

Genome wide association study based on pathway analysis relate to wool traits in Zandi sheep breed

ABSTRACT

Identifying of genomic regions and major genes with large effects on economically important traits, has been one of the important goal to sheep breeding. The present study aimed to conduct a genome wide association studies (GWAS) based on pathway analysis for identifying the loci associated with wool quality traits between of coarse wool Zandi sheep (96 Animal) and 240 fine wool Rambouillet sheep (240 Animal) were used. After quality control steps, Genome wide association study was performed using GEMMA software. Using the *biomaRt2* R package the SNP were assigned to genes and gene enrichment analysis was performed with the KOBAS platform. We identified different sets of candidate genes (Biological function) related to wool traits: *TMTC3* (keratinocytes production), *POUIF1* and *NLGN1* (keratinocytes differentiation), *PLCE1* (development of epidermal), *SPHKAP* (hair follicle development), *LOC101117971* and *LOC101118971* (insulin secretion and prolactin signaling). According to pathway analysis, 15 pathways from gene ontology and biological pathways were associated with the wool traits ($P < 0.05$). Some of the biological pathways related to regulation pathways of intracellular transfer proteins and epidermal cell differentiation played an important role in the growth of hair follicles. However, it is necessary to further investigate the genes related to the genomic regions obtained from the genome scan studies. Using these findings could potentially be useful for genetic selection in the breeding programs and can be used to understand the genetic mechanism controlling this trait.

Keywords: Candidate gene, Fiber diameter, Genome Scan, Staple length, Sheep

مطالعه پویش ژنومی مرتبط با صفات کیفی پشم – مقایسه بین گوسفندان نژاد پشم ضخیم با پشم ظریف

چکیده

شناسایی مناطق ژنومی و ژن‌های بزرگ اثر مؤثر بر صفات مهم اقتصادی برای افزایش بازده انتخاب و بهبود عملکرد تولیدی از مهم‌ترین اهداف اصلاح نژادی در پرورش گوسفند است. هدف از این پژوهش، شناسایی مناطق ژنومی و ژن‌های کاندیدای مرتبط با صفات کیفی پشم با مقایسه بین گوسفندان نژاد زندی پشم ضخیم (۹۶ رأس) و گوسفندان نژاد رامبویه پشم ظریف (۲۴۰ رأس) بوسیله مطالعه پویش ژنومی و آنالیز مسیر بود. پس از اجرای مراحل مختلف کنترل کیفیت داده‌ها، ارزیابی پویش ژنومی برای صفات مورد بررسی در نرم‌افزار GEMMA انجام شد. آنالیز مسیر با استفاده از بسته نرم افزاری *biomaRt2* ژن‌های معنی‌داری شناسایی گردید و تفسیر مجموعه ژنی بوسیله برنامه KOBAS با هدف شناسایی عملکرد بیولوژیکی ژن‌های نزدیک به مناطق انتخابی و ژن‌های کاندیدا انجام شد. در این پژوهش ژن‌هایی با (عملکردهای بیولوژیکی) مرتبط با صفات مورد مطالعه شامل *TMTC3* (تولید کراتینوسیت‌ها)، *POUIF1*، *NLGN1* (رشد و تفرق کراتینوسیت‌ها)، *PLCE1* (رشد و توسعه اپیدرمال)، *SPHKAP* (رشد و توسعه فولیکول‌های مو)، *LOC101117971* و *LOC101118971* (سیگنال‌دهی پرولاکتین و ترشح انسولین) شناسایی شدند. در آنالیز غنی‌سازی مجموعه ژنی تعداد ۱۵ مسیر هستی شناسی ژنی و مسیر KEGG با صفات الیاف مرتبط بودند ($P < 0.05$). از این بین، مسیرهای تنظیم پروتئین‌های انتقالی داخل سلولی و تفرق سلولی اپیدرمال نقش مهمی در رشد فولیکول‌های مو و تحریک کلاژن داشتند. به هر حال بررسی بیشتر ژن‌های مرتبط با مناطق ژنومی بدست آمده از مطالعات جستجوی پویش ژنومی ضروری است. استفاده از یافته‌های این تحقیق می‌تواند باعث تسریع در پیشرفت ژنتیکی برنامه‌های اصلاح نژادی گوسفند با هدف الیاف ظریف‌تر شود.

کلید واژه‌ها: پویش ژنوم، ژن کاندیدا، طول استاپل، قطر الیاف، گوسفند

مقدمه

در برنامه‌های اصلاح نژادی گوسفند جهت بهبود تولید پشم و کیفیت آن صفاتی مانند وزن پشم، میانگین قطر الیاف، طول دسته الیاف، نسبت کمپ و ضریب تنوع آنها مهمترین ملاک‌های انتخاب هستند (Zhao *et al.*, 2021). نرخ پیشرفت در صفات پشم نسبت به سایر صفات از جمله رشد سرعت کندی داشته است. اکثر صفات مرتبط با خصوصیات کمی و کیفی پشم دارای وراثت پذیری متوسط و بالا هستند، در نتیجه اگر در برنامه‌های اصلاح نژادی ملاک انتخاب قرار گیرند، می‌توان شاهد تغییرات قابل ملاحظه‌ای در این صفات در طی نسل‌های آینده بود (Wang *et al.*, 2014). شاخص‌های اصلاح نژادی برای صفات مرتبط با کمیت و کیفیت پشم به ترتیب روی دو صفت وزن پشم شسته شده¹ و قطر الیاف پشم گوسفندان نژاد مرینوس در کشور استرالیا انجام شده است، بطوریکه انتخاب باعث کاهش قطر الیاف پشم این نژاد از ۱۹/۷ به ۱۸/۳ میکرومتر شده بود (Taylor *et al.*, 2017). ولی مشکلاتی وجود دارد که روند انتخاب و بهبود صفات بر پایه داده‌های فنوتیپی را کند می‌نماید. اندازه‌گیری صفات پشم مانند استحکام الیاف یا اندازه‌گیری آنها مشکل بوده و یا پرهزینه است. همچنین به خاطر همبستگی‌های ژنتیکی منفی بین بعضی صفات تولیدی پشم با صفات تولید مثل و رشد موجب شده است تا تمایل به مطالعه ژن‌های کاندیدا و مکان‌های ژنومی کنترل کننده صفات کمی (QTL) در بین محققین برای صفات پشم افزایش پیدا کند و از این طریق تولید پشم و صفات کیفی پشم تولیدی بهبود پیدا نمایند (Ren *et al.*, 2016).

با توجه به بررسی منابع، تا به امروز براساس آخرین نسخه AnimalQTLdb (نسخه ۴۹) (<https://www.animalgenome.org/cgi-bin/QTLdb/OA/index/OA/index>, 28 Dec, 2022)، تحقیقات اندکی برای شناسایی QTL صفات کمی و کیفی پشم در گونه گوسفند گزارش شده است. به طوریکه تنها ۱۹۷ QTL مؤثر بر فاکتورهای مختلف تأثیرگذار بر صفات کمی و کیفی پشم انجام شده است که عمده مطالعات مکان‌یابی QTLها براساس نشانگرهای ریزماهورهای تمرکز یافته است که فاصله نشانگرها بسیار طولانی (چندین cM) و در نتیجه فاصله اطمینان و صحت QTLهای شناسایی شده پایین می‌باشد (Hu *et al.*, 2022).

مطالعات پویش کل ژنومی به منظور شناسایی ارتباط نشانگرهای SNP و یک صفت با استفاده از تراشه‌های با تراکم بالا در سطح ژنوم انجام می‌شود و هدف آن پیدا کردن جهش‌های مؤثر یا سببی بر فنوتیپ یک صفت تولیدی، تولیدمثل و یا بیماری می‌باشد. این اطلاعات می‌تواند برای انتخاب به کمک نشانگر مفید بوده و به درک معماری ژنتیکی صفات مورد مطالعه کمک نماید (Gebreselassie *et al.*, 2019). در مطالعات پویش کل ژنومی به طور معمول از تصحیح بنفرونی² برای تعیین آستانه معنی داری و جلوگیری از خطای نوع اول استفاده می‌گردد. یکی از ایرادات تحقیقات مطالعات پویش ژنومی در نظر گرفتن آستانه معنی‌داری برای جلوگیری از بروز خطای نوع اول است. در حالی که پرهیز از خطای نوع اول سبب افزایش خطای نوع دوم و در نظر نگرفتن SNPهای دارای اثر معنی‌دار پایین‌تر از آستانه می‌شود (Marjanovic & Calus, 2020).

یک جایگزین مناسب برای حل این مشکل، انجام مطالعات پویش کل ژنومی بر مبنای مسیر³ با استفاده از تجزیه و تحلیل مجموعه‌های ژنی است (Wang *et al.*, 2011). در واقع در این روش به جای انجام تجزیه برای یک SNP یا یک ژن، ارتباط بین صفت مورد مطالعه و واریانت‌های ژنتیکی در یک دسته یا گروه ژنی که به طور عملکردی با هم مرتبط هستند بررسی می‌شود. به عبارت دیگر ارتباط آماری بین یک مجموعه ژنی معنی‌دار با فنوتیپ، مورد آزمون قرار می‌گیرد (Wang *et al.*, 2011).

پس از اینکه در سال ۲۰۰۹ اولین تراشه SNP در گوسفند در قالب پروژه Sheep HapMap طراحی شد، تحقیقات موفقیت آمیز مختلفی با هدف شناسایی مناطق ژنومی و ژن‌های کاندیدای مرتبط با صفات پشم در گوسفند انجام شده است. اولین مطالعه روی گوسفندان مرینوس با شناسایی ژن‌های کاندیدای مؤثر بر صفات کیفی به طور موفقیت آمیزی انجام و ژن‌های *YWHAZ*، *KRTCAP3*، *TSPEAR*، *PIK3R4*، *KIF16B*، *PTPN3*، *GPRC5A*، *DDX47*، *TCF9*، *EPHA5*، *TPTE2* و *NBEA* به عنوان کاندیدا گزارش شدند (Wang *et al.*, 2014).

¹ Clean fleece weight

² Bonferroni correction

³ Pathway-based analysis

مطالعه پویش ژنومی بر مبنای مسیر با استفاده از تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی روی خصوصیات لخته‌شدگی شیر، تولید پنیر و استحکام دلمه انجام شد. براساس این نتایج خروجی‌ها براساس مسیر، منجر به شناسایی ۲۱ طبقات مختلف عملکردی هستی‌شناسی ژن و ۱۷ مسیر بیوشیمیایی KEGG معنی‌دار مرتبط با این صفات شد که شامل ژن‌های کاندیدای *CSN1S1*، *LAP3*، *DGATI*، *RPL8* و *HSF1* بودند (Dadousis et al., 2017). همچنین در مطالعه‌ی پویش ژنومی روی صفات مرتبط با پشم تولیدی در سن یکسالگی گوسفندان نژاد مرینوس چینی انجام شده است. ژن‌های کاندیدای *RHPN2*، *UBE2E3*، *BMP1B*، *ESR1*، *SMAD1*، *PLCB1*، *SMAD2*، *INSR*، *PTGS2* و *KCNMA1* مرتبط با صفات کمی و کیفی الیاف شناسایی شده بود (Zhao et al., 2021).

پژوهشی بر پایه تجزیه و تحلیل غنی‌سازی مجموعه ژنی برای شناسایی جایگاه‌های ژنی مؤثر بر نرخ دو قلوزایی در گوسفندان بلوچی انجام شده بود، موفق به شناسایی ژن‌های کاندیدای *CTH*، *ANKRD13C*، *SRSF11*، *PTGER3*، *LDHB* و *LRRC40* و *NTRK2* شدند. نتایج نشان داد روش پویش ژنومی بر مبنای مسیر کارایی بالاتری برای یافتن مناطق ژنومی نسبت به آنالیز پویش ژنومی بر پایه تک‌نشانگری داشت (Esmaeili-Fard et al., 2021). اخیراً، مطالعه پویش کل ژنومی در گوسفندان Akkaraman ترکیه‌ای با هدف شناسایی ژن‌های کاندیدا و مسیرهای زیستی مرتبط با صفات پشم انجام شده است. نتایج این پژوهش ژن‌های کاندیدای *TRIM2*، *MND1*، *JLR2*، *RNF175*، *CEP290*، *TMTC3*، *RERE*، *SOX2*، *MORN1*، *RPM2*، *DMRT1* و *CAPSL* و *PRKCZ* را گزارش نمودند (Arzik et al., 2023).

گوسفند زندی یکی از نژادهای متوسط وزن و ممتاز کشور از نظر صفات تولیدی می‌باشد. کیفیت پشم تولیدی در این نژاد از نوع پشم ضخیم یا پشم قالی می‌باشد. مقاومت در برابر سرما و گرما (از ۲۰- درجه سانتیگراد در زمستان تا ۴۰ درجه سانتیگراد در تابستان)، شرایط سخت تغذیه‌ای و مراتع ضعیف و هوای بارانی و برفی در این گوسفند بسیار شایان توجه است (Ghafouri Kesbi et al., 2008). همچنین گوسفندان نژاد رامبویه یک نژاد غالب پرورشی در کشور آمریکا و سازگار یافته در سیستم‌های مرتعی مناطق خشک و نیمه‌خشک است. این نژاد بزرگترین نژاد دو منظوره با اهداف تولید پشم با قطر الیاف کمتر از ۲۵ میکرون و گوشت می‌باشد (Lupton et al., 2007). هدف از انجام پژوهش حاضر، پویش کل ژنومی و تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی و مسیرهای زیستی مرتبط با صفات کیفی پشم در مقایسه بین نژادهای گوسفند بومی زندی با نژاد رامبویه براساس پویش کل ژنوم بود.

مواد و روش‌ها

از ۲۰۰ گوسفند نژاد زندی حاضر در دو گله مرکز اصلاح نژاد خجیر واقع در تهران وابسته به سازمان جهاد کشاورزی استان تهران، مقدار ۵-۷ سی‌سی خون استحصال و همراه با ۰/۵ میلی لیتر EDTA در لوله‌های خلاء با pH ۸-۷/۵ نگهداری شد. از میان نمونه‌های اخذ شده پس از آنالیز شجره، حیواناتی که کمترین رابطه خویشاوندی را با هم داشته باشند به تعداد ۹۶ رأس انتخاب شدند. استخراج DNA با استفاده از روش بهینه یافته استخراج نمکی از خون کامل انجام شد. ژنوتیپ نمونه‌ها در کمپانی Neogen (<http://genomics.neogen.com>) در کشور آمریکا با استفاده از آرایه‌های Illumina OvineSNP50K BeadChip با بکارگیری پروتکل استاندارد شرکت ایلومینا (Illumina Inc., San Diego, CA, USA) تعیین ژنوتیپ شده بودند (Mohammadi, 2017). از تمام دام‌های انتخاب شده قبل از پشم‌چینی سالیانه نمونه برداری پشم انجام گرفت. ابتدا از هر نمونه ۳ استاپل به طور تصادفی جدا شده و طول هر یک به کمک خط کش و بر حسب سانتی‌متر مطابق با روش استاندارد شماره ۱۹۴۱ ایران اندازه‌گیری شد. سپس میانگین طول ۳ استاپل محاسبه و با تقریب ۲ میلی‌متر برای هر نمونه گزارش شد. نمونه‌های پشم گرفته شده در آزمایشگاه با مواد شوینده شسته شده و صفات کیفی الیاف طبق پروتکل استاندارد A.S.T.M شماره D2130-90 اندازه‌گیری شد. برای اندازه‌گیری میانگین قطر (MFD)، به طور تصادفی ۲ تا ۳ دسته الیاف از داخل نمونه بیرون کشیده و از قسمت میانی آنها به طول حدود ۱۵۰ تا ۲۰۰ میکرومتر برش‌هایی زده و آنها را با استفاده از گلیسرین بر روی لام تثبیت نمودیم و با استفاده از میکروسکوپ پروژکتینا قطر تارهای پشم بر حسب میکرومتر اندازه‌گیری شد. میانگین قطر الیاف هر نمونه بر حسب ۴۰۰ تار پشم محاسبه گردید.

همچنین در پژوهش حاضر از اطلاعات ژنوتیپی و رکوردهای فنوتیپی مرتبط با صفات میانگین قطر الیاف و طول استایل، ۲۴۰ رأس گوسفندان غیر خویشاوند مربوط به نژاد رامبویه با پشم ظریف کشور آمریکا استفاده گردید. داده‌های مورد استفاده در این پژوهش در پایگاه برخط EVA (<https://www.ebi.ac.uk/eva/>) به شماره‌های دسترسی (RJE58836 و ERZ15609617) در دسترس می‌باشد. آمار توصیفی صفات کیفی پشم مورد بررسی در (جدول ۱) نشان داده شده است.

جدول ۱. آمار توصیفی صفات کیفی پشم مورد بررسی در نژادهای مورد مطالعه

نژاد	صفت
رامبویه	میانگین قطر الیاف (μm)
۲۲/۶۴	۲۹/۸۵
۱/۵۱	۳/۲۵
۱۹/۰۱	۲۲/۴۰
۲۷/۲۰	۳۹/۰۴
۰/۰۸	۰/۳۲
۱۲/۰۷	۱۱/۲۵
۱/۴۳	۳/۹۲
۷/۸۰	۶/۰۰
۱۷/۷۷	۱۹/۰۰
۰/۰۷	۰/۴۰

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 نشانگرهایی که فراوانی آلل نادر در آنها کمتر از ۱ درصد بود حذف شدند. سپس نشانگرهایی که نرخ تعیین ژنوتیپ آنها در نمونه‌ها کمتر از ۹۰ درصد بود شناسایی و حذف شدند. در نهایت برای SNPهای باقیمانده آنهایی که بر روی کروموزوم‌های جنسی بودند و یا در تعادل هاردی-واینبرگ قرار نداشتند کنار گذاشته شدند. بدین منظور سطح احتمال 10^{-6} در نظر گرفته شد. مراحل مختلف فیلتراسیون با استفاده از نرم افزار PLINK (نسخه ۱/۹۰) انجام شد (Purcell et al., 2007). آنالیز مؤلفه‌های اصلی (PCA) پس از گذراندن مراحل کنترل کیفیت داده‌ها با استفاده از نرم افزار PLINK V1.9 محاسبه و گراف مربوطه توسط برنامه R نسخه (۴/۲/۱) ترسیم شد تا علاوه بر تفکیک نژادها، دام‌هایی که خارج از توده نژادی خود قرار دارند، برای حذف شناسایی شوند.

آنالیز پویش ژنومی بر پایه مدل خطی مختلط

جهت ارتباط فنوتیپ‌ها با ژنوتیپ‌ها از نرم افزار GEMMA (Zhou & Stephens, 2012) نسخه ۰/۹۸ استفاده شد. مدل مورد استفاده بر پایه مدل خطی مختلط بر پایه رابطه ۱ بود:

$$y = Wa + x\beta + Zu + e \quad \text{رابطه ۱}$$

که y : بردار مشاهدات فنوتیپی، a بردار اثرات ثابت شامل جنس، سال تولد، تیپ تولد و سال نمونه‌برداری، x : بردار ژنوتیپ‌های نشانگر تحت آزمون، u بردار اثرات پلی‌ژنیک، e بردار اثر باقی‌مانده‌های تصادفی و W ، β و Z ماتریس‌های طرح ارتباط دهنده. مقدار P-value با روش تصحیح بنفرونی با سطح معنی‌داری پنج درصد که برای عدم تعادل پیوستگی در سطح ژنوم تصحیح شده بود، محاسبه شد. به منظور تعیین سطح آستانه P-value تعداد SNP مستقل بر روی کروموزوم‌های اتوزومی محاسبه شد. برای این کار، از مجموعه SNPهای مجاور با مقادیر r^2 بالاتر از ۰/۲ استفاده شد در مجموع ۱۱۲۵۴ نشانگر SNP مستقل به دست آمد.

دو سطح آستانه احتمال شامل سطح احتمال معنی‌داری ($10^{-6} \times 4/442$) و سطح احتمال پیشنهادی ($10^{-5} \times 8/851$) برای P-value در سطح ژنوم مشخص شد. براساس نتایج P-value، پلات منهتن و پلات Q-Q برای صفات مورد بررسی با استفاده از بسته qqman نرم‌افزار R نسخه ۴/۲/۱ ترسیم شد.

آنالیز پویش ژنومی بر پایه تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی (GSEA)

آنالیز پویش ژنومی بر پایه تجزیه و تحلیل مجموعه‌های ژنی در سه مرحله انجام گردید:

۱- تعیین مکان SNPها با ژن‌ها: SNPهایی که مقدار P-value آنها کمتر از ۰/۰۵ بود با استفاده از بسته نرم افزاری *biomaRt2* (Durinck et al., 2009) در محیط R و با استفاده از ژنوم مرجع گوسفند (*Oar_v4.0*) به ژن‌هایی که نشانگر SNP مورد نظر در داخل آن ژن و یا ۵۰kb بالادست یا پایین دست آن ژن قرار داشت، ارتباط داده شدند (Clancey et al., 2019).

۲- ارتباط ژن‌ها به طبقات عملکردی و مسیرهای بیوشیمیایی: جهت تعیین طبقات عملکردی ژنی و مسیرهای متابولیکی و تنظیمی ژن‌های معنی‌دار از پایگاه‌های اطلاعاتی شامل GO (<http://www.geneontology.org>)، KEGG (<http://www.genome.jp/kegg>)، Panther (<http://www.pantherdb.org>) و Reactome (<http://www.reactome.org>) جهت تعیین طبقات عملکردی و مسیرهای بیوشیمیایی استفاده می‌نماید. علاوه بر این از پایگاه برخط DAVID (<http://david.abcc.ncifcrf.gov>) برای تعیین بهتر دست‌های ژنی استفاده گردید. در این مرحله فرض بر این است که ژن‌هایی که در یک طبقه عملکردی خاص از سه مجموعه ژنی هستی‌شناسی قرار می‌گیرند می‌توانند به عنوان یک گروه از ژن‌هایی که برخی از ویژگی‌های خاص و مشترک دارند.

۳- پویش کل ژنومی بر پایه آنالیز مسیر: ارتباط‌های معنی‌دار مسیرهای عملکردی مرتبط با صفات تولیدمثلی با استفاده از توزیع فوق هندسی^۴ و آماره Fisher's exact test مورد آزمون قرار گرفت. تجزیه غنی‌سازی مجموعه ژنی با استفاده از برنامه برخط KOBAS (Bu et al., 2021) انجام گردید. در نهایت برای تفسیر بهتر عملکرد ژن‌های به دست آمده از پایگاه‌های اطلاعاتی آنلاین GeneCards (<http://www.genecards.org>) و UniProtKB (<http://www.uniprot.org>) استفاده شد.

نتایج و بحث

نتایج مربوط به مراحل مختلف ویرایش داده‌های تعیین ژنوتیپ در جدول ۲ ارائه شده است. تعداد دو حیوان از نژاد زندگی خارج از توده نژادی خود در آنالیز PCA قرار داشتند که به منظور جلوگیری از ایجاد اربیبی در محاسبات حذف شد. پس از حذف داده‌های حیوانات حذف شده، داده‌های نژاد بومی با نژاد گوسفند رامبویه با هم ترکیب شدند که دارای ۳۱۷۹۴ نشانگر SNP مشترک بودند که پس از مرحله کنترل کیفیت در سطح SNP در نهایت، تعداد ۲۹۴۳۵ نشانگر SNP باقی ماندند که شامل اطلاعات ۲۶ جفت کروموزوم گوسفند بودند (جدول ۳).

جدول ۲. توصیف مراحل کنترل کیفیت داده‌های مربوط به نژادهای مختلف مورد بررسی در سطح افراد

رامبویه	زندگی	
۲۴۰	۹۶	تعداد حیوانات
۰	۰	حذف نمونه‌هایی با بیش از ۱۰ درصد ژنوتیپ از دست رفته
۰	۰	حذف نمونه‌هایی با بیش از IBS بالاتر از ۰/۹۹
۰	۲	حذف حیواناتی که در آنالیز PCA خارج از توده جمعیتی خود قرار داشتند
۲۴۰	۹۴	تعداد کل نمونه‌های باقی‌مانده

جدول ۳. خلاصه مراحل مختلف کنترل کیفیت نشانگرهای SNP برای آنالیزهای پویش ژنومی

تعداد نشانگر	مراحل مختلف کنترل کیفی
۳۱۷۹۴	تعداد کل نشانگرهای مورد بررسی
۱۵۲۸	حذف نشانگرها با حداقل فراوانی آلی کمتر از یک درصد

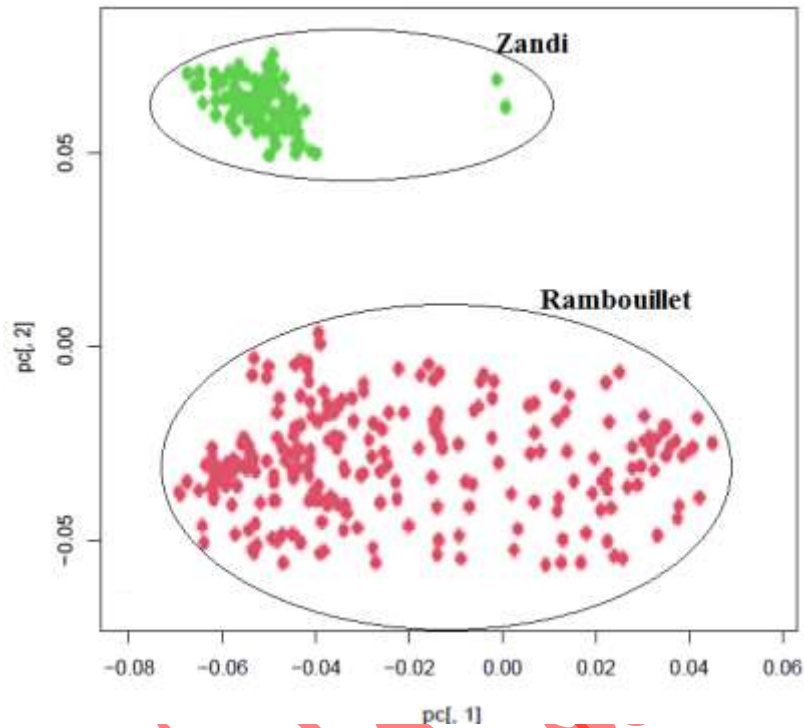
⁴Hypergeometric

۷۹۵	حذف نشانگرهای با نرخ فراخوانی کمتر از ۹۰ درصد در هر نمونه
۳۶	حذف نشانگرها با انحراف از تعادل هاردی-واینبرگ (10^{-6})
۲۹۴۳۵	تعداد نشانگرهای باقی مانده

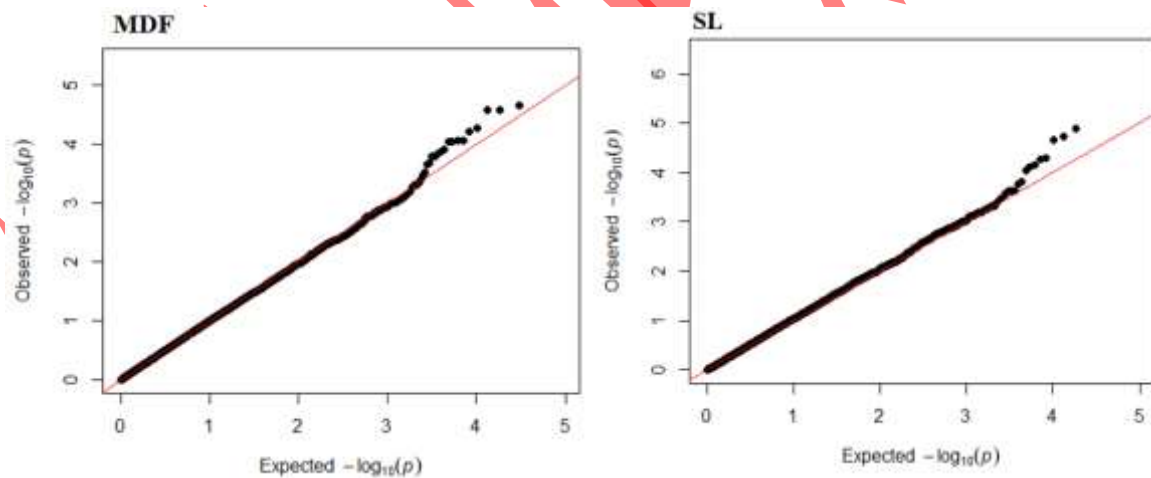
به منظور بررسی نحوه قرار گرفتن حیوانات در گروه‌های نژادی، تجزیه و تحلیل PCA با بکارگیری اطلاعات تعیین ژنوتیپی بین نژاد بومی زندی و نژاد رامبویه به عنوان راهی متمایز به منظور توصیف واگرایی مورد تجزیه و تحلیل قرار گرفت. نتایج آنالیز PCA نشان داد که با PC1 نژاد زندی و PC2 نژاد رامبویه به خوبی از یکدیگر تفکیک و جدا شدند و در هیچ نقطه‌ای با هم همپوشانی ندارند (شکل ۱). این PCA در مجموع حدود ۵/۷٪ واریانس (۳/۹٪ برای ویژه بردار ۱ و ۱/۸٪ برای ویژه بردار ۲ را توجیه می‌کنند). در نژادهای مختلف دیگر گوسفند، حیوانات فقط برپایه ویژه بردارهای یک و دو براساس نژاد و منطقه جغرافیایی دسته‌بندی شده‌اند (Kijas et al., 2012; Waaineina et al., 2022). در مطالعه‌ای که توسط Kijas et al. (2012) بر روی جمعیت‌های گوسفند دنیا انجام شده بود، PC1 و PC2 به ترتیب ۲/۹۸٪ و ۱/۴۴٪ واریانس را توجیه می‌کردند. در مطالعه پوشش ژنومی با هدف شناسایی نشانه‌های انتخاب مرتبط با صفات ذخیره بدن در سه نژاد گوسفندان بومی چینی (Tibetan, Altay, Han)، حیوانات برپایه PC1 و PC2 کاملاً از یکدیگر تفکیک شده و هر کدام از ویژه بردارهای ۱ و ۲ به ترتیب ۳/۳٪ و ۲/۷٪ از واریانس را توجیه می‌کردند (Zhao et al., 2020).

یکی از نکات مورد توجه در اجرای این پژوهش تعداد کم نمونه‌ها جهت انجام تحقیقات پیوستگی در سطح ژنوم بود. قبل از اجرای آنالیزهای پیوستگی با در نظر گرفتن اثرات ثابت معنی‌دار بر روی هر صفت، فاکتور تورم کنترل جمعیتی (لامبدا) در آنالیزهای پیوستگی محاسبه شد. فاکتور لامبدا از طریق آنالیز پیوستگی در PLINK برای هر دو صفت مورد مطالعه محاسبه شد که برای صفت میانگین قطر الیاف ۱/۰۱۶ و صفت طول استاپل ۱/۰۴۲ به دست آمد. همان طور که مشاهده می‌شود صفات مورد بررسی تقریباً برابر با ۱ بودند و نشان از عدم وجود لایه بندی جمعیتی برای اجرای آنالیزهای پیوستگی بود.

پلات‌های Q-Q و منهن برای دو صفت کیفی پشم مورد مطالعه در شکل‌های ۲ و ۳ ارائه شده است. در مجموع نتایج حاصل از این بخش نیز نشان دهنده عدم لایه‌بندی جمعیتی در مجموع حیوانات مورد مطالعه در این تحقیق بود. در تحقیقی که توسط روی صفات تولید پشم صورت گرفت نیز از اطلاعات نژاد مرینوس چینی استفاده شد و بررسی لایه بندی جمعیت با استفاده پلات‌های Q-Q هیچ لایه‌بندی‌ای را در تحلیل داده نشان نداد (Wang et al., 2014).



شکل ۱. خوشه‌بندی حیوانات براساس آنالیز PCA با استفاده از اطلاعات ژنوتیپی تمام حیوانات



شکل ۲. پلات‌های Q-Q برای صفات میان قطر الیاف (MFD) و طول استاپل (SL)

نتایج آنالیز GWAS برای صفت میانگین قطر الیاف با استفاده از مدل خطی مختلط پس از تصحیح بنفرونی وجود دو نشانگر SNP را در سطح احتمال پیشنهادی بر روی کروموزوم‌های شماره ۲ (ناحیه ۸۸۷۳۵۷ جفت بازی) و ۴ (ناحیه ۶۱۹۰۷۷۳۱ جفت بازی) را نشان داد (شکل ۲). با بررسی بیوانفورماتیکی این مناطق ژنومی مشخص گردید که ژن‌های موجود در این مناطق ارتباط مستقیمی با صفات کیفی پشم نداشتند، بطوریکه نقش این ژن‌ها بیشتر با صفات رشد و استرس حرارتی مرتبط بودند. همچنین برای صفت طول استاپل هیچ نشانگری در هر دو سطح آستانه معنی‌داری شناسایی نشد.

شناسایی ژن‌های کاندیدا در مناطق ژنومی حاصل از تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی

تعداد ۲۵۶۵۸ عدد ژن از ۲۷۰۵۴ ژن حاشیه‌نویسی شده در گوسفند به وسیله نشانگرهای SNP پوشش داده شد که در این میان تعداد ژن‌های معنی‌دار در این بررسی ۶۵۱ ژن بودند، یعنی حداقل یک نشانگر با P-value کمتر از ۰/۰۵ در داخل و یا در بالا یا پایین دست این ژن تا فاصله ۵۰ kb قرار گرفت. این ژن‌ها به‌عنوان ژن‌های معنی‌دار مرتبط با صفات پشم برای تجزیه GSEA-SNP انتخاب شدند. تعداد مجموعه‌های ژنی حاصل از پایگاه‌های داده‌ای مختلف شامل ۱۴۹ طبقات هستی‌شناسی (۹۲ طبقه فرآیند زیستی، ۲۶ طبقه اجزای سلولی و ۲۱ طبقه عملکرد مولکولی) به دست آمد و ۱۰ مسیر زیستی KEGG مشاهده شد. مسیرهای که بیشتر از سه ژن و کمتر از ۳۰۰ ژن داشتند گزارش شده‌اند.

جزئیات کامل ترم‌های هستی‌شناسی و مسیرهای بیوشیمیایی KEGG به همراه اسامی ژن‌های کاندیدا در جدول ۴ ارائه شده است. مسیر پاسخ به هورمون^۵ که جزء مسیرهای معنی‌دار مرتبط با میانگین قطر الیاف بدست آمد حاوی ۱۶ ژن در این مجموعه ژنی بود که از این بین ژن‌های کاندیدای *LOC105602796* و *LOC105607461* مرتبط با قطر الیاف می‌توان در نظر گرفت. این ژن‌ها جزئی از خانواده ابر ژنی *LOC* می‌باشند که دارای نقش انتقالی هستند. خانواده ژنی *LOC* که بیشتر در جمعیت‌های انسانی مورد تمایز و انتخاب قرار گرفته‌اند، نقش‌های متنوعی در مکانیسم‌های زیستی دارند. این خانواده ژنی، نقش‌های بیولوژیکی بسیار گسترده‌ای از قبیل سیگنال‌دهی پرولاکتین، ترشح انسولین، جذب طیف وسیعی از مواد مغذی، سنتز هورمون تیروئید و مسیرهای متابولیکی مختلف دیگر دارند (UniProtKB). در مطالعه پویش ژنومی با هدف شناسایی مناطق ژنومی و ژن‌های کاندیدای مرتبط با صفات کیفی الیاف پشم در گوسفندان نژاد Akkaraman ترکیه‌ای، ژن *LOC101117971* و *LOC101118971* با صفت میانگین قطر الیاف و ضریب تغییرات قطر الیاف پشم گزارش شده است (Arzik et al., 2023). همچنین در مطالعه قبلی پویش ژنومی مرتبط با صفات کیفی پشم در گوسفندان رامبویه، ژن کاندیدای *LOC121818710* مرتبط با طول استاپل معنی‌دار گزارش شده بود (Becker et al., 2023).

از مسیرهای مهم و معنی‌دار مرتبط با میانگین قطر الیاف می‌توان به مسیرهای بیولوژیکی گلیکوزیلاسیون پروتئین^۶، اسمبلی سیلیومی غیر متحرک^۷ و تنظیم مثبت تکثیر سلولی^۸ اشاره نمود. ژن کاندیدای *TMTC3* که در مسیر هستی‌شناسی گلیکوزیلاسیون پروتئین بر روی کروموزوم ۳ در ناحیه ۱۳۳/۸ مگابازی گوسفند قرار دارد. این ژن از خانواده ژنی *TMTC* می‌باشد و در مطالعات مختلفی ارتباط آن با صفات کیفی پشم گزارش شده است (Morgan et al., 2018). در مطالعه پویش کل ژنومی در گوسفندان نژاد Akkaraman ترکیه، ارتباط معنی‌داری بین ژن *TMTC3* با طول استاپل گزارش شده است (Arzik et al., 2023).

از دیگر مسیرهای بیولوژیکی معنی‌دار مرتبط با میانگین قطر الیاف می‌توان به مسیر اسمبلی‌های پیچیده حاوی پروتئین^۹ اشاره نمود که جزء مسیرهای فرآیندهای زیستی مرتبط با تولید می‌باشد که از بین ۲۲ ژن معنی‌دار، از این بین، ژن کاندیدای *RHPN2* دارای نقش بیولوژیکی مستقیمی با صفات مرتبط با تولید داشت. ژن کاندیدای *RHPN2* که در مسیر بیوشیمیایی Rho pathway قرار دارد. این ژن نقش اساسی در تنظیم تفرق سلول‌های کراتینوسیت‌ها از مسیر سیگنال‌دهی Rho دارد (McMullan et al., 2003). در مطالعه پویش ژنومی مرتبط با صفات پشم در گوسفندان نژاد چینی، ژن کاندیدای *RHPN2* مرتبط با صفت میانگین قطر الیاف گزارش شده است (Zhao et al., 2021).

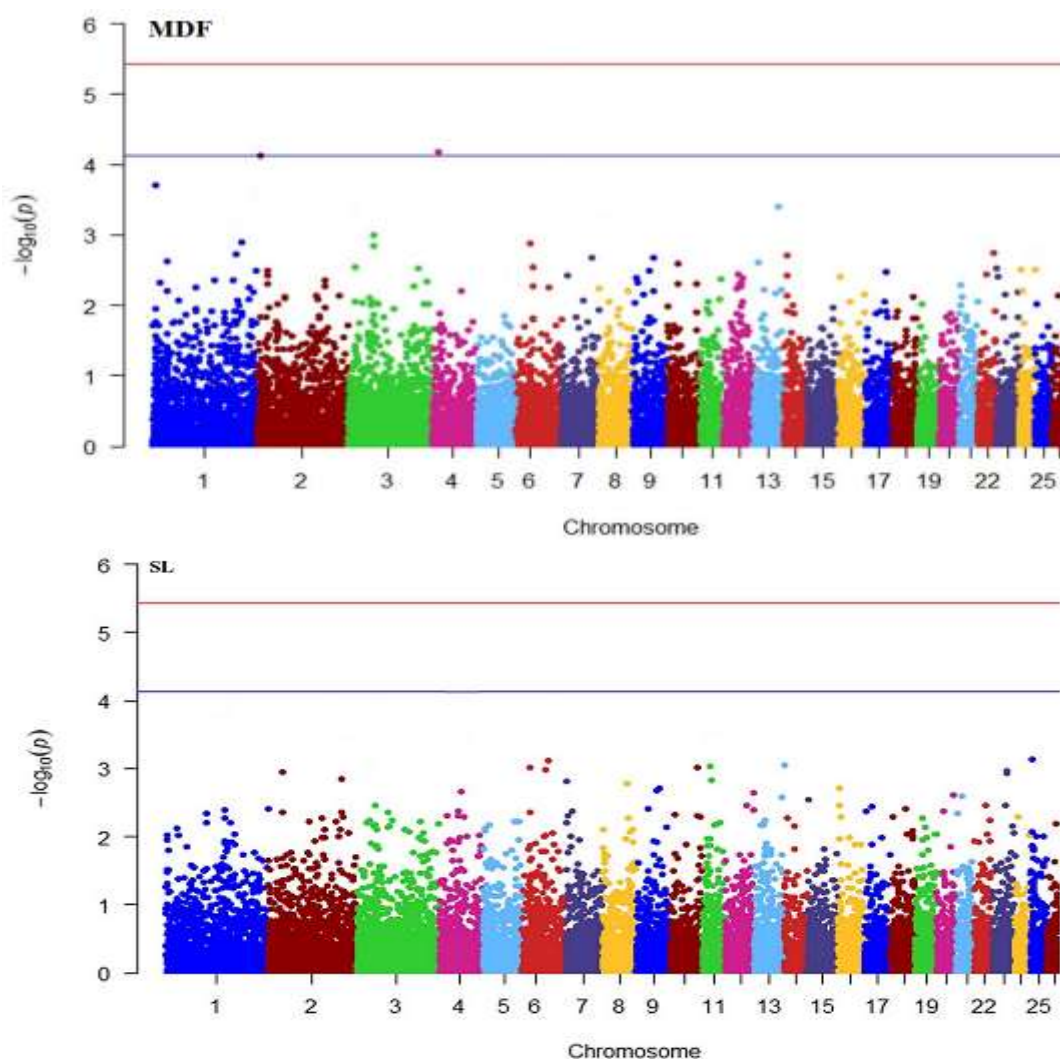
⁵ response to hormone

⁶ protein glycosylation

⁷ nonmotile cilium assembly

⁸ positive regulation of cell proliferation

⁹ protein-containing complex assemblies



شکل ۳. پلات‌های منهتن حاصل از مطالعات پویش کل ژنومی برای صفات کیفی پشم. میان قطر الیاف (MFD) و طول استاپل (SL)

از مسیرهای معنی‌دار مرتبط با طول استاپل می‌توان به فعالیت کانال کلسیمی^{۱۰} و تفرق سلول‌های اپیدرمال^{۱۱} اشاره نمود که از بین ژن‌های موجود در مسیرهای نامبرده ارتباط معنی‌داری بین ژن‌های کاندیدای *POUIF1* و *NLGN1* با صفات تولیدی پشم گزارش شده است. ارتباط معنی‌داری بین چندشکلی‌های موجود در ژن *POUIF1* با صفات مرتبط با پشم شامل تولید پشم در گوسفند (Sun et al., 2013) و طول دسته الیاف در بزهای کشمیری گزارش شده است (Lan et al., 2009). همچنین ژن کاندیدای *NLGN1* در رشد و تفرق کراتیونوسیت‌ها نقش کلیدی دارد (Gay et al., 2015).

شاید بتوان مسیر اتصال سلولی^{۱۲} که جزء هستی‌شناسی اجزای سلولی است را یکی از مهمترین مسیرهای مؤثر بر فرآیند کیفی الیاف پشم دانست. در این مسیر ژن کاندیدای *PLCE1* در مطالعات قبلی پویش ژنومی با صفت طول استاپل گزارش شده است. ژن *PLCE1* تنظیم‌کننده فعالیت فاکتور Ras guanine-exchange می‌باشد که نقش مهمی در رشد و توسعه اپیدرمال پوستی دارد. در مطالعه توالی‌یابی کل ژنوم در نژادهای مختلف گوسفند (Merino, Qinghai, Aohan)، ژن *PLCE1* مرتبط با میانگین قطر الیاف و طول استاپل گزارش شده است (Zhao et al., 2021).

¹⁰ Calcium channel activity

¹¹ epidermal cell differentiation

¹² Cell junction

از مسیرهای KEGG معنی‌دار مرتبط با طول استایل به مسیر بیوسنتز گلیکوزآمینوگلیکان - هیپران سولفات^{۱۳} اشاره نمود. از بین ۲۴ ژن معنی‌دار در این مسیر ژن کاندیدای SPHKAP در مطالعات قبلی پوشش ژنومی در ارتباط با صفات کیفی الیاف در نژاد Inner Mongolia Cashmere مرتبط با صفت طول تار الیاف گزارش شده است (Wang *et al.*, 2021).

با بررسی منابع انجام شده، پژوهش حاضر یکی از اولین مطالعات پوشش کل ژنومی مرتبط با برخی صفات کیفی پشم در مقایسه گوسفندان با پشم ضخیم و ظریف بوده است، لذا در تجزیه و تحلیل‌ها تلاش گردید از آزمون‌های پر کاربرد و سختگیرانه برای شناسایی مسیرهای زیستی معنی‌دار استفاده گردد. به هر حال باید توجه کرد که یکی از محدودیت‌های اجرای مطالعات پوشش ژنومی، کافی نبودن تعداد رکوردهای فنوتیپی و جمعیت غیر خویشاوند برای آنالیز پیوستگی می‌باشد که منجر به افزایش خطای نوع اول یا به عبارتی شناسایی نشانگرهای معنی‌دار مثبت کاذب می‌شود. با توجه به تعداد کم نمونه مورد استفاده در پژوهش حاضر نیز باید در استفاده از ژن‌های کاندیدای شناسایی شده مرتبط با صفات پشم در برنامه‌های اصلاحی با احتیاط عمل کرد. البته می‌توان با ادغام داده‌های تحقیقات مشابه جدید و استفاده از آنالیزهای آماری جامع‌تر برای تأیید نتایج پژوهش حاضر استفاده کرد. به طور کلی مسیرهای شناسایی شده در این تحقیق می‌تواند ما را به سمت ژن‌های مؤثر بر صفات هدایت کند و بررسی بیشتر نواحی مهم ژنومی شناسایی شده با استفاده از آزمون‌های آزمایشگاهی مختلف می‌تواند در تأیید نتایج بدست آمده در این پژوهش مؤثر باشند.

جدول ۴. تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی معنی‌داری ($P < 0.05$) مرتبط با صفات کیفی پشم مورد بررسی

نام و کد طبقات هستی‌شناسی	نام مسیر هستی‌شناسی	تعداد ژن‌های کاندیدا	ژن‌های کاندیدای موجود	ارزش -P تصحیح شده (FDR)
فرآیندهای زیستی GO: 0009725	response to hormone	۱۷	<i>RARG, EEF1A2, ERCC3, BIN1, TOP2A, CTNBL1, NET1, STK10, PTK6, TAOK3, TNK1, MTOR, NEK6, LOC105602796, DAPK1, RBM47, LOC105607461</i>	۰/۰۰۹
GO: 0090263	positive regulation of canonical Wnt signaling pathway	۱۴	<i>ADAM17, TNS4, SYNPO2, LIMD1, FERMT2, WASF1, TNS4, WWOX, CLASP1, GSN, SYNE2, TLN2, EPB41L5, AFAP1</i>	۰/۰۲۵
GO: 1905515	Nonmotile cilium assembly	۱۸	<i>STK10, PTK6, TAOK3, TNK1, MTOR, NEK6, DAPK1, PRKCA, SPHK1, NKX3-1, BCAR1, LGI2, SEPSECS, PI4K2B, ZCCHC4, ANAPC4, SLC34A2, SEL1L3</i>	۰/۰۴۱
GO:0070613	Regulation of protein processing	۱۵	<i>CAPN14, SPRED2, WDPCP, HEY1, MYC, ID2, TCF23, STAT3, ANO4, E2F6, TIGD2, HERC3, HERC6, PKD2, CCNY</i>	۰/۰۱۳
GO:0009913	epidermal cell differentiation	۱۹	<i>STX3, UNC13B, MAFG, FANCL, BRAF, LAMA5, RAPSN, CHRN1, NPNT, GSTCD, INTS12, POU1F1, TET2, CENPE, BDH2, SLC9B2, SLC9B1, CISD2, MANBA</i>	۰/۰۱۴
GO: 0097711	ciliary basal body-plasma membrane docking	۱۷	<i>CD44, PTK2B, CSPG4, RIPK2, LAMA5, ATXN10, ABLIM3, ATP6V1D, EHD3, UBA6, GNRHR, TMPRSS11D, TMPRSS11A, TMPRSS11F, YTHDC1, UGT2B10, MGC152010</i>	۰/۰۰۸
GO: 0008284	positive regulation of cell proliferation	۴	<i>PRKCZ, FAM19A5, TMEM17, MAFF</i>	۰/۰۱۹

¹³Glycosaminoglycan biosynthesis - heparan sulfate

0-28	<i>GRIN2D, CACNG4, GRM4, GRIN3A, DAGLA, HOMER3, GRID1</i>	7	Actin polymerization or depolymerization	GO:0097305
0-32	<i>CACNB2, CACNB4, CACNA2D1, CACNA2D3, ZSWIM5, HECTD3, PTCH2, NLGN1, PLK3, RPS8, KIF2C, ARMH1, TMEM53, RNF220, ERI3, DMAP1, UROD, EIF2B3, CALCR</i>	19	Calcium channel activity	GO:0035556
0-11	<i>CACNB2, PPP3R2, CACNB4, CACNA2D1, STMN1, CACNA2D3, RHPN2, SGCE, PEG10, PON1, PON2, PON3, ASB4, PDK4, DYNC11</i>	15	protein-containing complex assemblies	GO: 0031333
0-29	<i>ATF6B, ADCY3, ADCY5, ADCY8, ADCY9, CACNA1D, CAMK2A, GLP1R, KCNMA1, TMTC3, PRKCB, RYR2, PPP1R9A</i>	13	protein glycosylation	GO: 0006486
<hr/>				
0-03	<i>SAMD8, FKBP8, PLCE1, CACNB2, PPP3R2, CACNB4, CACNA2D1, GNAS, ITPR1, SPARA48, ZBPB, VWC2, ABCA13, UPP1SUN3, CLDN12, FZD1</i>	16	Cell junction	اجزای سلولی GO:0030054
0-41	<i>SEMA3D, TACC3, CMAS, ATF6B, ADCY3, ADCY5, ADCY8, ADCY9</i>	8	Photoreceptor cell cilium	GO: 0097733
<hr/>				
0-01	<i>PPP3R2, KCNMA1, ITPR1, PDE3A, NDC1, GLIS1, DMRTB1, LRP8, MAGOH, CPT2, CLS1A7, PODN, SCP2, POLN, TNFSF4, CD44, GALNTL5</i>	17	Cytokine-cytokine receptor interaction	مسیرهای KEGG oas04060
0-24	<i>SLC8A3, ITPKB, PPP3R2, ITPR1, SLC8A1, GRM1, ANKIB1, TMBIM7, RBM48, CDK6, GATAD1, PEX1, SAMD9, VPS50, HEPACAM2, TFPI2, BET1, GNGT1, GNG11, COL1A2, CASD1, SPHKAP, BTF3L4, TXNDC12</i>	24	Glycosaminoglycan biosynthesis - heparan sulfate heparin	oas00534

^۱ ژن‌های کاندیدا مرتبط با صفت به صورت پر رنگ نشان داده شده‌اند.

نتیجه‌گیری

در انتها بیان می‌گردد با توجه به عملکرد بیولوژیکی ژن‌های کاندیدا و مسیرهای شناسایی شده در این پژوهش، به نظر می‌رسد این ژن‌ها در بروز فنوتیپی صفات مرتبط با کیفیت پشم نقش مؤثری داشته باشند، با توجه به این موضوع، می‌توان کارآیی روش تجزیه و تحلیل غنی‌سازی مجموعه ژنی برای پویش ژنومی صفات مهم اقتصادی را مورد تأیید قرار داد. نهایتاً علی‌رغم اهمیت و مفید بودن این بررسی در پیشبرد کارهای اصلاح نژادی کشور، با توجه به اینکه این کار جزء اولین بررسی‌ها در ارزیابی ژنومی صفات مرتبط با کیفیت پشم در گوسفند در این سطح می‌باشد، نتایج این تحقیق از دیدگاه علمی در سطح کشور می‌تواند اهمیت زیادی داشته باشد.

منابع

محمدی، ح. (۱۳۹۶). مطالعه پیوستگی در سطح ژنوم جهت شناسایی مناطق کاندیدای مرتبط با برخی از صفات مهم اقتصادی و کاوش ژنومیکی نشانه‌های انتخاب در گوسفندان نژاد زندگی. رساله دکتری، دانشگاه تبریز، دانشکده کشاورزی.

REFERENCES

1. Arzik Y, Kizilaslan M, Behrem S, White SN, Piel LMW & Cinar M.U. (2023). Genome-Wide Scan of Wool Production Traits in Akkaraman Sheep. *Genes*, 14, 713.
2. Becker GM, Woods JL, Schauer CS, Stewart WC & Murdoch BM. (2023). Genetic association of wool quality characteristics in United States Rambouillet sheep. *Frontiers in Genetics*, 13, 1081175.
3. Bu D, Luo H, Huo P, Wang Z, Zhang S, He Z, Wu Y, Zhao L, Liu L & Guo J. (2021). KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Research*, 49, 317–325.
4. Clancey E, Kiser JN, Moraes JGN, Dalton JC, Spencer TE & Neibergs HL. (2019). Genome-wide association analysis and gene set enrichment analysis with SNP data identify genes associated with 305-day milk yield in Holstein dairy cows. *Animal Genetics*, (3), 254-258.
5. Dadousis C, Pegolo S, Rosa GJM, Gianola D, Bittante G & Cecchinato A. (2017). Pathway-based genome-wide association analysis of milk coagulation properties, curd firmness, cheese yield, and curd nutrient recovery in dairy cattle. *Journal of Dairy Science*, 100, 1223-1231.
6. Durinck S, Spellman PT, Birney E & Huber W. (2009). Mapping identifiers for the integration of genomic datasets with the R/bioconductor package biomaRt. *Nature Protocols*, 4, 1184–1191.
7. Esmaeili-Fard SM, Gholizadeh M, Hafezian SH & Abdollahi-Arpanahi R. (2021). Genome-wide association study and pathway analysis identify NTRK2 as a novel candidate gene for litter size in sheep. *PLoS One*, 1, e0244408.
8. Gay DL, Yang CC, Plikus MV, Ito M, Rivera C, Treffeisen E, Doherty L, Spata M, Millar SE & Cotsarelis G. (2015). CD133 expression correlates with membrane betacatenin and E-cadherin loss from human hair follicle placodes during morphogenesis. *Journal of Investigative Dermatology*, 1, 45–55.
9. Ghafouri Kesbi F, Eskandarinasab MP & Shahir MH. (2008). Estimation of direct and maternal effects on body weight in Zandi sheep using random regression models. *Archives Animal Breeding*, 51 (3), 235–246.
10. Gebreselassie G, Berihulay H & Jiang L, Ma Y. (2019). Review on Genomic Regions and Candidate Genes Associated with Economically Important Production and Reproduction Traits in Sheep (*Ovis aries*). *Animals (Basel)*, 1, 33.
11. Hu ZL, Park CA & Reecy JM. (2022). Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Research*, 50(D1), D956-D961.
12. Kijas, J, Lenstra JA, Hayes B, Boitard S, Porto Neto LR, 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 Biological*, 10(2), e1001258.
13. Lan XY, Shu JH, Chen H, Pan CY, Lei CZ & Wang X. (2009). A PstI polymorphism at 3'UTR of goat POU1F1 gene and its effect on cashmere production. *Molecular Biology Reports*, 36, 1371–1374.
14. Lupton CJ, Huston JE, Craddock BF, Pfeiffer FA & Polk WL (2007). Comparison of three systems for concurrent production of lamb meat and wool. *Small Ruminant Research*, 72 (2-3), 133–140.
15. Marjanovic J & Calus MPL. (2020). Factors affecting accuracy of estimated effective number of chromosome segments for numerically small breeds. *Journal of Animal Breeding and Genetics*, 138, 151-160.
16. Mohammadi H. (2017). Genome-wide association study for identify candidate regions with some important economic related traits and genomic scan of selective sweeps in Zandi sheep breed. Ph.D. thesis of Animal Breeding and Genetics, University of Tabriz, pp. 181.

17. Morgan MD, Pairo-Castineira E, Rawlik K, Canela-Xandri O, Rees J, Sims D, Tenesa A & Jackson IJ. (2018). Genome-Wide Study of Hair Colour in UK Biobank Explains Most of the SNP Heritability. *Nature Communications*, 9, 5271.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly M & Sham PC. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 3, 559-575.
19. Ren X, Yang GL, Peng WF, Zhao YX, Zhang M, Chen ZH, Wu FA, Kantanen J, Shen M & Li MH. (2016). A genome-wide association study identifies a genomic region for the polycerate phenotype in sheep (*Ovis aries*). *Scientific Reports*, 6, 21111.
20. Sun W, Ni R, Yin JF, Musa HH, Ding T & Chen L. (2013). Genome array of hair follicle genes in lambskin with different patterns. *PLoS One*, 8, e68840.
21. Taylor P, Mortimer S, Bird-Gardiner T & Atkins K. (2017). Merino breeding objectives & selection indexes to increase wool profit. (Ed. BJ McGuirk) pp. 579–592. (Australian Wool Corporation: Melbourne).
22. Waiteina RW, Okeno TO, Ilatsia ED & 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.
23. Wang L, Jia P & Wolfinger RD. (2011). Gene set analysis of genome-wide association studies: Methodological issues and perspectives. *Genomics*, 98, 1–8.
24. Wang Z, Zhang H, Yang H, Wang S, Rong E, Pei W, Li H & Wang N. (2014). Genome-wide association study for wool production traits in a Chinese Merino sheep population. *PLoS One*, 9, e107101.
25. Wang FH, Zhang L, Gong G, Yan XC, Zhang LT, Zhang FT, Liu HF, Lv Q, Wang ZY, Wang RJ, Zhang YJ, Wang ZX, Liu ZH, He LB, Su R, Zhao YH & Li JQ. (2021). Genome-wide association study of fleece traits in Inner Mongolia Cashmere goats. *Animal Genetics*, 52(3), 375-379.
26. Zhao F, Deng T, Shi L, Wang W, Zhang Q, Du L, Wang L. (2020). Genomic Scan for Selection Signature Reveals Fat Deposition in Chinese Indigenous Sheep with Extreme Tail Types. *Animals (Basel)*, 10(5), 773.
27. Zhao H, Guo T, Lu Z, Liu J, Zhu S, Qiao G, Han M, Yuan C, Wang T, Li F, Zhang Y, Hou F, Yue Y & Yang B. (2021). Genome-wide association studies detects candidate genes for wool traits by re-sequencing in Chinese fine-wool sheep. *BMC Genomics*, 1, 127.
28. Zhou X & Stephens M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44, 821.

Genome-Wide Association Study for wool quality traits- A comparison of coarse wool and fine wool sheep breeds

Extended Abstract

Introduction

Over recent years, advances in DNA-based marker technology have made it possible to identify genomic regions or quantitative trait loci (QTLs) underlying complex traits, such as fleece traits, in sheep. One way to identify new loci and confirm existing QTL is genome-wide association studies (GWAS). Furthermore, the identification of gene loci with major impacts on economically important traits is one of the most important goals of sheep breeding. It was hypothesized that QTL-assisted selection and genomic regions affecting production traits increase the efficiency of selection and improve production output. The present study aimed to conduct a genome wide association studies and gene-set enrichment analysis for identifying the loci associated with wool quality traits.

Materials and Methods

In order to identify the genomic region and candidate genes, genomic information of coarse wool Zandi sheep (96 Animal) and 240 fine wool Rambouillet sheep (240 Animal) were used. We measured and recorded two wool production traits, including mean fiber diameter (MFD) and staple length (SL). Genomic DNA extraction from sheep blood was performed by the applying a modified salting out protocol and genotyping of the Sheep SNPChip 50 K SNP Bead from Illumina Inc. The genomic information of foreign breed was extracted from the EVA database. Quality control was conducted using the Plink software. The markers or individuals were excluded from the further study based on the following criteria: unknown chromosomal or physical location, call rate <0.90, missing genotype frequency >0.01, minor allele frequency (MAF) < 0.01, and a P-value for Hardy–Weinberg equilibrium test less than 10^{-6} . After the quality control of the data, Genome wide association study was performed with wool traits using GEMMA software. Using the *biomaRt2* R package, the SNP were assigned to genes if they were within the genomic sequence of the gene or within a flanking region of 50 kb up- and downstream of the gene and bioinformatics analysis was implemented to identify the biological pathways performed in GO, KEGG, DAVID and PANTHER databases. The GO database designates biological descriptors to genes based on attributes of their encoded products and it is further partitioned into 3 components: biological process, molecular function, and cellular component. Finally, a Fisher's exact test was performed to test for overrepresentation of the significant genes for each gene-set. In the next step, a bioinformatics analysis was implemented to identify the biological pathways performed in GeneCards databases.

Results and discussion

The result from genomic control showed weak population stratification with for woot traits in between population Zandi and Rambouillet sheep. We identified different sets of candidate genes related to wool traits: *TMTC3*, *POU1F1*, *NLGN1*, *PLCE1*, *SPHKAP*, *LOC101117971* and *LOC101118971*. Some of the found genes, are consistent with some of the previous studies related to wool production traits. According to pathway analysis, 15 pathways from gene ontology and biological pathways were associated with the wool traits ($P < 0.05$). Some of the genes were found are consistent with some previous studies and to be involved biological pathways related to hair follicle development, keratinocytes differentiation, synthesises an enzyme of threonine kinases, development of epidermal and Wnt signaling pathway.

Conclusion

In total, this study supported previous results from GWAS of wool quality traits, also revealed additional genomic regions in the sheep genome associated with these economically important traits. Using these findings could potentially be useful for genetic selection in the breeding programs and can be used to understand the genetic mechanism controlling this trait.

عبد فاطمہ استاد