت های متنوع یک ژن در نژادهای گوسفند و بز است. به نظر می‌رسد که فرآیند حفاظت اگزون‌ها با میزان مشارکت آنها در فرآیند جایگزینی در هر دو نژاد گوسفند و بز مرتبط باشد. مسیر مولکولی PI3K-Akt در رشد سلولی و متابولیسم و توسعه تخمدان ها دراین گونه ها نقش مهمی دارد.

[DOI: 10.22034/jast.25.3.581]