

Whole genome scan of selection markers in Holstein, Najdi and Holstein-Najdi crossbreds

MR Zargar^{1*} M Tagi Nasiri Beygi² and MB Zandi³

Received: March 7, 2024

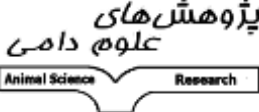

Accepted: November 2, 2024

¹ PhD Student, Department of Animal Science, agriculture university of Ramin Khuzestan

² Professor, Department of Animal Science, Faculty of Animal Science and Food Industry, Ramin Khuzestan University

³ Assistant professor, Department of Animal Science, Zanjan University

*Corresponding author: mohamadzargar@gmail.com

	<p>Journal of Animal Science Research / vol.34 No.4/ 2024/pp 107-120 https://animalscience.tabrizu.ac.ir</p>	
<p>© 2009 Copyright by Faculty of Agriculture, University of Tabriz, Tabriz, Iran This is an open access article under the CC BY NC license (https://creativecommons.org/licenses/by-nc/2.0/) DOI: 10.22034/as.2024.60843.1734</p>		

Introduction: In order to protect native breeds to maintain genetic diversity, it is necessary to keep the animal in its place of residence because the cost of protecting genetic resources is high and it is not possible to maintain all genetic resources in laboratories and research institutions. The aim of this research was to identify selection traits in Holstein, Najdi and their hybrids in the herds of Khuzestan province.

Material and Methods: For this purpose, the genomic information of 209 cattle samples including 21, 128 and 60 samples from Holstein, Najdi breeds and their Holstein-Najdi hybrids were used, which were genotyped using Illumina SNP30K arrays. First, animals with more than 5% missing genotype were excluded from the next steps, and monomorphic genotypes and genotypes with a minimum allelic frequency of less than 5%, as well as loci with more than 5% missing genotype, were excluded. Hardy-Weinberg equilibrium was checked in all loci and loci with P-value less than 6-10 were excluded from the data set. In this study, the signs of selection were investigated using two series of methods, multi-population (F_{ST} , hapFLK) and the other one is single-population (iHS, nSL). To determine the signs of selection in the studied cattle breeds, F_{ST} values for each SNP marker were calculated using the unbiased theta estimator method by Plink and R software. Also, in order to identify the signs of selection in the studied breeds, the FLK Single Marker haplotype expansion method (hapFLK) test was also used. iHS and nSL statistics have been used as intra-population statistics in order to investigate the areas carrying selection signs. To calculate these statistics for each SNP, VCFtools and Selscan version 2.0 were used, respectively (Szpiech 2014).

Results and Discussion: After data editing and imputation of 662,428 SNPs from 329 Najdi, Holstein and their hybrids, and after removing 3,129 SNPs with Hardy-Weinberg 6-10 test and 68,116 SNPs with MAF 0.05, 591,181 SNPs remained for further analysis. In order to identify selection signals at the genome level, the numerical value of fixation index (F_{ST}) was used. Only 0.1% of the regions of the genome were selected as selection marker where all had high values. Finally, three regions on the genome were selected for further analysis that passed the threshold. Three regions were located on autosomal chromosomes. These three regions, where the stabilization index had a higher numerical value of 0.38, were located on chromosomes 16, 21, and 8, respectively. A threshold of 0.1% of the high hapFLK value of the population was used to identify the regions under selection by

the hapFLK method. Finally, the region of 22774764 to 23377643 base pairs of chromosome 14 and the region of 33981177 to 34039961 base pairs of chromosome 26 had a selection mark at the specified threshold, which was used to identify genes in the next step. The iHS statistic was used to identify signs of selection for alleles that are polymorphic in the population, but have not yet reached the fixation stage. iHS less than zero indicates that the respective haplotype carries the derived allele, and iHS greater than zero indicates that the haplotype in question carries the ancestral gene. In this research, the markers with the highest amount of iHS on chromosomes 3, 7 and 10 were identified as selection markers. By using the nSL statistic which was calculated by Selscan software in the Linux environment and drawing the Manhattan plot diagram of this statistic by R software, it was determined that in the Holstein and Holstein*Najdi hybrids population, regions on chromosome 3 and in the Najdi population regions on chromosome 4 and 15 have been selected. Some of the areas identified with these statistics are directly and indirectly related to growth traits, reproduction, immune system and other traits related to them.

Conclusion: In this research, four statistics, F_{ST} , hapFLK, iHS and nSL, were used to identify the signs of selection in Holstein, Najdi cattle populations and their hybrids. The results of F_{ST} and hapFLK statistics, which examine the populations in a multi-population way, showed that the signs of selection detected by these two statistics were on chromosomes 8, 14, 16, 21 and 26, and chromosome 16 was common in each of the two strains. Among the areas identified with these statistics, they are directly or indirectly related to the characteristics of milk production and reproduction in cows. The results of the iHS and nSL statistics, which examine the populations as a single population, showed that the signs of selection detected by these two statistics were on chromosomes 3, 4, 7, 10 and 15, which the chromosome No. 3 was common in every two races. Some of the areas identified with these statistics are directly and indirectly related to growth traits, reproduction, immune system and other traits related to them. Considering that the statistics used are different in terms of the implementation method and the parameters used (the first two are multi-population based on allele frequency and the second two are single-population based on linkage disequilibrium), consequently their results did not match. But from the total of 59 genes identified by the four methods used, two genes, IQCA1, ACKR3, were jointly identified by iHS and nSL statistics. In addition, it can be concluded that the selection during the past years has been generally focused on traits that have been effective in increasing the production efficiency in addition to increasing the resistance of livestock to the environmental conditions of the place of residence.

Keywords: Holstein, Najdi, Selection markers, hapFLK and iHS

پویش کل ژنوم نشانه‌های انتخاب در گاو هلشتاین، نجدی و آمیخته‌های هلشتاین-نجدی

محمدرضا زرگر^{۱*} محمدتقی نصیری بیگی^۲ محمدباقر زندی^۳

تاریخ دریافت: ۱۴۰۲/۱۲/۱۷ تاریخ پذیرش: ۱۴۰۳/۸/۱۲
^۱ دانشجوی دکتری، گروه علوم دامی دانشکده علوم دامی و صنایع غذایی، دانشگاه کشاورزی و منابع طبیعی رامین خوزستان
^۲ استاد، گروه علوم دامی دانشگاه کشاورزی و منابع طبیعی رامین خوزستان
^۳ استادیار، گروه علوم دامی دانشگاه زنجان
*مسئول مکاتبه: mohamadzargar@gmail.com

چکیده:

زمینه مطالعاتی: برای محافظت از نژادهای بومی به منظور حفظ تنوع ژنتیکی، نگهداری حیوان در محل زندگی خود است چرا که هزینه محافظت از ذخایر ژنتیکی بالا بوده و امکان حفظ همه ذخایر ژنتیکی در آزمایشگاه‌ها و موسسات تحقیقاتی ممکن نیست. هدف: از این تحقیق شناسایی نشانه‌های انتخاب در گاو هلشتاین، نجدی و آمیخته‌های آن‌ها در گله‌های استان خوزستان بود. روش کار: به این منظور از اطلاعات ژنومی ۲۰۹ نمونه گاو شامل ۲۱، ۱۲۸ و ۶۰ نمونه به ترتیب از نژادهای هلشتاین، نجدی و آمیخته‌های هلشتاین-نجدی آن‌ها استفاده شد که با بکارگیری آرایه‌های Genotyping Array 30K تعیین ژنوتیپ شدند. در این مطالعه نشانه‌های انتخاب با استفاده از دو سری از روش‌ها مورد بررسی قرار گرفت که در دو روش (FST و Hapflk) به صورت چند جمعیتی و دوروش دیگر (iHS, nSL) به صورت تک جمعیتی عمل می‌کند. نتایج: در دو روش اول که چند جمعیتی بودند نقاطی بر روی کروموزوم‌های ۸، ۱۴، ۱۶ و ۲۴ شناسایی شدند که کروموزوم ۱۶ در هر دو روش مشترک بود. برخی از مناطق شناسایی شده با آماره‌های FST و Hapflk به صورت مستقیم و یا غیر مستقیم با صفات تولید شیره و تولید مثل در گاو در ارتباط هستند. در دو روش دیگر که تک جمعیتی بودند نقاطی بر روی کروموزوم‌های شماره ۳، ۴، ۷، ۱۰ و ۱۵ بودند که کروموزوم شماره ۳ در هر دو روش مشترک بود. برخی از مناطق شناسایی شده با آماره‌های nSL و iHS به طور مستقیم و غیر مستقیم با صفات رشد، تولید مثل، سیستم ایمنی و سایر صفات مرتبط با آن‌ها ارتباط دارند. ولی از مجموع ۵۹ ژن شناسایی شده توسط چهار روش مورد استفاده دو ژن ACKR3, IQCA1، به طور مشترک توسط آماره iHS و nSL شناسایی شدند. نتیجه گیری نهایی: ضمناً می‌توان نتیجه گرفت که انتخاب طی سالیان گذشته به طور کلی بر روی صفاتی متمرکز شده است که ضمن افزایش مقاومت دام‌ها به شرایط محیطی محل زندگی بر روی افزایش راندمان تولید نیز موثر بوده است.

واژه‌های کلیدی: هلشتاین، نجدی، نشانه‌های انتخاب، hapFLK و his

مقدمه

سازگاری با محیط تکامل یافته‌اند (فروغ عامری و همکاران ۲۰۱۵) بهترین راه برای محافظت از نژادهای بومی به منظور حفظ تنوع ژنتیکی، نگهداری حیوان در محل زندگی خود است چرا که هزینه محافظت از ذخایر ژنتیکی بالا بوده و امکان حفظ همه ذخایر ژنتیکی در

به دلیل تغییرات وسیع آب و هوایی در سرتاسر کشور ایران، نژادهای گاو در طول زمان با این شرایط آب و هوایی خود را سازگار کرده‌اند. لذا نژادهای بومی مختلفی از شمال تا جنوب و از شرق تا غرب کشور برای

یک راه تفسیر افزایش شاخص تثبیت، کاهش هتروزیگوسیتی در میان جمعیت‌ها نسبت به آنچه که در حالت طبیعی مورد انتظار است، می‌باشد. تحت شرایط خنثی، میزان شاخص تثبیت به وسیله رانش ژنتیکی و مهاجرت تعیین می‌شود، اما سازگاری منطقه‌ای می‌تواند میزان تمایز جمعیت‌ها را در لوکوس‌های خاص شتاب دهد و منجر به مقادیر بالاتری از شاخص تثبیت شود (ثابتی و همکاران ۲۰۰۲).

همچنین استفاده از آزمون هاپلوتایپ با دامنه گسترده^۳ برای شناسایی نشانه‌های انتخاب پیشنهاد شده است. پیشرفت‌های اخیر در زمینه ژنتیک مولکولی در کنار توسعه روش‌های آماری و محاسباتی و طراحی تراشه‌های تراکم بالا، این امکان را فراهم آورده است که بتوان کل ژنوم را به طور همزمان مورد بررسی قرار داد. همچنین بیشتر نژادهای تثبیت شده گاو را می‌توان با استفاده از نشانگرهای ژنتیکی تشخیص داد (کوهن و همکاران ۲۰۱۱).

مقایسه بین گروه‌های نژادی این پتانسیل را دارد که مناطق ژنومی را که موثر بر صفات پیچیده هستند بدون نیاز به تجهیزات پیچیده و جمع آوری سوابق فنوتیپی، مورد مطالعه قرار داده و به شناسایی ژن‌های کاندید و درک مکانیسم‌های بیولوژیکی کنترل صفات کمک کند (استلا و همکاران ۲۰۱۰). مطالعات مختلفی در مورد شناسایی نشانه‌های انتخاب در گونه‌های مختلف مانند انسان (ثابتی و همکاران ۲۰۰۶)، گاو (بویتارد و همکاران ۲۰۱۶)، گوسفند (کیجاس و همکاران ۲۰۱۴) و بز (بریتو و همکاران ۲۰۱۷) انجام شده است.

اگر جمعیت‌های نمونه دارای اندازه موثر نابرابر یا دارای ساختار سلسله مراتبی باشند، کاوش ژنوم بر اساس مقادیر F_{ST} معمول می‌تواند نتیجه اریب به همراه داشته باشد. برای تصحیح اثر رانش و کوواریانی، هنگام تشخیص مکان‌های تحت انتخاب، آمار FLK معرفی شد

آزمایشگاه‌ها و موسسات تحقیقاتی ممکن نیست (هلسینگر و وایر ۲۰۰۹). بدون شک مکان‌هایی از ژنوم که با صفات ایمنی و سازگاری مرتبط بوده‌اند، با استفاده از انتخاب طبیعی و تاحدی به وسیله انتخاب مصنوعی تحت تاثیر انتخاب قرار گرفته‌اند (فروغ عامری و همکاران ۲۰۱۵). با بررسی ظاهری گاوهای نواحی مختلف کشور هم می‌توان ویژگی‌هایی را یافت که نشان از سازگاری این دام با محیط طبیعی دارد (فروغ عامری و همکاران ۲۰۱۵). به عنوان مثال، در نژاد سیستانی به دلیل کمبود منابع خوراک دام، کوهان گاو به عنوان محلی برای ذخیره انرژی دیده می‌شود. همچنین نژادهایی نظیر نجدی که در مناطق گرم دیده می‌شوند، دارای اندامی کشیده و پوستی با غدد عرق فراوان بوده، لذا تبخیر سطح پوست آن‌ها نسبت به نژادهایی نظیر سرابی بیشتر است (فروغ عامری و همکاران ۲۰۱۵).

با تولید و توسعه فن‌آوری‌های توالی‌یابی DNA، این امکان به وجود آمده است که با دقت و صحت بالاتری پژوهشگران بتوانند اطلاعات ارزشمندی در ارتباط با منشاء و ساختار ژنوم گونه‌های مختلف بدست بیاورند. در صورتی که انتخاب طبیعی یا مصنوعی برای صفتی خاص به نفع ناحیه‌ای از ژنوم صورت گیرد، این امکان نیز وجود دارد که در طول زمان سایر چندشکلی‌ها فراوانی‌شان کاهش یافته یا ناپدید شده باشند (فروغ عامری و همکاران ۲۰۱۵). در نتیجه یک هاپلوتایپ خاص که حاوی چندین ژن است، ممکن است تنها هاپلوتایپ یا مهمترین هاپلوتایپ موجود در طبیعت باشد که به آن نشانه انتخاب^۱ یا رانش ناشی از انتخاب^۲ می‌گویند (اسمیت و هایج ۱۹۷۴ و دوو و همکاران ۲۰۱۰).

برای بررسی سطح تمایز بین جمعیت‌ها و در نتیجه شناسایی نشانه‌های انتخاب، اولین بار استفاده از شاخص تثبیت (F_{ST}) پیشنهاد شد (ویر و کوکرهام ۱۹۸۴).

³ Long-Range Haplotype

¹ Selection signatures

² Selective sweep

(بونهام و همکاران ۲۰۱۰)، و چند سال بعد روش HapFLK معرفی گردید (فاریلو و همکاران ۲۰۱۳)، که هاپلوتیپ‌های تحت انتخاب را بر اساس آزمون FLK شناسایی کرده و اریب ناشی از ساختار سلسله مراتبی جمعیت‌ها را برطرف می‌کند. هدف از تحقیق حاضر شناسایی نشانه‌های ژنومی انتخاب در مقایسه گاو هلشتاین، نجدی و آمیخته‌های این دو نژاد با استفاده از روش‌های مبتنی بر فراوانی اللی (HapFLK، F_{ST}) و مبتنی بر عدم تعادل لینکاژی (iHS و nSL) بود.

مواد و روش‌ها

جمع‌آوری نمونه و استخراج DNA

در این مطالعه نمونه‌گیری از ایستگاه پشتیبانی گاونجدی استان خوزستان و گله های مردمی واقع در شهرستان‌های دشت‌آزادگان، بندرماهشهر و هندیجان به تعداد ۳۲۹ راس شامل ۱۲۸ راس گاو نجدی، ۱۴۱ راس گاو هلشتاین و ۶۰ راس آمیخته از سنین و جنس‌های مختلف بین دو نژاد فوق استفاده شده است. براساس

داده‌های شجره‌ای و با پرسش از صاحبان دام‌ها نمونه‌هایی مورد انتخاب قرار گرفتند که دارای رابطه خویشاوندی نبوده و یا روابط خویشاوندی آن‌ها حداقل باشد. در مجموع ۳۲۹ راس گاو از گله‌های ایستگاه پشتیبانی گاو نجدی و اقماری گاو نجدی خوزستان خونگیری شدند (جدول ۱). پس از مراحل استخراج DNA که برای کلیه دام‌ها با استفاده روش نمکی در آزمایشگاه انجام شد، تعیین کمیت و کیفیت نمونه‌ها از دو روش ژل آگارز برای کمیت و دستگاه نانودراپ برای کیفیت نمونه‌ها استفاده شد. نمونه‌های DNA بعد از استخراج و گذراندن مرحله آماده‌سازی با استفاده از آرایه‌های Genotyping Array 30K در شرکت ایلومینا و با استفاده از پروتکل استاندارد ایلومینا ژنوتیپ شدند، که در مجموع ۳۰۱۰۶ نشانگر SNP خوانش شد. سپس نشانگرهای SNP بدست آمده از نمونه برای داشتن نتایج صحت بالا با روش‌های مختلف ایمپوت شدند که با انجام فرایند ایمپوتیشن از تعداد ۳۰ هزار نشانگر SNP به تعداد ۶۶۲۴۲۶ نشانگر SNP بدست آمد.

Table 1: Location and size of samples used

Location	Number	Breed
Shoshtar	141	Holstien
Shoshtar	60	Cross
Shoshtar, Hendijan, Mahshahr, Dastazadegan	128	Najdi

ویرایش داده‌ها

اولین گام در انتخاب ژنومی و همچنین مطالعات گسترده ژنوم، ویرایش داده‌های ژنوتیپی است. ویرایش داده‌های ژنوتیپی اجازه می‌دهد داده‌های غیر اطلاع‌دهنده و خطاهای ژنوتایپینگ مورد استفاده برای خواندن توالی DNA حذف شوند. این مرحله بسیار مهم است زیرا با حذف SNP‌هایی که در افزایش صحت برآوردها سهم نیستند، هزینه‌های محاسباتی را کاهش می‌دهد و ثبات

برآورد اثرات SNP‌های باقیمانده را افزایش می‌دهد (ون‌رادن و همکاران ۲۰۰۸). ویرایش داده‌ها با استفاده از نرم افزار Plink 1.07 انجام شد (پورسل و همکاران ۲۰۰۷). ابتدا حیوانات با بیش از ۵٪ ژنوتیپ از دست رفته از مراحل بعدی کنار گذاشته شد و ژنوتیپ‌های مونومورف و ژنوتیپ‌های دارای حداقل فراوانی آلی کمتر از ۵٪ و نیز جایگاه‌های با بیش از ۵٪ ژنوتیپ از دست رفته^۳

³ SNP Call Rate

¹ Animal Call Rate

² Minor Allele Frequency

از آماره‌های iHS و nSL به عنوان آماره‌های درون جمعیتی استفاده شده است. برای محاسبه این آماره‌ها برای هر SNP، به ترتیب با استفاده از برنامه VCFtools و نسخه Selscan 2.0 (شیپچ و هرماندز ۲۰۱۴) استفاده شد. در مطالعه کنونی، نسخه ۵/۲ برنامه Beagle (برونینگ و همکاران ۲۰۲۱) برای فازبندی هاپلوتایپی با استفاده از گزینه‌های burnin=5 و iteration=20 اجرا شد. نرمال‌سازی آماره‌های محاسبه شده برای هر SNP با استفاده از برنامه norm موجود در بسته Selscan صورت گرفت.

نتایج و بحث

ویرایش داده‌ها

پس از ویرایش و جانمایی داده‌ها از تعداد ۶۶۲۴۲۸ SNP از ۳۲۹ راس گاو نجدی، هلاشتاین و آمیخته‌های حاصل از آن‌ها و پس از حذف ۳۱۲۹ SNP با تست هاردی واینبرگ 10^{-6} و ۶۸۱۱۶ SNP با MAF 0.05 در نهایت ۵۹۱۱۸۱ SNP برای آنالیزهای بعدی باقی ماند. به منظور شناسایی سیگنال‌های انتخاب در سطح ژنوم از ارزش عددی شاخص تثبیت (F_{ST}) استفاده شد. تنها ۰/۱ درصد مناطقی از ژنوم که در آن تمام نشانگرها دارای ارزش بالایی بودند، به عنوان نشانه‌های انتخاب شناسایی و تعیین شدند (شکل ۱). حد آستانه در نظر گرفته شده سخت‌گیرانه بود و فقط مناطقی که بیشترین تمایز جمعیتی را در بین سه جمعیت نشان می‌دهند، انتخاب شدند. این حد آستانه در پژوهش‌های مختلف از ۰/۱ (مرادی و همکاران ۲۰۱۲) تا ۲/۵ درصد (قنبری و همکاران ۲۰۱۲) متغیر بوده است. در نهایت سه ناحیه بر روی ژنوم به منظور بررسی‌های بعدی انتخاب شدند که حد آستانه را رد کرده بودند.

حذف شدند. تعادل هاردی واینبرگ^۱ در کلیه جایگاه‌ها مورد بررسی قرار گرفت و جایگاه‌های با P-value کمتر از 10^{-6} از مجموعه داده‌های کنار گذاشته شدند.

تجزیه‌های آماری

برای تعیین نشانه‌های انتخاب در نژادهای گاو مورد مطالعه ارزش‌های F_{ST} برای هر نشانگر SNP به روش برآوردگر ناریب تتا (ویر و کوکرهام ۱۹۸۴) توسط نرم افزار Plink و R محاسبه شد. F_{ST} با استفاده از رابطه ۱ برآورد می‌شود. (مک ایچرن و همکاران ۲۰۰۹):

$$F_{ST} = \frac{H_T - H_S}{H_S} \quad \text{رابطه ۱}$$

در رابطه ۱ H_T و H_S به ترتیب هتروزیگوسیتی مورد انتظار در کل جمعیت و هتروزیگوسیتی داخل زیرجمعیت‌ها است. F_{ST} یک روش ارزیابی تفرق جمعیتی بر پایه داده‌های چند شکلی ژنتیکی است. ارزش‌های به دست آمده از تجزیه و تحلیل این آماره می‌توانند بین ۰ (بدون تفاوت) تا ۱ (تفاوت کامل، که هر جمعیت برای آلل متفاوتی تثبیت شده‌اند) متغیر باشند. در روش تتا، تفاوت اندازه جمعیت و خطای نمونه‌گیری در نظر گرفته می‌شود. به این دلیل روش از F_{ST} دارای دقت بیشتری است (بونهام و همکاران ۲۰۱۰). همچنین به منظور شناسایی نشانه‌های انتخاب در نژادهای مورد مطالعه از روش بسط هاپلوتایپی روش FLK Single Marker (بونهام و همکاران ۲۰۱۰) با عنوان آزمون hapFLK (فارلیو و همکاران ۲۰۱۳) نیز استفاده شد.

این آماره به وسیله اسکریپت پایتون که در صفحه وب hapFLK به آدرس پاورقی ارائه شده، محاسبه شد. استفاده از این روش تا حد زیادی قدرت تشخیص نشانه‌های انتخاب را با استفاده از داده‌های حاصل از تراشه با تراکم بالا افزایش می‌دهد (فارلیو و همکاران ۲۰۱۳). به منظور بررسی مناطق حامل نشانه‌های انتخاب،

²<https://forge-dga.jouy.inra.fr/projects/hapflk>

¹ Hardy-Weinberg tests

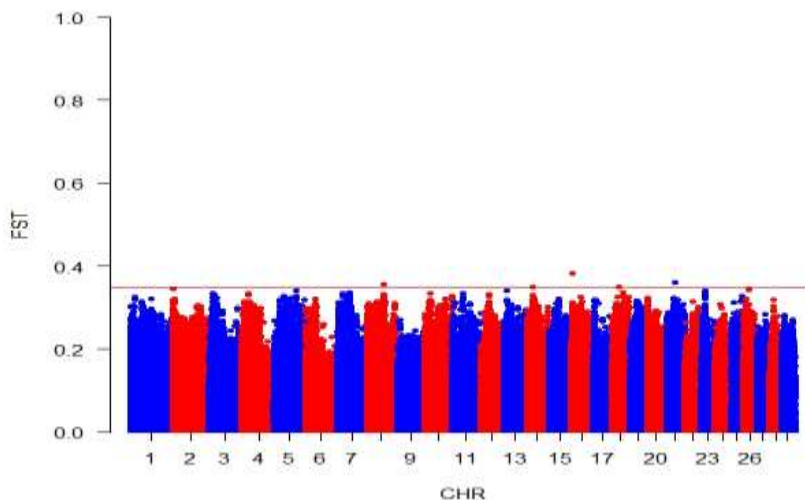


Figure 1: Distribution of stabilization index (F_{ST}) of the studied populations of Holstein, Najdi and their Holstein*Najdi hybrids

نتایج حاصل از بررسی F_{ST} در پایگاه داده‌های ENSEMBL در جدول ۲ نشان داده شده‌اند.

سه ناحیه بر روی کروموزوم‌های اتوزومی قرار داشت. این سه منطقه که در آن‌ها شاخص تثبیت ارزش عددی بالاتر ۰/۳۸ را داشت، به ترتیب بر روی کروموزوم‌های ۱۶، ۲۱ و ۸ قرار داشتند (شکل ۱).

Table 2 Regions under positive selection and genes identified by F_{ST} in these regions based on the ENSEMBL database

Reference	QTL	Gene	position on the genome	Chromosome
Zimin et al. 2009 and) (Leo et al. 2017	Reproduction, milk production, stress response, intracellular regulation and immune system	SOX13, ETNK2, REN, KISS1, GOLT1A, PLEKHA6, PPP1R15B, PIK3C2B	1867583-2156845	16
Smith et al. 2001 and) Glykina et al. 2019 and (Tripathi et al. 2012	Stress response, sex hormone regulation, gene expression regulation	CYP11A1, STRA6, ISLR, PML, STOML1, LOXL1, GZMB, STXBP6	34328574 - 35306047	21
(Zimin et al. 2009	Cell development, reproduction, fertility, growth and development, milk production and RNA processing and DNA repair	ZCCHC7, GRHPR, ZBTB5, U6, POLR1E, FBXO10, TOMM5, FRMPD1, TRMT10B, EXOSC3, DCAF10, SLC25A51,	61209453 - 62142447	8

ZBTB5, U6, POLR1E, FBXO10, TOMM5, FRMPD1, TRMT10B, EXOSC3, DCAF10, SLC25A51 در منطقه مورد انتخاب شناسایی شدند، که

همانطور که در جدول ۲ مشخص است بر روی کروموزوم شماره ۶ ژن‌های ZCCHC7, GRHPR،

ناحیه ۷۱،۱ مگاجفت بازی کروموزوم ۱ گزارش شده است (کمپر و همکاران ۲۰۱۴). با استفاده از روش شاخص تثبیت، نشانه‌هایی از انتخاب در نواحی میانی کروموزوم ۲ پیدا شده است (برندس و همکاران ۲۰۰۹ و قنبری و همکاران ۲۰۱۱). این نواحی با بازدهی خوراک و چربی بین ماهیچه‌ای در گاوهای گوشتی مرتبط بوده‌اند (برندس و همکاران ۲۰۰۹). ژن میواستاتین نیز بر روی کروموزوم شماره ۲ در ناحیه ۶/۲۲ مگاجفت باز یافت شده است و مرتبط با کمیت و کیفیت گوشت می‌باشد (اسماعیل زاده و همکاران ۲۰۰۸). در این مطالعه نیز نواحی پیشنهادی بر روی آن‌ها مشاهده شد، اما به دلیل اینکه زیر حد آستانه بودند، مورد بررسی قرار نگرفتند.

برای شناسایی مناطق تحت انتخاب با روش hapFLK از آستانه ۰/۱ درصد ارزش hapFLK بالای جمعیت استفاده شد. در نهایت ناحیه ۲۲۷۷۴۷۶ تا ۲۳۳۷۷۶۴ جفت باز کروموزوم ۱۴ و ناحیه ۳۳۹۸۱۱۷ تا ۳۴۰۳۹۹۶۱ جفت باز کروموزوم ۲۶ در آستانه مشخص شده دارای نشانه انتخاب بود، که این ناحیه برای شناسایی ژن‌ها در مرحله بعدی استفاده شد (شکل ۲).

ناحیه برگزیده بر روی کروموزوم ۱۴ در پایگاه اطلاعاتی ENSEMBL Biomart بررسی شده و ژن‌های کد کننده پروتئین موجود در این ناحیه شناسایی شدند. این ناحیه شامل ژن‌های XKR4, TMEM68, TGS1, LYN, CHCHD7, PLAG1, MOS, RPS20 بود، که بر روی صفات تنظیم بیان ژن، پاسخ به ایمنی، تکامل سلول، تولید شیر و چربی بدن، سم زدایی و آنتی‌اکسیدان و مکانیسم داخل سلولی موثر است (زیمین و همکاران ۲۰۰۹ و باستین و همکاران ۲۰۱۴ و هویو و همکاران ۲۰۲۰). همچنین بر روی کروموزوم ۲۶ نیز یک ژن به نام HAP2 شناسایی شد که بر روی تمایز سلولی و بیان ژن موثر است (اسمیت و همکاران ۲۰۰۱).

بر روی صفات تکامل سلول، تولید مثل، باروری، رشد و نمو، تولیدشیر موثر هستند (زیمین و همکاران ۲۰۰۹). بر روی کروموزوم شماره ۱۶ هم ژن‌های SOX13, ETNK2, REN, KISS1, GOLT1A, PLEKHA6, PPP1R15B, PIK3C2B در منطقه مورد انتخاب شناسایی شدند که بر روی صفات تولید مثلی، تولید شیر، پاسخ به استرس، تنظیمات درون سلولی و سیستم ایمنی موثر هستند (زیمین و همکاران ۲۰۰۹ و لیو و همکاران ۲۰۱۷). بر روی کروموزوم شماره ۲۱ نیز ژن‌های CYP11A1, STRA6, ISLR, PML, STOML1, LOXL1, GZMB, STXBP6 در منطقه مورد انتخاب شناسایی شدند که بر روی صفات پاسخ به استرس، تنظیم هورمون جنسی، تنظیم بیان ژن موثر هستند (اسمیت و همکاران ۲۰۰۱ و گلیاکینا و همکاران ۲۰۱۹ و تریپاتی و همکاران ۲۰۱۲). ژن‌هایی مانند DGAT1 که در کنترل صفات تولیدی نقش موثری دارند، بر روی کروموزوم ۱۴ واقع شده‌اند، اما در این پژوهش نشانه‌ای از انتخاب که حد آستانه را رد کند، بر روی این کروموزوم مشاهده نشد. یک دلیل احتمالی می‌تواند این باشد که در نژادهای بومی ایران همانند نژادهای اصیل انتخاب شدیدی برای صفات تولیدی نظیر شیر و چربی انجام نشده است و لذا فراوانی آلل‌های این ژن در حد متوسط باقی مانده است. مخبر و همکاران (۲۰۱۵) با مطالعه پویش نشانه‌های انتخاب بر روی ژنوم ۱۲۱ راس گاومیش خوزستانی و ۲۷ راس گاومیش مازندرانی به وسیله آرایه ژنومیکی 90K، در مجموع ۲۳ مکان ژنومی که مقدار برآوردگر ناریب F_{ST} آن‌ها بالاتر از ۰/۱ درصد حد بالای توزیع تجربی F_{ST} به عنوان نشانه‌های انتخاب گزارش کردند. بعد از انطباق ژنومی این نواحی، با ژنوم گاو، ۶۴ ژن و ۲۷ QTL پیدا شد که عموماً در مسیر بیولوژیکی مرتبط با اهلی شدن حیوانات و تولید شیر بودند (مخبر و همکاران ۲۰۱۵). ژن مرتبط با شاخ‌داری یا بی‌شاخی، در

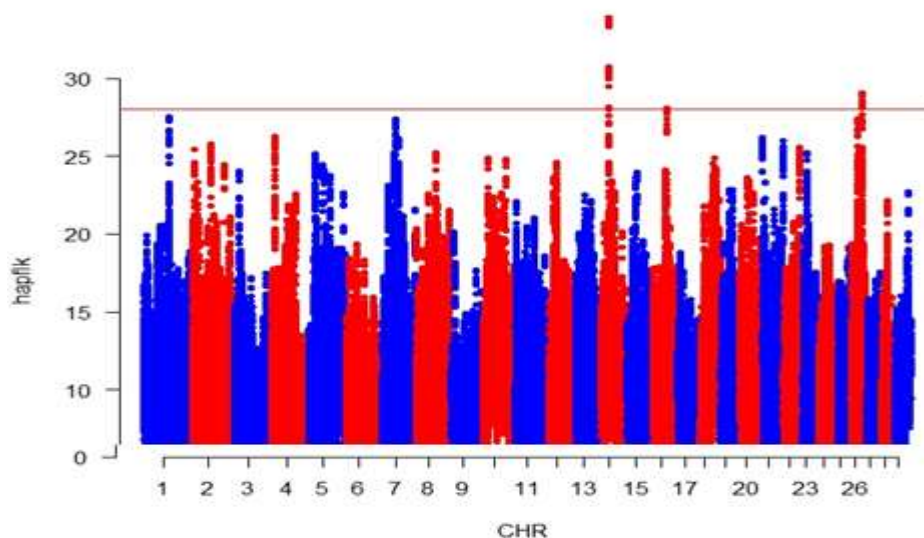


Figure 2: Distribution of hapFLK values of the studied populations of Holstein, Najdi and their Holstein*Najdi hybrids

نتایج حاصل از بررسی hapFLK در پایگاه داده‌های ENSEMBL در جدول شماره ۳ نشان داده شده‌اند.

Table 3 Regions under positive selection and genes detected by hapFLK in these regions based on the ENSEMBL database

reference	QTL	Gene	position on the genome	Chromosome
Zimin et al. 2009) and Bastin et al. 2014 and Hugh et (al. 2020	Regulation of gene expression, immune response, cell development, milk and body fat production, detoxification and antioxidant and intracellular mechanism	XKR4, TMEM68, TGS1, LYN, RPS20, MOS, PLAG1, CHCHD7	22774764-23377643	14
(Zimin et al. 2009)	Reproduction, immune response, general health, milk production and stress response	RERE, SLC45A1, ERFFI1, PARK7, TNFRSF9, UTS2, VAMP3, CAMTA1	44759576-45759576	16
(Smith et al. 2001	Gene expression and cell differentiation	HABP2		26

بر روی کروموزوم شماره ۱۶ نیز ژن‌های RERE, SLC45A1, ERFFI1, PARK7, TNFRSF9, UTS2, VAMP3, CAMTA1 شناسایی شدند که بر روی صفات تولید مثلی، تولید شیر پاسخ به ایمنی و استرس نقش دارند (زیمین و همکاران ۲۰۰۹).

همچنین بر روی کروموزوم شماره ۲۶ ژن HABP2 شناسایی شده که بر روی بیان ژن و تمایز سلولی موثر است (زیمین و همکاران ۲۰۰۹). در این پژوهش رتبه بندی

همانطور که در جدول فوق مشخص است بر روی کروموزوم شماره ۱۴ ژن‌های XKR4, TMEM68, TGS1, LYN, RPS20, MOS, PLAG1, CHCHD7 شناسایی شدند، که بر روی صفات تنظیم بیان ژن، پاسخ به ایمنی، تکامل سلول، تولید شیر و چربی بدن، سم زدایی و آنتی‌اکسیدان و مکانیسم داخل سلولی تاثیر دارند (زیمین و همکاران ۲۰۰۹ و باستین و همکاران ۲۰۱۴ و هویو و همکاران ۲۰۲۰).

با استفاده از آماره iHS در جمعیت هلشتاین و آمیخته بر روی کروموزوم شماره ۱۰ و در جمعیت نجدی بر روی کروموزوم‌های شماره ۳، ۷ و ۱۰ مناطقی تحت انتخاب قرار گرفته‌اند.

هاپلوتیپ یک پارچه برای کلیه نشانگرها با آماره‌های iHS و nSL در سه جمعیت هلشتاین، آمیخته و نجدی به طور جداگانه محاسبه شد. سپس داده‌های حاصل که ارزش آن‌ها درصدک بالای ۹۹٫۹۹٪ قرار داشتند به عنوان مناطق کاندیدای انتخاب شناسایی شدند (شکل ۴).

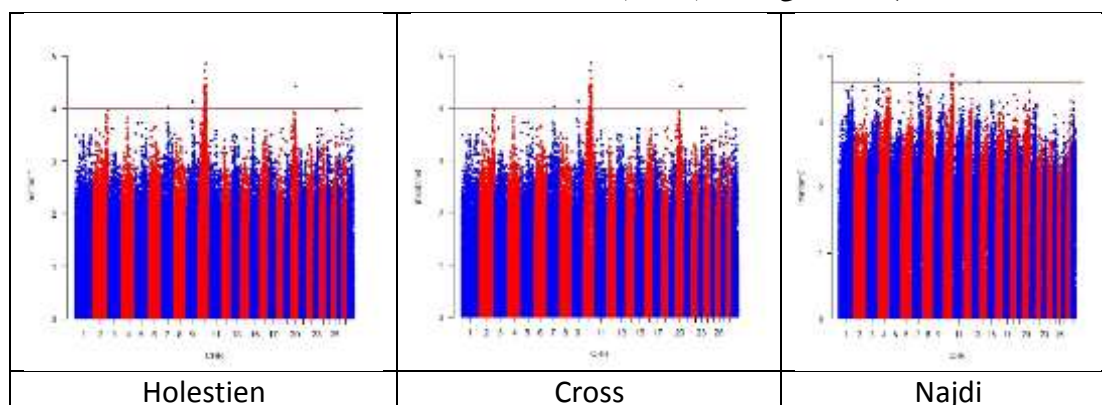


Figure 3: Manhattan plot of iHS statistics for Holstein, Najdi and Holstein*Najdi hybrids populations

اجدایی است (وویت و همکاران ۲۰۰۶). در این تحقیق نشانگرهایی که دارای بیشترین مقدار iHS بر روی کروموزوم‌های ۳، ۷ و ۱۰ بودند به عنوان نشانه‌های انتخاب شناسایی شدند. نتایج حاصل از بررسی iHS در پایگاه داده‌های ENSEMBL در جدول ۴ نشان داده شده‌اند.

آماره iHS برای شناسایی نشانه‌های انتخاب برای آللهایی که در جمعیت چند شکل بوده، ولی هنوز به مرحله تثبیت نرسیده‌اند استفاده شد (هاسون و همکاران ۱۹۹۴ و ثابتی و همکاران ۲۰۰۲ و تانگ و همکاران ۲۰۰۷). iHSهای کمتر از صفر مبین این است که هاپلوتیپ مربوطه حامل آلل مشتق شده است و iHS بزرگتر از صفر نشان می‌دهد که هاپلوتیپ مورد نظر حامل ژن

Table 4 Regions under positive selection and genes identified by iHS in these regions based on the ENSEMBL database

reference	QTL	Gene	position on the genome	Chromosome
(Eskamoto et al. 2020) and Teo et al. 2003 (and Zimin et al. 2009	Cell differentiation, immune system, growth and intracellular balance	FBN1, DUT, SLC12A1, CTXN2, MYEF2, SLC24A5, SEMA6D	61877451- 62877451	10
(Zimin et al. 2009)	Milk production, immune system, gene expression	AGAP1, GBX2, ASB18, IQCA1, ACKR3	115327907 - 116327907	3
(Zimin et al. 2009)	Immune system, neurotransmission	FAM114A2, MFAP3, GRIA1	64398601- 65398601	7

(Zimin et al. 2009)	Muscle contraction, stress regulation	KCNH5, RHOJ, GPHB5, PPP2R5E, WDR89	74869685-75869685	10
---------------------	---------------------------------------	------------------------------------	-------------------	----

تأثیر دارند. ضمناً بر روی کروموزوم شماره ۱۰ این جمعیت ژنهای *KCNH5, RHOJ, GPHB5, PPP2R5E, WDR89* که بر روی صفات انقباض عضلات، تنظیم استرس تأثیر دارند (زیمین و همکاران ۲۰۰۹). با استفاده از آماره nSL که توسط نرم افزار Selscan در محیط لینوکس محاسبه شد و ترسیم نمودار منتهن پلات این آماره توسط نرم افزار R مشخص شد که در جمعیت هلشتاین و آمیخته بر روی کروموزوم ۳ و در جمعیت نجدی بر روی کروموزوم ۴ و ۱۵ مناطقی تحت انتخاب قرار گرفته‌اند (شکل ۴). نتایج حاصل از بررسی آماره nSL در پایگاه داده‌های ENSEMBL در جدول ۵ نشان داده شده است.

پس از بررسی در سایت ENSEMBL همانطور که در جدول فوق مشخص است بر روی کروموزوم شماره ۱۰ در جمعیت هلشتاین و آمیخته ژنهای *FBN1, DUT, SLC24A5, SEMA6D, SLC12A1, CTXN2, MYEF2*، شناسایی شدند که بر روی صفات تمایز سلولی، سیستم ایمنی، رشد و نمو و تعادل درون سلولی تأثیر دارند (تو و بوچانان ۲۰۰۳ و زیمین و همکاران ۲۰۰۹). در جمعیت نجدی بر روی کروموزوم شماره ۳ ژنهای *AGAP1, GBX2, ASB18, IQCA1, ACKR3* شناسایی شدند که بر روی صفات تولید شیر، سیستم ایمنی، بیان ژن تأثیر دارند. همچنین بر روی کروموزوم شماره ۷ این جمعیت ژنهای *FAM114A2, MFAP3, GRIA1* شناسایی شدند که بر روی صفات سیستم ایمنی، انتقال عصبی

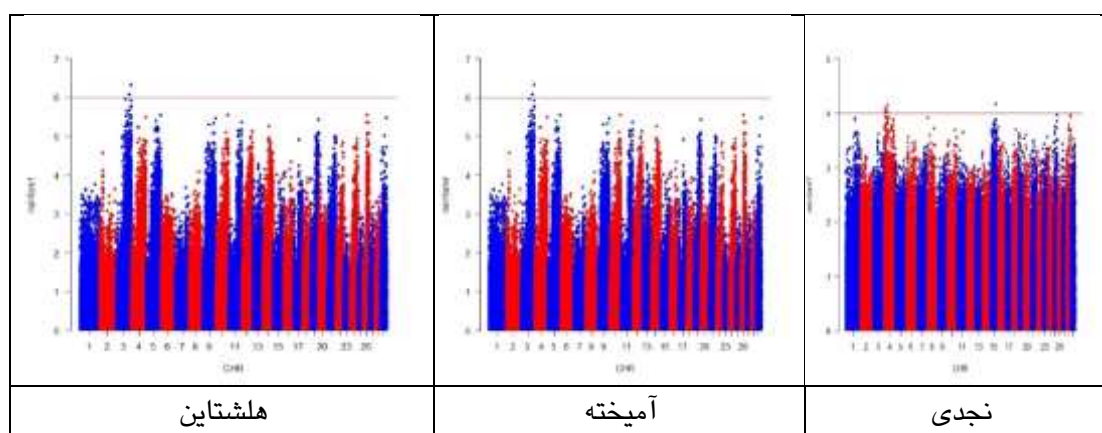


Figure 4: Manhattan plot of nSL statistics for Holstein, Najdi and Holstein*Najdi hybrids populations.

Table 5 Regions under positive selection and genes identified by nSL in these regions based on the ENSEMBL database

reference	QTL	Gene	position on the genome	Chromosome
(Smest et al. 2001)	stress response, immune system, transcriptional regulation	IQCA1, ACKR3, COPS8, COL6A3	11578374-11678374	3
(Zimin et al. 2009)	Gene expression regulation, cell cycle regulation, gonadal function	NDUFA4, SNORA62, PHF14, THSD7A	18789636-19789636	4

(Zimin et al. 2009)	Embryonic growth and development, fat metabolism and gene expression	METTL15	58890033 - 59809003 3	15
---------------------	--	---------	--------------------------------	----

بودند که کروموزوم ۱۶ در هر دوروش مشترک بود برخی از مناطق شناسایی شده با این آماره ها به صورت مستقیم و یا غیر مستقیم با صفات تولید شیرو تولید مثل در گاو در ارتباط هستند.

نتایج حاصل از آماره های iHS و nSL که جمعیت ها را به صورت تک جمعیتی مورد بررسی قرار می دهند نشان داد که نشانه های انتخاب شناسایی شده با این دو آماره بر روی کروموزوم های شماره ۳، ۴، ۷، ۱۰ و ۱۵ بودند که کروموزوم شماره ۳ در هر دوروش مشترک بود برخی از مناطق شناسایی شده با این آماره های به طور مستقیم و غیر مستقیم با صفات رشد، تولید مثل، سیستم ایمنی و سایر صفات مرتبط با آنها ارتباط دارند. با توجه به اینکه آماره‌های مورد استفاده از نظر روش اجرا و پارامترهای مورد استفاده با هم متفاوت هستند) دوتای اول چند جمعیتی و مبتنی بر فراوانی اللی و دوتا ی دوم تک جمعیتی و مبتنی بر عدم تعادل لینکاژی (در نتیجه نتایج آنها با هم مطابقت نداشت. ولی از مجموع ۵۹ ژن شناسایی شده توسط چهار روش مورد استفاده دو ژن *IQCA1*, *ACKR3*، به طور مشترک توسط آماره *iHS* و *nSL* شناسایی شدند. ضمناً می‌توان نتیجه گرفت که انتخاب طی سالیان گذشته به طور کلی بر روی صفاتی متمرکز شده است که ضمن افزایش مقاومت دام‌ها به شرایط محیطی محل زندگی بر روی افزایش راندمان تولید نیز موثر بوده است.

پس از بررسی در سایت ENSEMBL همانطور که در جدول فوق مشخص است بر روی کروموزوم شماره ۳ جمعیت هلشتاین و آمیخته ژن‌های *IQCA1*, *ACKR3*, *COPS8*, *COL6A3* بر روی صفات پاسخ به استرس، سیستم ایمنی، تنظیم رونویسی تاثیر دارند (Smith, Grosse et al. 2001). همچنین مشخص شد که بر روی کروموزوم شماره ۴ جمعیت نجدی ژن‌های *NDUFA4*, *SNORA62*, *PHF14*, *THSD7A* که بر روی صفات تنظیم بیان ژن، تنظیم چرخه سلولی، عملکرد غدد جنسی تاثیر دارند شناسایی شد. ضمناً در جمعیت نجدی بر روی کروموزوم شماره ۱۵ ژن *METTL15* که بر روی صفات رشد و توسعه جنینی، متابولیسم چربی و بیان ژن تاثیر دارند شناسایی شد (زیمین و همکاران ۲۰۰۹).

نتیجه گیری کلی

در تحقیق حاضر از چهار آماره *F_{ST}*، *hapFLK*، *iHS* و *nSL* که برای شناسایی نشانه‌های انتخاب در جمعیت‌های گاو هلشتاین، نجدی و آمیخته‌های حاصل از آنها استفاده شد. نتایج حاصل از آماره های *F_{ST}* و *hapFLK* که جمعیت ها را به طور چند جمعیتی بررسی می‌کنند نشان داد که نشانه‌های انتخاب شناسایی شده با این دو آماره بر روی کروموزوم‌های ۸، ۱۴، ۱۶، ۲۱ و ۲۶

منابع مورد استفاده

- Barendse W, Harrison BE, Bunch RJ, Thomas MB and Turner LB, 2009. "Genome wide signatures of positive selection: the comparison of independent samples and the identification of regions associated to traits." *BMC Genomics* 10(1): 1-15.
- Bastin B, Houser A, Bagley C, Ely K, Payton R, Saxton A, Schrick F, Waller J and Kojima C 2014. "A polymorphism in *XKR 4* is significantly associated with serum prolactin concentrations in beef cows grazing tall fescue." *Animal Genetics* 45(3): 439-444.
- Boitard S, Boussaha M, Capitan A, Rocha D and Servin B 2016. "Uncovering adaptation from sequence data: lessons from genome resequencing of four cattle breeds." *Genetics* 203(1): 433-450.

- Bonhomme M, Chevalet C, Servin B, Boitard S, Abdallah J, Blott S and SanCristobal M 2010. "Detecting selection in population trees: the Lewontin and Krakauer test extended." *Genetics* **186**(1): 241-262.
- Brito L F, Kijas J W, Ventura R V, Sargolzaei M, Porto-Neto L R, Cánovas A, Feng Z, Jafarikia M and Schenkel F S 2017. "Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers." *BMC Genomics* **18**: 1-20.
- Browning B L, Tian X, Zhou Y and Browning S R 2021. "Fast two-stage phasing of large-scale sequence data." *The American Journal of Human Genetics* **108**(10): 1880-1890.
- Daetwyler H D, Pong-Wong R, Villanueva B and Woolliams J A 2010. "The impact of genetic architecture on genome-wide evaluation methods." *Genetics* **185**(3): 1021-1031.
- Esmailizadeh A, Bottema C, Sellick G, Verbyla A, Morris C, Cullen N and Pitchford W 2008. "Effects of the myostatin F94L substitution on beef traits." *Journal Of Animal Science* **86**(5): 1038-1046.
- Fariello M I, Boitard S, Naya H, SanCristobal M and Servin B 2013. "Detecting signatures of selection through haplotype differentiation among hierarchically structured populations." *Genetics* **193**(3): 929-941.
- Frogameri N, Asadifozi M and Esamelizade A 2015. "Scanning the whole genome of eight breeds of native Iranian cattle to identify selection markers" *Journal of Animal Production* **18**(2): 201-213.
- Glyakina A V, Strizhov N I, Karpov M V, Dovidchenko N V, Matkarimov B T, Isaeva L V, Efimova V S, Rubtsov M A, Novikova L A and Donova M V 2019. "Ile351, Leu355 and Ile461 residues are essential for catalytic activity of bovine cytochrome P450scc (CYP11A1)." *Steroids* **143**: 80-90.
- Hou J, Qu K, Jia P, Hanif Q, Zhang J, Chen N, Dang R, Chen H, Huang B and Lei C 2020. "A SNP in PLAG1 is associated with body height trait in Chinese cattle." *Animal Genetics* **51**(1): 87-90.
- Hudson R R, Bailey K, Skarecky D, Kwiatowski J and Ayala F J 1994. "Evidence for positive selection in the superoxide dismutase (Sod) region of *Drosophila melanogaster*." *Genetics* **136**(4): 1329-1340.
- Kemper K E, Saxton S J, Bolormaa S, Hayes B J and Goddard M E 2014. "Selection for complex traits leaves little or no classic signatures of selection." *BMC Genomics* **15**(1): 1-14.
- Kijas J W, Porto- Neto L, Dominik S, Reverter A, Bunch R, McCulloch R, Hayes B J, Brauning R, McEwan J and Consortium I S G 2014. "Linkage disequilibrium over short physical distances measured in sheep using a high- density SNP chip." *Animal Genetics* **45**(5): 754-757.
- Kuehn L, Keele J, Bennett G, McDanel T, Smith T, Snelling W, Sonstegard T and Thallman R 2011. "Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project." *Journal Of Animal Science* **89**(6): 1742-1750.
- Liu H, Xu G, Yuan Z, Dong Y, Wang J and Lu W 2017. "Effect of kisspeptin on the proliferation and apoptosis of bovine granulosa cells." *Animal Reproduction Science* **185**: 1-7.
- MacEachern S, Hayes B, McEwan J and Goddard M 2009. "An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (*Bos taurus*) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic cattle." *BMC Genomics* **10**(1): 1-19.
- Manunza A, Cardoso T F, Noce A, Martínez A, Pons A, Bermejo L A, Landi V, Sánchez A, Jordana J, Delgado J V, Adán S, Capote J, Vidal O, Ugarte E, Arranz J J, Calvo J H, Casellas J and Amills M 2016. "Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK." *Scientific Reports* **6**(1): 1-10.
- Mokhber M, Shahrabak M M, Sadeghi M, Shahrabak H M and Williams J 2015. "Genome-wide survey of signature of positive selection in Khuzestani and Mazandrani buffalo breeds." *Iranian Journal Of Animal Science (IJAS)* **46**(2): 119-131.
- Moradi M H, Nejati-Javaremi A, Moradi-Shahrabak M, Dodds K G and McEwan J C 2012. "Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition." *BMC Genetics* **13**(1): 1-15.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A, Bender D, Maller J, Sklar P, De Bakker P I and Daly M J 2007. "PLINK: a tool set for whole-genome association and population-based linkage analyses." *The American Journal Of Human Genetics* **81**(3): 559-575.

