

Genomic comparison of Chios and Frizarta dairy sheep breeds for somatic cell count using F_{ST} unbiased estimator and hapFLK methods

H. Mohammadi¹, A. H. Khaltabadi Farahani^{2*}

1. Assistant Professor, Department of Animal Sciences, Faculty of Agriculture and Environmental Sciences, University of Arak, Arak, Iran

2. Associate Professor, Department of Animal Sciences, Faculty of Agriculture and Environmental Sciences, University of Arak, Arak, Iran

(Received: 06-10-2024 – Revised: 29-11-2024 – Accepted: 30-11-2024)

Introduction: The principal aim of the dairy sheep industry worldwide is to produce high-quality milk. Mastitis is an inflammatory disease in dairy animals that occurs in response to infectious factors and has substantial impacts on animal health, and economic profitability. Somatic cell count has been used as an indirect method to control mastitis. Genetic resistance to mastitis involves linked biological mechanisms resulting from differences in mastitis response, which activate and regulate different levels of the immune response. Over the last decade, interest in identifying genes or genomic regions targeted by selection has grown. Identifying selection signatures can provide valuable insights into the genes or genomic regions that are or have been under selection pressure, which in turn leads to a better understanding of genotype-phenotype relationships. This study aimed to identify effective genes and genomic regions under positive selection in two sheep breeds using selection signature methods. For this purpose, F_{ST} and hapFLK analyses were performed using the genome-wide single nucleotide polymorphisms (SNPs).

Materials and methods: In this study, data from 585 sheep of different breeds were used to identify genomic regions under selection related to the somatic cell counts in milk. Illumina ovine Bead Chip 50K was used to determine the genotype of the samples. The genomic information of sheep breeds was extracted from the Zenodo database. Quality control on genotyped samples was performed using PLINK v1.9. SNPs with a minor allele frequency (MAF) of less than 0.02 and those with a call rate of less than 0.97 were excluded. In addition, individuals with more than 10% missing genotype data were removed. SNPs that did not conform to Hardy–Weinberg equilibrium ($P < 10^{-6}$) were also eliminated. Following these quality control measures, 41,673 SNPs from 585 sheep were retained for further analysis. To identify the signatures of selection, two statistical methods of F_{ST} and hapFLK were used under the software packages F_{ST} and hapFLK, respectively. The candidate genes were identified using PLINK v1.9 software and the Illumina gene list in R by SNPs located in the 0.01 percentile of F_{ST} and hapFLK values. In addition, the latest published version of the animal genome database was used to define QTLs associated with economically important traits at identified loci. The GeneCards (<http://www.genecards.org>) and UniProtKB (<http://www.uniprot.org>) databases were also used to interpret the function of the obtained genes.

Results and discussion: Using the F_{ST} approach, nine genomic regions on chromosomes 2, 5, 7, 10, 11, 13 (two regions), 17, and 22 were identified. The identified candidate genes associated with somatic cell count in these genomic regions included *IL11RA*, *CDC16*, *CARD14*, *BTRC*, *OTUD4*, *COL23A1*, *LACTB*, and *PRELID3B*. Some of the genes located in the identified selection regions were associated with the immune system, innate immune response, inflammation response, cancer disease, and milk production. Some of these genes in the selected regions were consistent with previous studies. The investigation of reported QTLs showed that these regions are related

* Corresponding author: amfarahanikh@gmail.com



to QTLs of important economic traits, including milk somatic cell count, udder height and depth, clinical mastitis, bovine tuberculosis susceptibility, and heat tolerance. Moreover, the results of hapFLK statistics in this research led to the identification of six genomic regions on chromosomes 3, 4, 5, 7, 10, and 13. The identified candidate genes associated with the somatic cell count in these genomic regions included *FAM49A*, *CDK6*, and *DLGAP5*. Bioinformatics analysis demonstrated that some of these genomic regions overlapped with known genes related to innate immune and various cancers.

Conclusions: Different genes that emerged in the aforementioned regions can be considered candidates for selection based on their function. It was found that most of the selected genes were consistent with some previous studies and were involved in production traits. However, further investigation is recommended to determine the exact function of the identified genes and QTLs. These areas should also be confirmed by other independent studies using larger samples. In general, the data from this study may be used in research on genomic selection and genomic regions associated with mastitis in dairy sheep and in further reviews and evaluations for the improvement of dairy sheep production.

Keywords: hapFLK test, F_{ST} statistics, Dairy sheep, Candidate gene, Mastitis

Conflicts of interest: The authors declare no conflicts of interest.

Funding: The authors received no specific funding for this project.

How to cite this article:

Mohammadi, H., & Khaltabadi Farahani, A. H. (2024). Genomic comparison of Chios and Frizarta dairy sheep breeds for somatic cell count using F_{ST} unbiased estimator and hapFLK methods. *Animal Production Research*, 13(4), 35-49. doi: 10.22124/ar.2024.28618.1853



مقایسه ژنومی نژادهای گوسفند شیری کایاس و فریزارته از نظر شمار سلول‌های بدنی شیر با استفاده از روش‌های برآوردگر ناریب F_{ST} و hapFLK

حسین محمدی^۱، امیر حسین خلت‌آبادی فراهانی^{۲*}

۱- استادیار، گروه علوم دامی، دانشکده کشاورزی و محیط زیست، دانشگاه اراک
۲- دانشیار، گروه علوم دامی، دانشکده کشاورزی و محیط زیست، دانشگاه اراک

(تاریخ دریافت: ۱۴۰۳/۰۷/۱۵ - تاریخ بازنگری: ۱۴۰۳/۰۹/۰۹ - تاریخ پذیرش: ۱۴۰۳/۰۹/۱۰)

چکیده

از شمار سلول‌های بدنی به‌عنوان روشی غیرمستقیم برای کنترل ورم پستان استفاده می‌شود. هدف این پژوهش، مقایسه ژنومی دو نژاد گوسفند شیری کایاس و فریزارته از نظر شمار سلول‌های بدنی شیر بود. برای این منظور از اطلاعات ژنوتیپی ۵۸۷ رأس دام با کمترین خویشاوندی و تعیین ژنوتیپ شده با آرایه‌های 50K گوسفندی استفاده شد. برای شناسایی مناطق ژنومی از دو آزمون آماری برآوردگر ناریب F_{ST} و hapFLK استفاده شد. ژن‌های کاندیدا با استفاده از SNPهایی که در بازه ۰/۰۱ درصد ارزش بالای این دو آزمون واقع شده بودند شناسایی شدند. نتایج حاصل از آماره تتا به شناسایی نه ناحیه ژنومی روی کروموزوم‌های ۲، ۵، ۷، ۱۰، ۱۱، ۱۳ (دو ناحیه)، ۱۷ و ۲۲ منجر شد. ژن‌های کاندیدای شناسایی شده شامل ژن‌های *CARD14*، *CDC16*، *IL11RA*، *PRELID3B* و *LACTB*، *COL23A1*، *OTUD4*، *BTRC*، *DLGAP5* و *CDK6* بودند و عملکردهای متفاوتی با سیستم ایمنی، پاسخ ایمنی ذاتی، پاسخ التهابی، سرطان و تولید شیر داشتند. نتایج حاصل از آماره hapFLK به شناسایی شش ناحیه ژنومی روی کروموزوم‌های ۳، ۴، ۵، ۷، ۱۰ و ۱۳ منجر شد. ژن‌های کاندیدای شناسایی شده شامل *FAM49A*، *CDK6* و *DLGAP5* بودند و عملکردهای متفاوتی در ایمنی ذاتی و انواع سرطان داشتند. با توجه به اینکه تحقیقات اندکی در رابطه با شمار سلول‌های بدنی شیر در گوسفندان نژاد شیری انجام شده است، نتایج این تحقیق می‌تواند اطلاعات ارزشمندی در رابطه با شمار سلول‌های بدنی شیر ارائه دهد. به هر حال، به‌منظور درک بهتر ژن‌های مؤثر بر این صفت مهم ضروری است پژوهش‌های گسترده‌تر با تعداد نمونه‌های بیشتری انجام شود.

واژه‌های کلیدی: آزمون hapFLK، آماره F_{ST} ، ژن کاندیدا، گوسفند شیری، ورم پستان

* نویسنده مسئول: amfarahanikh@gmail.com

مقدمه

در پژوهشی با هدف شناسایی مناطق ژنومی و ژن‌های کاندیدای مرتبط با صفات تولید شیر با استفاده از آماره F_{ST} و با مقایسه گوسفندان نژاد شیری با نژادهای گوشتی، ژن‌های کاندیدای *CSN1*، *OSBPL8*، *CSN2*، *ST3GAL1*، *CSN3*، *SLC35A3* و *VPS13B* مرتبط با تولید شیر و ژن‌های کاندیدای *CCDC152*، *DPY19L1* و *NT5DC1* مرتبط با تداوم شیردهی گزارش شدند (Rezvannejad et al., 2022). در مطالعه پویس ژنومی دیگری که با هدف شناسایی نشانه‌های انتخاب در نژادهای مختلف گوسفندان شیری با تولید بالا و تولید پایین انجام شده بود، ژن‌های کاندیدای *GHR*، *ARNT*، *CTSS*، *CTSK*، *FCGR3A*، *ROR1*، *SLC29A4* و *TNRC18* مرتبط با تولید و ترکیبات شیر گزارش شده است (Li et al., 2023a).

در مطالعات شناسایی نشانه‌های انتخاب بر اساس آماره F_{ST} ، اگر جمعیت‌های نمونه دارای اندازه مؤثر نابرابر یا دارای ساختار سلسله مراتبی باشند، کاوش ژنومی بر اساس مقادیر F_{ST} معمول می‌تواند نتیجه‌گیری اریب به‌همراه داشته باشد. برای تصحیح اثر رانش و کوواریانس هنگام تشخیص مکان‌های مورد انتخاب، آماره hapFLK ارائه شد (Fariello et al., 2013).

نژاد کایاس مورد استفاده در پژوهش حاضر، یکی از بهترین نژادهای گوسفند شیری کشور یونان از نظر صفات تولید شیر و تداوم شیردهی، دارای تولید شیر بالا و عملکرد تولیدی مطلوبی در سیستم‌های پرورشی صنعتی و نیمه-صنعتی است (Argyriadou et al., 2020). تولید شیر این نژاد در یک دوره شیردهی ۱۷۰ روزه برابر با ۳۰۷ لیتر و درصد چربی و پروتئین شیر نیز به ترتیب ۶/۲٪ و ۵/۵٪ است (Basdagianni et al., 2019). گوسفندان نژاد کایاس به تولید شیر بالا و شمار سلول‌های بدنی پایین معروف هستند (Argyriadou et al., 2023). همچنین، نژاد فریزرته، یکی از بهترین نژادهای گوسفند متوسط وزن و ممتاز کشور یونان از لحاظ صفات تولید و ترکیبات شیر گوسفندی است. طول شیردهی در این نژاد طی یک دوره شیردهی ۱۹۰ روزه برابر با ۲۵۶ کیلوگرم است و درصد چربی و پروتئین شیر به ترتیب ۶/۲۸٪ و ۵/۶۱٪ است. گوسفند نژاد فریزرته به تولید شیر با پروتئین بالا شهرت دارد و شیر تولیدی این نژاد برای ساخت پنیر بسیار مناسب است (Kominakis et al., 2017).

ورم پستان یک بیماری التهابی در غدد پستانی است که در پاسخ به آسیب فیزیکی و یا عفونت با میکروارگانیسم‌های بیماری‌زا مانند *Staphylococcus*، *Escherichia coli* و *Streptococcus pyogenes aureus* ایجاد می‌شود (Oner et al., 2021). این بیماری نه تنها منجر به ضررهای مستقیم به دلیل هزینه‌های درمان، دامپزشکی و مرگ و میر می‌شود، بلکه منجر به کاهش تولید و کیفیت شیر، چالش‌های تولیدمثلی، رفاه دام و افزایش نرخ حذف می‌شود (Argyriadou et al., 2023). ورم پستان بالینی در گله‌های گوسفندان شیری شایع بوده و به کاهش تولید شیر منجر می‌شود. در گله‌های گوسفند نژاد شیری وال دل بلیس ایتالیایی، کاهش تولید شیر ناشی از ورم پستان به‌طور میانگین برابر با ۱۶ درصد بود (Sutera et al., 2018). میزان عفونت پستانی با افزایش شمار سلول‌های بدنی در شیر تعیین می‌شود. بنابراین، شمار سلول‌های بدنی به‌عنوان معیار انتخاب غیرمستقیم برای بیماری ورم پستان استفاده می‌شود (Leitner et al., 2008).

در صورتی که جهش جدیدی سبب افزایش درجه شایستگی فرد شود، انتخاب باعث می‌شود این افراد مشارکت بیشتری در تشکیل نسل بعد داشته باشند. در نتیجه، فراوانی آلل جهش یافته بسته به سهم آن در افزایش درجه شایستگی به‌سرعت افزایش می‌یابد، که منجر به انتخاب مثبت در ژنوم می‌شود. انتخاب در جهت افزایش فراوانی جهش‌های ایجاد شده که در برخی از جمعیت‌ها مفید بوده‌اند، باعث ایجاد نشانه‌هایی همچون کاهش تنوع ژنتیکی، افزایش فراوانی آلل جهش یافته، افزایش عدم تعادل پیوستگی و توسعه ساختار هاپلوטיפی در سطح ژنوم شده‌اند. این نشانه‌های ژنومی که در اثر انتخاب روی ژنوم باقی می‌مانند، نشانه‌های انتخاب نامیده می‌شوند. یکی از مزایای اصلی ردیابی نشانه‌های انتخاب، امکان اجرای آن تنها با استفاده از اطلاعات مولکولی در غیاب رکوردهای فنوتیپی دقیق است. بر اساس ساز و کار عملکرد شناسایی نشانه‌های انتخاب در سطح ژنوم، روش‌های متفاوتی برای شناسایی نشانه‌های انتخاب وجود دارد (Ben Jemaa et al., 2023; Mohammadi, 2024).

یکی از مهم‌ترین آماره‌های شناسایی نشانه‌های انتخاب، آماره تمایز جمعیتی یا شاخص تثبیت (F_{ST}) است.

آمده، اطلاعات فنوتیپی مرتبط با صفت شمار سلول‌های بدنی در نژادهای مورد مطالعه به دو گروه مجزا تقسیم‌بندی شدند.

تجزیه آماری شناسایی نشانه‌های انتخاب: در این پژوهش برای شناسایی نشانه‌های انتخاب مرتبط با شمار سلول‌های بدنی، دام‌ها به دو گروه نژادی شامل نژاد شیری کایاس با تولید شیر بالا و شمار سلول‌های بدنی پایین و نژاد شیری فریزارته با تولید شیر بالا و شمار سلول‌های بدنی بالا دسته‌بندی شدند. برای شناسایی نواحی ژنومی مورد انتخاب از دو آزمون تمایز جمعیتی F_{ST} و hapFLK استفاده شد. آزمون تمایز جمعیتی F_{ST} به دو روش رایت (Wright, 1965) و روش برآوردگر ناریب تتا (θ) (Weir and Cockerham, 1984) محاسبه می‌شود. محاسبه‌های مربوط به دو روش رایت و تتا در محیط R برنامه‌نویسی شدند. برای شناسایی بهتر سیگنال‌های انتخاب در سطح ژنوم، به‌جای در نظر گرفتن ارزش عددی هر SNP از میانگین‌گیری ارزش‌های عددی SNP‌های مجاور به روش Creeping Window (CW) با طول ۱۰ نشانگر استفاده شد (Kim et al., 2015). و تنها ۰/۰۱ درصد مناطقی از ژنوم که در آن تمام نشانگرهای مجاور ارزش بالایی داشتند به‌عنوان نشانه‌های انتخاب شناسایی و تعیین شدند (Kijas et al., 2012; Saravanan et al., 2021). برای ترسیم گراف ارزش‌های Win10 در مقابل موقعیت ژنومی از بسته نرم‌افزاری ggplot در برنامه R (نسخه ۴/۲/۰) استفاده شد.

علاوه بر این، جهت بررسی الگوی انتخاب مثبت در سطح ژنوم از روش hapFLK نرم‌افزار hapFLK V1.4 (Fariello et al., 2013) استفاده شد. تعیین فازهای هاپلوتیپی برای کل ژنوم هر یک از جمعیت‌های مورد مطالعه، با نرم‌افزار fastPHASE V.1.2.3 انجام شد. در این برنامه از روش بیشینه درستی برای تخمین ژنوتیپ‌های گمشده (امپیوت) و بازسازی هاپلوتیپ‌ها استفاده می‌شود (Scheet and Stephens, 2006). فایل‌های ورودی برای نرم‌افزار hapFLK شامل دو فایل داده برای فازهای هاپلوتیپی که برای هر رأس دام تعیین شده و فایل نقشه‌های SNP مورد استفاده است. بنابراین، ارزش‌های P مقادیر hapFLK با استفاده از رابطه زیر نرمال شدند:

$$\frac{hapFLK_{adj}}{SD(raw\ hapFLK)} = \frac{raw\ hapFLK - mean(raw\ hapFLK)}{SD(raw\ hapFLK)}$$

انتخاب برای صفات مهم اقتصادی در گوسفند طی نسل‌های گذشته می‌تواند منجر به ایجاد نشانه‌های انتخاب در سطح ژنوم شده باشد که بررسی و ردیابی این نشانه‌ها می‌تواند در شناسایی این مناطق و ژن‌های کاندیدای مؤثر بر تولید و ترکیبات شیر مفید باشد. بنابراین، هدف از پژوهش حاضر، مقایسه ژنومی دو نژاد گوسفند شیری کایاس و فریزارته مرتبط با صفت شمار سلول‌های بدنی با استفاده از آماره‌های F_{ST} و hapFLK برای پویا نشانه‌های انتخاب مثبت بود.

مواد و روش‌ها

در پژوهش حاضر، از اطلاعات ژنوتیپی و فنوتیپی ۵۸۷ رأس گوسفند نژادهای شیری کایاس و فریزارته با کمترین روابط خویشاوندی استفاده شد. نمونه‌ها با استفاده از آرایه‌های شرکت ایلومینا Ovine SNP50 Bead Chip برای ۵۴۲۴۱ نشانگر تعیین ژنوتیپ شده بودند. داده‌های تعیین ژنوتیپ شده همراه با اطلاعات فنوتیپی مورد استفاده این پژوهش در پایگاه برخط Zenodo (<https://zenodo.org/deposit>) با شماره دسترسی ۸۳۴۴۵۴۶ در دسترس است.

برای فیلتراسیون داده‌های ژنومی از معیارهای فراوانی نمونه‌های تعیین ژنوتیپ شده، نرخ تعیین ژنوتیپ نشانگرها در هر نمونه و فراوانی آلل نادر (MAF) استفاده شد. در ابتدا، نمونه‌هایی که فراوانی تعیین ژنوتیپ در آنها کمتر از ۹۰ درصد بود، شناسایی و حذف شدند. در مرحله بعد، نشانگرهای دارای فراوانی آلل نادر کمتر از دو درصد، حذف شدند. سپس، نشانگرهایی که نرخ تعیین ژنوتیپ آنها در نمونه‌ها کمتر از ۹۷ درصد بود شناسایی و حذف شدند. در نهایت، از SNP‌های باقیمانده، آنهایی که در تعادل هاردی-واینبرگ نبودند به‌عنوان معیاری از خطای تعیین ژنوتیپ کنار گذاشته شدند. بدین منظور، سطح احتمال 10^{-6} در نظر گرفته شد. مراحل مختلف فیلتراسیون با نرم‌افزار PLINK (نسخه ۱/۹۰) انجام شد (Chang et al., 2015). سپس، داده‌های کنترل کیفی شده با نشانگرهای SNP مشترک بین دو نژاد با دستور merge- نرم‌افزار PLINK1.9 با هم ادغام شدند.

به‌منظور بررسی ساختار جمعیتی دام‌های مورد مطالعه و شناسایی نحوه قرار گرفتن حیوانات در گروه‌های نژادی خود و تقسیم‌بندی گروه‌ها از تجزیه PCA استفاده شد. این تجزیه، با استفاده از تابع prcomp در محیط نرم‌افزار R (نسخه ۴/۲/۰) انجام شد. سپس، بر اساس نتایج به‌دست

باقی ماندند که شامل اطلاعات ۲۶ جفت کروموزوم گوسفند بودند (جدول ۲). پس از حذف، داده‌های تعیین ژنوتیپ شده نژاد گوسفندان کایاس و فریزارته با هم ترکیب شدند که دارای ۴۱۶۷۳ نشانگر SNP مشترک بودند. نتایج تجزیه PCA نشان داد که با PC1، نژاد کایاس و با PC2، نژاد فریزارته به خوبی از یکدیگر تفکیک و جدا شدند و در هیچ نقطه‌ای با هم هم‌پوشانی نداشتند (شکل ۱).

در مجموع، تجزیه PCA حدود ۵/۱٪ واریانس (۳/۷٪) برای ویژه بردار ۱ و ۱/۴٪ برای ویژه بردار ۲ را توجیه می‌کنند. در سایر مطالعات، حیوانات فقط بر پایه ویژه بردارهای ۱ و ۲ بر اساس نژاد و منطقه جغرافیایی دسته‌بندی شده‌اند (Kijas et al., 2012; Waaineina et al., 2022). در مطالعه-ای که به وسیله Kijas et al. (2012) روی جمعیت‌های گوسفند دنیا انجام شده بود، PC1 و PC2 به ترتیب ۲/۹۸٪ و ۱/۴۴٪ واریانس را توجیه می‌کردند.

با توجه به اینکه انتخاب، علاوه بر جهش سودمند، جایگاه‌های مجاور آن‌را هم تحت تأثیر قرار می‌دهد (Akey et al., 2002)، در این پژوهش برای شناسایی بهتر سیگنال‌های انتخاب در سطح ژنوم به جای در نظر گرفتن ارزش عددی هر SNP، میانگین‌گیری ارزش‌های عددی SNP‌های مجاور به روش CW با طول ۱۰ نشانگر SNP صورت گرفت و تنها ۰/۱ درصد مناطقی از ژنوم که در آن تمام نشانگرهای مجاور دارای ارزش بالایی باشند، به عنوان نشانه‌های انتخاب شناسایی و تعیین شدند (شکل ۲). آستانه تعیین شده سخت‌گیرانه بوده و فقط مناطقی از ژنوم که بیشترین تمایز جمعیتی را در بین نژادها نشان می‌دهند، انتخاب شدند. این حد آستانه در مطالعات مختلف از ۰/۱ درصد (Patiabadi et al., 2024) تا پنج درصد (Hu et al., 2023) متغیر بود.

بخش‌هایی از ژنوم که ارزش F_{ST} بالایی دارند نشان‌دهنده تمایز بین دو گروه نژادی مورد بررسی در آن مناطق ژنومی در اثر انتخاب مثبت هستند که به دنبال انتخاب‌های طبیعی و مصنوعی طی نسل‌های مختلف برای جایگاه‌های مورد نظر به وجود آمده است. نتایج به دست آمده نشان داد که در چندین منطقه ژنومی، SNP‌های مجاور دارای تفرق جمعیتی بالایی بودند که شامل نه ناحیه ژنومی روی کروموزوم‌های ۲، ۵، ۷، ۱۰، ۱۱، ۱۳ (دو منطقه)، ۱۷ و ۲۲ بودند و در صدک ۹۹/۹۹ کل ارزش‌های تتا قرار داشتند و معنی‌دار بودند (شکل ۲).

در رابطه بالا، $raw\ hapFLK$ ، $mean\ (raw\ hapFLK)$ و $SD\ (raw\ hapFLK)$ به ترتیب ارزش اولیه، میانگین و خطای استاندارد همه مقادیر hapFLK در کل ژنوم است. تنها ۰/۱ درصد حد بالای ارزش hapFLK به عنوان نشانه انتخاب در نظر گرفته شد (Waaineina et al., 2022). سپس، با استفاده از بسته نرم افزاری ggplot در برنامه R (نسخه ۴/۲/۰)، گراف منتهن پلات ارزش‌های hapFLK به دست آمد.

ژن‌های گزارش شده در این مناطق و ۵۰۰ کیلو جفت‌باز اطراف آنها از پایگاه اطلاعاتی برخط Genes 113 BioMart Database Ensemble به وسیله برای تفسیر بهتر عملکرد ژن‌های به دست آمده از پایگاه‌های اطلاعاتی برخط GeneCards (<http://www.genecards.org>) و UniProtKB (<http://www.uniprot.org>) استفاده شد. با توجه به اینکه تفسیر عملکردی (Annotation) ژنوم گوسفند هنوز به طور کامل تکمیل نشده است، پایگاه‌هایی که برای توصیف ساز و کارهای بیوشیمیایی و شبکه ژنی و شناسایی QTL موجود هستند، گونه گوسفند را یا قبول نمی‌کنند و یا اینکه ناقص هستند، لذا از ژنوم گاو جهت بررسی اینکه مناطق ژنومی شناسایی شده حاوی ژن خاصی هستند استفاده شد. مناطق ارتولوگوس ژنوم گاو با استفاده از جستجو در پایگاه اطلاعاتی برخط Genes 113 BioMart Database Ensemble به وسیله برای تفسیر عملکردی (Annotation) ژنوم گوسفند هنوز به طور کامل تکمیل نشده است، پایگاه‌هایی که برای توصیف ساز و کارهای بیوشیمیایی و شبکه ژنی و شناسایی QTL موجود هستند، گونه گوسفند را یا قبول نمی‌کنند و یا اینکه ناقص هستند، لذا از ژنوم گاو جهت بررسی اینکه مناطق ژنومی شناسایی شده حاوی ژن خاصی هستند استفاده شد. مناطق ارتولوگوس ژنوم گاو با استفاده از جستجو در پایگاه اطلاعاتی برخط Genes 113 BioMart Database Ensemble مورد بررسی قرار گرفت. در نهایت، برای بررسی اینکه آیا مناطق مورد نظر با QTL‌های شناسایی شده هم‌پوشانی دارند یا خیر، از مرکز اطلاعاتی برخط QTL (<http://i.animalgenome.org/jbrowse/>) براساس آخرین نسخه به روز شده (Release 52, Dec 27, 2023) استفاده شد.

نتایج و بحث

آمار توصیفی مرتبط با تولید شیر و شمار سلول‌های بدنی نژادهای مورد مطالعه پژوهش حاضر در جدول ۱ ارائه شده است. نتایج مربوط به مراحل مختلف ویرایش داده‌های تعیین ژنوتیپ در جدول ۲ ارائه شده است. پس از مراحل مختلف کنترل کیفی در سطح SNP، در نهایت ۴۳۹۹۸ و ۴۷۴۵۱ نشانگر SNP به ترتیب در نژادهای کایاس و فریزارته

جدول ۱- آمار توصیفی نژادهای مورد بررسی در پژوهش حاضر

Table 1. Descriptive statistics of breeds in this study

Breed	Sample size	Average milking lengths (day)	Average milk production (kg)	Somatic cell count (cells/mL)
Chios	288	189	292.5	263266
Frizarta	299	168	278.4	595116

جدول ۲- خلاصه مراحل مختلف کنترل کیفیت نشانگرهای SNP برای تجزیه نشانه‌های انتخاب

Table 2. Summary of quality control steps for single nucleotide polymorphism (SNP) markers used for signature of selection analysis

Item	Chios	Frizarta
Number of animals	288	299
Excluding animals with call rate < 0.90	0	0
Excluding animals that were clustered outside of their population mass in PCA analysis	0	0
Excluding SNPs $\leq 2\%$ minor allele frequency (MAF) over all animals	5905	1678
Excluding SNPs $\leq 97\%$ call rate over all animals	1945	1999
Excluding SNPs with deviation from HW equilibrium ($P < 0.000001$)	236	244
Unmapped or located in non-autosomal	2157	2144
Remaining animals	288	299
Remaining SNPs	43998	47451

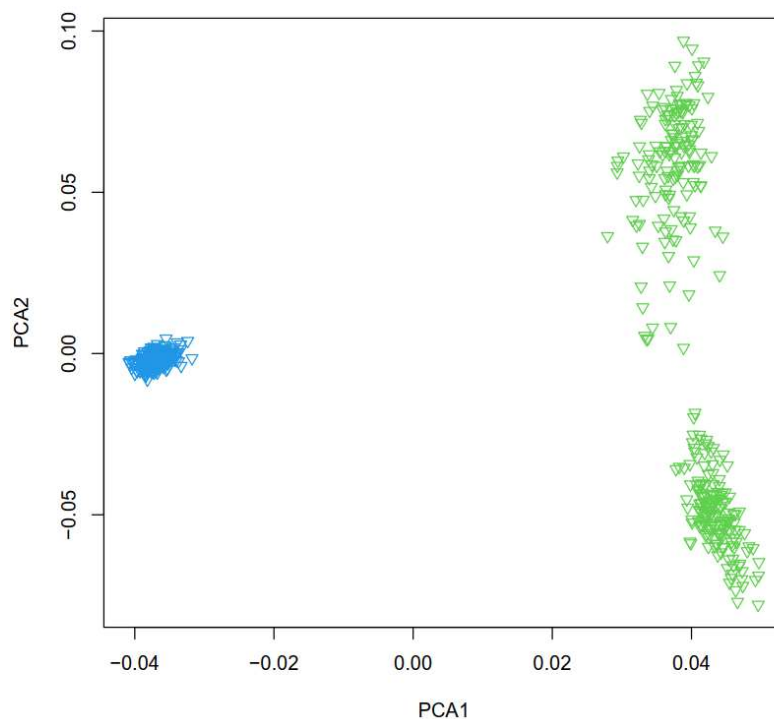


Fig. 1. Animals clustered on the basis of principal components analysis (PCA) using genotyping information of all animals. Green color: Chios; Blue color: Frizarta

شکل ۱- خوشه‌بندی حیوانات براساس تجزیه PCA با استفاده از اطلاعات ژنوتیپی تمام حیوانات. رنگ سبز: نژاد کایاس، رنگ

آبی: نژاد فریزارته

ژنومی زیر انتخاب، ژن‌های موجود در این مناطق شناسایی شده و عملکرد بیوشیمیایی آن‌ها بررسی شدند. در مجموع، ۱۳۷ ژن که در ارتباط با منطقه انتخابی بودند از این راه

از روش نشانه‌های انتخاب و آماره تتا تاکنون برای شناسایی مناطق ژنومی مرتبط با صفات مهم اقتصادی مختلف در دام‌های اهلی استفاده شده است. پس از شناسایی نواحی

اولین منطقه ژنومی شناسایی شده مرتبط با شمار سلول‌های بدنی روی کروموزوم ۲ در ناحیه ۳۶/۴۶-۳۷/۴۰ مگابازی قرار داشت. از بین ژن‌های موجود در این ناحیه، ژن کاندیدای *IL1IRA* بیشترین ارتباط را با سیستم ایمنی داشت. ژن *IL1IRA* در پاسخ ایمنی نقش اساسی دارد و به‌عنوان گیرنده اینترلوکین-۱۱ عمل می‌نماید که برای ارسال سیگنال به گیرنده‌های آنتی ژن سلول‌های T و B مورد نیاز است (<http://www.genecards.org>). در پویش ژنومی با هدف شناسایی نشانه‌های انتخاب مرتبط با صفات تولیدی در گاوهای نژاد Xia'nan، ژن کاندیدای *IL1IRA* مرتبط با سیستم ایمنی گزارش شده است (Song et al., 2024).

تعیین شدند. ژن‌های مشخص شده در جدول ۲، عملکردهای متفاوتی با سیستم ایمنی، پاسخ ایمنی ذاتی، پاسخ التهابی، سرطان و تولید شیر داشتند که می‌توانند به‌طور مستقیم و غیرمستقیم با صفت شمار سلول‌های بدنی مرتبط باشند. به‌علت عدم پیشرفت کافی در مهار بیماری ورم پستان به‌وسیله روش‌های سنتی دامپزشکی، امروزه بیشترین توجه به روش‌های طبیعی بر پایه ایمنی طبیعی دام‌ها علیه عفونت غدد پستانی است. شمار سلول‌های بدنی در یک میلی‌لیتر شیر که به‌طور عمده برای ارزیابی وضعیت سلامت غدد پستانی استفاده می‌شود می‌تواند در نتیجه عفونت‌های باکتریایی در غدد پستانی افزایش یابد. به همین خاطر، ژن‌های کاندیدای ورم پستان بالینی، یک یا چند نقش در سیستم ایمنی میزبان دارند.

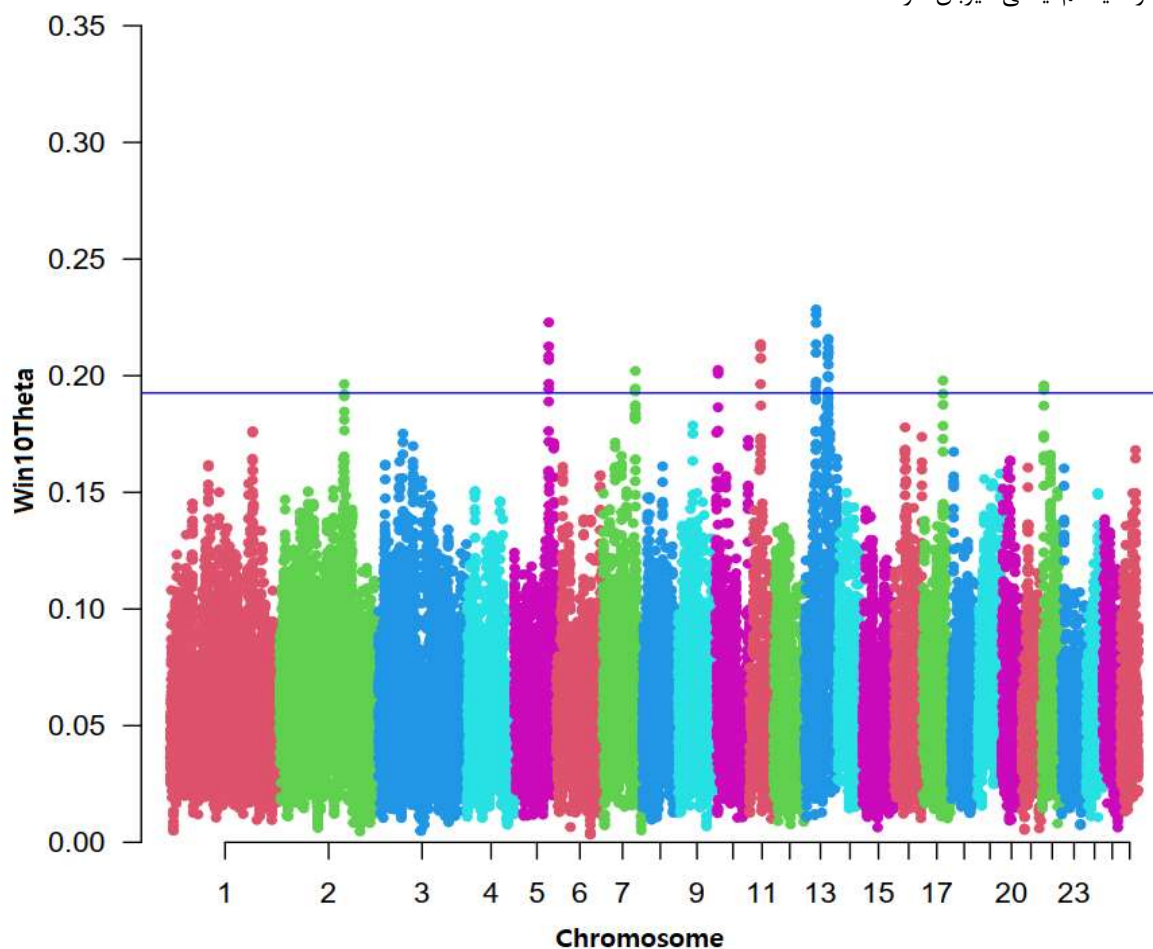


Fig. 2. Distribution of win10 F_{ST} statistic values in study population. The SNP position on different chromosomes shown on the X-axis, and F_{ST} values are plotted on the Y-axis

شکل ۲- توزیع ارزش‌های Win10 آماره F_{ST} در سطح ژنومی جمعیت‌های مورد مطالعه. موقعیت کروموزومی SNP‌ها روی محور X و ارزش‌های F_{ST} آنها روی محور Y نمایش داده شده است

منفی به ویژه بین تولید شیر و مقاومت به ورم پستان، کاهش مقاومت دام به بیماری‌هایی مانند ورم پستان مشاهده شده است. بنابراین، دور از انتظار نبود که در پژوهش حاضر، برخی از ژن‌های کاندیدای مرتبط با تولید و ترکیبات شیر شناسایی شوند. مهمترین منطقه ژنومی شناسایی شده با استفاده از هر دو آماره در پژوهش حاضر در ناحیه ۳۶/۴۹-۳۷/۴۴ مگابازی کروموزوم ۵ قرار داشت. از بین ژن‌های موجود در این ناحیه، ژن کاندیدای *COL23A1* بیشترین ارتباط را با تولید شیر داشت. در مطالعه پویش ژنومی با هدف شناسایی ژن‌های کاندیدای مرتبط با تولید و تداوم شیردهی در گاو میش‌های Murrah انجام شده بود، ژن کاندیدای *COL23A1* در ارتباط با صفت تولید شیر گزارش شد (Lázaro et al., 2024). همچنین، در مطالعه پویش ژنومی با استفاده از رویکرد تک مرحله‌ای در گاو میش‌های مصری نشان داده شد که ژن *COL23A1*، ۰/۴۱ درصد از واریانس ژنتیکی تولید شیر را توجیه می‌کند (Abdel-Shafy et al., 2020).

دومین منطقه ژنومی شناسایی شده با استفاده از آماره F_{ST} تا در پژوهش حاضر مرتبط با تولید شیر در ناحیه ۴۴/۷۹-۴۵/۷۹ مگابازی کروموزوم ۷ قرار داشت که از بین ژن‌های موجود در این ناحیه، ژن *LACTB* بیشترین ارتباط را داشت. ژن کاندیدای *LACTB* نقش کلیدی در سوخت و ساز لیپیدها در بدن داشته و در تولید اسیدهای چرب موجود در شیر نقش دارد (http://www.genecards.org). در مطالعه پویش ژنومی با هدف شناسایی ژن‌های کاندیدای مرتبط با ترکیبات شیر در بزهای نژاد Karachai، ژن کاندیدای *LACTB* مرتبط با ساخت چربی شیر در سلول‌های غدد پستانی گزارش شده است (Selionova et al., 2024).

سومین منطقه ژنومی شناسایی شده با استفاده از هر دو آماره در پژوهش حاضر مرتبط با تولید شیر در ناحیه ۹۱/۹۰-۵۷/۵۶ مگابازی کروموزوم ۱۳ قرار داشت که از بین ژن‌های موجود در این ناحیه، ژن *PRELID3B* بیشترین ارتباط را داشت. ژن کاندیدای *PRELID3B* دارای نقش کلیدی در تنظیم ساخت و انتقال چربی در بدن است (Miliara et al., 2019). در مطالعه پویش ژنومی که با هدف شناسایی ژن‌های مرتبط با تولید شیر در بزهای شیری انجام شده است، ژن کاندیدای *PRELID3B* مرتبط با چربی شیر گزارش شد (Zhao et al., 2024).

دومین منطقه ژنومی شناسایی شده در پژوهش حاضر در ناحیه ۸۵/۸۸-۸۶/۸۷ مگابازی کروموزوم ۱۰ قرار داشت. از بین ژن‌های موجود در این ناحیه، ژن کاندیدای *CDC16* بیشترین ارتباط را با شمار سلول‌های بدنی داشت. ژن *CDC16* به همراه پروتئین اتصال‌دهنده کلسیم و پروتئین اتصال‌دهنده لیپوساکاریدها در پاسخ ایمنی ذاتی نقش دارند (Shi et al., 2018). در مطالعه پویش ژنومی با هدف شناسایی ژن‌های کاندیدای مرتبط با پاسخ ایمنی به ویروس نیوکاسل، ژن کاندیدای *CDC16* گزارش شده است (Habimana et al., 2021).

سومین ناحیه ژنومی شناسایی شده روی کروموزوم ۱۱ در ناحیه ۵۱/۳۸-۵۲/۱۸ مگابازی قرار داشت. از بین ژن‌های موجود در این منطقه، ژن کاندیدای *CARD14* بیشترین ارتباط را با پاسخ سیستم ایمنی داشت. ژن *CARD14* واسطه اصلی پاسخ التهابی است و نقش کلیدی در جذب نوتروفیل‌ها، تعدیل سلول‌های تومور و القای التهاب دارد (http://www.genecards.org). در مطالعات مختلف پویش ژنومی، ارتباط معنی‌داری بین ژن *CARD14* با پاسخ التهابی و تنظیم ایمنی بدن گزارش شده است (Israel and Mellett, 2018, Wang et al., 2018).

منطقه ژنومی دیگر مرتبط با سیستم ایمنی در ناحیه ۸۲/۸۲-۱۳/۱۲ مگابازی از کروموزوم ۱۷ شناسایی شد. ژن کاندیدای *OTUD4* در این ناحیه قرار داشت. ژن *OTUD4* بخشی از مسیر فعال‌سازی در نوتروفیل‌هایی است که کمپلکس‌های ایمنی را در مکان‌های آسیب دیده یا عفونی محدود کرده‌اند (http://www.genecards.org). در مطالعه پویش ژنومی در خوک‌های نژاد Duroc، ژن کاندیدای *OTUD* مرتبط با سیستم ایمنی گزارش شده است (Li et al., 2024).

آخرین منطقه ژنومی مهم دیگری که در پژوهش حاضر با استفاده از آماره تتا و مرتبط با سیستم ایمنی شناسایی شد، ژن کاندیدای *BTRC* در ناحیه ۲۱/۹۰-۲۲/۹۰ مگابازی از کروموزوم ۲۲ بود. در پویش ژنومی در نژادهای مختلف گاو شیری با هدف شناسایی ژن‌های کاندیدای مرتبط با بیماری ورم پستان، ژن کاندیدای *BTRC* گزارش شده است (Cai et al., 2018). از سوی دیگر و از آنجایی که برنامه‌های اصلاح نژادی دام‌های شیری طی دهه‌های اخیر به‌طور عمده در جهت افزایش تولید شیر و ترکیبات شیر مانند درصد چربی و پروتئین بوده است، به دلیل همبستگی ژنتیکی

نشاندگی در شکل ۳ نشان داده شده است. نتایج به دست آمده نشان داد که در چندین منطقه ژنومی، SNPهای مجاور دارای تفرق جمعیتی بالایی بودند که شامل شش ناحیه ژنومی روی کروموزومهای ۳، ۴، ۵، ۷، ۱۰ و ۱۳ بودند و در صدک ۹۹/۹۹ کل ارزشهای hapFLK قرار داشتند و معنی دار نیز بودند (شکل ۳).

بررسی QTLهای شناسایی شده در مناطق ژنومی مورد نظر نشان داد که روی کروموزومهای ۵، ۷، ۱۰، ۱۱، ۱۷ و ۲۲، QTLهای گزارش شده مرتبط با صفات مهم اقتصادی بیماری ورم پستان و توبرکلوزیس، شمار سلولهای بدنی، صفات ساختاری سیستم پستانی و مقاومت به کهنه و گرما قرار دارند (جدول ۳).

شناسایی نشانه‌های انتخاب با استفاده از روش hapFLK نمودار منهن بر اساس ضرایب hapFLK در هر جایگاه

جدول ۳- نواحی ژنومی مورد انتخاب و ژنهای شناسایی شده در این مناطق ژنومی مرتبط با صفت شمارش سلولهای بدنی با استفاده از آماره F_{ST}

Table 3. Genomic regions containing selection signatures related to somatic cell count trait, and genes (QTL) reported in these genomic regions using F_{ST} method

Number of chromosomes	Genomic region	Genes including in these region	QTL reported in these region	(PubMed ID)
2	36461488: 37406488	<i>RASEF, CCL21, CCL19, CCL27, IL11RA, GALT, SIGMARI, ARID3C, DCTN3, RPP25L, CNTFR, ENHO, DNAIL1, FAM219A, MYORG, NUDT2, KIF24, UBAP1, RPL23, DCAF12</i>	-	-
5	36492388:37447388	<i>MXD3, U6, LMAN2, RGS14, SLC34A1, PFN3, FI2, GRK6, PRR7, DBN1, PDLIM7, DOK3, DDX41, FAM193B, TMED9, B4GALT7, N4BP3, RMND5B, NHP2, HNRNPAB, PHYKPL, COL23A1, CLK4</i>	Milk somatic cell count	21831322
7	44797111: 45797111	<i>USP3, CA12, APH1B, RAB8B, LACTB, TPM1, TLN2, bta-mir-190a</i>	Tick resistance	28619006
10	85883815:86873815	<i>CUL4A, RASA3, ATP4B, GRK1, TMEM255B, GAS6, CDC16, UPF3A, CHAMP1, LAMP1, GRTP1, DCUN1D2, TMC03, TFDP1</i>	Udder height and depth	21831322
11	51385928:52188928	<i>RPTOR, ENDOV, NPTX1, RNF213, SLC26A11, SGSH, CARD14, EIF4A3, GAA, CCDC40, TBC1D16, CBX4, CBX8, CBX2, ENPP7</i>	Clinical mastitis	25151887
13	61377331:62332331	<i>POFUT1, KIF3B, ASXL1, NOL4L, COMMD7, DNMT3B, MAPRE1, SUN5, BPIFB2, BPIFB6, BPIFB3, BPIFB4, BPIFA2C, BPIFA2B</i>	Teat length	21831322
13	56904112:57917112	<i>ZNF831, PRELID3B, ATP5F1E, TUBB1, CTSZ, NELFCD, GNAS, NPEPL1, STX16, APCDD1L</i>	Clinical mastitis	23647142
17	12826730:13821730	<i>VAPB, RAB22A, CIMIP1, SMAD1, OTUD4, ABCE1, ANAPC10, ANAPC10, HHIP</i>	Bovine tuberculosis susceptibility	26960806
22	21907407:22907407	<i>BTRC, POLL, DPCD, FBXW4, FGF8, NPM3, OGA, KCNIP2, ARMH3, HPS6, LDB1, PPRC1, SNORD22, NOLC1, ELOVL3, PITX3, GBF1, NFKB2, PSD, FBXL15, CUEDC2, MFSD13A, ACTR1A, SUFU</i>	Heat tolerance	28365122

اولین منطقه ژنومی شناسایی شده مرتبط با شمار سلول‌های بدنی روی کروموزوم ۳ در ناحیه ۲۴/۲۱-۲۵/۲۳ مگابازی قرار داشت. از بین ژن‌های موجود در این ناحیه، ژن کاندیدای *FAM49A* بیشترین ارتباط را با صفت مورد مطالعه داشت. ژن *FNDC3B* دارای نقش تنظیمی مثبت در تکثیر فیبروبلاست‌ها و تشکیل بافت همبند است (<http://www.genecards.org>). مطالعه پوشش کل ژنومی با هدف شناسایی مناطق ژنومی مرتبط با صفات ساختاری پستان در گاوهای هلشتاین فرانسوی انجام شده بود. ژن کاندیدای *FAM49A* مرتبط با صفات ارتفاع پستان و نوک پستان گزارش شد (Gonzalez et al., 2020). در مطالعات قبلی در خصوص صفات پستانی، همبستگی ژنتیکی بین شمار سلول‌های بدنی با اتصالات پستانی در گوسفند گزارش شده است (Casu et al., 2010).

از روش نشانه‌های انتخاب و روش hapFLK تاکنون جهت شناسایی مناطق ژنومی مرتبط با صفات مهم اقتصادی مختلف در دام‌های اهلی استفاده شده است. در پژوهشی با هدف شناسایی نشانه‌های انتخاب مرتبط با سازگاری، پاسخ سیستم ایمنی، تولید و تولیدمثل در نژادهای مختلف بز با استفاده از آماره hapFLK نشان داده شد که نواحی ژنومی روی کروموزوم‌های ۳، ۴، ۱۰، ۱۵، ۲۲ و ۲۶ دارای نشانه‌های انتخاب هستند. این مناطق ژنومی شامل ژن‌های کاندیدای *PCK*، *MSTI*، *PDE4B*، *HYALI* بودند که نقش کلیدی در سازگاری، سیستم ایمنی و تولید شیر داشتند (Waineina et al., 2022). پس از شناسایی نواحی ژنومی مورد انتخاب، ژن‌های موجود در این مناطق، شناسایی شده و عملکرد بیوشیمیایی آن‌ها بررسی شدند. در مجموع، ۶۱ ژن که در ارتباط با منطقه انتخابی بودند با این روش تعیین شدند (جدول ۴).

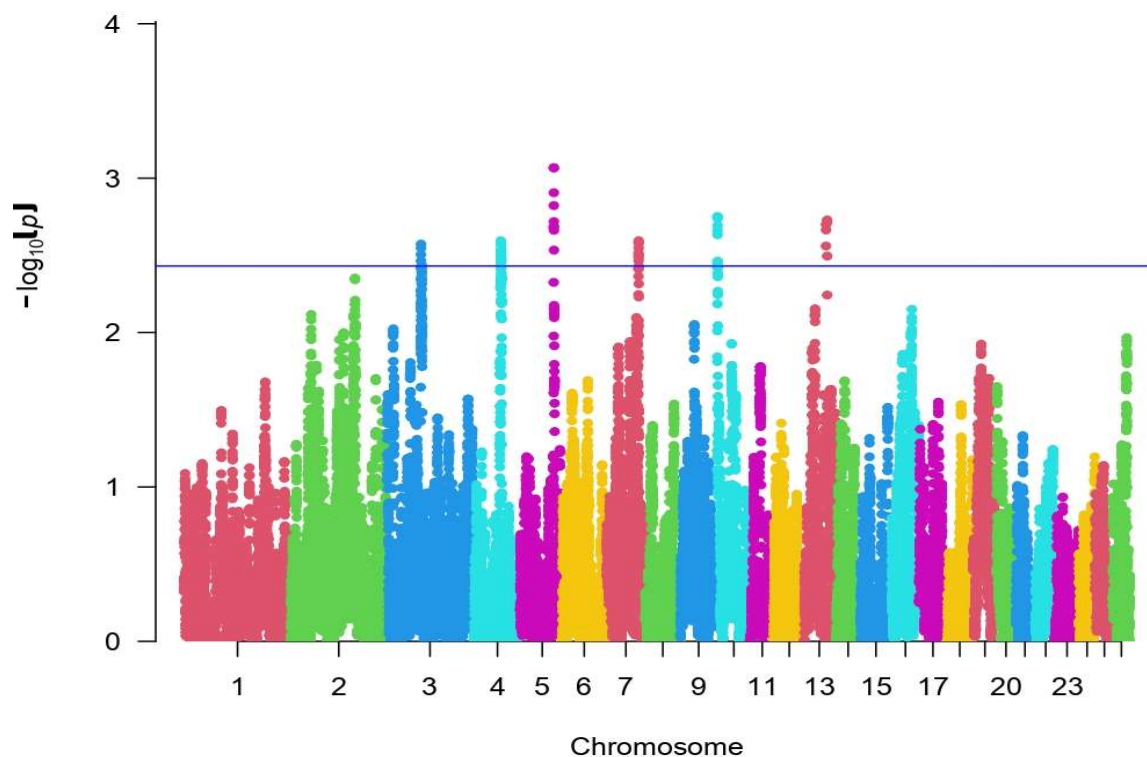


Fig. 3. Distribution of hapFLK statistic values in study population. The SNP position on different chromosomes is shown on the X-axis and hapFLK values are plotted on the Y-axis

شکل ۳- توزیع ارزش‌های hapFLK در سطح ژنومی جمعیت‌های مورد مطالعه. با موقعیت کروموزومی SNP‌ها روی محور X و ارزش‌های hapFLK آنها روی محور Y نمایش داده شده است

ژنومی عمومی شرکت ایلومینا استفاده کردند در معرض اربیبی‌هایی مانند خطاهای تعیین ژنوتیپ و نمونه‌گیری هستند که می‌تواند منجر به نتایج مثبت کاذب شود (Islam *et al.*, 2019).

یکی از راه‌حل‌های توصیه شده برای جلوگیری از خطاهای ذکر شده جهت شناسایی مناطق ژنومی مورد انتخاب، استفاده از آستانه‌های سخت‌گیرانه برای مناطق ژنومی کاندیدا است (Gouveia *et al.*, 2014). در پژوهش حاضر برای شناسایی دقیق‌تر مناطق مورد انتخاب مثبت از آستانه ۰/۰۱ درصد ارزش‌های هر دو آماره F_{ST} و hapFLK بالای جمعیت استفاده شد که دامنه سخت‌گیرانه‌ای بود و دامنه آن از ۰/۰۱ درصد (Waiveina *et al.*, 2022) تا پنج درصد (Hu *et al.*, 2023) در مطالعات مختلف گزارش شده است. همچنین، از دلایل متفاوت بودن مناطق ژنومی شناسایی شده در پژوهش حاضر با سایر تحقیقات می‌توان به متفاوت بودن اندازه نمونه و کم بودن تعداد نژادهای مورد بررسی و روش‌های آماری متفاوت برای شناسایی مناطق ژنومی مورد انتخاب مثبت در این مطالعه اشاره کرد.

در این پژوهش، برخی از ژن‌ها و جایگاه‌های ژنومی یافت شده در گوسفند، جدید هستند. با توجه به عملکرد زیستی ژن‌های کاندیدای شناسایی شده در این تحقیق که منطبق با گزارشات پویش ژنومی پیشین در گونه‌های دیگر است، می‌توان بروز فنوتیپی صفت مورد مطالعه را توجیه کرد. می‌توان با ادغام داده‌های تحقیقات مشابه جدید و استفاده از تجزیه‌های آماری جامع‌تر برای تأیید نتایج پژوهش حاضر استفاده کرد. در این پژوهش به دلیل عدم دسترسی به رکوردهای فنوتیپی و اطلاعات ژنوتیپی مرتبط با صفات مورد مطالعه پژوهش حاضر در گوسفندان بومی کشور، از اطلاعات فنوتیپی و ژنوتیپی گوسفندان نژاد شیری یونانی استفاده شد. لذا استفاده از نتایج این تحقیق در جمعیت‌های گوسفندان بومی کشور، نیاز به مطالعه بیشتر دارد تا در این جمعیت‌ها نیز تأیید شوند. بررسی بیشتر نواحی مهم ژنومی شناسایی شده، به‌خصوص مناطقی که شامل ژن‌های کاندیدا با اثر مستقیم بر تولید شیر و سیستم ایمنی هستند، با استفاده از آزمون‌های آزمایشگاهی مختلف (بررسی چندشکلی در ژن‌های کاندیدا و یا بیان ژن) می‌تواند در تأیید نتایج به‌دست آمده در این پژوهش مؤثر باشند.

دومین منطقه ژنومی شناسایی شده در پژوهش حاضر در ناحیه ۹/۹۷-۱۰/۹۸ مگابازی کروموزوم شماره ۵ قرار داشت. از بین ژن‌های موجود در این ناحیه، ژن کاندیدای *CDK6* بیشترین ارتباط را با صفت شمار سلول‌های بدنی داشت. ژن *CDK6* جزء ژن‌های کلیدی در مسیر سیگنال‌دهی PI3K-Akt است که به‌عنوان واسطه در سیستم ایمنی ذاتی، نقش مؤثری دارد (http://www.genecards.org). در مطالعه ژنومی با هدف شناسایی مناطق ژنومی مرتبط با بیماری‌های رحمی، ژن کاندیدای *CDK6* گزارش شده است (May *et al.*, 2022). سومین ناحیه ژنومی شناسایی شده روی کروموزوم ۷ در ناحیه ۶۳/۳۹-۶۴/۴۸ مگابازی قرار داشت. از بین ژن‌های موجود در این منطقه، ژن کاندیدای *DLGAP5* بیشترین ارتباط را با صفت شمار سلول‌های بدنی داشت. این ژن نقش مهمی در ایمنی ذاتی، تفرق و تمایز سلولی و آپوپتوز در بدن دارد (http://www.genecards.org). در مطالعه پویش کل ژنومی مرتبط با سندرم پلی‌سیستیک تخمدانی، ارتباط معنی‌داری بین ژن کاندیدای *DLGAP5* در انسان با این سندرم گزارش شده است (Deng *et al.*, 2022). همچنین، در مطالعه پویش ژنومی مرتبط با سرطان سینه، ژن کاندیدای *DLGAP5* گزارش شده است (Li *et al.*, 2023b). با بررسی منابع انجام شده، پژوهش حاضر، یکی از اولین مطالعات شناسایی نشانه‌های انتخاب بر پایه دو رویکرد متفاوت مرتبط با صفت شمار سلول‌های بدنی در گوسفندان نژاد شیری بوده است. یکی از آزمون‌های پرکاربرد آماری شناسایی نشانه‌های انتخاب، آماره تمایز جمعیتی یا شاخص تثبیت است. از مشکلات اصلی این روش، لحاظ نکردن خطای نمونه‌گیری است که این مورد با استفاده از روش ناریب تتا تصحیح شده است. همچنین، در آماره hapFLK که مبتنی بر عدم تعادل پیوستگی و طول هاپلوتیپی است، مناطق ژنومی مورد انتخاب به‌وسیله فازهای هاپلوتیپی تشخیص داده می‌شوند. یکی از مشکلات اصلی در آماره hapFLK که منجر به شناسایی مناطق ژنومی کمتر از آماره‌های دیگر می‌شود ماهیت آن در شناسایی مناطق ژنومی حاصل از مهاجرت و تنگنای جمعیتی (Fariello *et al.*, 2013) بوده و نتایج آن تحت تأثیر اندازه نمونه است (Bertolini *et al.*, 2018). از طرف دیگر، مطالعات ژنومی که با هدف شناسایی نشانه‌های انتخاب مثبت از آرایه‌های

جدول ۴- نواحی ژنومی مورد انتخاب و ژن‌های شناسایی شده در این مناطق ژنومی مرتبط با شمار سلول‌های بدنی با استفاده از آماره hapFLK

Table 4. Genomic regions containing selection signatures related to somatic cell count and genes (QTL) reported in these genomic regions using hapFLK method

Number of chromosomes	Genomic region	Genes including in these region	QTL reported in these region	(PubMed ID)
3	24218842:25235842	<i>MYCN, CYRIA, FAM49A</i>	Heat tolerance	24362770
4	9971314:10986314	<i>MTERF1, AKAP9, CYP51A1, LRRD1, KRIT1, ANKIB1, TMBIM7, GATAD1, PEX1, RBM48, FAM133B, CDK6</i>	-	-
5	36492388:37447388	<i>MXD3, U6, LMAN2, RGS14, SLC34A1, PFN3, F12, GRK6, PRR7, DBN1, PDLIM7, DOK3, DDX41, FAM193B, TMED9, B4GALT7, N4BP3, RMND5B, NHP2, HNRNPAB, PHYKPL, COL23A1, CLK4</i>	Milk somatic cell count	21831322
7	63396753:64545753	<i>FERMT2, DDHD1, DUOX1, DUOXA1, DUOX2, SORD, DLGAP5</i>	-	-
10	2190772:3194772	<i>DIAPH3, U6</i>	Bovine respiratory disease susceptibility	27077383
13	56904112:57917112	<i>ZNF831, PRELID3B, ATP5F1E, TUBB1, CTSZ, NELFCD, GNAS, NPEPL1, STX16, APCDD1L, VAPB, RAB22A, CIMIP1</i>	Clinical mastitis	23647142

نتیجه‌گیری کلی

که در این مناطق ژنومی، چندین QTL مرتبط با صفات ساختاری پستان، ورم پستان و مقاومت به بیماری از جمله سل گزارش شده است. با توجه به اهمیت اقتصادی بیماری ورم پستان و تولید شیر در پرورش گوسفند، نتایج این پژوهش می‌تواند منابع اطلاعاتی ارزشمندی برای شناسایی ژن‌های کاندیدای تولید شیر، سیستم ایمنی و ساز و کار مولکولی مربوطه ارائه دهد. بنابراین، از این اطلاعات می‌توان در برنامه‌های اصلاح نژادی به‌منظور بهبود سیستم ایمنی و افزایش تولید شیر از راه انتخاب به‌کمک نشانگرهای ژنتیکی برای افزایش بازده انتخاب و بهبود کمیت و کیفیت شیر در گوسفند استفاده کرد.

در این مطالعه برای شناسایی نشانه‌های انتخاب مثبت مرتبط با شمار سلول‌های بدنی از آماره‌های تمایز جمعیتی ناریب تتا و hapFLK استفاده شد و نشانه‌های انتخاب در بخش‌هایی از ژنوم شناسایی شد که برخی از این مناطق در بررسی‌های پیشین نیز تأیید شده بودند. بررسی ژن‌ها در نواحی مورد انتخاب نشان داد ژن‌های موجود در این مناطق با عملکردهای متفاوتی همانند سیستم ایمنی، پاسخ ایمنی ذاتی، پاسخ التهابی، سرطان و تولید شیر مرتبط هستند. همچنین، بررسی مناطق ژنومی شناسایی شده در این تحقیق با QTL‌های شناسایی شده در گونه گاو نشان داد

فهرست منابع

- Abdel-Salam, S. A. M., Mekkawy, W., Hafez, Y. M., Zaki, A. A., & Abou-Bakr, S. (2011). Fitting lactation curve of Egyptian Buffalo using three different models. *Egyptian Journal of Animal Production*, 48, 119-133. doi: 10.21608/ejap.2011.94054
- Akey, J. M., Zhang, G., Zhang, K., Jin, L., & Shriver, M. D. (2002). Interrogating a high-density SNP map for signatures of natural selection. *Genome Research*, 12(12), 1805-1814. doi: 10.1101/gr.631202
- Argyriadou, A., Gelasakis, A. I., Banos, G., & Arsenos, G. (2020). Genetic improvement of indigenous Greek sheep and goat breeds. *Journal of the Hellenic Veterinary Medical Society*, 71(1), 2063.

- Argyriadou, A., Michailidou, S., Vouraki, S., Tsartsianidou, V., Triantafyllidis, A., Gelasakis, A., Banos, G., & Arsenos, G. (2023). A genome-wide association study reveals novel SNP markers associated with resilience traits in tow Mediterranean dairy sheep breeds. *Frontiers in Genetics, 14*, 1294573. doi: 10.3389/fgene.2023.1294573
- Basdagianni, Z., Sinapis, E., & Banos, G. (2019). Evaluation of reference lactation length in Chios dairy sheep. *Animal, 13*(1), 1-7. doi: 10.1017/S1751731118000769
- Ben Jemaa, S., Tolone, M., Sardina, M. T., Di Gerlando, R., Chessari, G., Criscione, A., Persichilli, C., Portolano, B., & Mastrangelo, S. (2023). A genome-wide comparison between selected and unselected Valle del Belice sheep reveals differences in population structure and footprints of recent selection. *Journal of Animal Breeding and Genetics, 140*(5), 558-567. doi: 10.1111/jbg.12779
- Cai, Z., Guldbandsen, B., Lund, M. S., & Sahana, G. (2018). Prioritizing candidate genes post-GWAS using multiple sources of data for mastitis resistance in dairy cattle. *BMC Genomics, 19*(1), 656. doi: 10.1186/s12864-018-5050-x.
- Casu, S., Sechi, S., Salaris, S. L., & Carta, A. (2010). Phenotypic and genetic relationships between udder morphology and udder health in dairy ewes. *Small Ruminant Research, 88*, 77-83. doi: 10.1016/j.smallrumres.2009.12.013
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience, 4*, 7. doi:10.1186/s13742-015-0047-8
- Deng, Y., Li, H., Song, Y., Cen, J., Zhang, Y., Sui, Y., Cui, D., Li, T. C., Xu, Y., Wang, C. C., Chung, P. W. J., & Tang, T. (2022). Whole genome transcriptomic analysis of ovary granulosa cells revealed an anti-apoptosis regulatory gene *DLGAP5* in polycystic ovary syndrome. *Frontiers in Endocrinology, 13*, 781149. doi: 10.3389/fendo.2022.781149
- Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M., & Servin B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics, 193*(3), 929-941. doi: 10.1534/genetics.112.147231
- Gonzalez, M., Villa, R., Villa, C., Gonzalez, V., Montano, M., & Medina, G., (2020). Inspection of real and imputed genotypes revealed 76 SNPs associated to rear udder height in Holstein cattle. *Journal of Advanced Veterinary and Animal Research, 7*, 234-241. doi: 10.5455/javar.2020.g415
- Habimana, R., Ngeno, K., Okeno, T. O., Hirwa, C. D., Keambou Tiambo, C., & Yao, N. K. (2021). Genome-wide association study of growth performance and immune response to Newcastle disease virus of indigenous chicken in Rwanda. *Frontiers in Genetics, 12*, 723980. doi: 10.3389/fgene.2021.723980
- Hu, M., Jiang, H., Lai, W., Shi, L., Yi, W., Sun, H., Chen, C., Yuan, B., Yan, S., & Zhang, J. (2023). Assessing genomic diversity and signatures of selection in Chinese red steppe cattle using high-density SNP array. *Animals, 13*, 1717. doi: 10.3390/ani13101717
- Israel, L., & Mellett, M. (2018). Clinical and genetic heterogeneity of CARD14 mutations in psoriatic skin disease. *Frontiers in Immunology, 9*, 2239. doi: 10.3389/fimmu.2018.02239.
- Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R., San Cristobal, M., Servin, B., McCulloch, R., Whan, V., McEwan, J., & Dalrymple, B. (2012). Genome wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology, 10*(2), e1001258. doi: 10.1371/journal.pbio.1001258
- Kim, H., Song, K. D., Kim, H. J., Park, W., Kim, J., & Lee, T. (2015). Exploring the genetic signature of body size in Yucatan miniature pig. *PLoS ONE, 10*, 4e0121732. doi: 10.1371/journal.pone.0121732.
- Kominakis, A., Hager-Theodorides, A.L., Saridaki, A., Antonakos, G., & Tsiamis, G. (2017). Genome-wide population structure and evolutionary history of the Frizarta dairy sheep. *Animal, 11*(10), 1680-1688. doi: 10.1017/S1751731117000428
- Lázaro, S. F., Tonhati, H., Oliveira, H. R., Silva, A. A., Scalez, D. C. B., Nascimento, A. V., Santos, D. J. A., Stefani, G., Carvalho, I. S., Sandoval, A. F., & Brito, L. F. (2024). Genetic parameters and genome-wide association studies for mozzarella and milk production traits, lactation length, and lactation persistency in Murrah buffaloes. *Journal of Dairy Science, 107*(2), 992-1021. doi: 10.3168/jds.2023-23284
- Leitner, G., Silanikove, N., & Merin, U. (2008). Estimate of milk and curd yield loss of sheep and goats with intramammary infection and its relation to somatic cell count. *Small Ruminant Research, 74*, 221-225. doi: 10.1016/j.smallrumres.2007.02.009
- Li, R., Zhao, Y., Liang, B., Pu, Y., Jiang, L., & Ma, Y. (2023a). Genome-wide signal selection analysis revealing genes potentially related to sheep-milk-production traits. *Animals, 13*, 1654. doi: 10.3390/ani13101654
- Li, Y., Wei, J., Sun, Y., Zhou, W., Ma, X., Guo, J., Zhang, H., & Jin, T. (2023b). *DLGAP5* regulates the proliferation, migration, invasion, and cell cycle of breast cancer cells via the JAK2/STAT3 signaling axis. *International Journal of Molecular Science, 24*(21), 15819. doi: 10.3390/ijms242115819
- Li, M., Pu, L., MacHugh, D. E., Tian, J., Wang, X., Zhao, Q., Shi, L., Gao, H., Yu, Y., Wang, L., & Zhao, F. (2024). Genome-wide association studies of novel resilience traits identify important immune QTL regions

- and candidate genes in Duroc pigs. *Journal of Integrative Agriculture*, 1, 2095-3119. doi: 10.1016/j.jia.2024.04.017
- Liu, S. Q., Xu, Y. J., Chen, Z. T., Li, H., Zhang, Z., Wang, Q. S., & Pan, Y. C. (2024). Genome-wide detection of runs of homozygosity and heterozygosity in Tunchang pigs. *Animal*, 18(8), 101236. doi: 10.1016/j.animal.2024.101236
- May, K., Sames, L., Scheper, C., & König, S. (2022). Genomic loci and genetic parameters for uterine diseases in first-parity Holstein cows and associations with milk production and fertility. *Journal of Dairy Science*, 105(1), 509-524. doi: 10.3168/jds.2021-20685
- Mohammadi, H. (2024). Identification of genomic regions under positive selection that control the type traits in different goat breeds. *Animal Production Research*, 13(2), 43-57. doi: 10.22124/ar.2024.27378.1827 [In Persian]
- Öner, Y., Serrano, M., Sarto, P., Iguácel, L. P., Piquer-Sabanza, M., Estrada, O., Juan, T., & Calvo, J. H. (2021). Genome-wide association studies of somatic cell count in the Assaf breed. *Animals*, 11, 1531. doi: 10.3390/ani11061531
- Patiabadi, Z., Razmkabir, M., EsmailzadehKoshkoiyeh, A., Moradi, M., Rashidi, A., & Mahmoudi, P. (2024). Whole-genome scan for selection signature associated with temperature adaptation in Iranian sheep breeds. *PLoS ONE*, 19(8), e0309023. doi: 10.1371/journal.pone.0309023
- Rezvannejad, E., Asadollahpour Nanaei, H., & Esmailzadeh, A. (2022). Detection of candidate genes affecting milk production traits in sheep using whole-genome sequencing analysis. *Veterinary Medicine and Science*, 8(3), 1197-1204. doi: 10.1002/vms3.731
- Saravanan, K. A., Panigrahi, M., Kumar, H., Parida, S., Bhushan, B., Gaur, G. K., Dutt, T., Mishra, B. P., & Singh, R. K. (2021). Genomic scans for selection signatures revealed candidate genes for adaptation and production traits in a variety of cattle breeds. *Genomics*, 113(3), 955-963. doi: 10.1016/j.ygeno.2021.02.009
- Selionova, M., Trukhachev, V., Aibazov, M., Sermyagin, A., Belous, A., Gladkikh, M., & Zinovieva, N. (2024). Genome-wide association study of milk composition in Karachai Goats. *Animals*, 14, 327. doi: 10.3390/ani14020327
- Shi, M. T., Zhang, Y., & Zhou, G. Q. (2018). The critical roles of TBC proteins in human diseases. *Yi Chuan*, 40, 12-21. doi: 10.16288/j.ycz.17-343
- Song, X., Yao, Z., Zhang, Z., Lyu, S., Chen, N., Qi, X., Liu, X., Ma, W., Wang, W., Lei, C., Jiang, Y., Wang, E., & Huang, Y. (2024). Whole-genome sequencing reveals genomic diversity and selection signatures in Xia'nán cattle. *BMC Genomics*, 25(1), 559. doi: 10.1186/s12864-024-10463-3
- Sutera, A. M., Portolano, B., Di Gerlando, R., Sardina, M. T., Mastrangelo, S., and Tolone, M. (2018). Determination of milk production losses and variations of fat and protein percentages according to different levels of somatic cell count in Valle del Belice dairy sheep. *Small Ruminant Research*, 162, 39-42. doi: 10.1016/j.smallrumres.2018.03.002
- Wang, M., Zhang, S., Zheng, G., Huang, J., Songyang, Z., Zhao, X., & Lin, X. (2018). Gain-of-function mutation of CARD14 leads to Spontaneous Psoriasis-like Skin inflammation through enhanced Keratinocyte response to IL-17A. *Immunity*, 49, 66-79.e5. doi: 10.1016/j.immuni.2018.05.012
- Waineina, R. W., Okeno, T. O., Ilatsia, E. D., & Ngeno, K. (2022). Selection signature analyses revealed genes associated with adaptation, production, and reproduction in selected goat breeds in Kenya. *Frontiers in Genetics*, 13, 858923. doi: 10.3389/fgene.2022.858923
- Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358-1370. doi: 10.1111/j.1558-5646.1984.tb05657.x
- Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, 1, 395-420. doi: 10.1111/j.1558-5646.1965.tb01731.x
- Zhao, J., Mu, Y., Gong, P., Liu, B., Zhang, F., Zhu, L., Shi, C., Lv, X., & Luo, J. (2024). Whole-genome resequencing of native and imported dairy goat identifies genes associated with productivity and immunity. *Frontiers in Veterinary Science*, 11, 1409282. doi: 10.3389/fvets.2024.1409282