

مطالعه ساختار و لایه‌بندی جمعیتی و ارتباط ژنومی هاپلوتیپی صفات کیفی پشم در گوسفندان نژاد زندی

حسین محمدی^۱، سید عباس رأفت^{۲*}، حسین مرادی شهر بابک^۳، جلیل شجاع^۲ و محمد حسین مرادی^۴

تاریخ دریافت: ۹۶/۳/۲۰ تاریخ پذیرش: ۹۶/۶/۲۷

^۱ دانش آموخته دکتری گروه علوم دامی، دانشکده کشاورزی، دانشگاه تبریز

^۲ استاد گروه علوم دامی، دانشکده کشاورزی، دانشگاه تبریز

^۳ استادیار گروه علوم دامی، پردیس کشاورزی و منابع طبیعی کرج، دانشگاه تهران

^۴ استادیار گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه اراک

*مسئول مکاتبه: Email: abbasrafat@hotmail.com

چکیده

زمینه مطالعاتی: شناسایی ژن‌های بزرگ اثر، مؤثر بر صفات مهم اقتصادی یکی از مهم‌ترین اهداف اصلاح نژادی در پرورش گوسفند است. **هدف:** این تحقیق به منظور بررسی ساختار و لایه‌بندی جمعیتی و شناسایی جایگاه‌ها و ژن‌های مرتبط با صفات کیفی پشم، از طریق مطالعه ارتباط ژنومی هاپلوتیپی (GWAS) با استفاده از تراشه SNP ژنوم گوسفند (Illumina SNPChip 50K Beadchip) در یک جمعیت گوسفند زندی بود. **روش کار:** برای هر دام، صفات کیفی شامل: طول استاپل (SL)، میانگین قطر الیاف (MFD)، ضریب تغییرات قطر الیاف (FDCV) و نسبت الیافی که مساوی یا بیشتر از ۳۰ میکرومتر قطر الیاف ($F \geq 30$) پشم اندازه‌گیری شد. پس از شناسایی اثرات ثابت معنی‌دار، مطالعه کنترل ژنومیکی (GC) با استفاده از آماره لامبدا و پویش ژنومی در نرم افزار PLINK نسخه 1.90 با استفاده از مدل خطی معمول (GLM) ارتباط هر یک هاپلوتیپ‌ها با صفات کیفی پشم ارزیابی و برای کنترل نرخ اشتباه از تصحیح بنفرونی استفاده شد. **نتایج:** یافته‌های حاصل از کنترل ژنومیک لایه‌بندی ضعیفی برای صفات کیفی پشم شامل SL، MFD، FDCV و $F \geq 30$ نشان داد که حاکی از عدم وجود اختلاط ژنتیکی در جمعیت مورد مطالعه است. دو ناحیه هاپلوتیپی پیشنهادی در داخل ژن‌های کاندیدا ERBB2 و GNAS واقع شدند که پیشتر به عنوان ژن‌های مؤثر بر رشد و توسعه مو و پوست در انسان گزارش گردیده بودند. **نتیجه‌گیری نهایی:** ژن‌های کاندیدای شناسایی‌شده، عملکرد مولکولی مرتبط با صفات الیاف پشم داشتند. انتخاب برای بهبود کیفی در پشم تولیدی در گوسفندان زندی با استفاده از این یافته‌ها، باعث تسریع در پیشرفت ژنتیکی خواهد گردید.

واژگان کلیدی: گوسفند زندی، لایه‌بندی جمعیت، مطالعات پویش ژنومی، ژن کاندیدا، کیفیت الیاف

مقدمه

از طرف دیگر از نظر کیفیت، پشم گوسفندان ایرانی جزو پشم‌های ضخیم محسوب می‌شوند و تنها در صنعت قالیبافی از آن استفاده می‌شود. بنابراین با توجه به تعداد گوسفند در ایران که بالغ بر ۵۰ میلیون رأس (فائو ۲۰۱۳)

تولید پشم در ایران به ازای هر رأس دام اغلب کمتر از ۱/۲ کیلوگرم است، در حالی که استاندارد تولید در دنیا به طور میانگین ۲/۲ تا ۵ کیلوگرم می‌باشد (فائو ۲۰۱۳).

(۲۰۱۲) و گوسفند (جانستون و همکاران ۲۰۱۱) گزارش شده‌اند.

با توجه به بررسی منابع، تا به امروز براساس آخرین نسخه AnimalQTLdb در ۲۷ آوریل ۲۰۱۷ (<http://cn.animalgenome.org/cgi->

bin/QTLdb/OA/index, 27 Aug, 2017)، تحقیقات اندکی برای شناسایی QTL صفات کمی و کیفی پشم گزارش شده است. به طوریکه تنها ۹۸ 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 (هیو و همکاران ۲۰۱۶).

با وجود این و با توجه به اینکه طراحی اولین سری آرایه‌های SNPChip 50K گوسفندی در سال ۲۰۰۹ تجاری شد. اولین مطالعه بررسی صفات کمی و کیفی الیاف پشم در سال ۲۰۱۴ توسط وانگ و همکاران روی گوسفندان مرینوس با شناسایی ژن‌های کاندیدای مؤثر بر میانگین قطر الیاف و ضریب تغییرات قطر الیاف و میزان تجمع به طور موفقیت آمیزی انجام و ژن‌های کاندید پروتئین‌های وابسته به کراتین نوع ۳ گزارش شدند. در کشور ایران هم براساس بررسی منابع، تنها یک گزارش با استفاده از آرایه‌های 50K در گوسفندان نژاد زل و لری بختیاری گزارش شده است (راستی فر و همکاران ۲۰۱۵). طبق بررسی‌ها تا به حال تحقیقی برای بررسی ارتباط صفت کیفی الیاف پشم در سطح ژنوم گوسفند زندی گزارش نشده است.

از سویی دیگر ساختارهای زیر جمعیتی درون جمعیت‌های مورد مطالعه باعث ایجاد آریبی در مطالعات GWAS می‌شود (واک‌هولدر و همکاران ۲۰۰۲). لایه‌بندی جمعیتی به دلیل نتایج مثبت دروغین (اشتباه نوع اول) چالشی برای مطالعات GWAS است. روش‌های متعددی برای تعیین ساختار ژنتیکی و لایه‌بندی جمعیتی وجود دارد. از

و حجم زیاد تولید پشم و نقش تولید این محصول دامی در اقتصاد کشور، انجام تحقیقاتی در راستای شناسایی عوامل ژنتیکی تأثیرگذار بر این صفات ضروری به نظر می‌رسد.

در برنامه‌های اصلاح نژادی گوسفند جهت بهبود تولید پشم و کیفیت آن صفاتی مانند وزن پشم، میانگین قطر الیاف، طول دسته الیاف و ضریب تنوع آنها مهمترین ملاک‌های انتخاب هستند. نرخ پیشرفت در صفات پشم نسبت به سایر صفات از جمله رشد سرعت کندی داشته است. اکثر صفات پشم دارای وراثت پذیری متوسط و بالا هستند، از اینرو اگر در برنامه‌های اصلاح نژادی ملاک انتخاب قرار گیرند. می‌توان شاهد تغییرات قابل ملاحظه‌ای در این صفات در طی نسل‌های آینده بود. اما مشکلاتی وجود دارد که روند انتخاب و بهبود صفات بر پایه داده‌های فنوتیپی را کند می‌نماید (محمدی و همکاران ۲۰۱۲). 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 مثلی 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 (محمدی و همکاران ۲۰۱۳b).

در طی سال‌های اخیر، تکنیک‌های بیولوژی مولکولی با سرعت شگرفی توسعه پیدا کرده و ورود نسل‌های دوم و سوم دستگاه‌های توالی‌یابی موجب گردیده است تا پروژه‌های مختلف توالی‌یابی DNA انسان و بسیاری از گونه‌های مختلف گیاهی و جانوری دیگر تکمیل شود و متعاقب آن پانل‌های مختلفی از آرایه‌های نانویی با تراکم متفاوت، زمینه‌های مختلف ژنتیک و اصلاح نژاد امکان بررسی تمام ژنوم و ژن‌های موجود در آنها همزمان بررسی شود. در طول چند سال گذشته اولین آزمایشات موفقیت آمیز از مطالعات پویش ژنومی (GWAS) در حیوانات اهلی از جمله گاو (دایت وایلر و همکاران ۲۰۰۸)، مرغ (آبشت و لمونت ۲۰۰۷)، اسب (سچرودر و همکاران

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 پروژکتیو قطر تارهای پشم بر حسب میکرومتر اندازه گیری شد. میانگین قطر الیاف هر نمونه برحسب ۴۰۰ تار پشم محاسبه گردید. همچنین هنگام قرائت قطر الیاف پشم با استفاده از میکروسکوپ پروژکتیو نوع تار پشم نیز یادداشت برداری گردید. نهایتاً از اطلاعات به دست آمده سایر صفات کیفی پشم شامل ضریب تغییرات قطر الیاف (CVFD) و درصد الیاف با قطر مساوی یا بیشتر از ۳۰ میکرومتر ($FD \geq 30$) محاسبه شد.

استخراج DNA با استفاده از روش بهینه یافته استخراج نمکی از خون کامل انجام شد. پس از اطمینان از کمیت و کیفیت بالای نمونه‌ها، غلظت آنها تا ۵۰ ng/μl جهت تعیین ژنوتیپ رقیق شد. ژنوتیپ نمونه‌ها در کمپانی Neogen (<http://genomics.neogen.com>) با استفاده از آرایه های Illumina OvineSNP50K BeadChip با بکارگیری پروتکل استاندارد شرکت ایلومینا (Illumina Inc., San Diego, CA, USA) تعیین ژنوتیپ شدند. این آرایه‌ها امکان تعیین ژنوتیپ همزمان ۵۴۲۴۱ جایگاه نشانگری با میانگین فاصله بین نشانگرها ۵۰/۹ کیلوباز را فراهم می‌کنند (Illumina OvineSNP50 BeadChip Datasheet).

برای فیلتراسیون داده‌های ژنوتیپینگ از معیارهای فراوانی ژنوتیپینگ نمونه‌ها، نرخ ژنوتیپینگ نشانگرها در هر نمونه و حداقل فراوانی آلی (MAF) استفاده می‌شود. در ابتدا نمونه‌هایی که فراوانی ژنوتیپینگ آنها کمتر از ۹۰٪ بود، شناسایی و حذف شد. این نمونه‌ها احتمال بیشتری داشته که با داده‌های گمشده همراه بوده و خطای ژنوتیپینگ در آنها بالا باشد. در مرحله بعد نشانگرهایی که حداقل فراوانی آلی در آنها کمتر از ۵٪ بود حذف شدند. سپس نشانگرهایی که نرخ ژنوتیپینگ آنها در نمونه‌ها کمتر از ۹۵٪ بود شناسایی و حذف شدند. در نهایت برای SNP‌های باقیمانده آنهایی که در

جمله مهمترین روش‌ها، روش کنترل ژنومیک (GC) و مقیاس چند بعدی (MDS) است که قادر به تعیین ساختار جمعیتی می‌باشند (محمدی و همکاران ۲۰۱۳a).

بنابراین هدف از این تحقیق بررسی ژنوم گوسفند زندگی با تعیین ساختار و لایه‌بندی جمعیتی و شناسایی مناطق ژنومی مرتبط با صفات کیفی پشم با استفاده از آرایه‌های 50K گوسفندی است. شناسایی این جایگاه‌ها از دیدگاه علمی و اقتصادی می‌تواند دارای اهمیت زیادی باشد.

مواد و روش‌ها

از ۳۰۰ گوسفند نژاد زندگی حاضر در دو گله مرکز اصلاح نژاد خجیر واقع در تهران وابسته به سازمان جهاد کشاورزی استان تهران خون گیری به عمل آمد. مقدار ۷-۵ سی سی استحصال شده همراه با ۰/۵ میلی لیتر EDTA در لوله های خلاء با pH ۷/۸-۵ ریخته شد. نمونه‌های گرفته شده بلافاصله بعد از شماره گذاری و با حفظ شرایط سرد به آزمایشگاه منتقل و تا زمان استخراج DNA در دمای ۴ درجه سانتیگراد نگهداری شد. از میان نمونه‌های اخذ شده پس از آنالیز شجره، حیواناتی که کمترین رابطه خویشاوندی را با هم داشتند به تعداد ۹۶ رأس انتخاب شدند.

از تمام دام‌های انتخاب شده در سن ۱۲ تا ۱۵ ماهگی، قبل از پشم چینی سالیانه، از پهلو سمت چپ حیوان بعد از دنده سیزدهم و به اندازه کف دست پایین‌تر از مهره‌ها، به ابعاد تقریبی ۵×۵ سانتی متر مربع 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 سانتی متر مطابق با روش استاندارد شماره ۱۹۴۱ ایران اندازه گیری شد. سپس میانگین طول ۳ استاپل محاسبه و با تقریب ۲ میلی متر برای هر نمونه گزارش شد. نمونه‌های پشم گرفته شده در آزمایشگاه با مواد شوینده شسته شده و پس از شستشو و خشک نمودن نمونه‌ها در شرایط آزمایشگاهی صفات کیفی الیاف طبق پروتکل استاندارد A.S.T.M شماره D2130-90 اندازه‌گیری شد (درستکار و همکاران ۲۰۱۳).

دهنده عدم وجود لایه بندی جمعیتی می‌باشد (پورسل و همکاران ۲۰۰۷).

برای تعیین موقعیت ژنومی SNPها در سطح ژنوم گوسفند از آخرین نسخه ژنومی گوسفند یعنی assembly Oar_v4.0) از پایگاه اطلاعاتی NCBI استفاده شد (<https://www.ncbi.nlm.nih.gov/genome/83>). به منظور شناسایی ژن‌های مرتبط با بخش‌های معنی‌دار هاپلوتیپی ژنوم، از اطلاعات پایگاه اطلاعاتی NCBI (<https://www.ncbi.nlm.nih.gov/genome?term=ovis%20aries>) استفاده شد. 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. در نهایت برای بررسی اینکه آیا مناطق مورد نظر با QTLهای شناسایی شده همپوشانی دارد یا خیر، از آخرین نسخه (Release 32, 27 April, 2017) مرکز اطلاعاتی آنلاین QTL (<http://i.animalgenome.org/jbrowse/>) مورد بررسی قرار گرفت.

نتایج و بحث

آمار توصیفی صفات مورد بررسی در جدول ۱ نشان داده شده است. همچنین صفات کیفی الیاف پشم مورد بررسی در این تحقیق که دارای توزیع نرمال نبودند قبل از انجام آنالیز ارتباط ژنومی هاپلوتیپی، برای بهبود توزیع خطای تصادفی تبدیل داده‌ها صورت گرفت (لگاریتم بر پایه ۱۰، لگاریتم طبیعی و سینوسی).

تعادل هاردی-واینبرگ قرار نداشتند به عنوان معیاری از خطای تعیین ژنوتیپ (تئو و همکاران ۲۰۰۷) کنار گذاشته شدند. به منظور سطح احتمال 10^{-6} در نظر گرفته شد که با استفاده از تصحیح بنفرونی به دست آمد. مراحل مختلف فیلتراسیون با استفاده از نرم افزار PLINK (v1.90; <http://pngu.mgh.harvard.edu/purcell/plink>) انجام شد (چانگ و همکاران ۲۰۱۵).

برای بررسی وجود اثرات ثابت معنی‌دار با استفاده از رویه GLM در نرم افزار R 3.3.0 عوامل ثابت سال تولد، فصل تولد، تیپ تولد و سال نمونه برداری پشم حیوان بررسی شدند. این اثرات ثابت در صورت معنی‌دار بودن به عنوان اثرات ثابت در تجزیه و تحلیل‌های پویش ژنومی در نظر گرفته شدند. در این تحقیق برای آزمون ارتباط هاپلوتیپ‌ها با صفات کیفی پشم با استفاده از نرم افزار PLINK انجام گرفت. هاپلوتیپ‌ها جهت آنالیز، با استفاده از نرم افزار PLINK و با در نظر گرفتن پنجره‌هایی متشکل از ۱۰، ۱۵ و ۲۰ نشانگر تشکیل شد. همچنین در این تحقیق برای بخش بندی بلوک‌های هاپلوتیپی از روش گابریل و همکاران (۲۰۰۲) برای هر کروموزوم مشخص گردید.

برای کنترل نرخ اشتباه از آزمون‌های مختلفی همچون بنفرونی، HOLM و SIDAK استفاده شد. همچنین در این تحقیق برای بررسی وجود یا عدم وجود لایه‌بندی جمعیتی با استفاده از دو روش کنترل ژنومیک با استفاده از فاکتور تورم کنترل جمعیتی (λ) و روش آزمون MDS با بررسی ارتباط میان افراد براساس ماتریس همبستگی IBS بین دو فرد، در نرم افزار PLINK و ترسیم پلات‌های Q-Q در نرم افزار SNPEVWin_3.2 (وانگ و همکاران ۲۰۱۲) بررسی شد.

روش کنترل ژنومیک برای برآورد لایه بندی جمعیتی با استفاده از آماره لامبدا و پلات Quantile-quantile Q-Q استفاده می‌گردد. آماره لامبدا که از تقسیم میانه مقادیر کای مربع مشاهده شده بر میانه مورد انتظار (۰/۴۵۶) حاصل می‌شود، فاکتور تورم کنترل جمعیتی می‌باشد. اگر مقدار آماره لامبدا کمتر یا مساوی یک باشد، نشان

جدول ۱- آمار توصیفی صفات کیفی پشم در گوسفندان زندی

Table 1- Descriptive statistics of wool quality traits in Zandi sheep

خطای استاندارد Standard error	حداکثر Maximum	حداقل Minimum	انحراف استاندارد Standard deviation	میانگین Mean	صفات Traits
0.035	19.00	6.00	3.92	11.25	طول استاپل Staple length (cm)
0.032	39.04	22.40	3.25	29.85	قطر الیاف Fiber diameter (μm)
0.764	68.35	19.7	7.84	43.12	ضریب تغییرات قطر الیاف Fiber diameter coefficient of variation (%)
0.998	43.10	12.04	10.42	27.04	نسبت از الیاف که مساوی یا بیشتر از ۳۰ میکرومتر Proportion of fiber that are equal or more than 30 μm (%)

جدول ۲- مشخصات نشانگرهای مورد استفاده در آنالیزها

Table 2- Summary of characteristic SNPs used in the analysis

میانگین فاصله بین (Kb) SNP Average SNP interval	تعداد SNP Call rate No. of cut SNP from call rate	تعداد SNP حذف شده از حداقل فراوانی آلی No. of cut SNP from MAF	تعداد نشانگر بعد از کنترل کیفی No. of SNP after QC	تعداد نشانگر قبل از کنترل کیفی No. of SNP before QC	طول (Mb) Length	کروموزوم Chromosome
53.05	88	526	5168	5694	275.41	1
61.01	73	436	4036	4472	248.97	2
60.03	60	581	3191	3772	224.00	3
56.02	56	570	2098	2668	119.22	4
60.01	54	568	1778	2346	107.84	5
59.02	52	523	1957	2480	116.89	6
54.02	50	266	1845	2111	100.01	7
54.07	50	305	1653	1958	90.62	8
61.09	49	487	1546	2033	94.58	9
59.03	48	279	1461	1740	86.38	10
62.01	47	184	991	1175	62.17	11
60.05	45	308	1301	1609	79.03	12
65.02	45	322	1270	1592	82.95	13
63.02	42	182	991	1173	62.57	14
63.01	39	305	1278	1583	80.78	15
61.04	38	306	1170	1476	71.69	16
60.02	35	171	1204	1375	72.25	17
57.02	34	172	1201	1373	68.49	18
63.01	31	241	954	1195	60.45	19
60.05	30	244	845	1089	51.05	20
59.03	23	142	841	983	49.99	21
54.01	22	175	935	1110	50.78	22
63.02	18	147	985	1132	62.28	23
66.03	17	119	628	747	41.98	24
54.05	18	74	824	898	45.22	25
60.07	14	84	728	812	44.05	26

لایه‌بندی جمعیتی که ناشی از تفاوت در فراوانی آللی زیر جمعیت‌ها به دلیل تفاوت ژنتیکی جد مشترک و همچنین میزان خویشاوندی می‌تواند نتایج یک مطالعه ارتباط ژنومی را با مشکل مواجه نماید چرا که در مطالعات پویش ژنومی فرض می‌شود که جامعه مورد مطالعه همگن می‌باشد (پیرسون و مانولیو ۲۰۰۸). در نظر گرفتن اثرات خویشاوندی می‌تواند منجر به کاهش نتایج مثبت دروغین شود. تا کنون روش‌های متعددی جهت آنالیز داده‌ای پویش ژنومی با در نظر گرفتن اثرات خویشاوندی ارائه شده است که آنالیز PCA با رگرسیون تصادفی، GBLUP، روش‌های آماری مختلف بی‌زی، EMMAX، MLMM و مقیاس بندی چند بعدی (MDS) از مهمترین این روش‌ها می‌باشند. در این تحقیق از آزمون MDS برای لحاظ نمودن خویشاوندی در آنالیز داده‌ها به عنوان یک راهکار در کنترل نتایج مثبت دروغین استفاده شد. این روش ابعادی که فاصله ژنتیکی مشاهده شده را توضیح می‌دهد بر اساس روش Identity By State شناسایی می‌نماید. در این تحقیق جهت بررسی نحوه تأثیر لحاظ نمودن داده‌های این آزمون در کنترل نتایج، مدل اولیه فاقد هرگونه اثرات ثابت معنی‌دار بوده است که منجر به شناسایی ۳ منطقه هاپلوپیتی روی ۳ کروموزوم شد. در حالیکه با در نظر گرفتن نتایج آزمون MDS و همچنین اثر سال تولد و سال نمونه برداری (اثرات ثابت معنی‌دار) به عنوان عواملی که می‌توانند نتایج مثبت دروغین ایجاد کنند، این تعداد جایگاه به ۲ منطقه هاپلوپیتی روی ۲ کروموزوم کاهش یافت. علاوه بر این، پلات‌های مربوط به Q-Q در شکل ۱ ارائه شده است. همانطوریکه در پلات‌ها مشخص است لایه بندی جمعیتی ضعیفی را نشان می‌دهد که به ترتیب برابر با ۱/۱۲۷، ۱/۱۰۱، ۱/۰۵۹ و ۱/۰۰۹ برای صفات MDF، FDCV، $FD \geq 30$ و SL بودند. به عبارت دیگر، هر گونه انحراف از خط نشان دهنده وجود اختلاط و ارتباط ژنتیکی می‌باشد. آزمون MDS که در این تحقیق به کار گرفته شد به طور کلی انعطاف پذیرتر از روش PCA، از دیگر روش‌های متداول در کنترل نتایج در مطالعات ژنومی می‌باشد. استفاده از روش PCA نیازمند این است که داده‌های مورد استفاده توزیع نرمال چند جمله‌ای داشته باشند در

نتایج حاصل از مراحل مختلف کنترل کیفی

در این تحقیق ۹۶ رأس از گوسفندان نژاد زندی با استفاده از آرایه‌های 50K گوسفندی تعیین ژنوتایپ شدند. ابتدا جنسیت ثبت شده در مقابل جنسیت برآورد شده مورد آزمون قرار گرفت که نتایج این بخش صحت کامل ژنوتیپینگ و جنسیت ثبت شده را نشان داد. همچنین میزان تکرار پذیری در ۳ نمونه مورد بررسی قرار گرفت و نتایج این نمونه‌ها با نتایج قبلی همخوانی کامل داشته و به این ترتیب از صحت بالای ژنوتیپینگ اطمینان حاصل شد. از مجموع ۵۴۲۴۱ نشانگر به کار رفته در این تحقیق، ۴۰۸۷۹ نشانگر توانستند مراحل مختلف کنترل کیفیت را بگذرانند. به طور کلی ۷۷۱۷ نشانگر به دلیل حداقل فراوانی آللی کمتر از ۰/۰۵، ۱۰۷۰ نشانگر به دلیل نرخ ژنوتیپینگ کمتر از ۹۵٪ در هر نمونه، ۱۱۷ نشانگر به دلیل عدم تعادل هاردی واینبرگ، ۱۵۴۲ نشانگر روی کروموزوم X و ۲۹۱۶ نشانگر با موقعیت ناشناخته و همچنین ۲ نمونه به دلیل فراوانی ژنوتیپینگ کمتر از ۹۵٪ حذف شدند. تعداد نشانگرها به ترتیب کروموزوم، تعداد نشانگرها به ازای هر کروموزوم قبل و بعد کنترل کیفیت، تعداد نشانگرهای حذف شده به تفکیک هر کروموزوم و میانگین فاصله بین جفت SNPها در جدول ۲ نشان داده شده است.

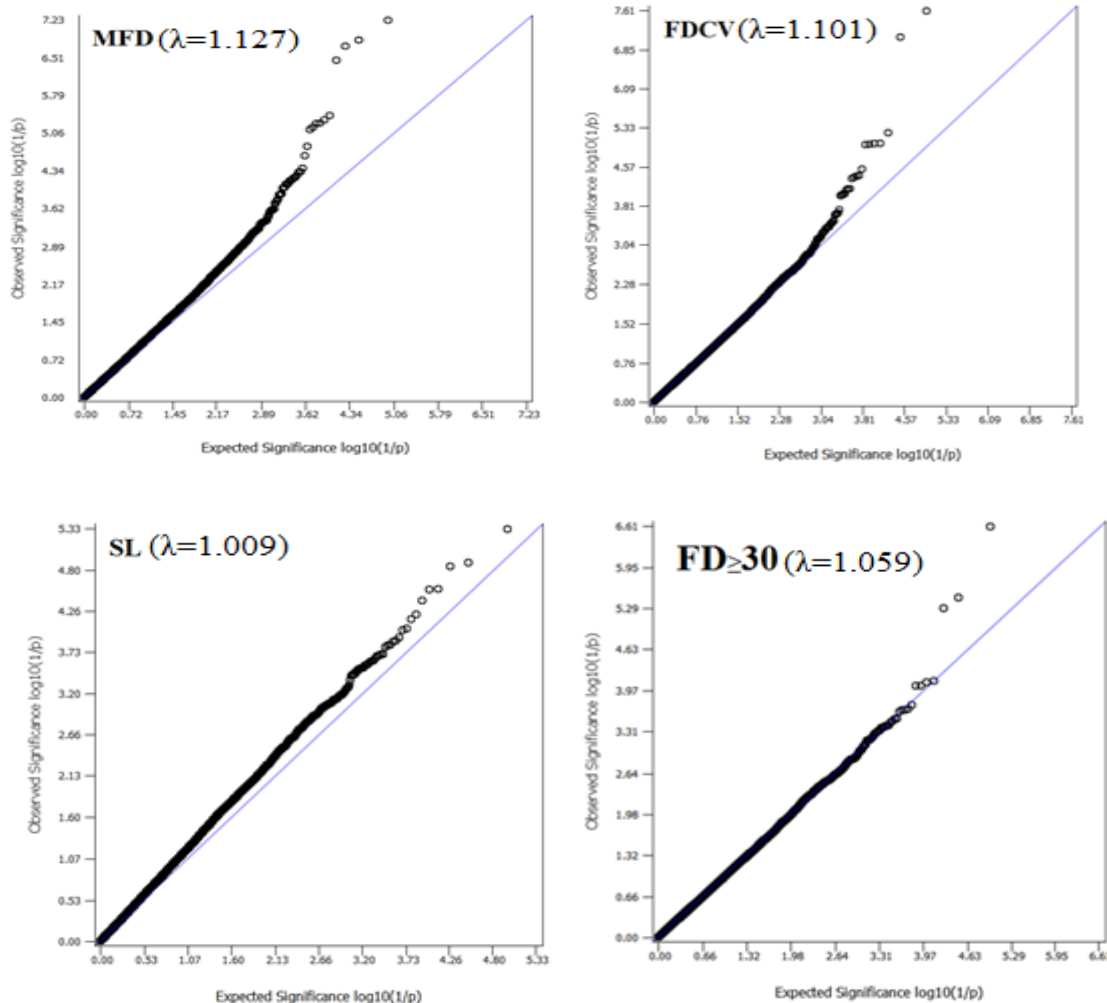
کنترل اشتباه نوع اول و تصحیح اثر لایه بندی

جمعیتی و میزان خویشاوندی

با افزایش تعداد نشانگر در مطالعات پویش ژنومی، حداقل سازی نتایج مثبت دروغین (اشتباه نوع اول) اهمیت بالاتری پیدا می‌کند. آزمون متعددی برای تصحیح نتایج پویش ژنومی وجود دارد که از مهمترین آنها می‌توان به آزمون‌های نرخ کشف خطای مثبت (بنجامین_هوشبرگ و بنجامین_یکوتیالی)، هولم، بنفرونی و سیداک به صورت تصحیح ارزش P به صورت تک مرحله‌ای اشاره کرد (چانگ و همکاران ۲۰۱۵). در تحقیق حاضر از تصحیح بنفرونی جهت حداقل سازی نتایج مثبت دروغین استفاده شد. بررسی نتایج این تحقیق نشان داد که نتایج تصحیح بنفرونی مشابه هولم و سختگیرانه‌تر نسبت به سیداک بود.

کوواریانس محاسبه شود که در این صورت نتایج کاربرد این آزمون در تصحیح لایه بندی جمعیتی برابر روش PCA می باشد (لی و یو ۲۰۰۸).

حالیکه آزمون MDS این محدودیت را ندارد. همچنین در آزمون PCA، ابتدا باید ماتریس کوواریانس محاسبه شود در حالیکه MDS برای هر نوع از تشابهات و فواصل می-تواند به کار رود که IBS یکی از این روشها می باشد. در یک حالت نیز آزمون MDS می تواند از روش ماتریس



شکل ۱- پلات‌های Q-Q برای ارزیابی لایه‌بندی جمعیتی در صفات قطر الیاف (FD)، ضریب تغییرات قطر الیاف (FDCV)، طول استاپل (SL) و نسبت الیاف که مساوی یا بزرگتر از ۳۰ میکرومتر ($F \geq 30$) در نژاد گوسفند زندی

Figure 1- Plots Q-Q for evaluation of stratification for Fibre diameter (FD), fiber diameter coefficient of variation (FDCV), Staple length (SL), and the proportion of fiber that are equal or more than 30 μ m in Zandi sheep breed

هاپلوتیپ‌ها ممکن است باعث شناسایی نواحی ژنومی با ارتباط معنی‌دار روی صفت مورد بررسی باشد که در آزمون ارتباط تک نشانگری شناسایی نمی‌شوند. با استفاده از نتایج آزمون GLM و سه عامل در آمیختگی شامل اثر سال تولد و سال نمونه برداری و لایه بندی

آزمون هاپلوتیپ‌ها

استفاده از آزمون هاپلوتیپ‌ها می‌تواند سبب بهبود قدرت آزمون‌های ارتباط ژنومی شود زیرا هاپلوتیپ‌ها نسبت به نشانگر انفرادی می‌توانند در LD بالاتری با QTL یا ژن‌های کاندیدا باشند. بنابراین استفاده از آزمون

ضریب تغییرات قطر الیاف دو قطعه هاپلو تپیی 1 Mb روی کروموزوم‌های ۱۱ و ۱۳ شناسایی شد (جدول ۳).

ژنتیکی به طور کامل همزمان به عنوان اثرات ثابت در آنالیز محاسباتی وارد شد. در این حالت ارتباط معنی-داری آزمون هاپلو تپیی‌ها و صفت میانگین قطر الیاف و

جدول ۳- ارتباط هاپلو تپیی‌ها با صفات تولیدی پشم

Table 3- Association haplotypes with wool production traits

P- value	صفت مرتبط Related-trait	اندازه پنجره (تعداد) (SNP)	نشانگر ۲- SNP-2	نشانگر ۱- SNP-1	کروموزوم- مکاباز Chr- Mb
1.55×10^{-5}	-	15	OAR1_144015756_X.1	OAR1_143087719.1	1-132
1.59×10^{-4}	-	20	OAR1_205123535.1	OAR1_204033889.1	1-189
2.34×10^{-4}	-	15	OAR1_278980576.1	OAR1_277983951.1	1-257
1.00×10^{-5}	-	20	OAR6_50026575.1	OAR6_49035950.1	6-44
4.12×10^{-3}	-	15	OAR6_79610534.1	OAR6_78519354.1	6-72
2.96×10^{-8}	FD	15	OAR11_42487494.1	s18734.1	11-39
2.45×10^{-7}	FD, FDCV	20	s45227.1	s55896.1	13-56
1.07×10^{-5}	-	15	OAR13_65432300.1	OAR13_64379923.1	13-59
6.11×10^{-5}	-	10	OAR15_84885513.1	OAR15_82519111.1	15-76
2.36×10^{-4}	-	15	OAR17_6510526.1	OAR17_5778603.1	17-5
4.28×10^{-3}	-	20	s42375.1	OAR18_17194010.1	18-17

سلول‌ها و فسفریله نمودن پروتئین در ارتباط هستند (Genecards).

در مطالعه‌ای با استفاده از تکنولوژی RNA-Seq با هدف تعیین تفرق بیان ژن‌های تعیین کننده میزان و درصد پروتئین شیر در اوج شیرواری در گاوهای شیری هلشتاین، ژن کاندیدای ERBB2 به عنوان یکی از ده ژن کاندیدای مؤثر بر تنظیم سنتز پروتئین شیر گزارش شده است (لی و همکاران، ۲۰۱۶).

دومین ناحیه هاپلو تپیی مؤثر بر میانگین قطر الیاف و ضریب تغییرات قطر الیاف بود، روی کروموزوم ۱۳ در ناحیه ۵۶ مگابازی شامل ۲۰ نشانگر و حاوی ۸ ژن بود که مهمترین مرتبط با صفات مورد مطالعه، ژن کاندیدا GNAS می‌باشد. ژن پروتئین اتصال- نوکلئوتیدی گوانین زیر واحد آلفا دارای نقش سیگنالینگ از طریق فعال نمودن آدنیلات سیکلازها (آنزیم کلیدی تبدیل آدنوزین تری فسفات به آدنوزین مونو فسفات حلقوی) می‌باشد و از این طریق در نتیجه باعث افزایش فعالیت مولکولهای آدنوزین مونو فسفات حلقوی cAMP می‌شود. cAMP یک پیام رسان ثانویه در بسیاری از واکنش‌های سلولی می‌باشد (Genecards). ژن GNAS همچنین در فرآیندهای

اولین ناحیه هاپلو تپیی معنی‌دار روی کروموزوم شماره ۱۱ در ناحیه ۳۹ مگابازی شامل ۱۵ نشانگر SNP و حاوی ۲۰ ژن مختلف بود که مهمترین ژن مرتبط با صفات مورد مطالعه، ژن کاندیدا ERBB2 می‌باشد.

ژن کاندیدای گیرنده تیروزین کیناز ۲ (ERBB2)، که HER2 هم شناخته می‌شود، عضوی از خانواده گیرنده تیروزین کینازها می‌باشد که ژن‌های کد کننده اعضای پروتئین‌های فاکتور رشد اپیدرمی (EGF) هستند. این پروتئین‌ها بواسطه تحریک رشد و تحریک انواع مختلفی از سلول‌ها از جمله سلولهای اپیدرمی پوست نقش کلیدی داشته و نوعی سیتوکین می‌باشند. فاکتور رشد اپیدرمی همچنین باعث افزایش ساخت پروتئین و RNA در سلول-های اپیدرمی در انسان می‌شود (Genecards).

ژن کاندیدای ERBB2 در انسان در رشد و توسعه مو از طریق استحکام نمودن میکروتوبول‌های سلول‌های کورتکس (لایه میانی مو) نقش دارد (UniProtKB). کورتکس بخش اصلی مو به شمار می‌رود و تعیین کننده استحکام، رنگ و بافت مو می‌باشد و نقش اصلی در اتصال های کراتین‌های مو دارد. علاوه براین، ژن کاندیدای ERBB2 در فرآیند بیولوژیکی تنظیم مثبت رشد

هاپلوتیپی معنی‌دار روی کروموزوم ۱۳ در منطقه‌ای از ژنوم تشخیص داده شد که قبلاً یک QTL مرتبط با میانگین قطر پشم گزارش شده بود. ویتزیکا و همکاران (۲۰۰۷) گزارش نمودند با تحقیق روی گوسفندان INRA 401 شامل ۱۳۱۰ بره حاصل از تلاقی ۳۰ قوچ و ۸۵۲ میش بود، ژنوتایپ نمودن ۴ میکروستلایت روی کروموزوم ۱۳، در ناحیه 64.8-64.9 Mb ارتباط معنی‌داری با میانگین قطر الیاف گزارش نمودند.

نتیجه‌گیری کلی

در این مطالعه با توجه به اینکه از دو آزمون GC و MDS برای بررسی ساختار جمعیتی گوسفند نژاد زندگی به کار برده شد، نتایج به دست آمده حاکی از عدم وجود لایه بندی و اختلاط ژنتیکی در این جمعیت مورد بررسی بود. علاوه بر این یافته‌های آزمون ارتباط ژنومی هاپلوتیپی منجر به شناسایی ژن‌های کانیدیا ERBB2 و GNAS گردید. ژن‌های کانیدای شناسایی شده، عملکرد مولکولی مرتبط با صفات الیاف پشم داشتند. انتخاب برای بهبود کیفی در پشم تولیدی در گوسفندان زندگی با استفاده از این یافته‌ها، باعث تسریع در پیشرفت ژنتیکی خواهد گردید.

سپاسگزاری

نویسندگان مقاله از خانم زینب عمرانی به دلیل کمک در اندازه‌گیری صفات پشم در آزمایشگاه کمال تشکر را دارند. همچنین از مرکز اصلاح نژاد و بهبود تولیدات دامی کشور به خاطر حمایت مالی پروژه و ایستگاه پرورش و اصلاح نژاد گوسفند زندگی استان تهران (ایستگاه خجیر) برای در اختیار گذاشتن حیوانات در اجرای این پروژه کمال تشکر را داریم.

بیولوژیکی مهمی مانند پاسخ سلولی به لیپولین، گلیکوننوژنزیس، تحریک کلاژن و تشکیل فولیکول‌ها مو در جنین نقش دارند (UniProtKB). همچنین ارتباط ژن GNAS در انسان با رشد و توسعه پوست به اثبات رسیده است (UniProtKB). در حیوانات، ارتباط معنی‌داری بین چند شکلی در ژن کانیدای GNAS با صفات میانگین خوراک مصرفی و وزن زنده، در ۱۳۸ رأس گاو گوشتی لیموزین گزارش شده است (مگی و همکاران، ۲۰۱۰).

بررسی QTL‌های گزارش شده در مناطق مورد نظر به ابعاد 1 Mb در اطراف هر هاپلوتیپ معنی‌دار نشان داد که در این مناطق جایگاه‌های مدنظر با QTL‌های گزارش شده برای صفات میانگین قطر الیاف و ضریب تغییرات قطر الیاف گزارش شده است.

بررسی QTL‌های مرتبط با کیفیت پشم در مناطق هاپلوتیپی شناسایی شده، نشان دهنده وجود یک QTL در ناحیه 66.6-70.6 cM روی کروموزوم ۱۱ مرتبط با ضریب تغییرات قطر الیاف (FDCV) در منطقه ژنومی شناسایی شده در این تحقیق بود. رولدان و همکاران (۲۰۱۰) با بررسی خصوصیات پشم در نژاد گوسفند مرینوس با استفاده از ۳۶ نشانگر ریزماهورای، منطقه-ای ژنومی را گزارش کردند که با ضریب تغییرات قطر الیاف (۱۷/۴۶٪ واریانس) در این نژاد در ارتباط معنی‌دار بود ($P < 0.05$). QTL گزارش شده در موقعیت 38.1-40.7 Mb قرار داشت که در محدوده هاپلوتیپ شناسایی شده روی کروموزوم شماره ۱۱ در این تحقیق بود. علاوه بر جایگاه مذکور، در ناحیه 9.8 Mb روی کروموزوم ۱۳، QTL در ارتباط با ضریب تغییرات قطر الیاف گزارش شده است که در تحقیق حاضر هاپلوتیپ معنی‌دار در ناحیه 23 Mb شناسایی و تشخیص داده شد. دومین ناحیه

منابع مورد استفاده

- Abasht B and Lamont SJ, 2007. Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population. *Animal Genetics* 38: 491–8.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM and Lee JJ, 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4:7.
- Daetwyler HD, Schenkel FS, Sargolzaei M and Robinson JA, 2008. A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. *Journal of Dairy Science* 91: 3225–3236.

- Dorostkar M, Shodja J, Rafat SA and Esfandyari H, 2013. Evaluation of some of wool fiber characteristics in Moghani sheep using image analysis method. *Animal Science Researches* 23: 37-45 (In Persian).
- Food and Agriculture Organization of the United Nations. 2013. www.FAOSTAT.com.
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J and Blumenstein B, 2002. The structure of haplotype blocks in the human genome. *Science* 296: 2225-2229.
- Hu ZL, Park CA and Reecy JM, 2016. Developmental progress and current status of the Animal QTLdb. *Nucleic Acids Research* 44: 827-833.
- Johnston SE, McEwan JC, Pickering NK, Kijas JW, Beraldi D, Pilkington JG, Pemberton JM and Slate J, 2011. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology* 20: 2555–2566.
- Li Q and Yu K, 2008. Improved correction for population stratification in genom wide association studies by identifying hidden population structures. *Genetic Epidemiology* 32: 215-226.
- Li C, Cai W, Zhou C, Yin H, Zhang Z, Looor JJ, Sun D, Zhang Q, Liu J and Zhang S, 2016. RNA-Seq reveals 10 novel promising candidate genes affecting milk protein concentration in the Chinese Holstein population. *Scientific Reports* 6: 26813.
- Magee DA, Berkowicz EW, Sikora KM, Berry DP, Park SD, Kelly AK, Sweeney T, Kenny DA, Evans RD, Wickham BW, Spillane C and Machugh DE, 2010. A catalogue of validated single nucleotide polymorphisms in bovine orthologs of mammalian imprinted genes and associations with beef production traits. *Animal* 4: 1958–1970.
- Mohammadi H, Moradi Shahrehabak M, Vatankhah M and Moradi shahrehabak H, 2012. Direct and maternal (co)variance components, genetic parameters, and annual trends for growth traits of Makooei sheep in Iran. *Tropical Animal Health and Production* 45: 185-191.
- Mohammadi H, Moradi Shahrehabak M and Sadeghi M, 2013a. Association between single nucleotide polymorphism in the ovine DGAT1 gene and carcass traits in two Iranian sheep breeds. *Animal Biotechnology* 24: 159–167.
- Mohammadi H, Moradi Shahrehabak M and Moradi Shahrehabak H, 2013b. Analysis of Genetic Relationship between Reproductive vs. Lamb Growth Traits in Makooei Ewes. *Journal of Agricultural Science and Technology* 15: 45-53.
- Pearson TA and Manolio T, 2008. How to interpret a genome-wide association study? *Journal of the American Medical Association* 299: 1335–1344.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D and Lantier H, 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. *The American Journal of Human Genetics* 81: 559-575.
- Rastifar M, Nejati-Javaremi A, Moradi MH and Abdollahi-Arpanahi R, 2015. Identification of genomic regions associated with wool diameter in Iranian sheep breeds. *Iranian Animal Science Journal* 46: 65-72. (In Persian).
- Roldan DL, Dodero AM, Bidinost F, Taddeo HR, Allain D, Zhang M, and Li S, 2010. Merino sheep: a further look at quantitative trait loci for wool production. *Animal* 4: 1330–1340.
- Schroder W, Klostermann A, Stock KF and Distl O, 2012. A genome-wide association study for quantitative trait loci of show-jumping in Hanoverian warmblood horses. *Animal Genetics* 43: 392–400.
- Teo YY, Fry AE, Clark TG, Tai ES and Seielstad M, 2007. On the usage of HWE for identifying genotyping errors. *Annals of Human Genetics* 71: 701-703.
- Vitezica ZG, Moreno CR, Lantier F, Lantier I, Schibler L, Roig A, François D, Bouix J, Allain D, Brunel JC, Barillet F and Elsen JM, 2007. Quantitative trait loci linked to PRNP gene controlling health and production traits in INRA 401 sheep. *Genetic Selection Evaluation* 39: 421–30.
- Wang S, Dvorkin D and Da Y, 2012. SNPEVG: a graphical tool for GWAS graphing with mouse clicks. *BMC Bioinformatics* 13:319.
- Wang Z, Zhang H, Yang H, Wang S, Rong E, Pei W, Li H and Wang N. 2014. Genome-wide association study for wool production traits in a Chinese Merino sheep population. *PLoS One* 9: e107101.

An assessment of population stratification and haplotype based Genome-wide association for wool quality traits in Zandi sheep breed

H Mohammadi¹, SA Rafat^{2*}, H Moradi Shahrebabak³, J Shodja³ and MH Moradi⁴

Received: June 10, 2017

Accepted: September 18, 2017

¹PhD Graduated Department of Animal Science, Faculty of Agricultural Sciences, University of Tabriz, Tabriz, Iran

²Professors, Department of Animal Science, Faculty of Agricultural Sciences, University of Tabriz, Tabriz, Iran

³Assistant Professor, Department of Animal Sciences, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

⁴Assistant Professor, Department of Animal Sciences, Faculty of Agriculture and Natural Resources, University of Arak, Arak, Iran

*Corresponding author: abbasrafat@hotmail.com

Introduction: Identifying of genes with large effects on economically important traits, has been one of the important goals in sheep breeding. Over the last decade, by the advent of genome-wide panels of single nucleotide polymorphisms (SNPs), it has become possible to identify and localize QTLs for complex traits in many livestock species. One important obstacle in association studies is the confounding effect of population structure. Because this effect generally increases in proportion to population size, population structure remains a major concern in association analyses. To date, 98 QTLs for wool traits have been reported via genome scan based on marker-QTL linkage analyses (<http://cn.animalgenome.org/cgi-bin/QTLdb/OA/index>, 27 Aug, 2017). Compared with traditional QTL mapping strategies, a GWAS has major advantages both in its power to detect causal variants with modest effects and in defining narrower genomic regions harboring causal variants for economically important traits. In this study, we assessed population stratification and performed a genome-wide association study (GWAS) of wool quality traits in Zandi sheep.

Material and methods: A total of 96 Iranian Zandi sheep was used in the study. The following four traits were analyzed: staple length (SL), mean fiber diameter (MFD), fiber diameter coefficient of variation (FDCV), and the proportion of fiber that was equal or more than 30 μm ($F_{\geq 30}$). Animals were genotyped using 50 K SNPChip panel. Quality control of the genotype data consisted in removing SNPs with a call rate less than 95 %, SNPs with a minor allele frequency (MAF) less than 5 %, SNPs with more than x % missing genotypes, and SNPs that deviated strongly from Hardy–Weinberg equilibrium ($P < 10^{-6}$). The PLINK was used to generate the matrix using the GLM algorithm. In this study, although the resources of this breed were very clear, we still examined the distribution of the test statistics obtained from the numerous association tests. We also assessed their deviation from the expected distribution of no SNPs being associated with the trait of interest using a quantile-quantile (Q-Q) plot, which is commonly used to analyze population stratification in GWAS. We use SNPEVG tool to show the (Q-Q) plot. The Bonferroni method was used to adjust for multiple testing from the number of SNP loci detected. We declared a significant SNP at the genome-wide significance level if the raw P-value was $0.05/N$, here N is the number of SNP loci tested in the analysis. The exact positions of the annotated genes were extracted from the latest sheep genome *Oar_v4.0* assembly along with the NCBI annotation release 102 of the sheep genome. To investigate if the significant SNPs detected in this study were within the range of previously identified QTL for relevant traits, we searched for meat or production QTL in the Animal QTLdb within a 1-Mb region on both sides of each significant haplotype.

Results and discussion: After quality control, 2 individuals were excluded, leaving 94 sheep for the association analysis. Additionally, we removed 1070 SNPs with call rates less than 95% and 7717

SNPs with MAF less than 0.05. A total of 40,879 SNPs passed these quality-control filters and were retained in the dataset. These SNPs were distributed across 26 autosomes, with the number of SNPs per chromosome ranging from 747 to 5694, and with a mean distance between adjacent SNPs ranging from 50.4 to 68.7 kb. Wool MFD was $29.85 \pm 0.03 \mu\text{m}$ with an individual sheep range of 22.4-39.04 μm . The overall coefficient of variation of fiber diameter was $43.12\% \pm 0.7\%$ with an individual sheep range of 19.7%–68.0%. The average percentages of fiber that had equal or more than 30 μm were $27.04 \pm 0.03\%$ with an individual sheep range of 12.04-43.10%. Average wool staple length was $11.25 \pm 0.03 \text{ cm}$ with an individual sheep range of 6-19 cm. The result from genomic control showed weak population stratification for SL, MDF, FDCV and $F \geq 30$ between populations of Zandi sheep. The genomic inflation factors (λ_{GC}) for the four traits were equal to 1.127, 1.101, 1.059, and 1.009 for MDF, FDCV, $FD \geq 30$ and SL, respectively. However, the Q-Q plots clearly showed there was no evidence of any systematic bias due to population structure or analytical approach in our case. Overall, two significant SNPs at the genome-wise level were identified for FD, and FDCV. No significant SNPs was identified for SL or $FD \geq 30$. Two Haplotype region within ERBB2 and GNAS genes previously reported in human growth and development hair and skin. Haplotypes were located with previously QTL reported to affect fiber diameter and fiber diameter coefficient of variation in Merino and INRA401 breed sheep. The functions of all of the above genes are directly or indirectly related to skin and hair development. Hair follicles are skin appendages and produce hair; therefore, we hypothesize that these genes control hair follicle development and fiber diameter trait.

Conclusion: The results of this study could provide a suite of novel SNP markers and candidate genes associated with wool traits and hence, may play an important role in understanding the biology of wool traits in fat-tailed sheep.

Keywords: Candidate gene, Genome-wide association study, Population stratification, Wool quality, Zandi sheep.