ژنتیک نوین
دوره یازدهم، شماره ۳، پاییز ۱۳۹۵
صفحه ٤٦٧ - ٤٥٩

شناسایی نواحی ژنومی تحت انتخاب با استفاده از نشانگرهای تک نوکلئوتیدی در گوسفندان مرینوس

A Genomic Scan for Detection of Selection Signatures using SNP Data in Australian Merino Sheep

مهدیه منتظری[۱]، مسعود اسدی‌فوزی*[۱]، علی اسمعیلی‌زاده کشکوئیه[۱]، محمدحسین فردوسی[۲]، جولیوس ون در ورف[۳]

۱- به‌ترتیب دانشجوی دکتری، دانشیار، استاد، ژنتیک و اصلاح دام، گروه علوم دامی، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان، کرمان، ایران
۲- محقق مرکز ژنتیک و اصلاح نژاد دانشگاه نیوانگلند، آرمیدل، استرالیا
۳- استاد، دانشکده کشاورزی و علوم روستائی دانشگاه نیوانگلند، آرمیدل، استرالیا

Montazeri M[1], Asadi Fozi M[*1], Esmailizadeh Koshkoieh A[2], Ferdosi MH[2], van der Werf J[3]

1- PhD Student, Associate Professor, Professor, Department of Animal Science, Faculty of Agriculture, Shahdi Bahonar University of Kerman, Kerman, Iran
2- Research Fellow, Animal Genetic and Breeding Unit, University of New England, Armidale, Australia
3- Professor, School of Environmental and Rural Science, University of New England, Armidale, Australia

* نویسنده مسئول مکاتبات، پست الکترونیکی:masadi@uk.ac.ir

(تاریخ دریافت: ۹٤/٥/۲۷ - تاریخ پذیرش: ۹٥/٦/۲۸)

چکیده

شناخت اساس ژنتیکی تنوع فنوتیپی، یکی از اهداف اصلی تحقیقات زیستی است و استفاده از حیوانات اهلی ابزاری سودمند برای پیشرفت در راستای این هدف می‌باشد. نواحی تحت انتخاب، نواحی‌ای از ژنوم هستند که دارای اهمیت عملکردی می‌باشند و تحت انتخاب طبیعی و مصنوعی قرار دارند. در این مطالعه، کاوش ژنومی با ٥٠٠٠٠ نشانگر تک نوکلئوتیدی به‌منظور شناسایی نواحی تحت انتخاب ۳۰۰۰ فرد گله‌های گوسفند مرینوس استرالیائی صورت گرفت. پنج ناحیه ژنومی روی چهار کروموزوم متفاوت به‌عنوان نواحی تحت انتخاب شناسایی شدند. این مناطق با ارزش نااریب شاخص تثبیت (تتا) بالای ۰/۱٤، روی کروموزوم‌های ٦، ۷، ۱۱ (دو ناحیه) و ۲٦ واقع شده‌اند. بعد از انطباق مناطق ژنومی تحت انتخاب با مناطق ژنومی گوسفند، نه ژن کاندیدا مرتبط با صفات تولید مثل و رشد شناسایی شد. در نهایت بررسی QTLهای گزارش شده نشان داد که این مناطق با QTLهای صفات مهم اقتصادی از جمله صفات مرتبط با لاشه، رشد و پشم در ارتباط می‌باشند. مطالعه بیش‌تر این نواحی در شناسایی ژن‌های کاندید برای صفات مهم اقتصادی در نژادهای گوسفند موثر خواهد بود.

واژه‌های کلیدی

تفرق جمعیتی
چند شکلی تک نوکلئوتیدی
کاوش ژنومی
نشانه‌های انتخاب

مقدمه

عدم تعادل فاز گامتی (LD) یکی از عوامل مهم در بررسی جهش‌های صفات مهم اقتصادی در حیوانات مزرعه است (Pollinger et al. 2007). عواملی چون جهش، مهاجرت، رانش و انتخاب باعث به وجود آمدن عدم تعادل فاز گامتی ژنتیکی می‌شود (Chevin et al. 2008). چنانچه یک جهش جدید باعث افزایش شایستگی افراد حامل آن نسبت به سایر افراد جامعه شود، انتخاب باعث می‌شود افرادی که دارای شایستگی بیش‌تری هستند در تشکیل نسل بعد مشارکت بیش‌تری داشته باشند. بدین ترتیب فراوانی واریانت جهش یافته، بسته به سهم آن در افزایش شایستگی به سرعت شروع به افزایش خواهد کرد (Sabeti et al. 2006). به عبارت دیگر، هنگامی که انتخاب مثبت به سمت افزایش فراوانی آلل مطلوب پیش می‌رود، جایگاه‌های نزدیک به این جایگاه اگر هم اثرشان خنثی باشد به علت پیوسته بودن با آلل مطلوب، تحت تاثیر قرار می‌گیرد که در نتیجه این پدیده، الگوی تنوع ژنتیکی و عدم تعادل فاز گامتی در جایگاه‌های اطراف این جهش انتخابی تغییر خواهد کرد. در ژنتیک جمعیت به این نواحی، نواحی تحت انتخاب[1] یا جاروی انتخاب[2] گفته می‌شود.

به‌دلیل اهمیت شناسایی ژن‌های مؤثر بر صفات مهم اقتصادی، جستجو برای ژن‌های مسئول در تنوع صفات کمی در دو مسیر فنوتیپ به ژنوم و ژنوم به فنوتیپ (مسیر مخالف) صورت می‌گیرد. در مسیر اول یعنی شناسایی ژن‌ها با مطالعات پویش کل ژنوم و ارتباط آن با فنوتیپ، نقشه‌یابی بر اساس LD، QTL یا ژن‌های کاندیدا انجام می‌گیرد. اما در مسیر دوم ارزیابی آماری داده‌های ژنومی برای شناسایی نواحی‌ای که در طول زمان هدف انتخاب قرار گرفته‌اند، مورد استفاده قرار می‌گیرد. اساس این روش شناسایی LD داخل و بین جمعیت‌ها که با فرضیه تئوری خنثی سازگار نیست می‌باشد (Qanbari et al. 2010). هم‌چنین به علت نداشتن شجره مناسب، رکوردهای دقیق برای سایر صفات و جمعیت‌های اصلاح نژادی با ساختار مشخص، شناسایی ژن‌های مؤثر بر صفات کمی از مسیر فنوتیپ به ژنوم بسیار مشکل و از اریبی بالایی برخوردار است. بنابراین شناخت ژن‌های مؤثر بر صفات کمی از مسیر دوم با دقت بالاتر و در زمان کوتاه‌تری صورت می‌گیرد.

پس از این‌که در سال ۲۰۰۹ اولین تراشه SNP در گوسفند در قالب پروژه Sheep Hap Map طراحی شد، تحقیقات موفقیت‌آمیز مختلفی در این زمینه در گوسفند و برخی حیوانات اهلی دیگر صورت گرفته است (Porto-Neto et al. 2014; Pan et al. 2013; Moradi et al. 2012). به عنوان مثال، گروهی در سال ۲۰۰۸ به مطالعه الگوی عدم تعادل فاز گامتی در سطح دو کروموزوم ۱۹ و ۲۹ گاوهای هلشتاین و آنگوس پرداختند. نتایج نشان داد که نواحی شناسایی شده در گاو هلشتاین اغلب با تولید شیر و اجزای آن و در گاوهای آنگوس با ژن‌های صفات لاشه مرتبط می‌باشند (Prasad et al. 2008). هم‌چنین، در پروژه گسترده‌ای در سال ۲۰۱٤ گروهی از دانشمندان جهت شناسایی نواحی تحت انتخاب در جمعیت گوسفندان سراسر جهان، بار دیگر پروژه Sheep Hap Map را مورد تجزیه و تحلیل قرار دادند (Fariello et al. 2014). نتایج پویش ژنومی سیگنال‌های انتخاب، ۳۱ ناحیه مرتبط با رنگ، ریخت‌شناسی استخوان، رشد و تولید مثل را آشکار نمود. شناسایی مناطق ژنومی دو نژاد گوسفند زل و لری بختیاری با حدود ٥٤۰۰۰ نشانگر SNP و بررسی تمایز جمعیتی با استفاده از دو روش F_{ST} ویر و کوکهام نشان داد که چندین منطقه ژنومی بر روی کروموزومهای ۲، ۵ و ۷ دارای شواهدی از انتخاب می‌باشند. بررسی QTLهای گزارش شده در مناطق اورتولوگوس گاوی نیز حاکی از آن بود که این مناطق با QTLهای صفات مهم اقتصادی مرتبط با لاشه و تولید مثل همپوشانی دارند (Moradi et al. 2012).

شناسایی نشانه‌های انتخاب، یکی از اهداف بزرگ در ژنتیک جمعیت می‌باشد و به‌دلیل شناخت مکان‌های ژنی دارای عملکرد و آشکار نمودن تنوع ژنتیکی زیربنای صفات کمی دارای اهمیت است. هم‌چنین شناسایی این نواحی، اطلاعات ارزشمندی در ارتباط با تکامل و مکانیسم‌های پایه که بر روی تکامل مولکولی تاثیر می‌گذارند را فراهم می‌نماید (Nielsen 2005; Biswas and Akey 2006). ژن‌هایی که دارای فراوانی بالا و هم‌چنین EHH[3] بالایی در مجاور آلل انتخابی باشند هدف انتخاب‌های مثبت بوده

[1] Signatures of Selection
[2] Selective Sweep
[3] Extended Haplotype Homozygosity

و به‌خصوص در دام‌ها، که علاوه بر انتخاب طبیعی تحت انتخاب مصنوعی نیز قرار دارند می‌توانند پاره‌ای از جایگاه‌های کاندیدا برای ژن‌های عمده اثر باشند (Hayes et al. 2009). به‌طور کلی، مناطقی از ژنوم که تحت انتخاب هستند بایستی اهمیت شایستگی و عملکردی داشته باشند، در غیر این‌صورت انتخاب آن‌ها را تحت تاثیر قرار نمی‌دهد. با توجه به مطالب بیان شده، این مناطق ژنومی عمدتاً با ژن‌های عمده اثر و صفات مهم اقتصادی همراه هستند و دارای اهمیت زیادی می‌باشند. هدف از این تحقیق، شناسایی مناطقی از ژنوم بود که در گله‌های گوسفند مرینوس استرالیایی طی سالیان متمادی به‌صورت طبیعی یا مصنوعی هدف انتخاب‌های مختلف قرار گرفته‌اند. این مطالعه می‌تواند مبنای انجام تحقیقات تکمیلی پویش کل ژنوم گله‌های مذکور به منظور شناسایی نواحی تحت انتخاب جهت بهبود راندمان برنامه‌های انتخاب و استراتژی‌های اصلاح نژادی آن باشد.

مواد و روش‌ها

در مجموع از ۳۰۰۰ نمونه خون گوسفند مرینوس استرالیایی از نه گله شامل هشت گله CRC[1] و یک گله SG[2] استفاده شد. این نمونه‌ها توسط دانشگاه نیوانگلند استرالیا جمع‌آوری شدند. لازم به ذکر است که مدیریت استخراج DNA و تعیین ژنوتیپ نیز توسط این دانشگاه انجام شد. در نهایت داده‌های ژنومی حاصله در اختیار تحقیق حاضر قرار گرفتند. پس از اطمینان از کمیت و کیفیت بالای DNA استخراج شده، نمونه‌ها با استفاده از آرایه‌های شرکت ایلومینا (Illumina Ovine SNPchip 50k) تعیین ژنوتیپ شدند (Illumina Inc., San Diego, USA). جهت اطمینان از کیفیت داده‌های حاصل از تعیین ژنوتیپ در تجزیههای نهایی، ویرایش داده‌ها و مراحل مختلف کنترل کیفیت[3] با استفاده از نرم‌افزار Excel و Plink (Purcell et al. 2007) بر روی داده‌های اولیه تعیین ژنوتیپ شده به‌عنوان معیاری از خطای تعیین ژنوتیپ، اعمال شد. مراحل کنترل کیفیت شامل حداقل فراوانی آللی (MAF[4])، نرخ خواندن یا داده گمشده به ازای هر فرد (Call rate)، نرخ ژنوتایپینگ نشانگرها در هر نمونه و بررسی تعادل هاردی-واینبرگ بود. در ابتدا نمونه‌هایی که فراوانی ژنوتایپینگ آن‌ها کمتر از ۹۵ درصد بود، شناسایی و حذف شدند. این نمونه‌ها احتمال بیش‌تری داشته که با داده‌های گمشده همراه بوده و خطای ژنوتایپینگ در آن‌ها بالا باشد. در مرحله بعد SNPهایی که در مجموع حیوانات دارای حداقل فراوانی آللی کمتر از یک درصد و حیوانات با بیش از پنج درصد ژنوتیپ از دست رفته، حذف شدند. در پایان نیز با در نظر گرفتن سطح احتمال 10^{-6} برای نشانگرهای باقیمانده، نشانگرهایی که در هر کدام از جمعیت‌های مورد مطالعه دارای تعادل هاردی-واینبرگ نبودند، به‌عنوان معیاری از خطای تعیین ژنوتیپ کنار گذاشته شدند.

تکنیک آماری تجزیه و تحلیل مولفه‌های اصلی (PCA[5])، روشی از تجزیه‌های چند متغیره آماری است که تعداد کمتری از عوامل را به‌نام مولفه‌های اصلی از میان عوامل اولیه گزینش می‌کند، به طوری که تعدادی از اطلاعات کم اهمیت حذف می‌شوند (Truxillo and Hamer 2007). در این تحقیق، تجزیه PCA با استفاده از مدل ۱ و دستور prcomp در محیط R صورت گرفت (R Core Team 2015).

$$PC_i = \sum w_{ij} x_i \qquad (۱)$$

در رابطه بالا PC_i معرف مولفه مورد نظر، w_{ij} ضریب مربوط به متغیرهای اولیه و x_i نیز iامین متغیر اولیه است. ضرایب w_{ij} طوری تخمین زده می‌شوند که اولین مؤلفه حداکثر واریانس داده‌ها را در نظر گرفته و دومین مؤلفه حداکثر واریانس در نظر گرفته نشده توسط مؤلفه اول را پیش بینی کرده و این روند ادامه می‌یابد تا آخرین مؤلفه، تمامی واریانس مورد نظر را در بر گیرد.

نتایج تحقیقات گسترده در طی چند سال اخیر نشان داده است که اهلی کردن و به دنبال آن شکل‌گیری نژادها و انتخاب طبیعی یا مصنوعی باعث ایجاد سیگنال‌های انتخاب در مناطق ژنوم شده است (Hayes et al. 2009; Stella et al. 2010; Qanbari et al. 2011). این جایگاه‌ها را می‌توان از طریق مقایسه تمایز جمعیتی در نژادهای مختلف با استفاده از روش‌های آماری مناسب شناسایی نمود. یک جمعیت در هر مرتبه‌ای که تقسیم شود،

[1] Cooperative Research Center
[2] SheepGENOMICS
[3] Quality Control
[4] Minor Allele Frequency
[5] Principal Component Analysis

هتروزیگوسیتی متفاوتی را نشان خواهد داد و میزان این تفاوت ژنی تحت تاثیر معیارهایی از قبیل ضریب هم‌خونی در افراد نسبت به جمعیت (F_{IS})، ضریب هم‌خونی در افراد نسبت به کل (F_{IT}) و ضریب هم‌خونی در زیرجمعیت‌ها (F_{ST}) قرار دارد. معیار F_{ST} (رابطه ۲) کاهش در هتروزیگوسیتی بر اثر تفرق آللی را نشان می‌دهد. در واقع این آماره بیانگر کاهش هتروزیگوسیتی در زیرجمعیت‌ها نسبت به آنچه که در حالت طبیعی مورد انتظار است، می‌باشد (Wright 1992).

$$F_{ST}= H_T - H_S/ H_T \quad (۲)$$

در رابطه فوق H_T و H_S به ترتیب میزان کل هتروزیگوسیتی و میانگین هتروزیگوسیتی در زیر جمعیت‌ها می‌باشد.

یکی از مشکلات اصلی روش Fst این می‌باشد که خطای نمونه-گیری را در نظر نمی‌گیرد که این مورد با استفاده از روش نااریب تتا (Weir and Cockerham 1984) تصحیح شده است. روش‌های Fst و تتا هر دو تفاوت در ژنوم را نشان می‌دهند. زمانیکه تعداد حیوانات در جمعیت‌های متفاوت، برابر باشد از روش Fst استفاده می‌شود اما در مواقعی که تعداد حیوانات جمیعت‌های مختلف، متفاوت باشند جهت لحاظ نمودن اریب ناشی از اندازه جمعیت-ها، روش تتا مورد استفاده قرار می‌گیرد (Weir and Cockerham, 1984). لذا برای بررسی نواحی تحت انتخاب در گله‌های مورد مطالعه، آماره نااریب تتا برای هر جایگاه در محیط R برآورد شد. جهت بررسی ژن‌ها و QTLهای احتمالی گزارش شده در گوسفند در مناطق ژنومی مورد نظر، ابتدا ژن‌هایی که تاکنون در نواحی مورد نظر گزارش شده بودند با استفاده از Biomart Ensemble (Hubbard et al. 2009) شناسایی شدند. سپس توالی ژنومی گوسفند در این ناحیه در فاصله‌ای که SNPها در صدک ۹۹/۹ بالای ارزش تتا واقع شده بودند، با استفاده از Sheep Genome Browser (Dalrymple et al. 2007) بدست آمد. برای تعیین موقعیت ژنومی SNPها در سطح ژنوم از OAR true chromosome (ver.1.0, as at 5/2008) مرکز CSIRO استرالیا استفاده شد. هم‌چنین توالی‌های اورتولوگوس بر روی ژنوم گاو با استفاده از جستجوی BLAT در UCSC Genome Browser (Baylor 4.0/bosTau4 Assembly, Oct.2007) شناسایی شد. در نهایت نیز جهت بررسی همپوشانی مناطق مورد نظر با QTLهای شناسایی شده، سامانه‌های اطلاعاتی http://www.animalgenome.Org/QTLdb/sheep.html و http://www.animalgenome.Org/QTLdb/cattle.html مورد استفاده قرار گرفت.

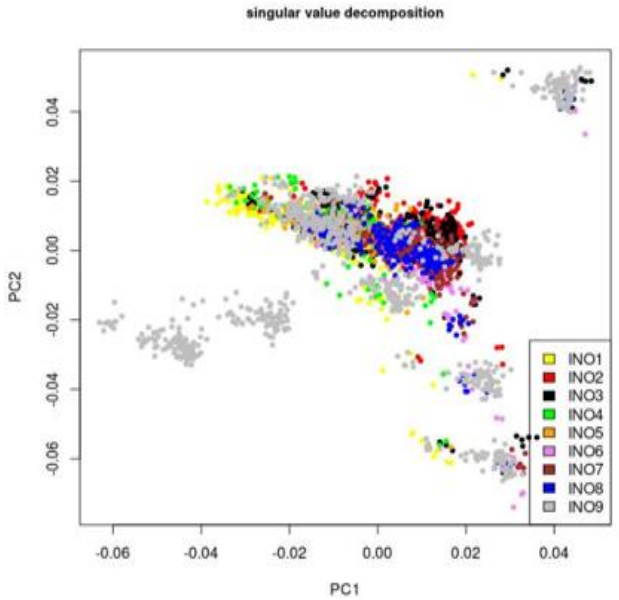
نتایج و بحث

در داده‌های اولیه ۳۰۰۰ فرد و حدود ٤٨٥٩٩ جایگاه نشانگری به ازای هر فرد وجود داشت اما پس از اجرای مراحل مختلف کنترل کیفیت داده‌های اولیه، مجموع ٤٤٢٨٩ نشانگر SNP و ٢٨٤٣ حیوان از کل جمعیت‌ها، جهت تجزیه‌های بعدی انتخاب شد. به طور کلی ٦٠٠ نشانگر SNP به‌دلیل فراوانی آللی کمتر از ۰/۰۱، ٣٤٤٢ نشانگر SNP به‌دلیل نرخ ژنوتایپینگ کمتر از ۹۵ درصد، ۱۵۷ فرد بدلیل نرخ داده گمشده به ازای هر فرد کمتر از ۹۵ درصد و ٢٦٨ نشانگر به‌دلیل انحراف از تعادل هاردی-واینبرگ ($P> ۱۰^{-٦}$) حذف شدند.

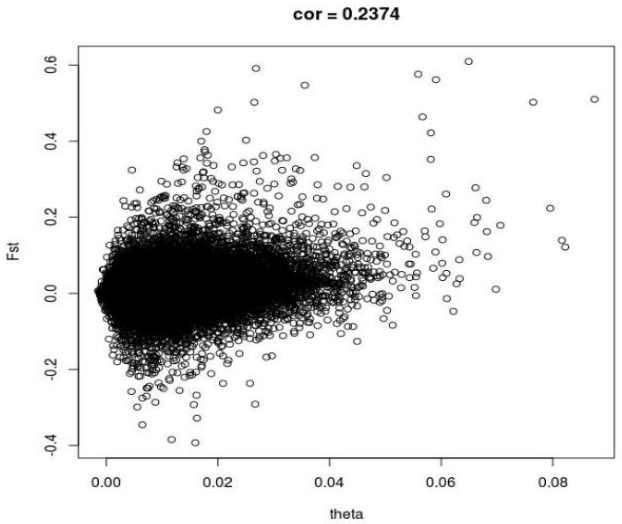
جهت بررسی ساختار جمعیت‌های مورد مطالعه، تجزیه PCA بر اساس همه اطلاعات SNPهای در دسترس مورد ارزیابی قرار گرفت. بدون در نظر گرفتن مولفه‌ی قبلی، با گذر از مولفه‌ی ابتدایی به سمت مولفه‌های انتهایی، هر مولفه واریانس کمتری را تشریح می‌کند. یعنی همیشه مولفه‌ی اصلی اول بیشترین واریانس و مولفه‌های آخر کمترین واریانس را شرح می‌دهند که در این صورت با حذف مولفه‌های آخر اطلاعات زیادی از دست نخواهد رفت (Jolliffe 2002). در مطالعه حاضر، کل نمونه‌ها به ٦ مولفه اصلی تفکیک شدند که به ترتیب مقادیر مولفه‌ها از بالا به پایین سیر نزولی داشت. در نتیجه سهم آنها در توضیح واریانس نیز کاهش یافت و بدین ترتیب مولفه اول بالاترین سهم از کل واریانس (۰/۱۳) و مولفه آخر نیز کمترین مقدار از کل واریانس (۰/۰۰۰۵) را توجیه نمودند. بخشی از واریانس پایین می‌تواند بدلیل تنوع بالای نمونه‌گیری باشد. در مطالعه‌ای روی گوسفندان سراسر دنیا، مولفه اول مقدار ۲/۹۸ درصد از کل واریانس و مولفه دوم و سوم به ترتیب حدود ۱/٤٤ و ۱/۱۹ درصد از واریانس را توجیه کردند (Kijas et al. 2012).

با وجود اینکه نتایج PCA نشان داد که تعدادی از افراد برخی جمعیت‌ها، خارج از دسته خود قرار دارند اما بیشتر افراد مورد مطالعه می‌توانند در ٦ دسته جداگانه خوشه‌بندی شوند (شکل ١). همانطور که در شکل ١ مشخص می‌شود، تقریبا تمام افراد گله-های CRC در یک دسته قرار دارند و با یکدیگر همپوشانی بالایی نشان می‌دهند که می‌تواند ناشی از یکسان بودن خط پدری در گله های مورد مطالعه باشد.

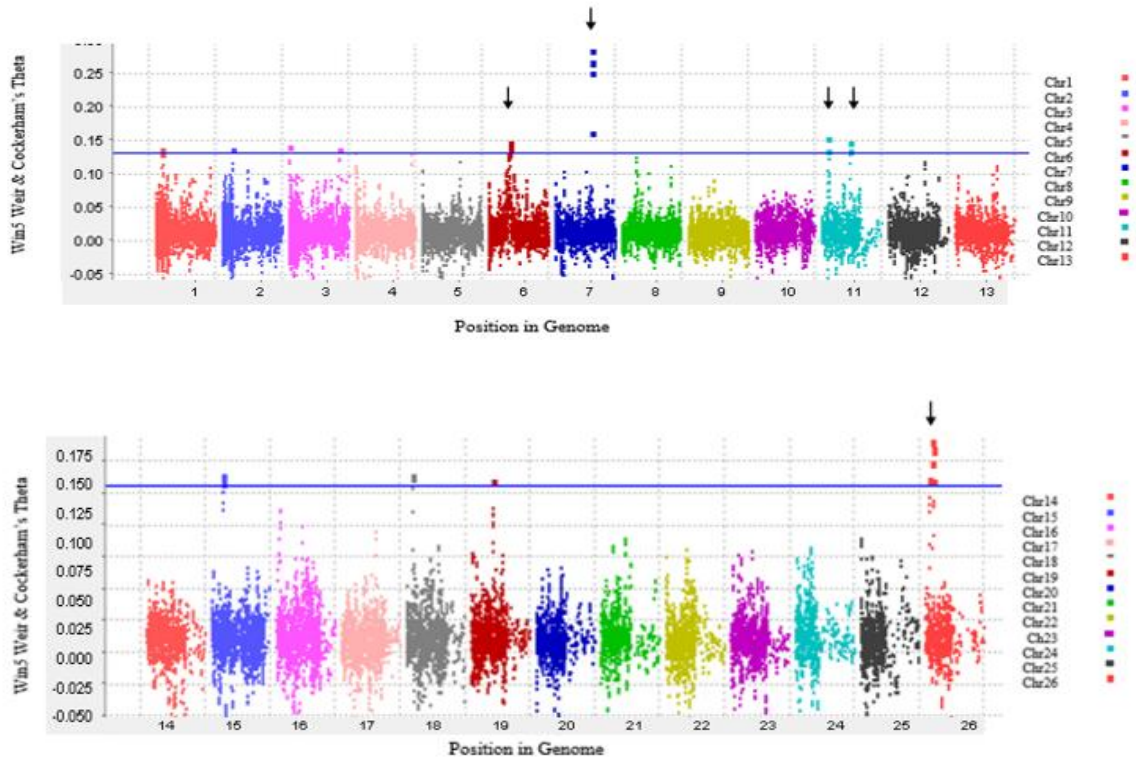
جهت مشخص نمودن جایگاه‌های تحت انتخاب، بر روش‌های مبتنی بر تنوع (بررسی سطح تمایز زیرجمعیتی) تمرکز شد. شکل ٢ همبستگی بین ارزش‌های شاخص تثبیت (Fst) و روش نااریب شاخص تثبیت (θ) را نشان می‌دهد. نتایج حاصله بیانگر همبستگی پایین (٠/٢٣) این دو آماره بود و انحراف معنی‌داری از یک نشان داد ($P< 0.05$). ارزیابی آماره Fst با روش Weir (1984) and Cockerham خیلی مفید می‌باشد، زیرا یک روش نااریب برای بررسی سطح تمایز ژنتیکی جمعیت‌های با تعداد نابرابر است. میانگین هر دو آماره شاخص تثبیت و تتا در بین نه جمعیت مورد مطالعه ٠/٠١ برآورد شد. پایین بودن میانگین این آماره‌ها گویای این مطلب است که این جمعیت‌ها از تنوع کافی برخوردار می‌باشند و عمل انتخاب می‌تواند به منظور کاهش هم‌خونی و در نتیجه کاهش بیماری‌های ژنتیکی نامطلوب صورت گرفته باشد. در مطالعات قبلی روی داده‌های ژنومی گوسفند و خوک، مقدار این آماره‌ها به ترتیب ٠/٠٢٤ و ٠/٣٧ گزارش شده است (Kijas et al. 2009; Yang et al. 2014). دلایل این تفاوت می‌تواند اریب-های ناشی از نمونه‌گیری وهمچنین تفاوت‌های ساختار ژنومی گوسفند و خوک باشد. پس از محاسبه ضرایب تتا در هر جایگاه نشانگری، با توجه به اینکه انتخاب علاوه بر جایگاه اصلی، سودمندی نشانگرهای مجاور آنرا هم تحت تاثیر قرار می‌دهد و از طرف دیگر نمایش بهتر سیگنال‌های انتخاب در سطح ژنوم، به جای در نظر گرفتن ارزش هر SNP، میانگین ارزش تتا مربوط به پنج SNP مجاور محاسبه و تحت عنوان ارزش Win5 تتا هر نشانگر در نظر گرفته شد و توزیع آن در سطح ژنوم مورد بررسی قرار گرفت (شکل ٣). مجموعا پنج ناحیه ژنومی با ارزش Fst بیشتر از حد آستانه (٠/١٤)، روی کروموزوم ٦ (در موقعیت ٤٠٩٥٥٩٢-٤١٤٧٦٤٩٧kb)، کروموزوم ٧ (در موقعیت ٦٤٨٧٥١٩١-٦٥٢١١٣٨١kb)، کروموزوم ١١ (در موقعیت ١٨٥٣٤٥٥٩-١٨٧٠١٤٢٨kb و ٥٢٢٦٣٥٤٢-٥٢٤٠٦٥٣٩) و کروموزوم ٢٦ (در موقعیت ٢١٠٢٥١١٨-٢١٤٧٢٧٠٩ kb) به عنوان نواحی تحت انتخاب شناسایی شدند (جدول ١). همانطور که در شکل ٣ مشخص می‌شود، بالاترین سیگنال انتخاب روی کروموزوم ٧ در موقعیت ٦٥/٠٦ Mb (OAR7_65066014; WinTheta= .283) می‌باشد.



شکل ١- خوشه‌بندی حیوانات بر اساس تجزیه و تحلیل مولفه‌های اصلی با استفاده از ژنوتیپ افراد (INO1-INO8 گله‌های CRC، INO9 گله SG)



شکل ٢- همبستگی بین ارزش‌های Fst و Fst نااریب ویر و کوکهام (تتا)



شکل ۳- توزیع ارزش‌های Win5 تتا در سطح ژنوم جمعیت‌های مورد مطالعه

جدول ۱- ژنهای گوسفند و QTLهای گزارش شده در مناطق ژنومی نشان دهنده تفرق بین جمعیت‌های مورد مطالعه

موقعیت روی ژنوم گوسفند (کروموزوم)	موقعیت روی ژنوم گاو (کروموزوم)	ژن	QTL در گوسفند	منبع QTL
(٦) ٤٠٩٥٥٩٢-٤١٤٧٦٤٩٧	(٦) ٤٠٠٩٧٣٤١-٤٤٢٧٨٥٤٠	-	وزن لاشه گرم افزایش وزن روزانه (تولد تا ٤٣ هفتگی) وزن بدن (کشتار)	Cavanagh et al. 2010 Cavanagh et al. 2010 Al-Mamun et al. 2015
(٧) ٦٤٨٧٥١٩١-٦٥٢١١٣٨١	(١٠) ٦٨٢٠٦٥٧٦-٧١٠٥٠٨٧٥	KTN1	طول استافل ضریب واریانس قطر فیبر	Ponz et al. 2001 Ponz et al. 2001
(١١) ١٨٥٣٤٥٥٩-١٨٧٠١٤٢٨	(١٩) ١٨٧٧٥٤١١-١٨٧٨١٦٥٥	OR4D1 MKS1 U6atac	وزن بدن (کشتار) مقدار اسید چرب اشباع مقدار پروتئین شیر	Al-Mamun et al. 2015 Karamichou et al. 2006 Ian and Ian 2005
(١١) ٥٢٢٦٣٥٤٢-٥٢٤٠٦٥٣٩	(١٩) ٥٤٤١٧٤٣٨-٥٤٤٥٧٠٤٤	RBFOX3 ENGASE C1QTNF1 CANT1	وزن بدن (کشتار) وزن بدن (مادری)	Al-Mamun et al. 2015 Al-Mamun et al. 2015
(٢٦) ٢١٠٢٥١١٨-٢١٤٧٢٧٠٩	(٧) ٣٤٣٣٢٠٠٩-٣٤٣٣٤٢٧٦	5S_rRNA	افزایش وزن روزانه تغییرات هماتوکریت مقدار انگل	Cavanagh et al. 2010 Marshall et al. 2013 Marshall et al. 2013

در مطالعه حاضر، حدود ۲۰ درصد جایگاه‌ها (۹۱۲۵ جایگاه) دارای ارزش تتا برابر با صفر یا زیر صفر بودند. ارزش‌های حاصل از Fst بین ۰ (بدون تفاوت ژنتیکی بین زیرجمعیت‌ها یا جمعیت اجدادی) تا ۱ (تفاوت کامل زیرجمعیت‌ها از یکدیگر یا از کل جمعیت) متغیر می‌باشد (Andres et al. 2014) و با توجه به اینکه ارزش Fst ویر و کوکهام، یک روش برآوردگر نااریب است، احتمال بدست آوردن ارزش‌های منفی نیز وجود دارد (Akey et al. 2009).

در مجموع پنج ناحیه ژنومی روی چهار کروموزوم متفاوت به عنوان نواحی تحت انتخاب شناسایی شدند. همان‌طور که در جدول ۱ مشاهده می‌شود، به جز موقعیت شناسایی شده روی کروموزوم ٦، بقیه نواحی حداقل با یک ژن مهم همپوشانی داشتند. به‌عنوان مثال ژن *MKS1* که در ساخت تاژک تاثیر دارد (Kyttala et al. 2006)، با موقعیت ۱۸۷۰۱٤۲۸–۱۸۵۳٤۵۵۹ روی کروموزم ۱۱ ارتباط دارد. عدم بیان این ژن منجر به بی تحرکی اسپرم و در نتیجه عقیم شدن فرد می‌شود و هم‌چنین در مجاری تنفسی مژک و تاژک برای جلوگیری از ورود گرد و غبار و ناهنجاری‌های تنفسی دام نقش بسزایی دارد. از طرفی ژن *KTN1* که با کروموزوم ۷ هم‌پوشانی دارد، در اتصالات سلولی و اندامک-های انتقال دهنده انرژی مانند دستگاه گلژی و میکروتوبول‌ها نقش دارد (Rao et al. 1998). پس می‌توان گفت این ژن نقش مهمی در اکثر فعالیت‌های بدن از قبیل تحرک، رشد، تولید گوشت و پشم دارد. با وجود این‌که ژن‌های دیگری در مجاورت مناطق تحت انتخاب شناخته شدند اما نقش بیولوژیکی یک‌سری از این ژن‌ها در گوسفند به‌طور کامل شناخته شده نیست. به هر حال جهت شناسایی نقش دقیق این ژن‌ها بایستی مطالعات پیوستگی و عملکردی بیشتری انجام گیرد.

بررسی QTLهای احتمالی شناسایی شده در مناطق ژنومی مورد نظر در گوسفند نشان داد که تقریبا تمامی جایگاه‌های مدنظر با QTLهای گزارش شده برای صفات مهم اقتصادی از جمله صفات مرتبط با تولید شیر (مقدار پروتئین شیر)، رشد (وزن بدن و افزایش وزن روزانه) و لاشه (وزن کشتار، وزن لاشه گرم) همپوشانی دارند (جدول ۱). با توجه به مطالعات قبلی، تنها دو QTL مرتبط با تولید پشم که با نواحی ژنومی مدنظر همپوشانی داشته باشد، گزارش شده‌است (کروموزوم ۷ در موقعیت Mb ٦٥/٠٦). انتظار می‌رفت با توجه به اهمیت این صفت در نژاد مرینوس، نشانه‌هایی از انتخاب که مرتبط با صفت تولید پشم باشند در این جمعیت‌ها مشاهده شود. لازم به ذکر است که بالاترین سیگنال انتخاب در موقعیت مذکور می‌باشد (شکل ۱)، که خود نشانی از تحت تاثیر انتخاب طبیعی یا مصنوعی قرار گرفتن نواحی ژنومی مرتبط با صفت تولید پشم از جمله ضریب واریانس قطر فیبر (CVFD) در این نژاد است. همبستگی ژنتیکی منفی بین CVFD و وزن یک سالگی (۰/۰٤ ± ۰/۲۲-) در گوسفندان مرینوس استرالیا گزارش شده‌است (Asadi Fozi et al. 2005).

یکی از ژن‌های مهم مرتبط با باروری در گوسفند، ژن *FecB* است که روی کروموزوم ٦ می‌باشد، اما همان‌طور که در جدول ۱ مشاهده می‌شود، در مطالعه حاضر موقعیت شناسایی شده روی این کروموزوم، با هیچ ژنی همپوشانی نداشت. یکی از مهم‌ترین دلایل این است که در کشور استرالیا روی این جمعیت‌ها عمدتا برای صفاتی از قبیل پشم و گوشت انتخاب صورت می‌گیرد. اما نتایج بررسی QTL حاکی از آن است که در این جایگاه در مناطق اورتولوگوس گاو دو QTL مرتبط با باروری (طول بارداری و عرض کفل) و در گوسفند چندین QTL مهم مرتبط با رشد و لاشه گزارش شده‌است (Boichard et al. 2003; Maltecca et al. 2008). اگر چه برخی دامداران نیز بدلیل مرگ و میر کمتر، تک قلوزایی را ترجیح می‌دهند اما چندقلوزایی و افزایش وزن از جمله صفات مهم اقتصادی در گوسفند می‌باشند که تحت تاثیر محیط و ژنتیک بوده و افزایش آن‌ها با سوددهی اقتصادی همراه است. به‌عنوان مثال افزایش نرخ تخمک‌ریزی منجر به افزایش درصد دو قلوزایی می‌شود و در نهایت باعث افزایش سودآوری در سامانه‌های پرورش گوسفند خواهد شد (Guan et al. 2007). در گذشته باور عمومی بر این بود که تولید مثل از جمله صفات کمی است که از نظر توارث، پلی‌ژنیک می‌باشد، یعنی فعالیت‌های تولید مثلی در گوسفند تحت تاثیر تعداد زیاد ژن قرار دارند. اما در سال‌های اخیر در کشورهای استرالیا و نیوزلند به خوبی مشخص شده‌است که تنوع در پدیده چندقلوزایی می‌تواند به‌واسطه تفرق ژن‌های کاندیدا مرتبط با تولید مثل و تخمک‌ریزی باشد (Hechert

(et al. 1992; Davis 2004; Davis 2005). این گزارش‌ها، گویای این مطلب است که انتخاب شدید طبیعی و مصنوعی برای صفات تولید پشم، گوشت و صفات تولید مثلی که از دیرباز برای پرورش دهندگان این نژادها اهمیت داشته، صورت گرفته است. لذا می‌توان وجود چنین نشانه‌هائی از انتخاب را ناشی از هر دو انتخاب مصنوعی و طبیعی دانست.

نتیجه گیری نهایی

در این مطالعه جهت شناسایی نشانه‌های انتخاب از آماره تتا استفاده شد. بررسی ژنها و QTLهای موجود در نواحی تحت انتخاب نشان داد که این ژنها با صفات تولید مثل، رشد، لاشه و پشم در ارتباط می‌باشند اما بررسی‌های پیوستگی و عملکردی بیشتری جهت شناسایی عملکرد ژنها و QTLهای مرتبط نیاز می‌باشد. در مجموع، نتایج این تحقیق می‌تواند منابع اطلاعاتی با ارزشی در جهت شناسایی ژن‌های متمایز کننده و شناسایی نواحی ژنومی کاندید برای بسیاری از صفات مهم اقتصادی موجود در گله‌های مورد مطالعه فراهم آورد. هم‌چنین از اطلاعات این تحقیق می‌توان در تحقیقات مرتبط با انتخاب ژنومیک، طراحی سیستم-های آمیزشی و مطالعات تکاملی استفاده نمود.

تشکر و قدردانی

نویسندگان مقاله از همکاری صمیمانه پروفسور جولیوس ون در ورف و دکتر محمد حسین فردوسی به دلیل در اختیار گذاشتن اطلاعات داده‌ها و کمک‌های ارزنده‌اشان، کمال تشکر را دارند.

منابع

Al-Mamun HA, Kwan P, Samuel A, Ferdosi MH, Tellamand R, Gondro C (2015) Genome-wide association study of body weight in Australian Merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. Genetic Selection Evolution 10: 47-66.

Akey JM (2009) Constructing genomic maps of positive selection in humans: Where do we go from here? Genome Research 19: 711-722.

Andres M, Turiégano E, Göpfert MC, Canal I, Torroja L (2014) The extracellular matrix protein artichoke is required for integrity of ciliated mechanosensory and chemosensory organs in Drosophila embryos. Genetics 196: 1091-1102.

Asadi Fozi M, Van der Werf JHJ, Swan AA (2005) The importance of accounting for maternal genetic effects in Australian fine-wool Merino breeding. Australian Journal of Agricultural Research 56: 789-796.

Boichard D, Grohs CA, Bourgeois F, Cerqueira F, Faugeras RA, Neau A, Rupp RA (2003) Detection of genes influencing economic traits in three French dairy cattle breeds. Genetics Selection Evolution 1: 77-101.

Biswas S, Akey JM (2006) Genomic insights into positive selection. Trends in Genomics 22: 437-446.

Chevin LM, Billiard S, Hospital F (2008) Hitchhiking both ways: Effect of two interfering selective sweeps on linked neutral variation. Genetics Society of America 108: 301-316.

Cavanagh CR, Jonas E, Hobbs M, Thomson PC, Tammen I, Raadsma HW (2010) Mapping quantitative trait loci (QTL) in sheep. III. QTL for carcass composition traits derived from CT scans and aligned with a meta-assembly for sheep and cattle carcass QTL. Genetic Selection Evolution 42: 1-14.

Davis GH (2004). Fecundity genes in sheep. Animal Reproduction Science 83: 247-253.

Davis GH (2005) Major genes affecting ovulation rate in sheep. Genetic Selection Evolution 37: 11-23.

Dalrymple BP, Kirkness EF, Nefedov M, McWilliam S, Ratnakumar A, Barris W, Zhao S, Shetty J, Maddox JF, O'Grady M (2007) Using comparative genomics to reorder the human genome sequence into a virtual sheep genome. Genome Biology 8: 152.

Fariello MI, Servin B, Tosser-Klopp G (2014). Selection signatures in worldwide sheep populations. PLOS 42: 1-30.

Guan F, Liu SR, Shi GQ, Yang LG (2007) Polymorphism of FecB gene in nine sheep breeds or strains and its effects on litter size, lamb growth and development. Animal Reproduction Science 99: 44-52.

Hayes BJ, Chamberlain AJ, Maceachern S, Savin K, McPartlan H, MacLeod I, Sethuraman L, Goddard ME (2009) A genome map of divergent artificial selection between Bos taurus dairy cattle and Bos taurus beef cattle. Animal Genetics 40: 176-184.

Hechert LL, Daley IJ, Griswold MD (1992) Structural organization of the follicle-stimulating hormone receptor gene. Endocrinology 156: 189-193.

Hubbard TJP, Aken BL, Ayling S, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Clarke L, Coates G, Fairley S, Fitzgerald S, Fernandez-Banet J, Gordon L, Graf S, Haider S, Hammond M, Holland R, Howe K, Jenkinson A, Johnson N, Kahari A, Keefe D, Keenan S, Kinsella R, Kokocinski F, Kulesha E, Lawson D, Longden I, Megy K, Meidl P, Overduin B, Parker A, Pritchard B, Rios D (2009) Ensembl .Nucleic Acids Research 37: 690-697.

Jolliffe I (2002) Principal component analysis. Second Edition, New York, USA, 20-28.

Ian WP, Ian RF (2005) Major genes and QTL influencing

wool production and quality: a review. Genetic Selection Evolution 37: 97-107.
Qanbari S, Pimentel E, Tetens J, Thaller G, Lichtner P (2010) A genome-wide scan for signatures of recent selection in Holstein cattle. Animal genetics 41: 377-389.
Qanbari S, Gianola D, Hayes B, Schenkel F, Miller S, Moore S, Thaller G, Simianer H (2011) Application of site and haplotype frequency based approaches for detecting selection signatures in cattle. BMC Genomics 12: 318-330.
Karamichou ERI, Richardson GR, Nute KP, Bishop SC (2006) Genetic analyses and quantitative trait loci detection, using a partial genome scan, for intramuscular fatty acid composition in Scottish Blackface sheep. Journal of Animal Science 84:3228-3238.
Kijas J, Lenstra JA, Hayes B, Boitard S, Neto LP, Cristobal MS, Servin B, McCulloch R, Whan V, Gietzen K (2009) Genome wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLOS Biology 10: 1-14.
Kijas, JW, Johannes A, Ben Hayes L, Boitard S, Porto-Neto LR, Cristobal MS, Servin B, McCulloch R, Whan V, Gietzen K, Paiva S, Barendse W, Ciani E, Raadsma H, McEwan J, Dalrymple B (2012) Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS Biology 10: 1-14.
Kyttala M, Tallila J, Salonen R, Kopra O, Kohlschmidt N, Paavola-Sakki P, Peltonen L, Kestilä M (2006) MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome. Nature Genetics 38: 155-157.
Maltecca C, Weigel KA, Khatib H, Cowan M, Bagnato A (2008) Whole-genome scan for quantitative trait loci associated with birth weight, gestation length and passive immune transfer in a Holstein x Jersey crossbred population. Animal genetics 40: 27-34.
Marshall K, Mugambi JM, Nagda S, Sonstegard TS, Van Tassell CP, Baker RL, Gibson JP (2013) Quantitative trait loci for resistance to Haemonchus contortus artificial challenge in Red Maasai and Dorper sheep of East Africa. Animal Genetics 44: 385-395.
Moradi MH, Nejati-Javaremi A, Moradi-Shahrbabak M, Dodds KG, McEwan JC (2012) Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. BMC Genetics13: 1-15.
Nielsen R (2005) Molecular signature of natural selection. Annual Review of Genetics 39: 197-218.
Pan D, Zhang S, Jiang J, Jiang L, Zhang Q, Liu JF (2013) Genome-Wide detection of selective signature in Chinese Holstein. Public Library of Science 8: 2-9.
Porto-Neto LR, Lee SH, Sonstegard TS, Van Tassell CP, Lee HK, Gibson JP, Gondro C (2014) Genome-wide detection of signatures of selection in Korean Hanwoo cattle. Animal Genetics 45: 180-190.
Pollinger JP, Bustamante CD, Fledel AA, Schmutz S, Gray MM, Wayne RK (2007) Selective sweep mapping of genes with large phenotypic effects. Genome Research 15: 1809-1819.
Ponz R, Moreno C, Allain D, Elsen JM, Lantier F, Lantier I, Brunel JC, Perez-Enciso M (2001) Assessment of genetic variation ex-pained by markers for wool traits in sheep via a segment mapping approach. Mammalian Genome 12: 569-572.
Prasad A, Schnabel RD, McKay SD, Murdoch B, Stothard P, Kolbehdari D, Wang Z, Taylor JF, Moore SS (2008) Linkage disequilibrium and selection signatures on chromosomes 19 and 29 in beef and dairy cattle. Animal Genetics 39: 597-605.
Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics 81: 559-575.
Rao PN, Yu H, Hodge R, Pettenati MJ, Sheetz MP. 1998. Assignment of the human kinectin gene (KTN1), encoding a kinesin-binding protein, to chromosome 14 band q22.1 by in situ hybridization. Cytogenetic Cell Genetics 79: 196-197.
R Development Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: the R Foundation for Statistical Computing. https://www.R-project.org.
Sabeti PC, Scheffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES (2006) Positive natural selection in the human lineage. Science 312: 1614-1620.
Stella A, Ajmone-Marsan P, Lazzari B, Boettcher P (2010) Identification of selection signatures in cattle breeds selected for dairy production. Genetics 185: 1451 - 1461.
Truxillo C, Hamer R (2007) Multivariate statistical methods: Practical research applications, course notes, SAS institute.
Weir BS, Cockerham CC (1984) Estimating F-statistic for the analysis of population stricter. Evolution 38: 1358-1370.
Wright S (1992) Coefficient of inbreeding and relationship. American Naturalist 56:330-338.
Yang S, Li X, Li K, Fan B, Tang, Z (2014) A genome-wide scan for signatures of selection in Chinese indigenous and commercial pig. BMC Genetic 15: 1-9.

Modern Genetics Journal

Vol 11, Number 3, Autumn 2016, Quaternary

A Genomic Scan for Detection of Selection Signatures using SNP Data in Australian Merino Sheep

Montazeri M[1], Asadi Fozi M[*1], Esmailizadeh Koshkoieh A[2], Ferdosi MH[2], van der Werf J[3]

1. PhD Student, Associate Professor, Professor, Department of Animal Science, Faculty of Agriculture, Shahdi Bahonar University of Kerman, Kerman, Iran.
2. Research Fellow, Animal Genetic and Breeding Unit, University of New England, Armidale, Australia.
3. Professor, School of Environmental and Rural Science, University of New England, Armidale, Australia.
* Corresponding Author, Email: masadi@uk.ac.ir

A B S T R A C T

Deciphering the genetic basis of phenotypic diversity is one of the main aims of biological research. Domestic animals are useful tool for making substantial progress towards this goal**.** Selection signatures are the regions of the genome that are functionally important and therefore have been under either natural or artificial selection. In this paper, a whole genome scan was performed using 3000 individuals, ~50000 SNP markers from nine populations of Australian Merino sheep, with the aim of identifying divergently selected regions of the genome. Five genomic regions on 4 chromosomes were identified as putatively harboring selective sweeps. These regions were located on chromosomes 6, 7, 11 (two areas) and 26. These selected genomic regions were surveyed to find encoding putative candidate genes and 9 genes were extracted from areas Ovine Genome v3.1 Assembly.Study of the reported QTL in these regions of the sheep genome showed that they overlapped with QTL of economically important traits such as carcass yield, growth and wool traits. Further investigation of these regions in validation studies will help to identify the candidate genes for economically important traits in sheep breeds.

Key Words

Genome Scan, Population Differentiation, Signatures of Selection, Single Nucleotide Polymorphism