

Identification of SNP markers related to important traits for Iranian native sheep using selection signature

Maryam Moghadam¹, Saeed Zerehdaran^{2*}, Mohammad Hossein Banabazi³,
Ali Javadmanesh⁴

¹ PhD student, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

² Professor, Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Email: zerehdaran@um.ac.ir

³ Researcher Sveriges lantbruksuniversitet (SLU), Swedish University of Agricultural Sciences, Swedish

⁴ Associate Professor, Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

Article Info

Article type:
Research Full Paper

Article history:

Received:
Revised:
Accepted:

Keywords:

Genome sequencing
Iranian native sheep
Selection signature
Sheep array

ABSTRACT

Background and objective: Identification of several genomic markers sequenced in farm animal enable us to study the structure of their genome. Samples from different parts of the world are usually used for sequencing genomic markers. Information on the sequence of specific markers of domestic animals of each geographical area will improve our knowledge about their genomic structure. The sequences of specific genome markers for Iranian native sheep were not yet identified. Present study aims to identify single-nucleotide polymorphism (SNPs) influencing important economic and physiological traits of Iranian domestic and wild sheep as candidate markers using selection signature.

Materials and methods: In order to identify significant genetic variants under selection in domestic and wild sheep, the whole genome sequence of 20 domestic (*Ovis aries*) and 14 wild sheep (*Ovis orientalis*) obtained from the Next Gene project was used. After doing quality control, variant under selection were detected using statistical methods consisting iHS and XP-EHH. Genetic variances for identified variants were estimated. The genetic variance of known variants in Merino, Suffolk, texel and Afshari were also estimated. The threshold for selecting variants based on iHS and XP-EHH methods were considered to be 99.9% in domestic sheep and 99.9% and 0.001% in wild sheep, respectively. Finally, variants with genetic variances higher than 0.25, were identified and suggested as candidate markers.

Results: In present study, 170 genomic variants with iHS and XP-EHH values higher than 3.83 and 3.44 in domestic sheep and 150 genomic variants with iHS values higher than 3.05 and XP-EHH values lower than -4.43 in wild sheep were identified and suggested as candidate markers. Genetic variances of these variants were 0.27 to 0.50 and the genetic variance of known variants in Merino, Suffolk, Texel and Afshari were from 0.02 to 0.50. This comparison shows the importance of identified variants in Iranian domestic and wild sheep. Most of identified variants in domestic and wild sheep associated with economically important traits including carcass, milk and wool related traits. The ontology of genes related to identified variants showed that these variants are located near genes related to metabolic process of production and reproduction traits

Conclusion: Present results showed that there are many significant genetic variants associated with economic traits in domestic and wild Iranian sheep which are not included in the commercial sheep arrays available in the

market. Identification of new SNPs related to economic traits in domestic and wild sheep may help to improve the studies related to genomic structure of Iranian sheep. These results together with similar studies could be efficiently used for producing SNP arrays designed for Iranian native sheep.

Cite this article: Moghadam, M., Zerehdaran, S., Banabazi, M.H., Javadmanesh, A. (2023). Identification of SNP markers related to important traits for Iranian native sheep using selection signature. *Journal of Ruminant Research*, 11(4), 73-92.



© The Author(s).

DOI:

Publisher: Gorgan University of Agricultural Sciences and Natural Resources

شناسایی نشانگرهای تک نوکلئوتیدی مؤثر بر صفات مهم گوسفندان بومی ایران با استفاده از ردپای انتخاب

مریم مقدم^۱، سعید زره داران^{۲*}، محمدحسین بنابازی^۳، علی جواد منش^۴

^۱دانشجوی دکتری اصلاح نژاد دام گروه علوم دامی دانشکده کشاورزی دانشگاه فردوسی مشهد، مشهد، ایران
^۲آستاذ گروه علوم دامی دانشکده کشاورزی دانشگاه فردوسی مشهد، مشهد، ایران، رایانامه: zerehdaran@um.ac.ir
^۳آستادیار و محقق دانشگاه کشاورزی سوئد (SLU)
^۴دانشیار گروه علوم دامی دانشکده کشاورزی دانشگاه فردوسی مشهد، مشهد، ایران

اطلاعات مقاله	چکیده
نوع مقاله: مقاله کامل علمی - پژوهشی	سابقه و هدف: شناسایی تعداد فراوان نشانگرهای ژنومی توالی یابی شده در ژنوم حیوانات اهلی ما را قادر به مطالعه ساختار ژنومی آنها می‌کند. معمولاً برای توالی یابی نشانگرهای ژنومی حیوانات اهلی از نمونه‌های جمع‌آوری شده از نقاط مختلف دنیا استفاده می‌شود. شناسایی نشانگرهای ژنومی اختصاصی حیوانات بومی هر منطقه جغرافیایی می‌تواند منجر به بهبود اطلاعات مربوط به ساختار ژنومی آن گونه شود. تا به حال، برای جمعیت گوسفندان بومی ایران نشانگرهای ژنومی اختصاصی شناسایی نشده است. هدف این پژوهش شناسایی واریانت‌های ژنتیکی مؤثر بر صفات مهم اقتصادی و فیزیولوژیکی گوسفندان اهلی و وحشی ایران و پیشنهاد آنها به عنوان نشانگر اختصاصی از طریق ردپای انتخاب بود.
واژه‌های کلیدی: آرایه ژنومی گوسفند توالی یابی ژنومی ردپای انتخاب نشانگر ژنومی	مواد و روش‌ها: برای شناسایی واریانت‌های ژنتیکی نواحی تحت انتخاب گوسفندان اهلی و وحشی، ابتدا اطلاعات توالی یابی کل ژنوم ۲۰ گوسفند اهلی (<i>Ovis aries</i>) و ۱۴ گوسفند وحشی (<i>Ovis orientalis</i>) از پروژه Next Gene گرفته شدند. سپس واریانت‌های ژنتیکی کنترل کیفیت شده و واریانت‌های ژنتیکی نواحی تحت انتخاب با استفاده از روش‌های آماری iHS و XP-EHH شناسایی شدند. برای این واریانت‌ها واریانس ژنوتیپی برآورد شد. علاوه بر آن، واریانس ژنوتیپی نشانگرهای ژنومی شناخته شده در گوسفندان نژادهای مریوس، سافوک، تکسل و افشاری جهت مقایسه با گوسفندان بومی ایران، نیز برآورد شدند. آستانه انتخاب واریانت‌های ژنومی برای ضرایب iHS و XP-EHH در گوسفندان اهلی کران ۰/۹۹۹ و در گوسفندان وحشی کران ۰/۹۹۹ و ۰/۰۰۱ در نظر گرفته شد. در نهایت از بین واریانت‌های شناسایی شده، آنهایی که واریانس ژنوتیپی بیشتر از ۰/۲۵ داشتند، به عنوان نشانگر اختصاصی این گونه‌ها پیشنهاد شدند.
	یافته‌ها: در این پژوهش ۱۷۰ واریانت ژنومی با ضرایب iHS و XP-EHH بالاتر از ۳/۸۳ و ۳/۴۴ در گوسفندان اهلی و ۱۵۰ واریانت ژنومی با ضرایب iHS بالاتر از ۳/۰۵ و XP-EHH پایین تر از ۴/۴۳- در گوسفندان وحشی شناسایی و انتخاب شدند. واریانس ژنوتیپی واریانت

های ژنومی شناسایی شده، در دامنه ۰/۲۷ تا ۰/۵۰ بود. واریانس ژنوتیپی نشانگرهای ژنومی شناخته شده در نژادهای مرینوس، سافوک و تکسل و نژاد بومی افشاری به صورت میانگین در دامنه ۰/۰۲ تا ۰/۵۰ برآورد شدند. این مقایسه بیانگر تاثیر قابل توجه واریانت‌های شناسایی شده در گوسفندان بومی ایران است. بسیاری از این واریانت‌ها در نواحی ژنتیکی مرتبط با صفات اقتصادی گوسفند نظیر صفات لاشه، شیر و پشم قرار داشتند. علاوه بر این، هستی شناسایی ژن های شناسایی شده مرتبط با واریانت‌های ژنومی پیشنهادی نشان داد، این واریانت‌ها در مجاورت نواحی ژنی مرتبط با فرآیندهای متابولیکی صفات تولیدی و تولیدمثلی هستند.

نتیجه‌گیری: مطالعه حاضر نشان داد که تعداد قابل توجهی نشانگر معنی‌دار مرتبط با صفات مهم اقتصادی و فیزیولوژیکی در گوسفندان اهلی و وحشی ایران وجود دارد که در آرایه‌های تجاری موجود در بازار در نظر گرفته نشده‌است. تشخیص نشانگرهای جدید می‌تواند به بهبود اطلاعات ما در خصوص ساختار ژنومی گوسفندان بومی ایران کمک کند. همچنین، این نتایج در کنار مطالعات مشابه می‌تواند به طور موثری در تولید آرایه‌های تجاری برای گوسفند ایرانی مورد استفاده قرار گیرد.

استناد: مقدم، م.، زره داران، س.، بنابازی، م.ح.، جواد منش، ع. (۱۴۰۲). شناسایی نشانگرهای تک نوکلئوتیدی مؤثر بر صفات مهم گوسفندان بومی ایران با استفاده از ردپای انتخاب. پژوهش در نشخوارکنندگان، ۱۱(۴)، ۹۲-۷۳.

DOI:



© نویسندگان.

ناشر: دانشگاه علوم کشاورزی و منابع طبیعی گرگان

مقدمه

گوسفند یکی از اولین حیواناتی است که ۷۰۰۰ سال قبل از میلاد مسیح در نواحی آسیای غربی و اروپا شرقی، اهلی شده است. گوسفندان وحشی آسیا و اروپا اجداد اولیه گوسفندان اهلی امروزی هستند. گوسفندان وحشی شامل دو رده موفلون و اورپال می شوند. گوسفندان موفلون به دو نژاد آسیایی و اروپایی تقسیم بندی می شوند. این نژادها به ترتیب موفلون آسیایی (*Ovis orientalis*) و موفلون اروپایی (*Ovis miusimon*) نامگذاری شده اند. موفلون آسیایی به لحاظ جثه کوچکتر از موفلون اروپایی است و جزء کوچکترین نژادهای گوسفندان وحشی دنیا می باشد. موفلون آسیایی خود به دو گروه زیرگونه ای تقسیم می شود که گروه شرقی آن گوسفند وحشی اورپال و گروه غربی آن گوسفند وحشی ارمنی نامیده می شود. گوسفندهای وحشی البرز مرکزی حاصل تلاقی بین قوچ و میش های ارمنی و اورپال هستند (Ryder, ۱۹۸۳). این گونه در شمال و شمال غرب ایران، نخجوان و ارمنستان دیده می شود.

اهلی شدن گوسفند باعث تغییر در ساختار مورفولوژی این گونه شده است. همچنین، انتخاب طبیعی و انتخاب توسط انسان باعث ایجاد ساختار ژنتیکی منحصر به فرد در گوسفندان هر منطقه جغرافیایی شده است (Ryder, ۱۹۶۴). به عنوان مثال، به مرور زمان فولیکول های تولیدکننده موهای کوتاه و ضخیم در گوسفندان اهلی در مقایسه با اجداد وحشی تحلیل رفته و فولیکول های موهای بلند و ظریف توسعه پیدا کرده اند. یکی دیگر از نتایج انتخاب در گوسفندان اهلی نسبت به اجداد وحشی آنها، ساختار دنبه است. گوسفندان وحشی دارای دم کوتاه هستند اما گوسفندان اهلی در نژادهای مختلف دارای نیم دنبه یا دنبه سنگین هستند (Mahal و همکاران، ۱۹۵۱).

هدف اصلی برنامه های اصلاح نژادی بهبود ساختار ژنتیکی حیوانات مزرعه ای از طریق افزایش صحت پیش بینی ارزش های اصلاحی و انتخاب بهترین حیوانات به عنوان والدین نسل بعد می باشد. در اصلاح نژاد سنتی، استفاده بهینه از اطلاعات شجره و رکوردهای فنوتیپی حیوانات اهلی نظیر گاو، گوسفند و بز راه رسیدن به این هدف بود (Goddard و Hayes, ۲۰۰۹). در سال های اخیر با شناسایی و توالی یابی نشانگرهای ژنومی در سراسر ژنوم حیوانات اهلی، مطالعات ژنومی جایگزین مطالعات اصلاح نژادی سنتی گذشته شده است (Meuwissen و همکاران، ۲۰۰۱). با توسعه روزافزون روش های توالی یابی ژنومی و شناسایی نشانگرهای جدید در ژنوم حیوانات، مطالعات ژنومی به طور چشم گیری بهبود یافته اند (Tayara و Chong, ۲۰۱۹). در مورد گوسفند و بز، توالی یابی ژنوم در سال ۲۰۰۷ شروع شد. به دنبال آن نشانگرهای چندشکلی تک نوکلئوتیدی در این حیوانات توالی یابی شدند. در نهایت، شرکت ایلومینا در سال ۲۰۰۹ به عنوان قسمتی از پروژه کنسرسيوم بين المللی ژنوم گوسفند، اولین آرایه برای گوسفند را تولید نمود (Rupp و همکاران، ۲۰۱۶). فناوری های توالی یابی جدید فرصتی تازه برای شناسایی واریانت های مؤثر بر صفات مهم اقتصادی را پدید آورده اند (Djari و همکاران، ۲۰۱۳). از طرفی با توجه به هزینه بالای استفاده از آرایه های با تراکم بالا، تمایل برای استفاده از آرایه های اقتصادی با تراکم پایین تر افزایش یافته است. آرایه ۵۰ هزار جفت باز برای گاو (Matukumalli و همکاران، ۲۰۰۹) و آرایه ۶۰ هزار جفت باز برای مرغ (Groenen و همکاران، ۲۰۱۱) مثال هایی از آرایه های با تراکم متوسط هستند. استفاده از این نوع آرایه ها که برای اکثر دام ها در دسترس هستند، هزینه تعیین ژنوتیپ را کاهش داده و محاسبات را ساده تر می کنند.

صفات اقتصادی برای تبدیل شدن به نشانگر ژنومی در گوسفندان بومی (اهلی و وحشی) ایران بود.

مواد و روش‌ها

داده‌های ژنومی: در این مطالعه داده‌های ژنومی گوسفندان اهلی و وحشی بومی ایران از پروژه Next Gene گرفته شدند. این پروژه با هدف برآورد تنوع زیستی گوسفند، بز و گاو با استفاده از روش‌های ریز آرایه و توالی‌یابی نسل جدید انجام شده‌است. توالی-یابی دام‌های اهلی و وحشی توسط شرکت ایلومینا صورت گرفت. پس از نمونه‌گیری و توالی‌یابی، داده‌ها در سال ۲۰۱۴ در بانک اطلاعاتی "ای ام ای" اروپا در دسترس عموم قرار گرفت. هدف مطالعه مذکور به طور عمده ارزیابی منابع ژنتیکی در نشخوارکنندگان کوچک در شمال غرب ایران بوده است. بخشی از داده‌های ژنومی پروژه Next Gene مربوط به توالی‌یابی ۲۰ گوسفند اهلی و ۱۴ گوسفند وحشی ایرانی می‌باشد. نمودار توزیع جغرافیای گوسفندان ایرانی این پروژه در شکل ۱ نشان داده شده است. کدهای دسترسی گوسفندان اهلی و وحشی این پروژه نیز به تفکیک برای هر نمونه در جدول ۱ ارائه شده‌است. داده‌ها به صورت دو فایل فشرده از پایگاه داده "ای ام ای" دانلود گردید. هر کدام از این فایل‌ها به طور جداگانه حاوی خوانش‌های رفت و برگشت می‌باشد. علاوه بر این، برای مقایسه واریانس ژنوتیپی واریانت‌های شناسایی شده با نشانگرهای ژنومی شناخته شده، داده‌های ژنوتیپی تراشه متراکم ۶۰۰ هزار جفت باز تجاری نژادهای مریوس، سافوک، تکسل و افشاری نیز از سایت "دبلیو دی دی ای" گرفته شدند.

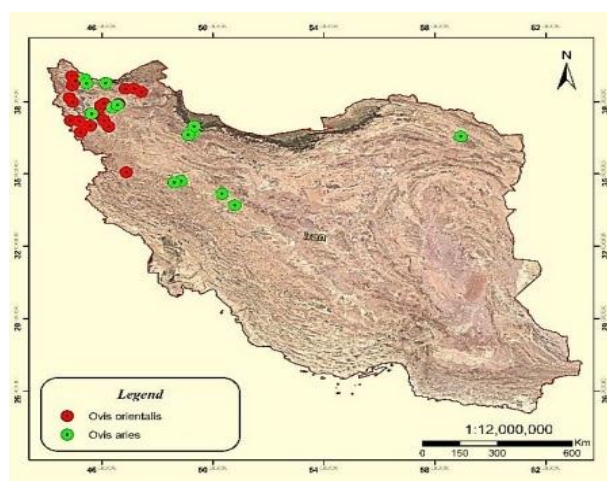
اگرچه افزایش تراکم نشانگرها در سرتاسر ژنوم (از ۳ هزار به ۵۰ هزار تا ۷۷۷ هزار) منجر به افزایش صحت پیش‌بینی‌های ژنومی می‌شود ولی این بهبود متناسب با افزایش تراکم آرایه‌ها نیست (Tosser-klopp و همکاران، ۲۰۱۴). علاوه بر آن باید در نظر داشت که استفاده از آرایه‌های با تراکم بالا، علاوه بر افزایش هزینه تعیین ژنوتیپ، پیچیدگی‌های محاسباتی بیشتری را نیز در پی دارد. ذکر این نکته نیز ضروری است که به دلیل تعداد زیاد نژادهای گوسفند شیری و گواشتی در مناطق جغرافیایی مختلف احتمال اینکه یک آرایه منحصر به فرد از نشانگرهای تک نوکلوتیدی در سرتاسر جهان استفاده شود، منطقی به نظر نمی‌رسد (Rupp و همکاران، ۲۰۱۶). انتخاب آرایه با تعداد مناسب نشانگرهای تک نوکلوتیدی و نیز ماهیت آن‌ها برای پیش‌بینی‌های مبتنی بر توالی‌یابی کل ژنوم یک چالش مهم پیش روی متخصصین اصلاح نژاد دام است. از این رو از راه‌کارهای مناسب برای توالی‌یابی نشانگرهای ژنومی با تراکم مناسب، استقبال می‌شود (Yang و همکاران، ۲۰۱۶). در طراحی و ساخت آرایه‌های تجاری موجود در بازار برای گوسفند از اطلاعات مربوط به توالی جایگاه‌های تک نوکلوتیدی گوسفندان بومی ایران استفاده نشده‌اند. بنابراین، ضروری است نشانگرهایی که به طور اختصاصی با صفات مهم تولیدی و تولیدمثلی این حیوانات ارتباط معنی‌داری دارند، شناسایی و در برنامه‌های اصلاحی نژادی مورد استفاده قرار گیرند. هدف این تحقیق شناسایی واریانت‌های مرتبط با

جدول ۱- کدهای دسترسی نمونه‌های گوسفندان اهلی و وحشی

Table 1- Accession number of wild and domestic sheep samples

Sample ID شماره نمونه	Accession No. کد دسترسی	Sample ID شماره نمونه	Accession No. کد دسترسی	Sample ID شماره نمونه	Accession No. کد دسترسی	Sample ID شماره نمونه	Accession No. کد دسترسی
IROA-B2-5037	ERS154864	IROA-D5-5081	ERS239051	IROA-F5-5051	ERS154868	IROO-D6-5104	ERS239059
IROA-B2-5296	ERS239046	IROA-D6-5152	ERS154866	IROA-G3-5095	ERS239057	IROO-E3-5492	ERS154533
IROA-B3-5134	ERS239047	IROA-D7-5033	ERS239052	IROO-D6-0004	ERS154529	IROO-E5-5146	ERS154531
IROA-B4-5190	ERS154865	IROA-E5-5157	ERS239053	IROO-C3-0001	ERS154526	IROO-F5-5079	ERS154532
IROA-B5-5295	ERS154863	IROA-E6-5351	ERS239054	IROO-C3-2743	ERS419579	IROO-J11-0602	ERS239060
IROA-B6-5139	ERS239048	IROA-E7-5036	ERS239055	IROO-D6-0002	ERS154527	IROO-J11-0905	ERS239061
IROA-C3-5212	ERS154869	IROA-F10-5068	ERS239056	IROO-D6-0003	ERS154528	IROO-K7-0639	ERS154871
IROA-C6-5187	ERS239049	IROA-F3-5142	ERS154867	IROO-D6-0005	ERS154530		
IROA-C7-5042	ERS239050	IROA-G4-5205	ERS154862	IROO-D6-0006	ERS239058		

IROA و IROO به ترتیب نمونه‌های گوسفندان اهلی و وحشی هستند.



شکل ۱- توزیع جغرافیایی نمونه‌های گوسفند که از مناطق مختلف ایران جمع آوری شدند. دایره‌های سبز و قرمز به ترتیب نشان دهنده گوسفندان اهلی و گوسفندان وحشی هستند.

Figure 1. Geographical distribution of sheep samples collected from different areas of Iran. Green and red circles representing domestic and wild sheep, respectively

مرتب‌سازی، علامت‌گذاری تکراری‌ها، هم‌ردیفی مجدد، حذف و اضافه بازهای آلی در طول ژنوم و درجه‌بندی مجدد کیفیت داده‌ها توسط نرم‌افزار GATK (نسخه ۱/۸) انجام شد. شناسایی نشانگرهای تک نوکلئوتیدی و هاپلوتایپ‌ها به همراه فیلترکردن آنها نیز توسط نرم‌افزار GATK (نسخه ۱/۸) انجام و در نهایت فایل‌هایی با ساختار وی سی اف تشکیل شد (Lühken و Bathke, ۲۰۲۱). سپس تمامی داده‌های

کنترل کیفیت داده‌های ژنومی: کنترل کیفیت و ویرایش داده‌های توالی‌یابی DNA اخذ شده از پروژۀ Next Gene توسط دو نرم‌افزار FastQC (نسخه ۰.۱۱.۸) و Trimmomatic (نسخه ۰.۳۶) انجام شد (Sewe و همکاران, ۲۰۲۲). هم‌ردیفی داده‌های توالی DNA بر روی ژنوم مرجع توسط نرم‌افزار BWA (نسخه ۰/۷/۱۷) انجام شد (Sachajko و همکاران, ۲۰۱۹). پردازش خوانش‌های نقشه‌یابی شده از جمله

خاصیت جهش تک نوکلئوتیدی واریانت‌های شناسایی شده، واریانس ژنوتیپی برای واریانت‌ها با استفاده از معادله آماری $2P_i(1-P_i)$ برآورد شد (VanRaden, ۲۰۰۸). در این معادله، P_i فراوانی آللی آلل مرجع واریانت‌ها است. علاوه بر واریانس ژنوتیپی واریانت‌های شناسایی شده در گوسفندان بومی ایران، واریانس ژنوتیپی برای نشانگرهای ژنومی شناخته شده گوسفندان خارجی نیز برآورد و با یکدیگر مقایسه شدند.

روش آماری iHS: روش iHS برای شناسایی نواحی تحت انتخاب درون جمعیت گوسفندان اهلی و وحشی در این مطالعه استفاده شد. این روش مناطقی از ژنوم را شناسایی می‌کند که در مقایسه با کل ژنوم به صورت قابل توجهی عدم تعادل پیوستگی بالاتری دارند. در این روش، انتگرال زیر منحنی در همه جایگاه‌های پیرامون واریانت ژنومی مرکزی تا هنگامی که به ۰/۰۵ می‌رسد محاسبه و ارزش‌های به‌دست آمده برای آلل‌های اجدادی و مشتق شده (غیراجدادی) با هم مقایسه می‌شوند. انتگرال زیر منحنی در دو طرف آلل مرکزی باهم جمع شده و بسته به این که مربوط به آلل اجدادی یا مشتق شده باشند به ترتیب هاپلوتیپ یکپارچه‌ی اجدادی (iHHA) و هاپلوتیپ یکپارچه‌ی مشتق شده (iHHD) نامیده می‌شوند. آن‌گاه ضریب iHS با استفاده از رابطه‌ی زیر استاندارد می‌شوند (Voight و همکاران، ۲۰۰۶).

$$iHS = \frac{\ln\left(\frac{iHHA}{iHHD}\right) - E_{ps}[\ln\left(\frac{iHHA}{iHHD}\right)]}{SD_{ps}[\ln\left(\frac{iHHA}{iHHD}\right)]} \quad \text{رابطه ۱.}$$

در رابطه ۱، iHHA هاپلوتیپ یکپارچه‌ی اجدادی، iHHD هاپلوتیپ یکپارچه‌ی مشتق شده، SD_{ps} انحراف استاندارد واریانت مرکزی و E_{ps} امید ریاضی فراوانی واریانت مرکزی است (Mohammadi و همکاران ۲۰۱۷).

ژنومی براساس فراوانی آللی کمتر از ۰/۰۵، نرخ ژنوتیپ گمشده واریانت ژنتیکی ۰/۰۰۱، نرخ ژنوتیپ گمشده حیوان ۱ و آزمون هاردی-وینبرگ با آستانه معنی‌داری 10^{-6} تصحیح شدند. به علاوه، با استفاده از نقشه‌ی ژنومی آرایه تجاری گوسفند با تراکم ۶۰۰ هزار، معرفی شده توسط کمپانی ایلومینا، واریانت‌های ژنومی در محدوده یک کیلو باز اطراف نشانگرهای شناخته شده، حذف شدند. برای کنترل کیفیت داده-های ژنوتیپی، ساخت فایل‌های ژنوتیپی، برآورد فراوانی آللی و ترسیم نمودار مولفه‌های اصلی از نرم افزار PLINK (نسخه ۱/۰۹) (Purcell و همکاران، ۲۰۰۷) و نرم افزار R (نسخه ۴/۲/۲) استفاده شد.

روش‌های آماری

جهت بررسی ارتباط جمعیتی بین نمونه‌های جمع‌آوری شده از گوسفندان بومی ایران در پروژه Next Gene از روش تجزیه به مولفه‌های اصلی استفاده شد. هدف از تجزیه به مؤلفه‌های اصلی، تجزیه واریانس موجود در داده‌های چندمتغیره به مؤلفه‌هایی است که به ترتیب علت بیشترین واریانس موجود در داده‌ها باشد. در این روش هر مؤلفه مستقل از مؤلفه‌های دیگر بوده و بین آنها همبستگی وجود ندارد. با توجه به اینکه دو مؤلفه اول (PC1 و PC2) جمعیت‌ها را به خوبی از یکدیگر تفکیک نمودند، از این دو معیار برای شناسایی ارتباط بین جمعیت‌ها استفاده شد. روش تجزیه مولفه‌های اصلی با استفاده از تابع prcomp در نرم‌افزار R (نسخه ۴/۲/۲) انجام شد.

برای شناسایی واریانت‌های ژنتیکی نواحی تحت انتخاب در گوسفندان بومی اهلی و وحشی از روش‌های آماری مبتنی بر عدم تعادل پیوستگی و طول بلوک‌های هاپلوتیپی درون جمعیتی (iHS) و بین جمعیتی (XP-EHH) استفاده شد. برای اطمینان از

برای شناسایی و انتخاب واریانت‌های ژنومی به‌عنوان نشانگر ژنومی پیشنهادی، آستانه انتخاب واریانت‌های ژنومی برای مقادیر ضرایب iHS و XP-EHH کران ۰/۹۹۹ در گوسفندان اهلی و کران ۰/۹۹۹ و ۰/۰۰۱ برای گوسفندان وحشی در نظر گرفته‌شد. در نهایت از بین واریانت‌های شناسایی‌شده، آنهایی که واریانس ژنوتیپی بیشتر از ۰/۲۵ داشتند، به عنوان نشانگرهای پیشنهادی گزارش شدند (Kijas و همکاران، ۲۰۰۹).

شناسایی جایگاه‌های صفات اقتصادی و ژن‌های مرتبط با واریانت‌های گوسفندان بومی ایران: ابتدا نقشه ژنومی جایگاه‌های صفات کمی گوسفند از سایت ژنوم حیوانی^۱ دانلود شد. سپس با استفاده از نرم افزارهای R (نسخه ۴/۲/۲) و PLINK (نسخه ۱/۰۹) جایگاه صفات کمی حاوی واریانت‌ها، شناسایی و شناسه جایگاه ذخیره شد (Zeraatpisheh و همکاران، ۲۰۲۳). در نهایت، برای گوسفند اهلی ۱۵ جایگاه و برای گوسفند وحشی ۱۰ جایگاه مرتبط با واریانت‌ها در چهار زیرگروه صفات وزن بدن و لاشه، تولید شیر، پشم و صفات تولید مثلی شناسایی شدند. برای شناسایی ژن‌های مرتبط با واریانت‌ها نیز لیست ژن‌های گوسفند از سایت مرکز ملی اطلاعات بیوتکنولوژی^۲ اخذ و سپس با نرم‌افزار، PLINK (نسخه ۱/۰۹)، ژن‌های مرتبط با واریانت‌ها شناسایی و ذخیره شد. به‌منظور شناسایی مسیرهای زیستی درگیر برای ژن‌های شناسایی شده، از تجزیه هستی‌شناسی استفاده شد (Lammers و همکاران، ۲۰۱۰). بدین منظور از نرم‌افزار ClueGo (نسخه ۲/۵/۸) برای شناسایی مسیرهای قابل توجه استفاده شد. این برنامه یک افزونه به نرم‌افزار Cytoscape است و به‌طور قابل توجهی تفسیرهای بیولوژیکی ژن‌ها را توسعه می-

روش آماری XP-EHH: این روش یکی از مهم‌ترین روش‌های آماری مبتنی بر عدم تعادل پیوستگی و طول بلوک هاپلوتایپی می‌باشد. آماره XP-EHH توان بالایی در شناسایی مناطق تثبیت شده دارد (Sabeti و همکاران، ۲۰۰۷). در این روش برای شناسایی مناطق متأثر از انتخاب نیازی به مشخص کردن آلل‌های اجدادی نیست و مناطق ژنومی متأثر از انتخاب مثبت بوسیله آلل‌هایی با EHH و فراوانی بالا تشخیص داده می‌شوند. در آزمون XP-EHH با مقایسه عدم تعادل پیوستگی و طول بلوک‌های هاپلوتایپی در طول ژنوم دو جمعیت، نواحی ژنومی تحت انتخاب هر جمعیت نسبت به جمعیت دیگر برآورد می‌شود. در این مطالعه، نواحی ژنومی تحت انتخاب در گوسفندان اهلی و وحشی نسبت به یکدیگر شناسایی شدند.

$$EHH = \frac{\sum_{i=1}^{h_x} \binom{ni}{2}}{\binom{na}{2} \binom{nA}{2}} \quad \text{رابطه ۲}$$

در رابطه ۲، na و nA به ترتیب تعداد بلوک‌های هاپلوتایپی است که در آن‌ها به ترتیب آلل‌های a و A وجود دارند و h_x تعداد هاپلوتیپ‌های مجزا را در یک منطقه ژنومی تا فاصله x نشان می‌دهد. بعد از برآورد EHH آماره XP-EHH از معادله زیر به دست می‌آید.

$$XP - EHH_{\text{score}} = \ln\left(\frac{I_A}{I_B}\right) \quad \text{رابطه ۳}$$

در معادله بالا I_A و I_B به ترتیب میزان انتگرال EHH با توجه به فاصله ژنتیکی در دو جمعیت A و B می‌باشند (Taheri و همکاران، ۲۰۲۲). مقادیر ضرایب این روش در دو جهت منفی و مثبت محاسبه می‌شود به نحوی که مقادیر مثبت بیانگر نواحی تحت انتخاب در جمعیت اهلی و مقادیر منفی بیانگر نواحی تحت انتخاب در جمعیت وحشی هستند. برای برآورد ضرایب iHS و XP-EHH از پکیج rehh (نسخه ۲) استفاده شد (Gautier و همکاران، ۲۰۱۷).

¹-<https://www.animalgenome.org>

²-<https://www.ncbi.nlm.nih.gov/gene>

اهلی از مناطق جغرافیایی متفاوت جمع‌آوری شدند. به احتمال زیاد نمونه‌های اهلی از نژادهای مختلف هستند که متاسفانه در اطلاعات پروژه Next Gene اشاره‌ای به نژاد آنها نشده است. بر اساس نتیجه تجزیه به مولفه‌های اصلی، فاصله ژنتیکی بین گونه‌های گوسفندان اهلی و وحشی بسیار زیاد است که در راستای اهداف این مطالعه می‌باشد.

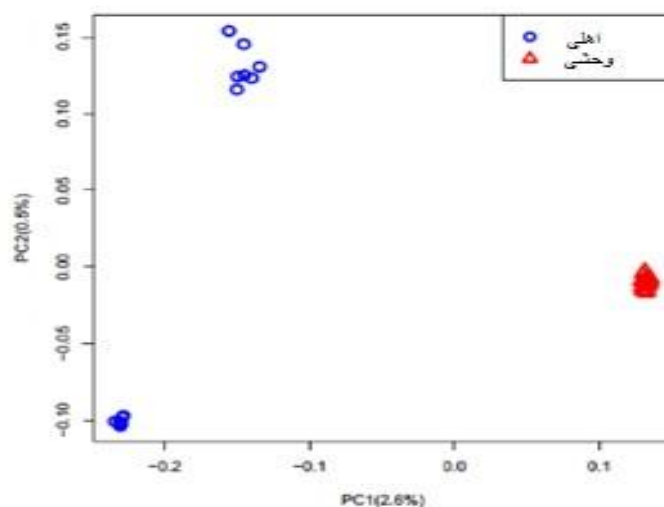
دهد (Bindea و همکاران، ۲۰۰۹). مسیرهایی که اعتبار آماری آن‌ها کمتر از ۰/۰۵ بود به عنوان مسیرهای تحت انتخاب معرفی شدند (Roshandel و همکاران، ۲۰۲۲). برای تجزیه و تحلیل چرخه‌های بیولوژیکی، از پایگاه‌های اطلاعات ژنی آنلاین کیوتو (KEGG)^۱ استفاده شد. این پایگاه شامل اطلاعات ژنومی، شیمیایی و فنوتیپی برای چرخه‌ها می‌باشد چرخه‌های KEGG شامل فرآیندهای متابولیکی، پردازش اطلاعات ژنتیکی، پردازش اطلاعات محیط زیست، اطلاعات مسیر سیگنالینگ و پروسه‌های سلولی می‌باشد. بهترین و کامل‌ترین اطلاعات در پایگاه اطلاعاتی کیوتو مربوط به چرخه‌های متابولیکی است (Aoki-Kinoshita و Kanehisa، ۲۰۰۷).

نتایج و بحث

داده‌های ژنومی اخذ شده از پروژه NEXT GENE شامل ۱۶۶۴۷۹۵۸۰۶ واریانت برای گوسفندان اهلی و ۱۶۶۴۶۱۶۹۸۳ واریانت برای گوسفندان وحشی بود. پس از کنترل کیفیت داده‌های ژنومی، به ترتیب ۱۵۷۳۵۶۵۵ و ۱۷۸۴۷۱۱۴ واریانت ژنتیکی برای گوسفندان اهلی و وحشی باقی ماند. بخشی از واریانت‌های حذف شده مربوط به نشانگرهای شناخته شده در گوسفندان خارجی بود. بنابراین کلیه واریانت‌های باقی‌مانده فاصله ژنتیکی حداقل یک کیلو بازی با نشانگرهای ژنومی شناخته شده داشتند.

نمودار تجزیه به مولفه‌های اصلی (شکل ۲) نشان داد، جمعیت گوسفندان وحشی در یک توده ژنتیکی کاملاً متراکم و گوسفندان اهلی در دو توده ژنتیکی با فاصله ژنتیکی قابل توجه و جدا از جمعیت وحشی قرار دارند. با توجه به شکل ۱، حیوانات وحشی از یک منطقه جغرافیایی در شمال غرب ایران و حیوانات

^۱ - Kyoto Encyclopedia of Genes and Genomes

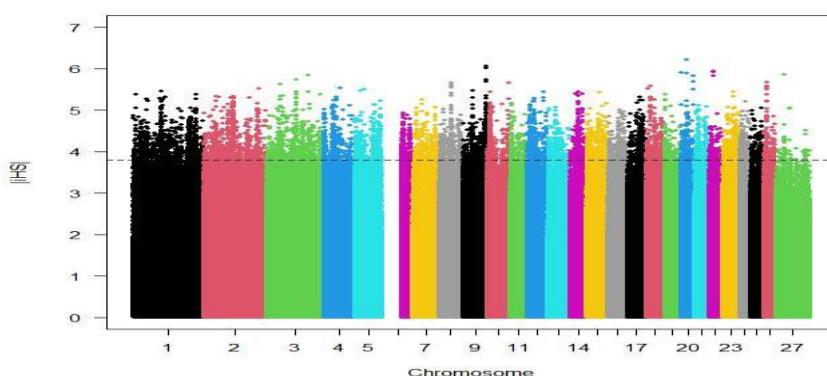


شکل ۲- تجزیه مولفه‌های اصلی گوسفندان اهلی و وحشی

Figure 2. PCA analysis for domestic and wild sheep

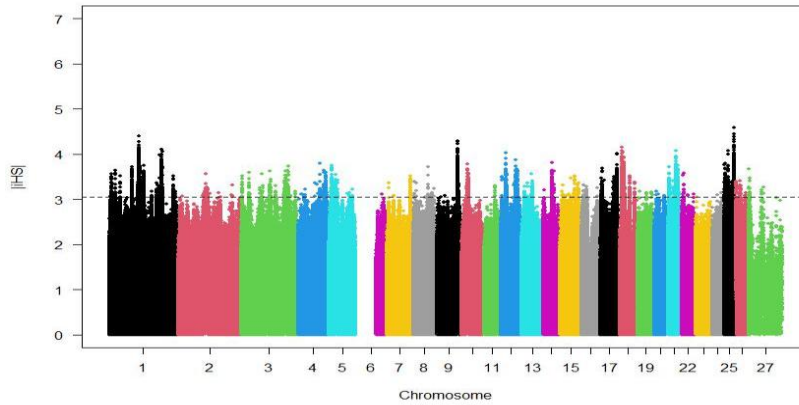
انتخاب برای گوسفندان اهلی ۳/۸۳ بود. در حیوانات وحشی کروموزوم‌های ۱، ۹، ۱۷ و ۲۵ بیشترین مقادیر ضریب iHS را نشان دادند. بیشه و کمینه مقادیر قدر-مطلق ضرایب iHS واریانت‌های ژنتیکی برای گوسفندان وحشی به ترتیب ۴/۸۵ و ۰/۰۰۳ برآورد شدند. آستانه انتخاب ۰/۹۹۹ واریانت‌های تحت انتخاب برای گوسفندان وحشی ۳/۰۵ بود (شکل ۴).

نتایج حاصل از روش آماری iHS برای گوسفندان اهلی و وحشی (شکل‌های ۳ و ۴) نشان داد که تقریباً در تمامی کروموزوم‌های گوسفندان اهلی نواحی ژنومی تحت انتخاب وجود دارند. بیشه و کمینه مقادیر قدرمطلق ضرایب iHS واریانت‌های ژنتیکی گوسفندان اهلی به ترتیب ۶/۴۲ و ۰/۰۰۹ برآورد شدند. آستانه انتخاب ۰/۹۹۹ واریانت‌های تحت



شکل ۳- توزیع ارزش‌های آماره iHS در سطح ژنگانی جمعیت گوسفندان اهلی با موقعیت کروموزومی نشانگرها روی محور X و ارزش‌های iHS استاندارد شده آن‌ها روی محور Y. خط ترسیم شده نشان‌دهنده آستانه انتخاب ۹۹/۹ درصد است.

Figure 3- Distribution of iHS statistic values at the genetic level of the domestic sheep population with the chromosomal position of SNPs on the X axis and their standardized iHS values on the Y axis. The drawn line indicates significance threshold of 99.9%.

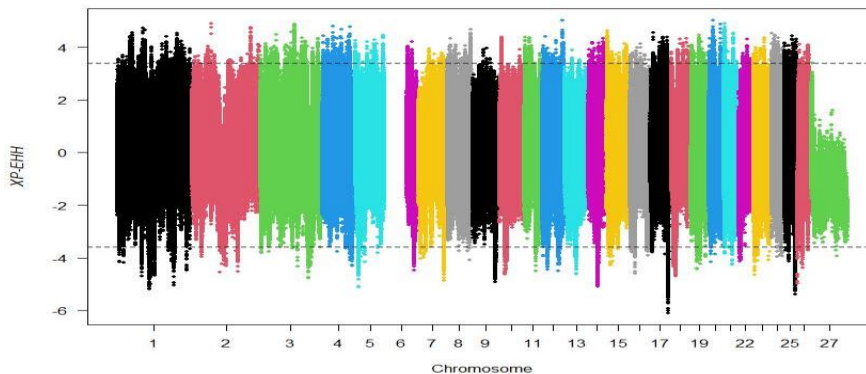


شکل ۴- توزیع ارزش‌های آماره iHS در سطح ژنگانی جمعیت گوسفندان وحشی با موقعیت کروموزومی نشانگرها روی محور X و ارزش‌های iHS استانداردشده آن‌ها روی محور Y. خط ترسیم شده نشان‌دهنده آستانه انتخاب ۹۹/۹ درصد است.

Figure 4- Distribution of iHS statistic values at the genetic level of the wild sheep population is shown with the chromosomal position of SNPs on the X axis and their standardized iHS values on the Y axis. The drawn line shows significance threshold of 99.9%.

انتخاب ۰/۰۰۱ برای گوسفندان وحشی ۴/۴۳- محاسبه شد. Eydivandi و همکاران (۲۰۲۰) در مطالعه تنوع ژنتیکی گوسفندان بومی ایران نشان دادند، به دلیل اثرات انتخاب در برنامه اهلی‌سازی گوسفندان ایرانی، فاصله ژنتیکی زیادی ناشی از اشتقاق نژادی بین گوسفندان اهلی و وحشی وجود دارد. نتایج شناسایی نواحی ژنومی تحت انتخاب مطالعه حاضر همانند مطالعه آنها نشان می‌دهد به لحاظ طول بلوک هاپلوتایپی بین گوسفندان اهلی و وحشی تفاوت وجود دارد.

نتایج حاصل از روش آماری XP-EHH بین گوسفندان اهلی و وحشی (شکل ۵) نشان داد، در گوسفندان اهلی (مقادیر مثبت) نسبت به گوسفندان وحشی تقریباً در تمامی کروموزوم‌ها نواحی ژنومی تحت انتخاب وجود دارند. مقادیر بیشینه و کمینه این ضریب به ترتیب ۴/۵۶ و ۰/۰۰۶ برآورد شدند. آستانه انتخاب ۰/۹۹۹ واریانت‌های تحت انتخاب برای گوسفندان اهلی ۳/۴۴ بود. در گوسفندان وحشی (مقادیر منفی) در کروموزوم‌های ۹، ۱۷ و ۲۵ بیشترین مقادیر ضریب XP-EHH برآورد شد. بیشینه و کمینه این ضریب به ترتیب ۰/۰۰۵- و ۵/۹۴- بود. آستانه



شکل ۵- توزیع مقادیر XP-EHH بین دو گونه اهلی و وحشی گوسفند. جمعیت بالای صفر مربوط به گونه اهلی و جمعیت پایین صفر مربوط به گونه وحشی. خط‌های ترسیم شده بالا و پایین به ترتیب نشان‌دهنده آستانه انتخاب ۹۹/۹ درصد در حیوانات اهلی و ۰/۰۱ درصد در حیوانات وحشی است.

Figure 5- Distribution of XP-EHH values between domestic and wild species of sheep. Above population (higher than zero) related to domestic species and bottom population (lower than zero) related to the wild type. The up and down drawn lines indicate significance threshold of 99.9% for domestic sheep and 0.01% for wild sheep, respectively.

۲۰۱۷). در نهایت برای گوسفند اهلی و وحشی به ترتیب ۱۷۰ و ۱۵۰ واریانت شناسایی شدند. برای تایید دامنه‌ی واریانس واریانت‌های انتخاب شده، واریانس ژنوتیپی واریانت‌های شناخته شده در نژادهای تکسل، سافوک، مرینوس و افشاری نیز برآورد شد. نتایج برآورد واریانس ژنوتیپی واریانت‌های شناخته شده، واریانس ژنوتیپی واریانت‌های انتخاب شده در این مطالعه را تأیید می‌کند (جدول ۲).

پس از برآورد ضرایب آماری iHS و XP-EHH برای دو جمعیت، ۴۵۱ واریانت در گوسفند اهلی و ۳۷۰ واریانت در گوسفند وحشی، انتخاب شدند. واریانس ژنوتیپی واریانت‌های ژنتیکی انتخاب شده در محدوده ۰/۰۵ تا ۰/۵۰ بودند. در مرحله بعد، واریانت‌های با واریانس ژنوتیپی بیش از ۰/۲۵ به عنوان واریانت‌های پیشنهادی برای آرایه گوسفندان ایرانی انتخاب و پیشنهاد شدند (Moradi و همکاران،

جدول ۲- واریانس ژنوتیپی واریانت‌های شناخته شده در نژادهای تکسل و سافوک و مریوس و افشاری و دیگر گوسفندان تحت مطالعه

Table 2- Genotypic variance of known SNPs in Texel, Suffolk, Merino, Afshari breeds and other studied sheep populations

نژاد	حداقل	چهارک اول	میانه	میانگین	چهارک سوم	حداکثر
Breed	Min	Q1	Median	Mean	Q3	Max
تکسل	0.04	0.22	0.38	0.33	0.47	0.50
مرینوس	0.02	0.31	0.42	0.38	0.48	0.50
سافولک	0.05	0.23	0.36	0.33	0.47	0.50
افشاری	0.03	0.27	0.41	0.36	0.48	0.50
گوسفندان اهلی	0.27	0.29	0.35	0.34	0.38	0.48
Domestic sheep						
گوسفندان وحشی	0.29	0.34	0.41	0.40	0.46	0.50
Wild sheep						

جدول ۳- فهرست ژن‌ها و صفات مرتبط در گوسفند اهلی

Table 3- List of genes and related traits in domestic sheep

نام ژن	صفت	
Gene name	Trait	
NTM, SLC7A2	افزایش وزن روزانه	
	Average daily gain	
LOC114116824, LOC101122398, FZD7, TIMD4, PCSK1, NTM, ADCYAP1	وزن بدن	
	Body weight	
NEGR1, ST6GALNAC3, LOC114116824, LOC101122398, LOC101102295, ADCYAP1	درصد چربی لاشه	وزن بدن و لاشه
	Carcass fat percentage	Body weight and carcass
FZD7, SLC39A11, ADCYAP1	وزن لاشه گرم	
	Hot carcass weight	
NEGR1, ST6GALNAC3, LOC114116824, LOC101122398, ADCYAP1	درصد گوشت لخم لاشه	
	Lean meat yield percentage	
NEGR1, ST6GALNAC3, LOC114116824, LOC101122398, SLC7A2	وزن عضله لاشه	
	muscle weight in carcass	
LOC114116824, LOC101122398, LOC101102295, FZD7, LOC101111617, MSRB3, SLC35F	درصد چربی شیر	
	Milk fat percentage	
LOC101102295, LOC101111617, SLC39A11, CD2AP, ADCYAP1	تولید چربی شیر	
	Milk fat yield	
LOC114116824, LOC101122398, LOC101102295, UBR1, LOC114118439	درصد پروتئین شیر	تولید شیر
	Milk protein percentage	Milk traits
LOC101111617, CD2AP	تولید پروتئین شیر	
	Milk protein yield	
LOC114116824, LOC101122398, LOC101111617, LOC105606356, CD2AP, ADCYAP1	تولید شیر	
	Milk Yield	
LOC105606356, CLIP1	امتیاز سلول‌های بدنی	
	Somatic cell score	
LOC101102295, LOC114117254, SMYD2	دفعات فحلی در سال	صفات تولید مثلی
	No. of estrus per year	Reproductive traits
LOC101113126	وزن پشم ناشور	
	Greasy fleece weight	پشم
LOC101113126, CDH23, CDH24, CDH25, CDH26, FRMPD2	میانگین قطر الیاف پشم	Wool
	Mean fiber diameter	

ژن VSNL1 در فرآیند تنظیم منفی ترشح هورمون‌ها نقش دارد. این ژن عضوی از زیرخانواده پروتئین‌های حس‌گر کلسیم عصبی است. پروتئین حاصل از بیان این ژن در سلول‌های گرانول مخچه بیان می‌شود و مسیرهای سیگنال درون سلولی سیستم عصبی مرکزی را با تنظیم فعالیت آدنیلیل سیکلاز تعدیل می‌کند. ژن FA2H نیز در تنظیم چرخه رشد مو، فعالیت اسید چرب آلفا هیدروکسیلاز، تمایز سلولی غدد چربی، تمایز سلول‌های شوان، فرآیند زیستی گلوکوزیل سرامید و نگهداری میلین سیستم عصبی مرکزی نقش دارد و برای عملکرد صحیح سیستم عصبی ضروری است (Hama, 2010). ژن FZD7 در تنظیم رشد اکتودرم و در بازسازی و نگهداری بافت ماهیچه اسکلتی نقش دارد (Aznar و همکاران، 2015).

با توجه به نتایج این مطالعه و بررسی جایگاه‌های صفات اقتصادی و ژن‌های مرتبط با واریانت‌های شناسایی شده، استفاده از واریانت‌های پیشنهادی در آرایه‌های تجاری می‌تواند قابلیت استفاده از آنها در گوسفندان اهلی را فراهم نماید. واریانت‌های ss1113654398, ss1113654985, ss1113654986, ss1113654988, ss1113654989, ss1113654994, ss1113654996, ss1113654997, ss1113654998, ss1113655010, ss1113655013, ss1113655015, ss1113655021, ss1113655024, ss1113655029, ss1113655039, ss1113655040, ss1113655044, ss1113655047, ss1113655068, ss1113655682, ss1113655917 بر روی کروموزوم یک و ss1123710180 بر روی کروموزوم ۲ به ترتیب با ژن‌های NEGR1 و FZD7 مرتبط هستند. این ژن‌ها صفاتی نظیر درصد چربی لاشه، درصد گوشت لخم لاشه، وزن لاشه گرم، وزن بدن و درصد چربی شیر را کنترل می‌کنند. مطالعات هستی‌شناسی ژن‌ها نشان داد،

با استفاده از پایگاه اطلاعاتی مرکز ملی اطلاعات زیست‌فناوری، مشخص گردید که واریانت‌های شناسایی شده در گوسفندان اهلی با ۳۳ ژن مرتبط بودند که این ژن‌ها نیز با ۱۵ صفت فیزیولوژیکی در ارتباط هستند. در گوسفندان وحشی نیز، واریانت‌های شناسایی شده با ۸ ژن مرتبط بودند که در عملکرد ۱۰ صفت تأثیرگذار هستند (جدول‌های ۳ و ۴).

در مطالعه بر روی ژن‌های مرتبط با واریانت‌های شناسایی شده در گوسفندان اهلی ژن MSR3B در فعالیت پپتید متیونین اکسید ردوکتاز درگیر است که در سوخت‌وساز اسیدآمینه متیونین نقش دارد. این اسید آمینه در سرعت وزن‌گیری، سلامت و ایمنی دام مؤثر است (Ahmed و همکاران، 2011, Kim & Gladyshev, 2004). ژن UBR1 در فرآیند تخریب پروتئین دخیل می‌باشد (Kwak و همکاران، 2004؛ Zenker و همکاران، 2005؛ Kume و همکاران، 2010؛ Matta-Camacho و همکاران، 2010). ژن SLC7A2 در فعالیت ناقل آرژنین نقش ایفا می‌کند و در تولید اکسید نیتریک در پاسخ التهابی و انتقال غشایی L-آرژنین نیز دخیل می‌باشد (Closs و همکاران، 1997). ژن CD2AP در تولید و تنظیم فاکتور رشد بتا و تنظیم اکتین در سلول‌های ماهیچه درگیر می‌باشد (Xie و همکاران، 2020). ژن CLIP1 در انتقال پروتئین به غشای سلول، تنظیم مثبت پلیمریزاسیون یا دپلیمریزاسیون میکروتوبول و فرآیندهای انتقال پروتئین به غشای پلازما دخیل است و در رشد سیستم عصبی نقش دارد (Larti و همکاران، 2015). ژن ADCYAP1 در اتصال گیرنده پلی‌پپتیدی فعال‌کننده آدنیلات سیکلاز هیپوفیز و فعالیت آن تأثیرگذار است. مطالعات انجام گرفته نشان می‌دهد که این ژن می‌تواند مستقیماً باعث افزایش ترشح دوپامین و در نتیجه مهار ترشح پرولاکتین در گوسفند شود (Anderson و همکاران، 2005).

رشد اکتودرم، بازسازی و نگهداری بافت ماهیچه‌های اسکلتی نقش دارند. به همین دلیل، این واریانت‌ها نسبت به سایر واریانت‌های شناسایی شده اهمیت بیشتری برای تبدیل شدن به نشانگر ژنومی برای گوسفندان اهلی ایران را دارند.

ژن FZD7 با چرخه‌های زیستی GO:2000383، GO:0014834، GO:0043403، GO:2000384، GO:0001705، GO:0001712، GO:0001715، GO:0042665، GO:0042666 ارتباط معنی‌داری ($P < 0.00001$) دارد. این چرخه‌های زیستی در تنظیم

جدول ۴- فهرست ژن‌ها و صفات مرتبط در گوسفند وحشی

Table 4- List of genes and related traits in wild sheep

نام ژن Gene name	صفت Trait
LOC101108006, CCDC190, LOC114109068	وزن بدن، تراکم استخوان، درصد چربی لاشه، وزن ماهیچه در لاشه، وزن استخوان در لاشه Body weight, Bone density, Carcass fat percentage, Muscle weight in carcass, Bone weight in carcass
UCK2	وزن بدن، تراکم استخوان، درصد چربی لاشه، وزن ماهیچه در لاشه، وزن استخوان در لاشه، کل نواحی ماهیچه ای بدن Body weight, Bone density, Carcass fat percentage, Muscle weight in carcass, Bone weight in the carcass, Total muscle areas of the body
USH2A	درصد چربی لاشه Carcass fat percentage
CIT, TMEM233	تراکم استخوان Bone density
YWHAZ	وزن لاشه گرم، وزن ماهیچه لاشه Warm carcass weight, Muscle weight in carcass

نتایج این مطالعه، ژن‌های UCK2 و CCDC190 بیشترین مشارکت را در بروز صفات در گوسفندان وحشی دارند.

مطالعات بر روی جایگاه صفات و ژن‌های مرتبط آنها در گوسفندان وحشی که با واریانت‌های انتخاب شده مرتبط هستند نشان داد، استفاده از واریانت‌های پیشنهادی در آرایه‌های تجاری می‌تواند قابلیت استفاده از آنها در گوسفندان وحشی را فراهم نماید. واریانت ss1151920929 بر روی کروموزوم ۱ مرتبط با ژن UCK2 و واریانت‌های ss1151829398، ss1151829412، ss1151829403، ss1151829401، ss1151829418، ss1151829425، ss1151829429، ss1151829444، ss1151829444، ss1151829447 بر روی کروموزوم ۱ مرتبط با ژن CCDC190 هستند. ژن‌های UCK2 و CCDC190 صفاتی نظیر وزن

در مطالعه بر روی ژن‌های مرتبط با واریانت‌های شناسایی شده در گوسفندان وحشی، ژن UCK2 از طریق کد کردن پروتئین پیریمیدین ریبونوکلوئوزید کیناز در فعالیت آنزیم نوکلئوزید کیناز دخیل است (Van Rompay و همکاران، ۲۰۰۱). ژن CIT در تنظیم مثبت سیتوکینز مؤثر می‌باشد و خاموشی این ژن، سیتوکینز و در نتیجه رشد سلولی را مختل می‌کند (Kawano و همکاران، ۲۰۱۴). ژن GRID1 نیز در کارایی سیستم عصبی از طریق تأثیر بر ساختار بنیادی غشای پس‌سیناپسی تأثیرگذار است (Nenadic و همکاران، ۲۰۱۲). ژن NTN4 در فرآیندهایی نظیر سیگنال‌دهی ماتریکس خارج سلولی، اتصال لامینین، مورفوژنز غدد بزاقی و بازسازی سلول‌های عصبی نقش دارد (Koch و همکاران، ۲۰۰۰). از عملکرد سایر ژن‌ها اطلاعاتی در دسترس نیست. با توجه به

مرتبط با صفات اقتصادی در جمعیت‌های گوسفند بومی ایران به‌منظور طراحی و ساخت آرایه برای این حیوانات بسیار ضروری به‌نظر می‌رسد. در تحقیق حاضر، بر اساس اطلاعات ژنوم گوسفندان اهلی و وحشی ایران، تعدادی نشانگر تک نوکلئوتیدی معنی‌دار مرتبط با صفات مهم اقتصادی و فیزیولوژیکی شناسایی شد که در آرایه‌های تجاری موجود در بازار لحاظ نشده‌اند. یافته‌های این مطالعه در کنار سایر مطالعات مشابه بر روی گوسفندان بومی کشور، می‌تواند در طراحی آرایه کم تراکم ویژه گوسفند ایرانی مورد استفاده قرار گیرد.

سپاسگزاری

این پژوهش با حمایت مالی دانشگاه فردوسی مشهد و گرنت شماره ۳/۴۶۴۰۹ انجام شد و از گروه علوم دامی دانشکده کشاورزی دانشگاه فردوسی برای در اختیار گذاشتن سخت افزار مورد نیاز جهت دانلود داده‌ها با حجم بالا و نیز آنالیز داده‌ها در اجرای این پروژه، تشکر و قدردانی می‌گردد.

بدن، تراکم استخوان، درصد چربی لاشه، وزن ماهیچه در لاشه، وزن استخوان در لاشه و کل نواحی ماهیچه-ای بدن را کنترل می‌کنند. نتایج هستی‌شناسی ژن‌ها نشان داد، ژن UCK2 با چرخه زیستی GO:0019206 ارتباط معنی‌داری ($P < 0.00001$) دارد. این چرخه زیستی با فعالیت نوکلئوزید کیناز مرتبط است. به همین خاطر این واریانت‌ها نسبت به سایر واریانت‌های پیشنهادی، برای تبدیل شدن به نشانگر ژنومی در گوسفند وحشی اهمیت بیشتری دارند.

نتیجه‌گیری کلی

در این بررسی، برای شناسایی نشان‌های انتخاب از روش‌های آماری iHS و XP-EHH استفاده شد و نشان‌های انتخاب در بخش‌هایی از ژنوم شناسایی شدند که در مطالعات پیشین بررسی نشده بود. بررسی ژن‌ها و جایگاه‌های موجود در نواحی تحت انتخاب نشان داد که این ژن‌ها با صفات مرتبط با ساختار بدنی، ایمنی، رشد، و تنظیم کارایی سیستم عصبی در ارتباط هستند. شناسایی جایگاه‌های تک نوکلئوتیدی

منابع

- Ahmed, Z.M., Yousaf, R., Lee, B.C., Khan, S.N., Lee, S., Lee, K., Husnain, T., UrRehman, A., Bonneux, S., Ansar, M., Ahmad, W., Leal, S.M., Gladyshev, V.N, Belyantseva, I.A., VanCamp, G., Riazuddin, S., Friedman, T.B & Riazuddin, S. (2011). Functional null mutations of MSRB3 encoding methionine sulfoxide reductase are associated with human deafness DFNB74. *The American Journal of Human Genetics*, 88: 19-29.
- Anderson, S., Kusters, D.H., Clarke, I.J., Pow, D.V., & Curlewis, J.D. (2005). Expression of pituitary adenylate cyclase activating polypeptide type 1 receptor (PAC1R) in the ewe hypothalamus: distribution and colocalization with tyrosine hydroxylase immunoreactive neurones. *Journal of Neuroendocrinology*, 17: 298-305.
- Aoki-Kinoshita, K., & Kanehisa, M. (2007). Gene annotation and pathway mapping in KEGG, *Methods in Molecular Biology*, 396: 71-91.
- Aznar, N., Midde, K.K., Dunkel, Y., Lopez-Sanchez, I., Pavlova, Y., Marivin, A., Barbazán, J., Murray, F., Nitsche, U., Janssen, K.P., Willert, K., Goel, A., Abal, M., Garcia-Marcos, M., & Ghosh, P. (2015). Daple is a novel non-receptor GEF required for trimeric G protein activation in Wnt signaling. *Elife*, 4: e07091.
- Bathke, J., & Lühken, G. 2021. OVarFlow: a resource optimized GATK 4 based open source variant calling work flow. *BMC Bioinformatics*, 22, 1-18.
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pagès, F., Trajanoski, Z., & Galon, J. (2009). ClueGO: a cytoscape plug-in to decipher

- functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, 25(8): 1091-1093.
- Closs, E.I., Gräf, P., Habermeier, A., Cunningham, J.M., & Förstermann, U. (1997). Human cationic amino acid transporters hCAT-1, hCAT-2A, and hCAT-2B: three related carriers with distinct transport properties. *Biochemistry*, 36: 6462-68.
- Djari, A.D., Esquerre, B., Weiss, F., Martins, C., Meersseman, M., Boussaha, C., Klopp, A., & Rocha D., 2013. Gene-based single nucleotide polymorphism discovery in bovine muscle using next-generation transcriptomic sequencing. *BMC Genomics*, 14(1): 307.
- Eydivandi, S., Sahana, G., Momen, M., Moradi, M.H., & Schönherz, A.A. (2020). Genetic diversity in Iranian indigenous sheep vis- à- vis selected exogenous sheep breeds and wild mouflon. *Animal Genetics*, 51(5), 772-787.
- Fan, B., Du, Z.Q., Gorbach, D.M. & Rothschild, M.F. (2010). Development and application of high-density SNP arrays in genomic studies of domestic animals. *Asian-Australasian Journal of Animal Science*, 23(7): 833-847.
- Gao, H. Wu, G., Spencer, T.E., Johnson, G.A., & Bazer, F.W. (2009). Select nutrients in the ovine uterine lumen. III. Cationic amino acid transporters in the ovine uterus and peri-implantation conceptuses. *Biology of Reproduction*, 80: 602-09.
- Gautier, M., Klassmann, A., & Vitalis, R. (2017). Rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Molecular Ecology Resources*, 17(1): 78-90.
- Goddard, M.E., & Hayes, B.J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programs. *Nature Reviews Genetics*, 10(6), 381-391.
- Groenen, M.A.M., Megens, H.J., Zare, Y., Warren, W.C., Hillier, L.W., Crooijmans, R.P.M. A., Vereijken, A., Okimoto, R., Muir, W.M. & Cheng, H.H. (2011). The development and characterization of a 60K SNP chip for chicken. *BMC Genomics*, 12(1):1-9.
- Hama, H. 2010. Fatty acid 2-Hydroxylation in mammalian sphingolipid biology. *Biochimica et Biophysica Acta*, 1801: 405-414.
- Helyar, S. J., Hemmer- Hansen, J., Bekkevold, D., Taylor, M.I., Ogden, R., Limborg, M.T., Cariani, A., Maes, G. E., Diopere, E., Carvalho, G.R., & Nielsen, E.E. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources*, 11: 123-136.
- Kawano, Y., Sahin, I., Moschetta, M., Wang, J., Manier, S., Glavey, S., Mishima, Y., Kokubun, K., Tsukamoto, S., Sacco, A., Anderson, K. Roccaro, A.M., Gray, N.S., & Ghobrial, I.M. (2014). Citron rho-interacting serine/threonine kinase (CIT) is a novel therapeutic target in multiple myeloma cells. *Blood*, 124: 3430.
- McCulloch, R., & McWilliam, S. (2009). A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. *PLoS ONE*, 4: 46-68.
- Kim, H.Y., & Gladyshev, V.N. (2004). Methionine sulfoxide reduction in mammals: characterization of methionine-R-sulfoxide reductases. *Molecular Biology of the Cell*, 15: 1055-64.
- Koch, M., Murrell, J.R., Hunter, D.D., Olson, P.F., Jin, W., Keene, D.R., Brunken, W.J., & Burgeson, R.E. (2000). A novel member of the netrin family, β -netrin, shares homology with the β chain of laminin: identification, expression, and functional characterization. *The Journal of Cell Biology*, 151: 221-34.
- Kume, K., Iizumi, Y., Shimada, M., Ito, Y., Kishi, T., Yamaguchi, Y., & Handa, H. (2010). Role of N- end rule ubiquitin ligases UBR1 and UBR2 in regulating the leucine- mTOR signaling pathway. *Genes to Cells*, 15: 339-49.
- Kwak, K.S., Zhou, X., Solomon, V., Baracos, V.E., Davis, J., Bannon, A.W., Boyle, W.J., Lacey, D.L., & Han, H. Q. (2004). Regulation of protein catabolism by muscle-specific and cytokine-inducible ubiquitin ligase E3 α -II during cancer cachexia. *Cancer Research*, 64: 8193-98.
- Lammers, G., Gilissen, C., Nillesen, S.T., Uijtdewilligen, P.J., Wismans, R.G., Veltman, J.A., Daamen, W.F. & Van Kuppevelt, T.H. (2010). High density gene expression microarrays

- and gene ontology analysis for identifying processes in implanted tissue engineering constructs. *Biomaterials*, 31(32): 8299-8312.
- Larti, F., Kahrizi, K., Musante, L., Hu, H., Papari, E., Fattahi, Z., Bazazzadegan, N., Liu, Z., Banan, M., Garshasbi, M., Wienker, T.F., Ropers, H., Galjart, N., & Najmabadi, H. (2015). A defect in the CLIP1 gene (CLIP-170) can cause autosomal recessive intellectual disability. *European Journal of Human Genetics*, 23: 331-36.
- Matta-Camacho, E., Kozlov, G., Li, F.F., & Gehring, K. (2010). Structural basis of substrate recognition and specificity in the N-end rule pathway. *Nature Structural & Molecular Biology*, 17: 1182-87.
- Mahal, G.S., Johnston, A., & Burns, R.H. (1951). Types and dimensions of fiber scales from the wool types of domestic sheep and wild life. *Textile Research Journal*, 21(2), 83-93.
- Matukumalli, L., Lawley, C., Schnabel, R., Taylor, J., Allan, M., Heaton, M., O'Connell, J., Moore, S., Smith, T., Sonstegard, T. & Van Tassell, C. (2009). Development and characterization of a high density SNP genotyping assay for cattle. *PLoS ONE*, 4: 4:e5350.
- Meuwissen, T.H., Hayes, B.J., & Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4), 1819-1829.
- Mohammadi, H., Rafat, S.A., Moradi Shahrehabak, H., Shodja, J. & Moradi, M.A. (2017). Genome-wide analysis for detection of loci under positive selection in Zandi sheep breed. *Iranian Journal of Animal Science*, 48(4): 533-548. (In Persian).
- Moradi, M.H., Farahani, A.H. & Nejati-Javaremi, A. (2017). Genome-wide evaluation of effective population size in some Iranian sheep breeds using linkage disequilibrium information. *Iranian Journal of Animal Science*, 48(1), 39-49. (In Persian).
- Nenadic, I., Maitra, R., Scherpiet, S., Gaser, C., Schultz, C.C., Schachtzabel, C., Smesny, S., Reichenbach, J.R., Treutlein, J., Mühleisen, T.W., Deufel, T., Cichon, S., Rietschel, M., Nöthen, M.M., Sauer, H., & Schölsser, R. G. (2012). Glutamate receptor delta 1 (GRID1) genetic variation and brain structure in schizophrenia. *Journal of Psychiatric Research*, 46(12), 1531-1539.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P. de Bakker, P.I.W., Daly, M.J., & Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3): 559-575.
- Ryder, M. L. (1964). The history of sheep breeds in Britain. *The Agricultural History Review*, 12(1), 1-12.
- Ryder, M. L. (1983). Sheep and man. London: Duckworth, 846 pp.
- Roshandel Ghalezo, M., Zerehdaran, S., & Javadmanesh, A. (2022). Selection signature and gene ontology for traits related to fat deposition in Afshari and Sunite fat-tailed sheep breeds. *Journal of Ruminant Research*.10: 47-62. (In Persian).
- Rupp, R., Mucha, S., Larroque, H., McEwan, J., & Conington, J. (2016). Genomic application in sheep and goat breeding. *Animal Frontiers*, 6(1):39-44.
- Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., & et al. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449(7164): 913-918.
- Sachajko, M., Herudzińska, M., Jaskowski, J.M., Szczepanek, J., Buszewska-Forajta, M., Feng, Y., Kumar D., & Pareek, C.S. (2019). Data set for transcriptome analysis of liver in cattle breeds. *Translational Research in Veterinary Science*, 2(2), 51-56.
- Sewe, S.O., Silva, G., Sicut, P., Seal, S.E., and Visendi, P. (2022). Trimming and validation of Illumina short reads using trimmomatic, trinity assembly, and assessment of RNA-seq data. *Methods in Molecular Biology*, 2443: 211-232.
- Taheri, S., Zerehdaran, S., and Javadmanesh, A. (2022). Genetic diversity in some domestic and wild sheep and goats in Iran. *Small Ruminant Research*, 209:106641.
- Tayara, H., and Chong, K.T. (2019). Improving the quantification of DNA sequences using evolutionary information based on deep learning. *Cells*, 8(12), 1635.

- Tang, K., Thornton, K.R., & Stoneking, M. (2007). A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology*, 5: 171.
- Tosser-Klopp, G., Bardou, P., Bouchez, O., Cabau, C., Crooijmans, R., Dong, Y., & et al. 2014. Design and characterization of a 52K SNP chip for goats. *PLoS ONE*, 9(1): 86227.
- VanRaden, P.M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11), 4414 - 4423.
- Voight, B.F., Kudravalli, S., Wen, X., & Pritchard, J.K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*, 4: 72.
- Wu, C., Li, J., Xu, X., Xu, Q., Qin, C., Liu, G., Wei, C., Zhang, G., Tian, K., & Fu, X. (2022). Effect of the FA2H gene on cashmere fineness of Jiangnan cashmere goats based on transcriptome sequencing. *BMC Genomics*, 23: 1-14.
- Wu, X.L., Xu, J., Feng, G., Wiggans, G.R., Taylor, J.F., He, J., Qian, C., Qiu, J., Simpson, B., Walker, J., and Bauck, S. (2016). Optimal design of low-density SNP arrays for genomic prediction: algorithm and applications. *PLoS ONE*, 11(9): e0161719.
- Xie, W., Chen, C., Han, Z., Huang, J., Liu, X., Chen, H., Zhang, T., Chen, S., Chen, C., Lu, M., Shen, X. & X. (2020). CD2AP inhibits metastasis in gastric cancer by promoting cellular adhesion and cytoskeleton assembly. *Molecular carcinogenesis*, 59: 339-352.
- Yang, J., Li, W.R., Lv, F.H., He, S.G., Tian, S.L., Peng, W.F., & et al. (2016). Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular Biology and Evolution*, 33(10), 2576-2592.
- Zenker, M., Mayerle, J., Lerch, M.M., Tagariello, A., Zerres, K., Durie, P.R. & et al. (2005). Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). *Nature Genetics*, 37: 1345-1350.
- Zeraatpisheh, Y., Zerehdaran, S., & Javadmanesh, A. (2023). Investigation of metabolic pathways related to the QTLs of parasite resistance trait in sheep genome using gene network and gene ontology. *Veterinary Research and Biological Products*, 36: 83-90. (In Persian).

