Vol 26, No. 1, Spring 2024

University of Tehran Press

Journal of Animal Production

Online ISSN: 2382-994X

Homepage: https://jap.ut.ac.ir/

Identification of selection signatures associated with Iranian sheep compared to non-Iranian Romanov breed using whole genome sequencing data

Abbas Mirzapour-Abibagloo[1] | Nemat Hedayat[2]✉ | Reza Khalkhali-Evrigh[3] | Reza Seyedsharifi[4]

1. Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: amirza@student.uma.ac.ir
2. Corresponding Author, Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: nhedayat@uma.ac.ir
3. Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: r.khalkhali@uma.ac.ir
4. Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran. E-mail: sharifi_r@uma.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 17 september 2023
Received in revised form
12 February 2024
Accepted 13 February 2024
Published online 15 March 2024

Keywords:
Iranian Sheep
Romanov
Signature of Selection
Whole-Genome Sequencing

ABSTRACT

Introduction Iranian sheep breeds, due to the climate diversity of country, show high diversity and have acquired high adaptability. Compromise with poor quality feed, tolerance of adverse weather and manageable body size are among the factors that probably caused sheep to adapt to different climates. Heretofore, several studies have been carried out in the field of identification of selection signatures in the different native breeds based on SNP-chip data. However, the use of whole genome data can provide researchers with more information about the differences between breeds and their genetic capacities. Identifying and evaluating the effects of climate on the genome of native breeds of Iranian sheep can be effective in designing breeding and conservation strategies. The aim of the present study was to identify the signs of selection related to Iranian sheep compared to the Romanov breed at the genomic level.

Materials and Methods For present study, we used the whole-genome sequencing data related to 43 Iranian and non-Iranian sheep available in the NCBI database. These reads, after performing quality control, were aligned to the sheep reference genome by BWA program. Here, RealignerTargetCreator and IndelRealigner commands available in the GATK program were used to realign around insertions and deletions. Then, the HaplotypeCaller algorithm was used to identify the variants of all samples in ERC GVCF mode. Further, using GenotypeGVCFs module, the variants of all samples were simultaneously identified and finally a VCF file containing raw variants was created. Using the SelectVariants command of the GATK program, all SNPs were separated from other variants. After applying multiple quality filters, high-quality SNPs were extracted and only bi-allelic SNPs present in autosomal chromosomes were used for downstream analysis. Putative selection signatures were identified by using two methods including Fst and XP-EHH. Genes located in positively selected genomic regions were extracted using BEDtools program and the GTF file related to the sheep genome. Gene ontology (GO) analysis was performed on selected genes by "g:Profiler" web-based tool.

Results and Discussion Here, Fst and XP-EHH methods were used to identify the signatures of selection related to Iranian sheep in comparison with Romanov sheep. After converting Fst values to ZFst, 958 genomic windows containing 907 protein-coding genes were detected that had scores above the threshold (ZFst > 3.35). GO analysis on 907 genes identified by the ZFst method led to the identification of 157 significant GO terms in the field of biological processes. In addition, 26 significant terms related to molecular functions and 5 significant terms related to cellular components were also identified. The number of genomic windows identified by the XP-EHH method was 953, which contained a total of 311 protein-coding genes. Among identified genes for each method, 29 genes were detected by both methods as signatures of selection for Iranian sheep. From the GO analysis of 29 common genes, no significant term was obtained. However, these genes were involved on traits related to improving milk fat quality (PCCB), fertility (SPATA5, RAB35 and DICER1), muscle growth and development (NF1, AKAP6 and HDAC9), body weight (FBXL3, GRID2 and ADAMTS17), adaptability to harsh desert and mountain condition (BMPR2 and NF1) and also, milk related traits (EXOC6B).

Conclusion The results showed that Iranian sheep were probably selected to adapt to dry desert areas and improve the quality of meat and milk. The gradual accumulation of such information in different populations will improve the understanding and knowledge of researchers and breeders and will help them to implement efficient breeding programs.

Cite this article: Mirzapour-Abibagloo, A., Hedayat, N., Khalkhali-Evrigh, R., & Seyedsharifi, R. (2024). Identification of selection signatures associated with Iranian sheep compared to non-Iranian Romanov breed using whole genome sequencing data. *Journal of Animal Production*, 26 (1), 1-13.
DOI: https://doi.org/10.22059/jap.2024.365469.623758

© The Author(s).
Publisher: The University of Tehran Press.
DOI: https://doi.org/10.22059/jap.2024.365469.623758

دوره بیست و ششم، شماره اول، بهار ۱۴۰۳

انتشارات دانشگاه تهران

تولیدات دامی

شاپا الکترونیکی: ۲۳۸۲-۹۹۴x

Homepage: https://jap.ut.ac.ir/

شناسایی نشانگرهای انتخاب مرتبط با گوسفندان ایرانی در مقایسه با نژاد غیرایرانی رومانف با استفاده از داده‌های توالی‌یابی کل ژنومی

عباس میرزاپور آبی‌بگلو[۱] | نعمت هدایت[۲]✉ | رضا خلخالی ایوریق[۳] | رضا سیدشریفی[۴]

۱. گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: amirza@student.uma.ac.ir

۲. نویسنده مسئول، گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: nhedayat@uma.ac.ir

۳. گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: r.khalkhali@uma.ac.ir

۴. گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: sharifi_r@uma.ac.ir

اطلاعات مقاله

نوع مقاله: مقالهٔ پژوهشی

تاریخ دریافت: ۱۴۰۲/۰۶/۲۶

تاریخ بازنگری: ۱۴۰۲/۱۱/۲۳

تاریخ پذیرش: ۱۴۰۲/۱۱/۲۴

تاریخ انتشار: ۱۴۰۲/۱۲/۲۵

کلیدواژه‌ها:

توالی‌یابی کل ژنوم

رومانف

گوسفندان ایرانی

نشانه‌های انتخاب

چکیده

به‌دلیل تنوع اقلیمی بالا در ایران، گوسفندان بومی کشور، از تنوع چشم‌گیری برخوردار هستند. به‌نظر می‌رسد تفاوت‌های اقلیمی، ردپایی در ژنوم نژادهای بومی مناطق مختلف جهان به جای گذاشته باشند. هدف از مطالعه حاضر، شناسایی نشانه‌های انتخاب مرتبط با گوسفندان ایرانی در مقایسه با نژاد غیرایرانی رومانف در سطح ژنوم می‌باشد. به این منظور، از داده‌های توالی‌یابی‌شده کل ژنوم گوسفندان ایرانی و غیرایرانی موجود در پایگاه داده‌ای NCBI استفاده گردید. این توالی‌ها، پس از سنجش و پالایش کیفی، به ژنوم مرجع گوسفند هم‌طراز شدند. پس از شناسایی تنوع‌های ژنومی، برای شناسایی نواحی تحت انتخاب مثبت در ژنوم گوسفندان ایرانی، از دو روش Fst و XP-EHH استفاده شد. پس از استخراج ژن‌های موجود در نواحی تحت انتخاب، با استفاده از نرم‌افزار BEDtools و فایل GTF مرتبط با ژنوم گوسفند، شرح‌نویسی عملکردی این ژن‌ها با استفاده از آنالیز هستی‌شناسی ژن صورت پذیرفت. براساس نتایج به‌دست‌آمده، به‌ترتیب ۹۰۷ و ۳۱۱ ژن کدکننده پروتئین توسط روش‌های Fst و XP-EHH شناسایی شدند که تعداد ۲۹ ژن بین این دو روش مشترک بودند. ارزیابی‌های بیش‌تر نشان داد که تعدادی از این ژن‌ها، در صفات مرتبط با بهبود کیفیت چربی شیر (PCCB)، باروری (SPATA5، RAB35 و DICER1)، رشدونمو ماهیچه‌ای (NF1، AKAP6 و HDAC9)، وزن بدن (FBXL3، GRID2 و ADAMTS17)، سازگاری به نواحی سخت بیابانی و کوهستانی (BMPR2 و NF1) و شیر (EXOC6B) دخیل هستند. نتایج نشان داد که گوسفندان ایرانی احتمالاً برای سازگاری به مناطق خشک بیابانی و ارتقای کیفیت گوشت و شیر موردانتخاب قرار گرفته‌اند.

استناد: میرزاپور آبی‌بگلو، عباس؛ هدایت، نعمت؛ خلخالی ایوریق، رضا و سیدشریفی، رضا (۱۴۰۳). شناسایی نشانگرهای انتخاب مرتبط با گوسفندان ایرانی در مقایسه با نژاد غیرایرانی رومانف با استفاده از داده‌های توالی‌یابی کل ژنومی. *نشریه تولیدات دامی*، ۲۶ (۱)، ۱-۱۳.

DOI: https://doi.org/10.22059/jap.2024.365469.623758

ناشر: مؤسسه انتشارات دانشگاه تهران.

© نویسندگان.

۱. مقدمه

گوسفند از اولین حیواناتی است که حدود ۱۰ هزار سال پیش در منطقه‌ای به نام هلال حاصلخیز اهلی شد. از آن زمان تاکنون، این گونه حیوانی پراکندگی خود را در سطح جهان به‌شدت گسترش داده است. سازگاری با مواد غذایی کم‌کیفیت، تحمل آب‌وهوای نامساعد و اندازه بدن قابل‌مدیریت، از جمله عواملی هستند که احتمالاً باعث سازگاری گوسفند به اقلیم‌های مختلف شده‌اند (Kijas *et al*., 2009). گوسفندان در ابتدا فقط برای تولید گوشت پرورش داده می‌شدند، اما از پنج هزار سال پیش، از این گونه حیوانی برای اهداف دیگر پرورشی نیز استفاده شد. طبق آمار و ارقام منتشرشده در پایگاه سازمان خواروبار جهانی (FAO)، آسیا با دارابودن حدود ۵۱۲ میلیون راس، بیش‌ترین جمعیت گوسفند جهان را در اختیار دارد. در کشورهای توسعه‌یافته، مصرف‌کنندگان مواد غذایی، خوراکی را که دربرگیرنده بسیاری از احتیاجات غذایی‌شان است بر سایر اقلامی که سلامتی آن‌ها را به‌خطر می‌اندازد، ترجیح می‌دهند. بنابراین، همواره خواستار محصولاتی هستند که باکیفیت بالای خود ضامن سلامتی آن‌ها باشد. افزایش جهانی جمعیت، سطح درآمد و شهرنشینی، احتمالاً از دلایل دیگری هستند که جمعیت دام‌های جهان را تحت‌تأثیر قرار می‌دهند.

تاکنون مطالعات مختلفی در زمینه شناسایی نواحی تحت انتخاب نژادهای بومی انجام گرفته است. اما استفاده از داده‌های کل ژنوم می‌تواند با وسعت و دقت بیش‌تری در شناسایی و ارزیابی نشانه‌های انتخاب کمک‌کننده باشد. هم‌چنین استفاده از داده‌های کل ژنومی می‌تواند اطلاعات بیش‌تری درباره تفاوت بین نژادها و ظرفیت‌های ژنتیکی آن‌ها در اختیار پژوهش‌گران قرار دهد. تلاش برای یافتن مناطق ژنومی حاوی نشانگرهای انتخابی و ژن‌های منتخب توسط این داده‌ها می‌تواند در اصلاح نژاد گوسفندان بومی مؤثر واقع شود. شرایط اقلیمی یکی از مهم‌ترین نیروهایی به‌شمار می‌رود که در تکامل گونه‌ها و نژادهای مختلف تأثیرگذار است. تفاوت‌های اقلیمی در مناطق مختلف جهان، تأثیرات شگرفی بر نواحی خاصی از ژنوم موجودات بومی آن مناطق گذاشته که شناخت آن‌ها می‌تواند در زمینه اصلاح و حفاظت از نژادهای بومی آن مناطق مهم باشد.

۲. پیشینه پژوهش

انسان به‌عنوان تأثیرگذارترین گونه بر گونه‌های دیگر به‌شمار می‌رود. در واقع، انسان‌ها، تعدادی از حیوانات را رام کرده، بر آن‌ها مسلط شده و تغییرات بارزی را در ساختار ژنومی و فنوتیپی آن‌ها ایجاد کرده‌اند. عواملی مانند جغرافیا، آب‌وهوا و فرهنگ، نحوه تعامل انسان‌ها با حیوانات را تحت تأثیر قرار داده و منجر به پیداش اکوتیپ‌ها و نژادهای متنوعی از حیوانات اهلی در قمست‌های مختلف جهان شده است. امروزه با توسعه فناوری‌هایی مانند توالی‌یابی کل ژنوم و ریزآرایه‌های SNP به‌همراه پیشرفت در زمینه بیوانفورماتیک، امکان شناسایی نشانه‌های انتخاب در ژنوم موجودات مختلف فراهم شده است تا تأثیر محیط و انسان بر ژنوم گونه‌های مختلف، موردارزیابی قرار گیرد.

تاکنون مطالعات متعددی در زمینه شناسایی نشانه‌های انتخاب در گونه‌ها و نژادهای مختلف جانوری صورت گرفته است که منجر به تولید اطلاعات ارزشمندی در زمینه شناخت این حیوانات شده است. در مطالعات مربوط به شناسایی نشانه‌های انتخاب، معمولاً از گروه‌هایی برای مقایسه استفاده می‌شود که از نظر عواملی مانند جغرافیا، اقلیم و یا اهداف تولیدی، متفاوت باشند. به‌طور مثال، در یک پژوهش به‌منظور بررسی و شناسایی ژن‌های مرتبط با گوشت دنده گوسفند، از ۴۰ گوسفند Hu که گوشت دنده بزرگی داشتند به‌همراه ۴۰ گوسفند Hu دیگر که گوشت دنده کم‌تری داشتند استفاده شد. در نهایت تعداد هفت SNP موجود در شش ژن به‌عنوان کاندید انتخاب شدند. این شش ژن شامل ALS2، ST6GAL2، LOC105611989، PLXNA4، DPP6 و COL12A1 بودند. نتایج نشان دادند

که جهش‌های SNP موجود در سه ژن LOC105611989، DPP6 و COL12A1 به‌صورت معنی‌داری با صفت گوشت دنده در ارتباط بودند و می‌توان آن‌ها را به‌عنوان نشانگرهای مولکولی جهت اصلاح این صفت درنظر گرفت (Zhao *et al*., 2012).

در مطالعه دیگر و به‌منظور بررسی تأثیر جغرافیا بر ژنوم، تعداد ۱۲ جمعیت گوسفند اتیوپیایی با هدف مطالعه پراکندگی آن‌ها در سطح کشور و شناسایی مناطقی از ژنوم که تحت‌تأثیر جغرافیای این کشور تغییر یافته‌اند، توالی‌یابی شد. نتایج حاصل از آنالیز ساختار ژنتیکی این نمونه‌ها توسط PCA و ADMIXTURE حاکی از آن بود که بین جمعیت‌ها اختلاف ژنتیکی وجود دارد. براساس نتایج PCA، همه نمونه‌ها مطابق با پراکنش جغرافیایی خود دسته‌بندی شدند. نتایج این بررسی نشان داد که سازگاری محیطی این جمعیت‌ها به عوامل مرتبط با بارندگی بیش‌تر از سازگاری آن‌ها با عواملی مانند ارتفاع و دمای بالا می‌باشد. هم‌چنین شواهدی از سازگاری این جمعیت‌ها به شرایط مرتفع نیز شناسایی شد (Wiener *et al*., 2021). نتایج این گونه مطالعات نشان‌دهنده قدرت داده‌های ژنومی برای شناسایی پایه‌های ژنتیکی مرتبط با تنوعات فنوتیپی در بین گونه‌ها و نژادهای مختلف در سطح جهان می‌باشد. در این راستا، این مطالعه با هدف بررسی و شناسایی نشانگرهای انتخابی، در سطح کل ژنوم، بین گوسفندان بومی ایرانی و نژاد خارجی رومانف انجام گرفت. در واقع، نژاد رومانف در این مطالعه نماد گوسفندانی است که در شرایط اقلیمی و با توانایی‌های تولیدی متفاوتی نسبت به گوسفندان ایرانی تکامل یافته و اصلاح شده است.

در یک مطالعه، به‌منظور شناسایی نشانه‌های انتخاب در گوسفندان تبتی پانو، از گوسفندان نژاد اولو به‌عنوان مرجع و برای انجام مقایسه ژنومی استفاده شد. این مطالعه با استفاده از تعداد ۱۰ نمونه به‌ازای هر نژاد و با استفاده از داده‌های توالی‌یابی کل ژنوم صورت پذیرفت. پرورش در ارتفاعات بالای فلات تبت، باعث عادت‌پذیری گوسفندان پانو به محیط‌هایی سرد با سطح اکسیژن پایین شده است. شناسایی نشانه‌های انتخاب در نژاد پانو با استفاده از دو روش Fst و تنوع نوکلئوتیدی انجام شد. یکی از ژن‌هایی مهمی که در نژاد پانو، به‌عنوان نشانه انتخاب شناسایی شد، ژن MITF بود. ازآنجایی‌که به‌دلیل جو رقیق‌تر در ارتفاعات بالایی مانند فلات تبت اشعه ماورای بنفش شدیدتر می‌باشد، لذا این ژن احتمالاً در عادت‌پذیری گوسفندان پانو نسبت به این شرایط کارساز بوده است. از ژن‌های کاندید دیگر که به‌عنوان نشانه‌های انتخاب شناسایی شدند، می‌توان به ژن‌های دخیل در عادت‌پذیری به سطح اکسیژن پایین، رشد و توسعه بدن، مقاومت به بیماری‌ها و صفات تولیدمثلی اشاره کرد (Shi *et al*., 2023).

در یک پژوهش برای یافتن ژن‌های مرتبط با صفت شیردهی در گوسفندان، از دو گروه شامل گروه با تولید شیر بالا (شامل شش گروه ژنتیکی مختلف) و گروه با تولید شیر پایین (شامل پنج نژاد مختلف) استفاده شد. در مطالعه مذکور، از سه روش Fst، تنوع نوکلئوتیدی و نرخ هتروزیگوسیتی بهره گرفته شد. در نهایت مشخص شد که چندین ژن از جمله FCGR3A، CTSK، CTSS، ARNT، GHR، SLC29A4، ROR1 و TNRC18 جزو ژن‌های دخیل در صفات مرتبط با شیردهی هستند. پیش‌تر ثابت شده بود که تعدادی از این ژن‌ها از جمله FCGR3A، ARNT، CTSK، CTSS و GHR در دوران شیردهی گاوها و گاومیش‌ها دارای بیان بالایی می‌باشند. از بین ژن‌های مذکور، با استفاده از روش آزمایشگاهی RT-qPCR ثابت شد که ژن FCGR3A یکی از اصلی‌ترین نشانه‌های انتخاب بوده و با صفت شیردهی ارتباط دارد. این ژن در سلامت پستان نقش داشته و با جلوگیری از عفونت‌های پستانی، در تداوم شیردهی حیوان مهم می‌باشد. ازآنجایی‌که بیان بالای این ژن نشانه یک بیماری جدی در سیستم پستانی است، لذا این ژن در سیستم پستانی سالم به میزان کمی بیان می‌شود که مطابق با نتایج RT-qPCR بود (Li *et al*., 2023).

۳. روش‌شناسی پژوهش

۳. ۱. نمونه‌ها و طراز خوانش‌ها به ژنوم مرجع

برای بررسی تفاوت‌های ژنومی گوسفندان ایرانی با نژاد غیرایرانی رومانف از ۴۳ داده کل ژنومی استفاده شد. این داده‌ها شامل نژادهای قره‌گل پنج نمونه)، افشاری (پنج نمونه)، مغانی (سه نمونه)، ماکویی (سه نمونه)، بلوچی (سه نمونه)، شال (سه نمونه)، خاکستری (سه نمونه)، قزل (سه نمونه) و رومانف (۱۵ نمونه) بودند که از پایگاه داده‌ای NCBI به‌دست آمدند. به‌منظور سنجش کیفیت داده‌های خام گردآوری‌شده، از نرم‌افزار FastQC (نسخه ۱۱/۹) استفاده شد. در گام بعدی، جایگاه‌ها و خوانش‌های کم‌کیفیت، توسط نرم‌افزار Trimmomatic (نسخه ۰/۳۹) تصحیح شدند. ازآنجایی‌که معمولاً کیفیت بازهای توالی‌یابی‌شده در ابتدا و انتهای خوانش‌ها[1]، کم‌تر از بخش‌های دیگر است، لذا اگر نمره کیفی ابتدا[2] و انتهای خوانش‌ها[3] کم‌تر از پنج بود، این نواحی حذف شدند. برای کنترل کیفی کل خوانش نیز از روش پنجره‌ای[4] استفاده شد به‌طوریکه هر خوانش در پنجره‌های پنج بازی بررسی شده و اگر میانگین نمره هر پنجره کم‌تر از ۲۰ بود، پنجره مذکور حذف گردید. در نهایت، خوانش‌هایی که حداقل طول[5] آن‌ها در اثر پالایش کیفی به زیر ۴۰ باز رسیده بود به‌طور کامل از آنالیزهای بعدی کنار گذاشته شدند. در برخی از نمونه‌های موردمطالعه، آلودگی آداپتوری مشاهده شد که توسط گزینه ILLUMINACLIP شناسایی و حذف شدند. با استفاده از دستور MEM در برنامه BWA (نسخه ۷/۱۷) همه خوانش‌های پالایش‌شده در ژنوم مرجع گوسفند (GCF_016772045.1) هم‌تراز شدند. فایل‌های خروجی ایجاد شده، با به‌کارگیری نرم‌افزار Picard (http://broadinstitute.github.io/picard)، مرتب و تمامی خوانش‌های مضاعف‌شده، توسط الگوریتم MarkDuplicates نرم‌افزار مذکور حذف گردید. از دو دستور RealignerTargetCreator و IndelRealigner موجود در برنامه GATK (نسخه ۳/۷) برای هم‌ترازی مجدد خوانش‌ها در اطراف حذف و اضافه‌ها (ایندل‌ها) استفاده شد.

۳. ۲. شناسایی و پالایش کیفی تنوع‌ها

در این مرحله، برای این‌که مقایسه و مطالعه بین نژادهای مختلف امکان‌پذیر باشد، تمامی نمونه‌ها اعم از گوسفندان ایرانی و نژاد رومانف در قالب یک فایل VCF قرار گرفتند. بدین منظور، در ابتدا و با کمک الگوریتم HaplotypeCaller برنامه GATK، تنوع‌های همه افراد در حالت ERC GVCF و به‌صورت انفرادی فراخوانی شدند. سپس، با استفاده از ماژول GenotypeGVCFs، به شناسایی همزمان تنوع‌های همه نمونه‌ها پرداخته شد و درنهایت یک فایل VCF حاوی تنوع‌های خام مربوط به همه ۴۳ نمونه، ایجاد گردید. با کمک دستور SelectVariants برنامه GATK، SNPها از سایر تنوع‌ها جدا گردیده و با استفاده از برنامه‌های BCFtools، VCFtools و GATK مورد پالایش کیفی قرار گرفتند. به این صورت که همه SNPهایی که نمره کیفی نرمالایزشده براساس عمق خوانش[6] < ۲/۰، رشته فیشر[7] (معیاری برای سنجش خطای توالی‌یابی)> ۶۰/۰، کیفیت نگاشت[8] < ۴۰/۰، MQRankSum < -12.5، ReadPosRankSum < -8.0 و SOR > 3.0 داشتند، از ادامه آنالیز کنار گذاشته شدند. در نهایت، با به‌کاربردن برنامه VCFtools همه SNPهایی که از

1. Reads
2. LEADING:5
3. TRAILING:5
4. SLIDINGWINDOW:5:20
5. MINLEN:40
6. QD< 2.0
7. FS> 60.0
8. MQ< 40.0

معیارهای روبه‌رو تبعیت نمی‌کردند، حذف شدند؛ حداقل پوشش[1] ۵، فراوانی آللی جزئی[2] ۰/۰۵، حداقل تعداد آلل[3] ۲، حداکثر تعداد آلل[4] ۲، حداکثر جایگاه‌های گمشده[5] ۰/۸ و آستانه تعادل هاردی–واینبرگ[6] ۰/۰۰۱. لازم به ذکر است منظور از حداکثر جایگاه‌های گمشده ۰/۸ این است که برای این‌که یک SNP در فایل نهایی قرار گرفته و حذف نشود باید حداقل در ۸۰ درصد افراد موردمطالعه شناسایی شده باشد. آستانه تعادل هاردی–واینبرگ ۰/۰۰۱ نیز در واقع بیانگر یک مقدار p-value به‌ازای هر SNP می‌باشد که پایین‌تر از این آستانه نشان‌دهنده عدم تعادل هاردی–واینبرگ برای آن SNP بوده و موجب حذف آن می‌شود. برای ادامه روند آنالیزی فقط از SNPهای موجود در کروموزوم‌های اتوزومی استفاده شد.

۳. ۳. ساختار جمعیتی

برای بررسی ساختار جمعیتی نمونه‌های موردمطالعه، از آنالیز تحلیل مؤلفه‌های اصلی[7] (PCA) و ترسیم درخت فیلوژنتیکی استفاده شد. برای دستیابی به نتایج آنالیز PCA، ابتدا SNPهای شناسایی‌شده توسط گزینه indep-pairwise 100 50 0.1 نرم‌افزار PLINK (نسخه ۱/۹) موردپالایش قرار گرفتند. در این گزینه پنجره‌های دارای ۱۰۰ SNP با قطعات همپوشان دارای ۵۰ SNP ایجاد می‌شود. حد آستانه برای همبستگی (عدم تعادل لینکاژی) بین هر دو SNP موجود در هر پنجره، روی مقدار ۰/۱ تنظیم شد. با اجرای این دستور، یکی از جفت SNPهایی که همبستگی بالاتر از ۰/۱ داشته باشند، حذف خواهد شد. در ادامه، همه مؤلفه‌های آنالیز PCA توسط نرم‌افزار PLINK محاسبه شده و دو مؤلفه اول، با استفاده از بسته نرم‌افزاری ggplot2 موجود در نرم‌افزار R، ترسیم شد. هم‌چنین یک درخت فیلوژنتیکی توسط نرم‌افزار PLINK بر پایه روش NJ (Neighbor Joining) ایجاد گردید. نتایج حاصل، با استفاده از نرم‌افزار FigTree (نسخه ۱/۴/۴) ترسیم شد.

۳. ۴. شناسایی نشانه‌های انتخاب

در مطالعه حاضر، به‌منظور شناسایی نواحی تحت انتخاب از روش‌های Fst و XP-EHH استفاده شد. در هر دو روش، پنجره‌هایی به طول ۵۰ هزار باز ایجاد شد که Step Size یا نواحی همپوشان آن‌ها، دارای ۲۵ هزار باز بودند. روش Fst توسط نرم‌افزار VCFtools اجرا شد و برای نرمال‌کردن آن، همه این مقادیر توسط فرمول (Fst – μFst)/σFst به ZFst تبدیل شدند. در این معادله μFst بیانگر میانگین Fstها بوده و σFst انحراف از استاندارد آن‌ها را نشان می‌دهد. برخلاف روش Fst که به‌ازای هر پنجره ژنومی، یک مقدار Fst گزارش می‌دهد. در روش XP-EHH به‌ازای هر SNP یک عدد محاسبه می‌شود. در گام بعدی، اعداد محاسبه‌شده برای هر SNP با استفاده از برنامه norm که بخشی از نرم‌افزار Selscan (نسخه ۲/۰) است، نرمال‌سازی شدند. برای پنجره‌بندی XP-EHH در اندازه ۵۰ هزار بازی با نواحی همپوشان ۲۵ هزار بازی از یک برنامه دست‌نویس با زبان پایتون استفاده شد. برای جلوگیری از انتخاب نواحی کاذب، پنجره‌ای حاوی کم‌تر از ده SNP حذف شدند. لازم به ذکر است که پیش از اجرای روش XP-EHH برای فازبندی هاپلوتیپ‌ها از نرم‌افزار Beagle (نسخه ۵/۲) با گزینه‌های burnin=5 و iterations=20 استفاده گردید. در نهایت، پنجره‌های ژنومی که در دامنه ۱ درصد بالایی ZFst و XP-EHH قرار داشتند، به‌عنوان نواحی تحت انتخاب احتمالی در نظر گرفته شدند

1. MinDP
2. Maf
3. Min-alleles
4. Max-alleles
5. Max-missing
6. Hwe
7. Principle Component Analysis (PCA)

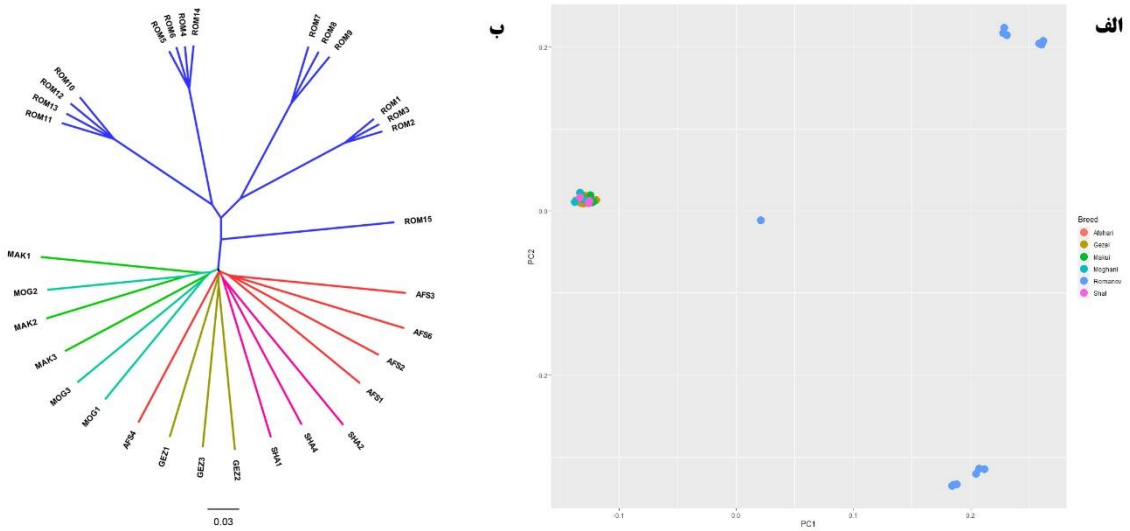
(Khalkhali-Evrigh *et al*., 2022). برای تعیین پنجره‌های ژنومی مدنظر، پس از محاسبه ZFst و XP-EHH برای هر کدام از پنجره‌های مذکور، با استفاده از فرمول percentile در برنامه اکسل، پنجره‌هایی که در دامنه عددی ۱ درصد بالایی هرکدام از دو روش ZFst (پنجره‌های ژنومی با اعداد بزرگ‌تر از ۳/۳۵) و XP-EHH (پنجره‌های ژنومی با اعداد بزرگ‌تر از ۱/۹۶) قرار داشتند، موردانتخاب قرار گرفتند.

۳. ۵. شرح‌نویسی ژن‌های منتخب

پس از استخراج ژن‌های موجود در نواحی تحت انتخاب، با استفاده از نرم‌افزار BEDtools (نسخه ۲/۲۷/۱) و فایل GTF (ARS-UI_Ramb_v2.0_genomic) مرتبط با ژنوم گوسفند، شرح‌نویسی عملکردی این ژن‌ها با کمک آنالیز هستی‌شناسی ژن (GO) در ابزار تحت وب "g:Profiler" صورت پذیرفت. برای شناسایی عبارات معنادار GO و مسیرهای KEGG، از آستانه مقدار p تصحیح‌شده (با استفاده از روش Benjamini-Hochberg) کم‌تر و مساوی ۰/۰۵ استفاده شد.

۴. یافته‌های پژوهش و بحث

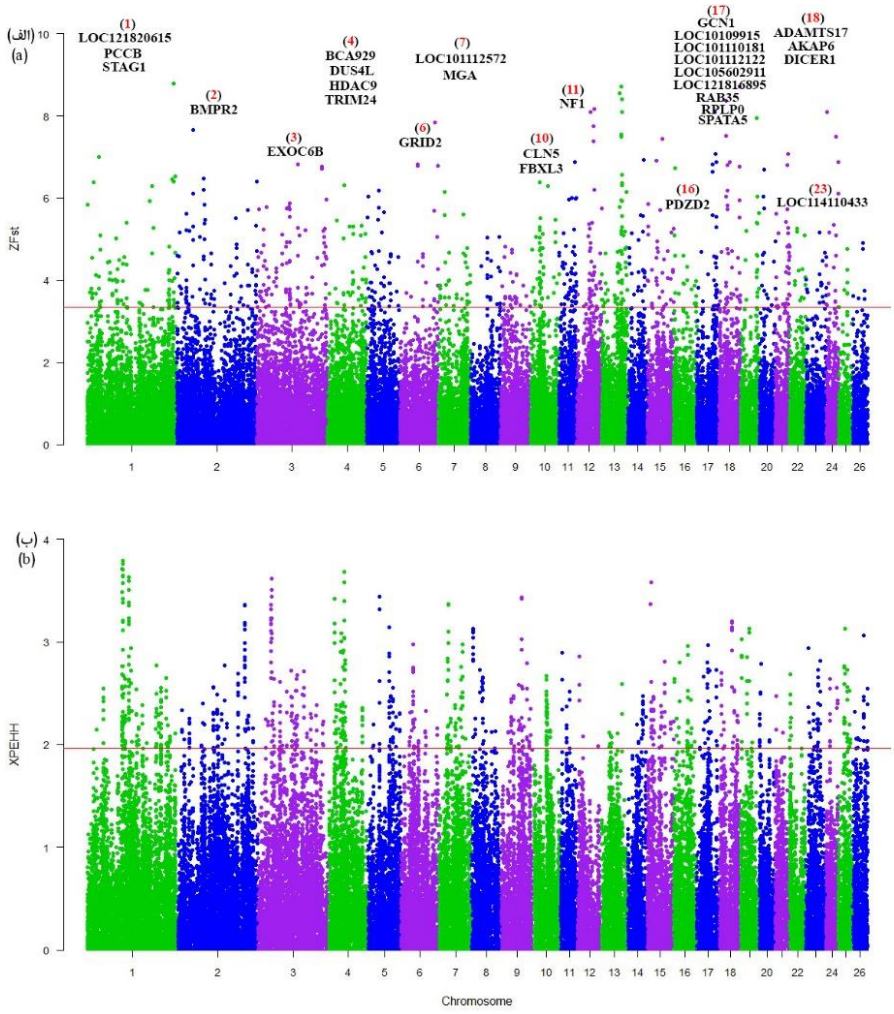
در مطالعه کنونی، پس از شناسایی و پالایش کیفی تنوع‌های ژنومی و جداسازی SNPهای دو آللی آتوزومی، در نهایت از تعداد ۱۵۹۸۱۴۸ SNP برای انجام آنالیزهای مربوطه استفاده گردید. براساس مؤلفه اول آنالیز PCA، گوسفندان رومانف به‌صورت واضحی از گوسفندان ایرانی جدا شدند (شکل ۱–الف). گوسفندان ایرانی با وجود این‌که متشکل از چندین نژاد مختلف بودند، به‌صورت منسجمی در یک گروه قرار گرفتند. نتایج حاصل از رسم درخت فیلوژنتیکی نیز موئد نتایج PCA بود (شکل ۱–ب).



شکل ۱. نتایج حاصل از PCA (الف) و درخت فیلوژنتیکی (ب) مربوط به نمونه گوسفندان موردمطالعه شامل گوسفندان ایرانی و نژاد غیرایرانی رومانف

برای شناسایی نشانگرهای انتخاب در گوسفندان بومی ایرانی در قیاس با گوسفند رومانف از دو روش Fst و XP-EHH استفاده شد. پس از تبدیل مقادیر Fst به ZFst، ۹۵۸ پنجره ژنومی حاوی ۹۰۷ ژن کدکننده پروتئین تشخیص داده شد که

نمراتی بالاتر از حد آستانه تعریف‌شده (ZFst > 3.35) داشتند. تعداد پنجره‌های ژنومی شناسایی‌شده توسط روش XP-EHH، ۹۵۳ عدد بود که در مجموع حاوی ۳۱۱ ژن کدکننده پروتئین بودند. از این بین، ۲۹ ژن منتخب مشترک شناسایی شد که در هر دو روش ZFst و XP-EHH به‌عنوان ژن‌های تحت انتخاب مثبت تعریف شده بودند (شکل ۲).



شکل ۲. نمودار منهتن پنجره‌های ژنومی شناسایی‌شده با استفاده از روش‌های ZFst (الف) و XP-EHH (ب)؛ خط قرمز نمایانگر حد آستانه برای شناسایی نشانه‌های انتخاب در دو روش مذکور می‌باشد. اعداد قرمز داخل پرانتز نمایانگر شماره هر یک از کروموزوم‌های حاوی ژن‌های مشترک می‌باشد.

برای شناسایی مسیرها و فرایندهای بیولوژیکی متأثر از این ژن‌ها، آنالیز هستی‌شناسی ژن با به‌کارگیری ابزار تحت وب gProfiler اجرا شد. آنالیز مذکور روی ۹۰۷ ژن شناسایی‌شده توسط روش ZFst منجر به شناسایی ۱۵۷ عبارت معنی‌دار در زمینه فرایندهای بیولوژیکی مانند تمایز سلول چربی (GO:0045444)، تنظیم تمایز سلول چربی (GO:0045598)، تنظیم مثبت فرایندهای متابولیکی (GO:0009893)، پاسخ به لیپید (GO:0033993)، پاسخ سلولی به لیپید (GO:0071396)، فرایندهای کاتابولیکی (GO:0009056)، توسعه سیستم ایمنی (GO:0002520)، پاسخ به استرس (GO:0006950) و توسعه جنینی

(GO:0009790) شد. به‌علاوه، ۲۶ عبارت معنی‌دار مرتبط با عملکردهای مولکولی به‌همراه پنج عبارت معنی‌دار مرتبط با اجزای سلولی نیز شناسایی شد (جدول ۱). هنگام بررسی ۳۱۱ ژن منتخب شناسایی‌شده توسط روش XP-EHH، به‌ترتیب شش، شش و ۱۵ اصطلاح معنی‌دار مرتبط با عملکردهای مولکولی، فرایندهای بیولوژیکی و اجزای سلولی شناسایی گردید (جدول ۲). از آنالیز هستی‌شناسی ۲۹ ژن مشترک، هیچ عبارت معنی‌داری حاصل نشد.

جدول ۱. پنج عبارت برتر آنالیز هستی‌شناسی ژن مرتبط با ژن‌های اختصاصی شناسایی‌شده توسط روش ZFst

دسته‌بندی	عبارت ژن آنتولوژی	توضیحات	مقدار p تصحیح‌شده
MF	GO:0003690	double-stranded DNA binding	۰/۰۰۷
MF	GO:0019825	oxygen binding	۰/۰۰۷
MF	GO:0005515	protein binding	۰/۰۰۷
MF	GO:1990837	sequence-specific double-stranded DNA binding	۰/۰۰۷
MF	GO:0043565	sequence-specific DNA binding	۰/۰۰۹
BP	GO:0048731	system development	۰/۰۰۰۱
BP	GO:0048518	positive regulation of biological process	۰/۰۰۰۱
BP	GO:0032502	developmental process	۰/۰۰۰۱
BP	GO:0007275	multicellular organism development	۰/۰۰۰۱
BP	GO:0048856	anatomical structure development	۰/۰۰۰۱
CC	GO:0005737	cytoplasm	۰/۰۰۰۱
CC	GO:0005833	hemoglobin complex	۰/۰۰۰۱
CC	GO:0005654	nucleoplasm	۰/۰۰۸۳
CC	GO:0005829	cytosol	۰/۰۳۰۹
CC	GO:1902494	catalytic complex	۰/۰۴۸۴

جدول ۲. پنج عبارت برتر آنالیز هستی‌شناسی ژن مرتبط با ژن‌های اختصاصی شناسایی شده توسط روش XP-EHH

دسته‌بندی	عبارت ژن آنتولوژی	توضیحات	مقدار p تصحیح‌شده
MF	GO:0004984	olfactory receptor activity	۰/۰۰۰۸
MF	GO:0038023	signaling receptor activity	۰/۰۱
MF	GO:0004930	G protein-coupled receptor activity	۰/۰۱
MF	GO:0060089	molecular transducer activity	۰/۰۱
MF	GO:0004888	transmembrane signaling receptor activity	۰/۰۱
BP	GO:0051716	cellular response to stimulus	۰/۰۰۵۲
BP	GO:0023052	Signaling	۰/۰۰۸۴
BP	GO:0007154	cell communication	۰/۰۰۸۴
BP	GO:0007165	signal transduction	۰/۰۱۵۳
BP	GO:0098742	cell-cell adhesion via plasma-membrane adhesion molecules	۰/۰۴۴۵
CC	GO:0034702	ion channel complex	۰/۰۱۹۷
CC	GO:1990351	transporter complex	۰/۰۲۹
CC	GO:0016020	Membrane	۰/۰۲۹
CC	GO:0016460	myosin II complex	۰/۰۲۹
CC	GO:0098862	cluster of actin-based cell projections	۰/۰۲۹

بررسی ۲۹ ژن مذکور نشان داد که اغلب این ژن‌ها روی صفاتی نظیر بهبود کیفیت چربی شیر (PCCB)، باروری (SPATA5، RAB35 و DICER1)، رشدونمو ماهیچه‌ای (NF1، AKAP6 و HDAC9)، وزن بدن (FBXL3، GRID2 و ADAMTS17)، سازگاری به نواحی سخت بیابانی و کوهستانی (BMPR2 و NF1) و هم‌چنین صفات مرتبط با شیر (EXOC6B) اثرگذار هستند. ژن PCCB مسئول تولید آنزیم بسیار مهمی است که در فرایند گلوکونئوژنز در کبد گاوها نقش دارد. در یک بررسی روی کروموزوم شماره ۱۳ خوک، مشخص شد که ژن PCCB آنزیمی میتوکندریایی به نام پروپیونیل کوآ کربوکسیلاز (Propionyl-CoA Carboxylase) را کد می‌کند که کربوکسیل‌دارکردن پروپیونیل کوآ و تبدیل آن به متیل‌مالونیل–کوآ را طی یک فرایند وابسته به ATP، تسریع می‌بخشد. این آنزیم در کاتابولیسم اسیدهای آمینه شاخه‌دار و اسیدهای چربی که زنجیره‌هایی با طول نامتعارف دارند، دخالت دارد (Van Poucke *et al*., 1997). در مطالعه‌ای که جهت شناسایی ژن‌های دخیل در تفاوت‌های نژادی موجود بین گوسفندان بومی چین صورت پذیرفت، ژن PCCB با رشته‌های هموزیگوت ژنوم گوسفندان دم‌بزرگ آلتای و هان، همپوشانی نشان داد (Liu *et al*., 2021). این ژن در صفات مرتبط با رشد و گوشت لخم این نژادها اثرگذار بود. ژن PCCB در بهبود کیفیت چربی شیر گوسفندان تیپ شیری نیز نقش عمده‌ای داشته و اثر مثبتی روی این صفت می‌گذارد؛ به این صورت که هاپلوتیپ [T;T;T;A]، اسیدهای چرب اشباع (PUFA) شیر را افزایش داده اما غلظت اسید چرب پالمیتیک آن را کاهش می‌دهد. در بررسی کل ژنومی قطعات هموزیگوت بزهای بومی چینی، این ژن به‌همراه ژن STAG1، با باروری و رشدونمو جنین در ارتباط بوده و احتمالاً در اهلی‌شدن این گونه جانوری نقش مهمی بازی کرده است (Li *et al*., 2022).

ژن SPATA5 نوعی آنزیم تجزیه‌کننده ATP را کد می‌کند که در فرایندهای توسعه‌ای سلولی و حفظ عملکرد و انسجام میتوکندریایی نقش عمده‌ای دارد. این ژن باعث تغییر شکل میتوکندریایی شده و آن را فشرده می‌کند (Sujit *et al*., 2020). فشرده‌شدن میتوکندری اقدامی لازم در فرایند اسپرماتوژنز بوده و برای تأمین انرژی اسپرم در حال حرکت ضروری می‌باشد. این ژن در سلول‌های اسپرماتوگونی و اسپرماتوسیت بیان بالایی داشته و در فرایندهای مرتبط با باروری و بالغ‌شدن سلول‌های اسپرماتوگونیا و تبدیل آن به اسپرم فعال، یک عامل بسیار مهم تلقی می‌شود (Guo *et al*., 2020). از طرفی، پژوهش‌گران دریافتند که متیل‌دارشدن بیش از حد ژن SPATA5 با نازایی همبستگی بالایی دارد. در این نوع نازایی که کم‌نطفگی نامیده می‌شود، به‌دلیل کم‌بودن تعداد سلول‌های اسپرماتوزوئید در مایع منی، باروری به‌شدت کاهش می‌یابد (Sujit *et al*., 2020). متیل‌دارشدن DNA یک فرایند بیوشیمیایی بوده و عامل مهمی در رشدونمو معمولی سلول‌های بدن در نظر گرفته می‌شود. ژن RAB35 از خانواده ژنی RAB بوده و در جابه‌جایی وزیکولی بین غشای پلاسما و اندوزوم نقش اساسی ایفا می‌کند. از دست‌رفتن عملکرد ژن RAB35 نیز باعث کاهش درون‌بری و اختلال در بیرون‌بری شده و همین امر ممکن است برای فرایندهای اسکلت‌زایی و گاستروﻻسیون مهم باشد. با کنار هم قراردادن این دو عامل، مشخص شد که احتمالاً این ژن تنظیم‌کننده مهم بافت جنینی در اوایل دوران رشدونمو جنین می‌باشد (Remsburg *et al*., 2021). حضور ژن DICER1 در بدن برای تولید miRNAها لازم می‌باشد. از طرفی میزان بیان miRNAها در نواحی تناسلی بالاست. پس به‌نظر می‌رسد فعالیت ژن DICER1 و میکروRNAهای تولیدی آن، برای رشدونمو نرمال بسیاری از اندام‌های بدن، به‌ویژه سیستم تولیدمثلی، ضروری باشد (Hong *et al*., 2008).

ژن NF1 با تنظیم منفی مسیر سیگنالی Ras، سازگاری گوسفندان تبتی به ارتفاعات بالا را افزایش می‌دهد (Wei *et al*., 2016). با کاهش بیان این ژن، مسیر سیگنالی Ras فعال شده و به این ترتیب سلول‌های خون‌ساز رشد نامعمولی خواهند داشت. ژن NF1 در گوسفند چاکا جزو کاندیدهایی بود که رشدونمو بافت ماهیچه‌ای را تحت تأثیر قرار می‌دهند (Cheng *et al*., 2020). در گوسفند فین‌شیپ، NF1 یکی از پنج ژن کاندید اثرگذار روی صفت تعداد بره در هر زایش

گزارش شده است (Xu *et al*., 2018). ژن AKAP6 که با نام AKAP ماهیچه‌ای نیز شناخته می‌شود، با تنظیم MEF2 باعث تسریع تمایز بافت ماهیچه‌ای می‌شود. میزان بیان این ژن در سلول‌های ماهیچه‌ای اسکلتی و قلبی بسیار بالا بوده و گزارش شده که باعث عضلانی‌ترشدن گوسفندان ایسلندی می‌شود (Guðmundsdóttir, 2015). طی مطالعه در زمینه شناسایی ژن‌های تأثیرگذار روی گوسفند نژاد زندی، AKAP6 جزء ژن‌هایی دسته‌بندی شد که در رشد و تمایز بافت ماهیچه‌ای نقش پررنگی داشتند (Mohammadi *et al*., 2020). ژن HDAC9 نیز می‌تواند با پروتئین MEF2 ارتباط داشته و از فعالیت آن بکاهد. این پروتئین، توسعه ماهیچه‌های اسکلتی را تحت کنترل خود داشته و در میزان بیان ژن‌های ماهیچه‌ای نقش اساسی دارد. ژن HDAC9، از کارایی فاکتور MEF2 کاسته و در نتیجه باعث می‌شود تا رونویسی ژن‌های ماهیچه‌ای و رشدونمو این بافت دچار ضعف شود (Haberland *et al*., 2007). این ژن در بررسی نواحی ژنومی گوسفند چاکا جزو کاندیدهایی بود که توسعه بافت ماهیچه‌ای را تحت‌تأثیر قرار می‌داد (Cheng *et al*., 2020). ژن FBXL3، از ژن‌های برتری گزارش شده است که به‌طور معنی‌داری صفت وزن بدن بزها را تحت‌تأثیر قرار می‌دهد (Zhang, 2021). در بررسی نواحی ژنومی دارای جایگاه‌های صفات کمی، پژوهش‌گران دریافتند که ژن GRID2 نیز با نواحی QTL مرتبط با صفت وزن بدن همپوشانی دارد (Hu *et al*., 2021). ژن ADAMTS17 یک متالوپروتئاز را کد می‌کند که در تخریب ماتریکس بیرون‌سلولی نقش داشته و در انسان رشد و صفت طول قد را تحت تأثیر قرار می‌دهد، بنابراین به‌نظر می‌رسد که این ژن جزو ژن‌های کاندیدی در نظر گرفته می‌شود که صفت رشد و اندازه گوسفندان را تحت‌تأثیر قرار می‌دهد. هم‌چنین گزارش شده است که این ژن در گاو (Lee *et al*., 2020) نیز با فرایندهای مرتبط با رشدونمو استخوان و ماهیچه در ارتباط است.

ژن BMPR2 از کاندیدهای دیگری است که در این مطالعه شناسایی گردید. این ژن، یکی از عوامل مرتبط با بیماری فشار خون ریوی وابسته به ارتفاع (APH) در نظر گرفته می‌شود (Serranito *et al*., 2021). هم‌چنین در مطالعه انجام‌گرفته روی گوسفندان تبتی، این ژن جزو ژن‌های منتخبی گزارش شد که صفات بدنی این گوسفندان را تحت‌تأثیر قرار داده و باعث سازگاری آن‌ها به محیط سخت بیابانی شده است (Yang *et al*., 2016).

ژن EXOC6B فاکتور عملکردی دیگری است که کارکرد و ریخت‌شناسی نوک پستان را تحت‌تأثیر قرار می‌دهد. محصول این ژن یک جزء برون‌خوار دارد که به همین دلیل ممکن است در شکل‌دهی مجاری دخالت داشته باشد (Pausch *et al*., 2012). چون این ژن جزئی از یک کمپلکس برون‌خوار پیچیده می‌باشد، احتمال دارد که در اتصال گرانول‌های انسولین به غشای پلاسمایی و هم‌چنین ترشح انسولین نقش اساسی داشته باشد.

۵. نتیجه‌گیری و پیشنهادها

شناسایی نشانه‌های انتخاب در یافتن پایه ژنتیکی مرتبط با تنوعات موجود در فنوتیپ نژادهای مختلف، کمک‌کننده می‌باشد. در مطالعه حاضر، ژنوم کامل گوسفندان بومی با گوسفند نژاد رومانف خارجی موردارزیابی قرار گرفت. پس از استخراج تنوع‌های محتمل، اقدام به شناسایی نشانه‌های انتخاب اثرگذار روی صفات مختلف گوسفندان بومی گردید و ژن‌های دخیل در تفاوت‌های موجود موردبررسی قرار گرفت. نتایج هستی‌شناسی ژن‌ها نشان داد که بسیاری از ژن‌های منتخب در زمینه‌هایی مانند بهبود کیفیت چربی شیر، باروری، رشدونمو ماهیچه‌ای، وزن بدن، سازگاری به نواحی سخت بیابانی و کوهستانی و صفات مرتبط با شیر دخالت دارند. ویژگی‌های ذکرشده از عوامل بسیار مهم بقا در شرایط سخت اقلیمی کشور ایران محسوب شده و احتمالاً ژن‌های مربوطه طی انتخاب‌های متوالی در جمعیت‌های بومی تثبیت شده‌اند. شناسایی نشانه‌های انتخاب در نژادهای مختلف، گزینه‌های انتخابی و آمیخته‌گری را در اختیار قرار می‌دهد. یافتن ارتباط

بین تنوعات فنوتیپی و دلایل ژنتیکی آن‌ها یکی از کارآمدترین روش‌ها برای انتخاب و آمیزش‌های هدفمند بین نژادهای مختلف می‌باشد.

۶. تشکر و قدردانی

از دانشگاه محقق اردبیلی جهت حمایت از انجام پژوهش حاضر، تشکر و قدردانی می‌گردد.

۷. تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد.

۸. منابع

Cheng, J., Zhao, H., Chen, N., Cao, X., Hanif, Q., Pi, L., Hu, L., Chaogetu, B., Huang, Y., Lan, X., & Lei, C. (2020). Population structure, genetic diversity, and selective signature of Chaka sheep revealed by whole genome sequencing. *BMC genomics*, 21, 1-10.

Guðmundsdóttir, Ó. Ó. (2015). *Genome-wide association study of muscle traits in Icelandic sheep* (Doctoral dissertation).

Guo, J., Zhong, J., Liu, G. E., Yang, L., Li, L., Chen, G., Song, T., & Zhang, H. (2020). Identification and population genetic analyses of copy number variations in six domestic goat breeds and Bezoar ibexes using next-generation sequencing. *BMC Genomics*, 21(1), 1-13.

Haberland, M., Arnold, M. A., McAnally, J., Phan, D., Kim, Y., & Olson, E. N. (2007). Regulation of HDAC9 gene expression by MEF2 establishes a negative-feedback loop in the transcriptional circuitry of muscle differentiation. *Molecular and Cellular Biology*, 27(2), 518-525.

Hong, X., Luense, L. J., McGinnis, L. K., Nothnick, W. B., & Christenson, L. K. (2008). Dicer1 is essential for female fertility and normal development of the female reproductive system. *Endocrinology*, 149(12), 6207-6212.

Hu, L., Zhang, L., Li, Q., Liu, H., Xu, T., Zhao, N., Han, X., Xu, S., Zhao, X., & Zhang, C. (2022). Genome-wide analysis of CNVs in three populations of Tibetan sheep using whole-genome resequencing. *Frontiers in Genetics*, 13, 971464.

Khalkhali-Evrigh, R., Hedayat, N., Ming, L., & Jirimutu. (2022). Identification of selection signatures in Iranian dromedary and Bactrian camels using whole genome sequencing data. *Scientific Reports*, 12(1), 9653.

Kijas, J.W., Townley, D., Dalrymple, B.P., Heaton, M.P., Maddox, J.F., McGrath, A., Wilson, P., Ingersoll, R.G., McCulloch, R., McWilliam, S., & Tang, D. (2009). A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PloS one*, 4(3), p.e4668.

Lee, Y. L., Bosse, M., Mullaart, E., Groenen, M. A., Veerkamp, R. F., & Bouwman, A. C. (2020). Functional and population genetic features of copy number variations in two dairy cattle populations. *BMC Genomics*, 21(1), 1-15.

Li, G., Tang, J., Huang, J., Jiang, Y., Fan, Y., Wang, X., & Ren, J. (2022). Genome-Wide Estimates of Runs of Homozygosity, Heterozygosity, and Genetic Load in Two Chinese Indigenous Goat Breeds. *Frontiers in Genetics*, 13.

Li, R., Zhao, Y., Liang, B., Pu, Y., Jiang, L., & Ma, Y. (2023). Genome-Wide Signal Selection Analysis Revealing Genes Potentially Related to Sheep-Milk-Production Traits. *Animals*, 13(10), 1654.

Liu, J., Shi, L., Li, Y., Chen, L., Garrick, D., Wang, L., & Zhao, F. (2021). Estimates of genomic inbreeding and identification of candidate regions that differ between Chinese indigenous sheep breeds. *Journal of Animal Science and Biotechnology*, 12, 1-14.

Mohammadi, H., Rafat, S. A., Moradi Shahrbabak, H., Shodja, J., & Moradi, M. H. (2020). Genome-wide association study and gene ontology for growth and wool characteristics in Zandi sheep. *Journal of Livestock Science and Technologies*, 8(2), 45-55.

Pausch, H., Jung, S., Edel, C., Emmerling, R., Krogmeier, D., Götz, K. U., & Fries, R. (2012). Genome-wide association study uncovers four QTL predisposing to supernumerary teats in cattle. *Animal Genetics*, 43(6), 689-695.

Remsburg, C., Testa, M., & Song, J. L. (2021). Rab35 regulates skeletogenesis and gastrulation by facilitating actin remodeling and vesicular trafficking. *Cells & Development*, 165, 203660.

Serranito, B., Cavalazzi, M., Vidal, P., Taurisson-Mouret, D., Ciani, E., Bal, M., Rouvellac, E., Servin, B., Moreno-Romieux, C., Tosser-Klopp, G., & Hall, S. J. (2021). Local adaptations of Mediterranean sheep and goats through an integrative approach. *Scientific Reports*, 11(1), p.21363.

Shi, H., Li, T., Su, M., Wang, H., Li, Q., Lang, X., & Ma, Y. (2023). Whole genome sequencing revealed genetic diversity, population structure, and selective signature of Panou Tibetan sheep. *BMC Genomics*, 24(1), 1-15.

Sujit, K. M., Singh, V., Trivedi, S., Singh, K., Gupta, G., & Rajender, S. (2020). Increased DNA methylation in the spermatogenesis-associated (SPATA) genes correlates with infertility. *Andrology*, 8(3), 602-609.

Van Poucke, M., Sjoberg, A., Mattheeuws, M., Van Zeveren, A., Bouquet, Y., Chowdhary, B. P., & Peelman, L. J. (1997). Mapping of the ATP2B2 and PCCB genes on porcine chromosome 13. *Mammalian Genome*, 8(11), 852.

Wei, C., Wang, H., Liu, G., Zhao, F., Kijas, J. W., Ma, Y., Lu, J., Zhang, L., Cao, J., Wu, M., & Wang, G. (2016). Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *Scientific Reports*, 6(1), 26770.

Wiener, P., Robert, C., Ahbara, A., Salavati, M., Abebe, A., Kebede, A., Wragg, D., Friedrich, J., Vasoya, D., Hume, D.A., & Djikeng, A. (2021). Whole-genome sequence data suggest environmental adaptation of Ethiopian sheep populations. *Genome Biology and Evolution*, 13(3), evab014.

Xu, S. S., Gao, L., Xie, X. L., Ren, Y.L., Shen, Z. Q., Wang, F., Shen, M., Eyþórsdóttir, E., Hallsson, J. H., Kiseleva, T., & Kantanen, J. (2018). Genome-wide association analyses highlight the potential for different genetic mechanisms for litter size among sheep breeds. *Frontiers in Genetics*, 9, 118.

Yang, J. I., Li, W. R., Lv, F. H., He, S. G., Tian, S. L., Peng, W. F., Sun, Y. W., Zhao, Y. X., Tu, X. L., Zhang, M., & Xie, X. L. (2016). Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular Biology and Evolution*, 33(10), 2576-2592.

Zhang, L., Wang, F., Gao, G., Yan, X., Liu, H., Liu, Z., Wang, Z., He, L., Lv, Q., Wang, Z., & Wang, R. (2021). Genome-wide association study of body weight traits in Inner Mongolia cashmere goats. *Frontiers in Veterinary Science*, 8, 752746.

Zhao, H., Guo, T., Lu, Z., Liu, J., Zhu, S., Qiao, G., Han, M., Yuan, C., Wang, T., Li, F., & Zhang, Y. (2021). Genome-wide association studies detects candidate genes for wool traits by re-sequencing in Chinese fine-wool sheep. *BMC Genomics*, 22, 1-13.