



## تخمین عدم تعادل پیوستگی و پویش کل ژنومی جهت شناسایی جایگاه‌های ژنی تحت انتخاب مرتبط با وزن بدن در گوسفندان نژاد زندی

حسین محمدی<sup>۱</sup>، سید عباس رأفت<sup>۲</sup>، حسین مرادی شهربابک<sup>۳\*</sup>، جلیل شجاع غیاث<sup>۴</sup>، محمد حسین مرادی<sup>۵</sup>

<sup>۱</sup> دانش‌آموخته دکتری گروه علوم دامی، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران.

<sup>۲</sup> استاد گروه علوم دامی، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران.

<sup>۳</sup> استادیار گروه علوم دامی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، تهران، ایران.

<sup>۴</sup> استاد گروه علوم دامی، دانشکده کشاورزی، دانشگاه تبریز، تبریز، ایران.

<sup>۵</sup> استادیار گروه علوم دامی، دانشکده کشاورزی، دانشگاه اراک، اراک، ایران.

تاریخ دریافت: ۱۳۹۶/۰۴/۲۵، تاریخ پذیرش: ۱۳۹۶/۰۵/۳۱

### چکیده

وزن بدن از مهمترین معیارهای تعیین کننده سود اقتصادی در صنعت پرورش گوسفند است. در مطالعات پیوستگی و انتخاب ژنومی تعیین میزان و گستره عدم تعادل پیوستگی (LD) در تشخیص تعداد نشانگر مورد نیاز و اندازه نمونه، اهمیت بسزایی دارد. علاوه بر این، انتخاب برای افزایش فراوانی موتاسیون‌های جدیدی که فقط در برخی از زیرجمعیت‌ها سودمند هستند باعث باقی گذاشتن نشانه‌هایی در سطح ژنوم می‌شود. اغلب این مناطق با ژن‌ها و QTL‌های مرتبط با صفات مهم اقتصادی در ارتباط هستند. بنابراین، هدف این تحقیق مطالعه گستره LD و شناسایی مناطقی از ژنوم گوسفندان نژاد زندی با وزن بدن متفاوت بود که در طی سال‌های مختلف به صورت مصنوعی یا طبیعی انتخاب شده‌اند. بدین منظور، از ۷۳ رأس گوسفند زندی نمونه خون تهیه شد و با استفاده از آرایه‌های SNPChip 50K شرکت ایلومینا ژنوتیپ ۵۴۲۴۱ نشانگر بر مبنای آخرین نسخه Oar\_4.0 ژنوم گوسفند تعیین شد. پس از مراحل کنترل کیفی، در نهایت ۴۰۸۷۹ نشانگر SNP مربوط به ۷۱ حیوان آنالیز شدند. مقدار LD با محاسبه آماره  $r^2$  بین تمام جفت جایگاه‌ها از طریق نرم افزار PLINK محاسبه شد. برای شناسایی نشانه‌های انتخاب بر پایه روش‌های عدم تعادل پیوستگی از آزمون هموزیگوسیتی هاپلوتایپی بسط داده شده جمعیت تلاقی (XP-EHH) استفاده شد. در این مطالعه گستره مفید عدم تعادل پیوستگی در 40K برابر با  $r^2 = 0.2$  برآورد شد. نتایج این تحقیق ۱۰ منطقه ژنومی روی کروموزوم‌های ۱، ۲، ۳، ۵، ۶، ۹، ۱۲، ۱۳، ۲۱ و ۲۲ را شناسایی کرد. تجزیه و تحلیل بیوانفورماتیکی نشان داد که برخی از این مناطق ژنومی با ژن‌های مؤثر بر صفات رشد، متابولیسم چربی، توسعه سیستم اسکلتی، متابولیسم انرژی و کیفیت گوشت مانند ژن‌های PKD2، FAM184B، FBP1، CAPN2، MEPE، SPP1، FGF19 همپوشانی دارند.

واژه‌های کلیدی: عدم تعادل پیوستگی، نشانه‌های انتخاب، آزمون XP-EHH، گوسفند زندی، وزن بدن.

گوسفندان استرالیایی نژاد بوردر لایسستر، مرینو و پول دورست در فاصله‌های کمتر از ۱۰Kb به ترتیب ۰/۳۴، ۰/۲۷ و ۰/۳۳ گزارش شده است (Al-Mamun *et al.*, 2015).

گوسفند به عنوان یکی از اولین دام‌های اهلی در حدود ۱۱ هزار سال پیش در ناحیه عراق و ایران امروزی برای تولید گوشت، پشم و شیر اهلی شد (Zhao *et al.*, 2016). بالغ بر ۲۶ نژاد گوسفند در ایران وجود دارد که با مناطق مختلف سازگار شده‌اند (Mohammadabadi & Sattai Mokhtari, 2013; Zamani *et al.*, 2015; Khodabakhshzadeh *et al.*, 2016b). در حال حاضر، تولید گوشت مهمترین دلیل پرورش گوسفند در ایران است و تولیدات دیگر مانند پشم، شیر و پوست در درجات بعدی اهمیت قرار دارند (Mohammadabadi & Sattai Mokhtari, 2013; Zamani *et al.*, 2015; Khodabakhshzadeh *et al.*, 2016b). از زمان اهلی شدن تا کنون نژادهای مختلف گوسفند در سراسر جهان هدف انتخاب‌های طبیعی و مصنوعی قرار گرفته‌اند که این انتخاب‌ها با ایجاد نشانه‌هایی در سطح ژنوم همراه است. LD یکی از عوامل مهم در بررسی جهش‌های صفات مهم اقتصادی در حیوانات مزرعه‌ای است عواملی مانند جهش، مهاجرت، رانش و انتخاب باعث بوجود آمدن عدم تعادل پیوستگی ژنتیکی می‌شود (Chevin *et al.*, 2008; Mousavizadeh *et al.*, )

عدم تعادل پیوستگی (LD) یا عدم تعادل فاز گامتی به ارتباط غیر تصادفی آل‌های دو جایگاه بر روی یک کروموزوم در یک گامت گفته می‌شود (Wall & Pritchard, 2003). معیار تعیین کننده موفقیت استفاده از مطالعات پویش ژنومی (GWAS) و انتخاب ژنومیک (GS) به مقدار LD بین نشانگرها و جایگاه‌های ژنومی مسبب صفات کمی در طول کل ژنوم بستگی دارد. اساس مطالعات ژنومی بر مبنای مقدار LD در جمعیت مورد مطالعه و پایداری فاز عدم تعادل پیوستگی بین نشانگر و جایگاه ژنومی صفات کمی (QTL) در خلال نسل‌ها است (Zhao *et al.*, 2014). مقدار LD ممکن است در بین جمعیت‌های مختلف بسیار متفاوت باشد. علت این اختلاف می‌تواند در اشتقاق جمعیت‌ها در سال‌های خیلی قبل، تغییر در اندازه مؤثر جمعیت و یا فاصله زیاد بین نشانگر و QTL باشد که منجر به شکستن LD جمعیت اولیه و تغییر آن در زیر جمعیت‌های حاصله می‌شود (De Ross *et al.*, 2008). بنابراین، محققان زیادی الگوی عدم تعادل پیوستگی را در حیوانات اهلی مختلف بررسی کرده‌اند (Niu *et al.*, 2016). به عنوان نمونه، عدم تعادل پیوستگی در سطح ژنوم گوسفند اسپانیایی دارای میانگین  $r^2$  برابر ۰/۳۲۹ در فاصله‌های کمتر از ۱۰Kb و در فاصله‌های ۲۰۰-۵۰۰ برابر با ۰/۰۶۱ بود (Garcia-Gamez *et al.*, 2012) در حالی که میزان LD در

همگام با افزایش پیشرفت‌های ژنتیکی، حفظ تنوع ژنتیکی یک جمعیت برای اطمینان از پاسخ به انتخاب بلند مدت، یک وظیفه مهم برنامه‌های اصلاح نژادی است ( *Khodabakhshzadeh et al.*, 2015a). برنامه‌های اصلاح نژادی بیشتر در جهت بهبود وضعیت تولیدی حیوانات اهلی عمل می‌کنند که توفیق این برنامه‌ها به مقدار تنوع ژنتیکی موجود در گله‌ها بستگی دارد ( *Khodabakhshzadeh et al.*, 2015a).

پس از اینکه در سال ۲۰۰۹ اولین تراشه SNP در گوسفند در قالب پروژه Sheep HapMap طراحی شد، تحقیقات موفقیت آمیز مختلفی با استفاده از روش XP-EHH در گوسفند و برخی از حیوانات اهلی دیگر صورت گرفته است ( *Qanbari et al.*, 2011; *Kim et al.*, 2015; *Utsunomiya et al.*, 2015; *Zhao et al.*, 2016; *Chen et al.*, 2016; *Iso-Touru et al.*, 2016). در پروژه گسترده‌ای در سال ۲۰۱۲ گروهی از دانشمندان جهت شناسایی نواحی تحت انتخاب مثبت در جمعیت گوسفندان سراسر جهان شامل ۷۴ نژاد مختلف، در قالب پروژه Sheep HapMap مورد تجزیه و تحلیل قرار دادند ( *Kijas et al.*, 2012). نتایج پوشش ژنومی سیگنال‌های انتخاب، ۳۱ ناحیه مرتبط با رنگ پوست، مورفولوژی استخوان، اندازه بدن، رشد و تولیدمثل را آشکار نمود. همچنین شناسایی مناطق ژنومی در نژادهای مختلف چینی با حدود ۵۴۰۰۰ نشانگر SNP و بررسی تمایز

(2009). 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. بنابراین زمانی که یک آلل سودمند در طی زمان‌های مختلف، هدف انتخاب مثبت قرار می‌گیرد باعث ایجاد نشانه‌هایی در سطح ژنوم می‌گردد که از طریق بررسی طیف فراوانی آللی و LD قابل شناسایی است ( *Sabeti et al.*, 2006). از طرفی استفاده از نشانگرهای ژنتیکی در انتخاب و اصلاح نژاد به طور مهیجی پیشرفت ژنتیکی را تسریع می‌کند ( *Mohammadi et al.*, 2008; *Javanmard et al.*, 2011) و مطالعه ساختار ژنتیکی نژادهای بومی برای حفاظت از منابع ژنتیکی لازم و ضروری است ( *Khodabakhshzadeh et al.*, 2016a). به طور کلی در بیشتر کشورهای جهان هدف از اصلاح نژاد گوسفند سودآوری از طریق بهبود یک یا چند صفت اقتصادی در جامعه است ( *Soufy et al.*, 2009). این بهبود می‌تواند از طریق انتخاب گوسفندان برتر به عنوان والدین نسل آینده بر اساس شایستگی ژنتیکی آنها برای صفات مورد نظر صورت گیرد ( *Khodabakhshzadeh et al.*, 2015a; *Khodabakhshzadeh et al.*, 2016a).

### مواد و روش‌ها

از ۳۰۰ رأس گوسفند نژاد زندی حاضر در دو گله مرکز اصلاح نژاد خجیر واقع در تهران وابسته به سازمان جهاد کشاورزی استان تهران خون‌گیری به عمل آمد. مقدار ۷-۵ سی سی EDTA در لوله‌های خلاء با pH ۵-۷ ریخته شد. نمونه‌های گرفته شده بلافاصله بعد از شماره‌گذاری 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 داشته و در دو انتهای توزیع صفت وزن از شیرگیری قرار داشتند به تعداد ۷۳ رأس انتخاب شدند. بدین منظور از اطلاعات مربوط به شجره و رکوردهای وزن تولد، شیرگیری، شش ماهگی، نه ماهگی و یکسالگی شامل ۸۳۶۶ رکورد که از سال ۱۳۷۱ تا ۱۳۹۰ در ایستگاه پرورش و اصلاح نژاد گوسفند زندی استان تهران (ایستگاه خجیر) جمع‌آوری شده بود، مورد استفاده قرار گرفت. پس از تجزیه و تحلیل داده‌ها و پیش‌بینی ارزش‌های اصلاحی حیوانات، دام‌هایی که در دو انتهای توزیع ارزش‌های اصلاحی بالا و پایین قرار داشتند انتخاب شدند. 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

جمعیتی با استفاده از روش XP-EHH نشان داد که چندین منطقه ژنومی بر روی کروموزومهای ۱، ۳، ۵، ۷، ۱۲، ۱۹، ۲۰ و ۲۳ دارای شواهدی از انتخاب می‌باشند. بررسی QTL‌های گزارش شده در مناطق کاندیدا نیز حاکی از آن بود که این مناطق با QTL‌های صفات مهم اقتصادی مرتبط با لاشه، تولیدمثل و متابولیسم چربی همپوشانی دارند (Zhao et al., 2016). ژن‌هایی که دارای فراوانی بالا و همچنین EHH بالایی در مجاور آلل انتخابی باشند هدف انتخاب‌های مثبت بوده و به خصوص در دام‌ها، که علاوه بر انتخاب طبیعی تحت انتخاب مصنوعی نیز قرار دارند می‌توانند پاره‌ای از جایگاه‌های کاندیدا برای ژن‌های عمده اثر باشند (Zhao et al., 2016). هدف از این تحقیق، بررسی ساختار گستره عدم تعادل پیوستگی و شناسایی مناطقی از ژنوم بود که در گوسفند نژاد زندی طی سالیان متمادی به صورت طبیعی و مصنوعی هدف انتخاب‌های مختلف قرار گرفته‌اند. همچنین معیار انتخاب در این ایستگاه براساس وزن از شیرگیری و برخی از مشخصات ظاهری بدن آنان انجام می‌شود. این مطالعه می‌تواند مبنای انجام تحقیقات تکمیلی پویش کل ژنوم گله مذکور به منظور شناسایی نواحی تحت انتخاب جهت بهبود راندمان برنامه‌های انتخاب و استراتژی‌های اصلاح نژادی آن باشد.

حذف شدند. در نهایت برای SNP های باقیمانده آنهایی که در تعادل هاردی-واینبرگ قرار نداشتند به عنوان معیاری از خطای تعیین ژنوتیپ (Teo et al., 2007) 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 (Purcell et al., 2007). برای محاسبه مقدار LD در بین نشانگرها از آماره  $r^2$  استفاده شد (William, 1974):

$$r^2 = \frac{(P_{11}P_{12} - P_{12}P_{21})^2}{P_{A1}P_{A2}P_{B1}P_{B2}}$$

که در آن  $P_{A1}$ ،  $P_{A2}$ ،  $P_{B1}$  و  $P_{B2}$  فراوانی های هر یک از آلل ها در دو جایگاه A و B است.  $P_{11}$ ،  $P_{12}$ ،  $P_{21}$  و  $P_{22}$  فراوانی هاپلوتیپ ها به ترتیب برای  $A1B1$ ،  $A1B2$ ،  $A2B1$  و  $A2B2$  می باشد. میزان LD برای هر کروموزوم اتوزومی به صورت جداگانه و در فواصل 0-10 kb، 10-25 kb، 25-50 kb، 50-100 kb، 100-500 kb، 0.5-1 Mb، 1-5 Mb و 5-10 Mb محاسبه شد.

جهت بررسی الگوی انتخاب مثبت در سطح ژنوم این دو جمعیت از روش XP-EHH استفاده شد. مهمترین آزمون های مبتنی بر پایه عدم تعادل پیوستگی و طول هاپلوتیپی، روش XP-EHH می باشد در این روش مناطق ژنومی تحت انتخاب بوسیله آلل های با EHH و فراوانی بالا را

رقیق شد. ژنوتیپ نمونه ها در کمپانی Neogen (<http://genomics.neogen.com>) در کشور آمریکا با استفاده از آرایه های Illumina BeadChip OvineSNP50K با بکارگیری پروتکل استاندارد شرکت ایلومینا (Illumina Inc., San Diego, CA, USA) تعیین ژنوتیپ شدند. این آرایه ها امکان تعیین ژنوتیپ همزمان ۵۴۲۴۱ جایگاه نشانگری با میانگین فاصله بین نشانگرها ۵۰/۹ Kb را فراهم می کنند (Illumina BeadChip OvineSNP50 Datasheet). برای فیلتراسیون داده های تعیین ژنوتیپ شده از معیارهای فراوانی تعیین ژنوتیپ شده نمونه ها، نرخ تعیین ژنوتیپ شده نشانگرها در هر نمونه و حداقل فراوانی آللی نادر (MAF) استفاده می شود. در ابتدا نمونه هایی که فراوانی تعیین ژنوتیپ شده آنها کمتر از ۹۰ درصد بود، شناسایی و حذف شد. این نمونه ها احتمال بیشتری داشته که با داده های گمشده همراه بوده و خطای تعیین ژنوتیپ شده در آنها بالا باشد. در مرحله بعد نشانگرهایی که حداقل فراوانی آللی در آنها کمتر از ۵ درصد بود حذف شدند، زمانی که فراوانی آللی نشانگرها کمتر از ۵ درصد باشد سبب کمتر برآورد شدن آماره  $r^2$  برای میزان LD بین جفت نشانگرها خواهد شد (Espigolan et al., 2013). سپس نشانگرهایی که نرخ تعیین ژنوتیپ شده آنها در نمونه ها کمتر از ۹۵ درصد بود شناسایی و

1. underestimate

استفاده شد  
<https://www.ncbi.nlm.nih.gov/genome/83>  
 (اما، از آنجا که فاصله ژنتیکی بین SNP ها به سانتی مورگان در دسترس نیست برای تعیین موقعیت ژنتیکی نشانگرها از تبدیل ساده  $1cM=1Mbp$  استفاده شد (Kijas et al., 2012).  
 فایل های خروجی XP-EHH با توجه به میانگین و انحراف معیار کل ارزش های به دست آمده استاندارد می شود که واسطه آن ارزش های نهایی دارای میانگین صفر و واریانس یک خواهد بود.  
 در این تحقیق برای ترسیم گراف منهن از نرم افزار Haploview v.4.2 استفاده شد (Barrett et al., 2005). به منظور شناسایی ژن های مرتبط با بخش های انتخاب ژنوم، ۵۰۰ Kb پیش و پس از نقاط آغاز و پایان توالی مرتبط مورد بررسی بیشتر قرار گرفت. با توجه به اینکه شرح نویسی ژنوم گوسفند هنوز تکمیل نشده است، پایگاه هایی که برای توصیف مکانیسم های بیوشیمیایی و شبکه ژنی موجود هستند، گونه گوسفند را قبول نمی کنند لذا از ژنوم انسان جهت بررسی اینکه مناطق ژنومی شناسایی شده حاوی ژن خاصی هستند، در مناطق اورتولوگوس ژنوم انسان با استفاده از جستجو در پایگاه اطلاعاتی آنالین Ensembl Genes 88 Database بوسیله BioMart (www.ensembl.org/biomart) مورد بررسی قرار گرفت. برای تفسیر بهتر عملکرد ژن های به دست آمده از پایگاه های اطلاعاتی آنالین GeneCards (http://www.genecards.org) و

تشخیص می دهد (Sabeti et al., 2007). جهت محاسبه آماره XP-EHH بایستی وضعیت اجدادی آلل های مختلف SNP های مورد استفاده مشخص شود. بدین منظور وضعیت اجدادی SNP ها از پروژه HapMap Sheep (http://www.sheephapmap.org) تهیه شد. در روش XP-EHH هاپلوتیپ ها در دو جمعیت با هم مقایسه می شوند تا تنوع در میزان نوترکیبی در قسمت های مختلف ژنوم نیز در نظر گرفته شود. جهت تعیین فازهای هاپلوتیپی برای کل ژنوم هر یک از جمعیت های مورد مطالعه، از نرم افزار fastPHASE V.1.2.3 استفاده شد. در این برنامه از روش حداکثر درستنمایی برای تخمین ژنوتیپ های گمشده و بازسازی هاپلوتیپ ها استفاده می شود (Scheet & Stephens, 2006). برنامه XP-EHH از وبسایت آزمایشگاه Pritchard (<http://hgdp.uchicago.edu/Software/>) به دست آمد. فایل های ورودی برای نرم افزار XP-EHH شامل دو فایل داده برای فازهای هاپلوتیپی برای هر جمعیت تعیین شده و فایل نقشه های SNP های مورد استفاده است. فایل نقشه ورودی حاوی ۵ ستون شامل نام SNP، موقعیت فیزیکی و موقعیت ژنتیکی و آلل هایی که بیانگر عدد یک و صفر در فایل هاپلوتیپ هستند می باشد. برای تعیین موقعیت ژنومی SNP ها در سطح ژنوم گوسفند از آخرین نسخه ژنومی گوسفند (assembly Oar\_v4.0) پایگاه اطلاعاتی NCBI

نشانگر به دلیل عدم تعادل هاردی واینبرگ، ۱۵۴۲ نشانگر روی کروموزوم X و ۲۹۱۶ نشانگر با موقعیت ناشناخته حذف شدند. همچنین ۲ نمونه به دلیل فراوانی تعیین ژنوتیپ شده کمتر از ۰.۹٪ حذف شدند. تعداد نشانگرها به ترتیب کروموزوم، میانگین و حداقل و حداکثر فاصله بین دو نشانگر SNP در جدول ۱ نشان داده شده است. بیشترین تعداد SNP بر روی کروموزوم شماره یک و کمترین آنها روی کروموزوم ۲۴ واقع شده است. میانگین فاصله بین دو SNP، ۵۹ Kb و حداقل و حداکثر فاصله بین دو SNP به ترتیب ۵/۱۴ Kb در داخل کروموزوم ۶ و ۲/۸۹ Mb در داخل کروموزوم ۲۲ قرار داشت.

نتایج آنالیز LD در بین جفت نشانگرها به تفکیک هر کروموزوم تا فاصله ۱۰ Mb بین جفت نشانگرها در جدول ۲ نشان داده شده است. با افزایش فاصله جفت نشانگرها مقدار LD کاهش یافته است و مقدار LD حدود ۰/۱۸ کمتر از 50 Kb توسعه یافته است. در فاصله کمتر از 10 Kb میانگین مقدار LD برابر با ۰/۳۴ بوده و بیشترین و کمترین میزان LD به ترتیب روی کروموزوم-های ۲۵ و ۱۷ مشاهده شد.

UniProtKB (<http://www.uniprot.org>) استفاده شد. جهت کاوش فرآیندهای بیولوژیکی و عملکردی ژن‌ها، آنتولوژی ژن و آنالیز مسیر (KEGG) با استفاده از نرم افزار آنالیز 6.8 DAVID (<http://david.abcc.ncifcrf.gov/>) استفاده شد. جهت ترسیم شبکه‌های ژنی به منظور بررسی ارتباط متقابل پروتئین‌های شناسایی شده در مقایسات مربوطه با یکدیگر از پایگاه اطلاعاتی STRING v10 (<http://string-db.org>) استفاده شد. در نهایت برای بررسی اینکه آیا مناطق مورد نظر با QTL‌های شناسایی شده همپوشانی دارد یا خیر، از مرکز اطلاعاتی آنالیز QTL (<http://i.animalgenome.org/jbrowse/>) مورد بررسی قرار گرفت.

## نتایج و بحث

پس از اجرای مراحل مختلف کنترل کیفیت داده‌ها، 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 آلی کمتر از ۵ درصد، ۱۰۷۰ نشانگر به دلیل نرخ تعیین ژنوتیپ شده کمتر از ۰.۹۵٪ در هر نمونه، ۱۱۷

جدول ۱- مشخصات نشانگرهای مورد استفاده در آنالیزها.

Table 1- Summary of characteristic SNPs used in the analysis.

کمترین فاصله بین SNP (kb)	بیشترین فاصله بین SNP (Mb)	میانگین فاصله بین SNP (Mb)	تعداد نشانگر	طول (Mb)	کروموزوم Chromosome
Shortest SNP interval	Longest SNP interval	Average SNP interval	No. of SNP	Length	
5.393	0.813	0.053	5168	275.41	1
5.735	0.784	0.061	4036	248.97	2
5.525	1.284	0.060	3191	224.00	3
5.204	0.457	0.056	2098	119.22	4
5.164	0.743	0.060	1778	107.84	5
5.141	1.672	0.059	1957	116.89	6
5.402	0.205	0.054	1845	100.01	7
5.184	0.404	0.054	1653	90.62	8
5.705	0.506	0.061	1546	94.58	9
5.585	2.019	0.059	1461	86.38	10
5.399	0.511	0.062	991	62.17	11
5.387	1.068	0.060	1301	79.03	12
5.577	0.803	0.065	1270	82.95	13
5.504	1.009	0.063	991	62.57	14
5.613	1.055	0.063	1278	80.78	15
5.792	0.524	0.061	1170	71.69	16
5.712	0.463	0.060	1204	72.25	17
5.565	0.575	0.057	1201	68.49	18
5.577	0.373	0.063	954	60.45	19
5.598	1.269	0.060	845	51.05	20
5.574	2.022	0.059	841	49.99	21
5.411	2.891	0.054	935	50.78	22
5.484	0.615	0.063	985	62.28	23
5.613	0.434	0.066	628	41.98	24
5.712	0.490	0.054	824	45.22	25
5.714	1.192	0.060	728	44.05	26



جدول ۲- آمار توصیفی میانگین  $r^2$  بین جفت نشانگرها تا فاصله ۱۰ مگاباز در ژنوم.

**Table 2- Statistical information for average  $r^2$  as distance between pairs of SNP up to 10 Mb for the genome.**

فاصله بین SNP Distance interval SNP								کروموزوم Chromosome
5-10 Mb	1-5 Mb	0.5-1 Mb	100-500 Kb	50-100 Kb	25-50 Kb	10-25 Kb	0-10 Kb	
0.031	0.061	0.079	0.089	0.105	0.186	0.212	0.347	1
0.021	0.061	0.070	0.083	0.114	0.175	0.283	0.354	2
0.031	0.061	0.071	0.093	0.114	0.184	0.256	0.338	3
0.020	0.051	0.079	0.083	0.112	0.193	0.269	0.352	4
0.030	0.051	0.050	0.083	0.124	0.178	0.245	0.353	5
0.031	0.051	0.065	0.093	0.122	0.192	0.284	0.395	6
0.031	0.051	0.060	0.072	0.108	0.163	0.263	0.339	7
0.030	0.061	0.079	0.092	0.157	0.197	0.290	0.401	8
0.010	0.051	0.068	0.082	0.104	0.214	0.282	0.355	9
0.022	0.047	0.072	0.073	0.111	0.203	0.282	0.343	10
0.029	0.044	0.059	0.072	0.132	0.156	0.197	0.371	11
0.020	0.045	0.078	0.092	0.105	0.162	0.170	0.332	12
0.030	0.051	0.080	0.093	0.117	0.176	0.185	0.367	13
0.030	0.051	0.079	0.092	0.109	0.167	0.202	0.352	14
0.020	0.049	0.079	0.092	0.104	0.158	0.181	0.350	15
0.020	0.049	0.070	0.082	0.106	0.191	0.254	0.371	16
0.020	0.050	0.072	0.089	0.159	0.186	0.195	0.256	17
0.020	0.049	0.077	0.082	0.093	0.132	0.144	0.291	18
0.030	0.049	0.079	0.082	0.123	0.164	0.172	0.383	19
0.030	0.050	0.070	0.087	0.190	0.192	0.242	0.319	20
0.030	0.050	0.072	0.087	0.140	0.161	0.199	0.348	21
0.032	0.049	0.072	0.087	0.130	0.159	0.164	0.287	22
0.032	0.050	0.078	0.082	0.155	0.192	0.211	0.303	23
0.032	0.043	0.077	0.083	0.133	0.156	0.191	0.284	24
0.042	0.068	0.088	0.099	0.162	0.242	0.257	0.417	25
0.011	0.056	0.078	0.083	0.119	0.139	0.155	0.276	26

مورد مطالعه، عدم تجانس LD در طول ژنوم، تراکم نشانگر و عدم تجانس آللیک QTL بستگی خواهد داشت. جمعیتی با مقدار بالای LD به تراکم نشانگری کمتر نیازمند می‌باشد و بالعکس (Meadows *et al.*, 2008). در این مطالعه برای

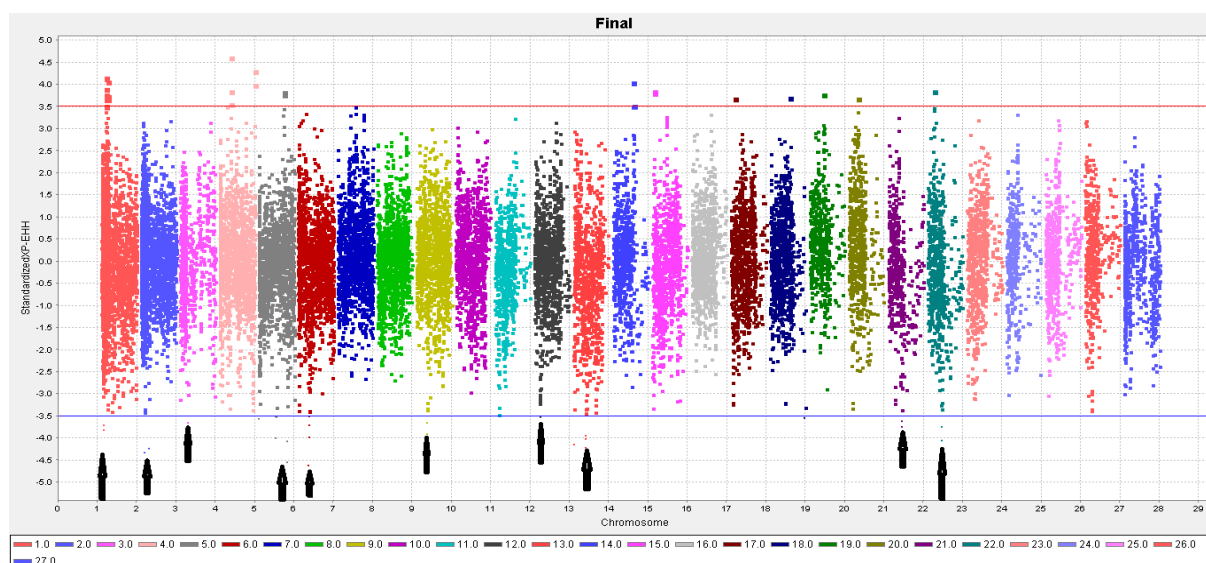
همچنین نیمه طول محاسبه شده برای  $r^2$  (فاصله‌ای که میانگین  $r^2$  در هر Mb از مقدار حداکثری خود به نصف می‌رسد) ۰/۰۷ بوده که متناظر با ۴/۷ Mb بود. مطالعات نشان می‌دهد که بازدهی نقشه‌یابی LD به میزان LD در جمعیت

LD در بین کروموزوم‌های مختلف در این تحقیق متفاوت بود. تفاوت در مقدار LD مابین کروموزوم‌ها در گوسفند توسط محققین مختلفی گزارش شده است (Liu et al., 2015; Gholizadeh 2017). دلیل آنرا می‌توان نرخ متفاوت نوترکیبی درون و بین کروموزوم‌ها، هتروزیگوسیتی، رانش ژنتیکی و تأثیر انتخاب صفات مهم اقتصادی مربوط دانست (Liu et al., 2017). به علاوه، در داخل یک کروموزوم نیز نرخ نوترکیبی از ناحیه سانترومیک به ناحیه تلوومیک افزایش می‌یابد. این موضوع می‌تواند منجر به تنوع در مقدار و وسعت عدم تعادل پیوستگی در نواحی مختلف ژنوم شود. سطح مفید عدم تعادل پیوستگی در واقع میزان تنوع جایگاه صفت کمی است که در نشانگر مشاهده می‌شود. Meuwissen et al. (2001) گزارش کردند برای افزایش صحت برآورد ارزش اصلاحی ژنومی (GEBV) تا ۰/۸۵ مقدار عدم تعادل پیوستگی براساس آماره  $r^2$  حداقل بایستی ۰/۲ باشد. با توجه به جمعیت حاضر تقریباً هر 40 Kb بایستی یک نشانگر وجود داشته باشد تا بتوان به میانگین  $r^2$  برابر با ۰/۲ رسید. آماره XP-EHH که مبتنی بر عدم تعادل پیوستگی و طول هاپلوتیپی می‌باشد، جهت شناسایی نواحی ژنومی تحت انتخاب مثبت مورد استفاده قرار گرفت. این روش مناطق ژنومی تحت انتخاب را بوسیله آل‌های اجدادی و فازهای هاپلوتیپی تشخیص می‌دهد. بعد از مرحله فراخوانی SNP‌های دارای اطلاعات اجدادی از

بررسی مقدار LD از آماره  $r^2$  (William 1974) استفاده شد. بررسی این آماره نشان داد که با افزایش فاصله نشانگرها، مقدار LD کاهش یافته است. کاهش مقدار  $r^2$  با افزایش فاصله نشان می‌دهد که چه مقدار نشانگر یا فنوتیپ در ابتدای مطالعه ژنومی مورد نیاز می‌باشد. در پژوهشی Przeworski & Pritchard (2001) گزارش دادند که به طور کلی اندازه نمونه باید به مقدار  $1/r^2$  افزایش یابد تا QTL ژنوتیپ نشده شناسایی شود. برتری دیگر  $r^2$  نسبت به آماره‌های دیگر علاوه بر قدرت پیش‌بینی نسبت به تعداد نشانگر این است که برخلاف آماره  $D'$  تحت تأثیر اریبی تعداد کم نمونه و یا فراوانی پایین آللی قرار نمی‌گیرد. نتایج این تحقیق نشان داد که مقدار  $r^2$  حدود ۰/۲۰ در گوسفند نژاد زندگی کمتر از 40 Kb توسعه یافته است که با نتیجه به دست آمده در گوسفند بلوچی ایرانی مطابقت دارد (Gholizadeh et al., 2014). گزارشات قبلی در گوسفند و گاو تا ۲۰ Mb مقدار LD را براساس آماره  $r^2$  گزارش داده‌اند و فرض می‌شود که بعد از ۱۰ Mb نوترکیبی رخ نمی‌دهد (De Roos et al., 2008). نتایج تحقیق حاضر کمتر از دامنه توسعه LD در گاوهای گوشتی می‌باشد که به طور میانگین 50 Kb گزارش شده است (Porto-Neto et al., 2014). به نظر می‌رسد پایین بودن مقدار LD نسبت به آنچه در نژادهای گاو گزارش شده است به طور عمده به اندازه مؤثر جمعیت مربوط می‌شود که در گوسفندان بالاتر می‌باشد. مقدار

وزن بدن و وزن لاشه و ترکیب لاشه بر روی کروموزوم‌های مختلف گویای این است که توارث چندژنی ذخیره گوشت و چربی وابسته به جنس نیست. در کل بررسی همه مناطق ژنومی تحت انتخاب با استفاده از پایگاه اطلاعاتی GeneCards و UniProtKB نشان داد که همه این مناطق با صفات رشد و متابولیسم چربی، توسعه نظام استخوان بندی، متابولیسم قند و انرژی و تولیدمثل مرتبط هستند و فهرست ژن‌هایی که درون یا نزدیک این جایگاه‌های کروموزومی قرار دارند در جدول ۳ ارائه شده است. شماری از ژن‌ها در این بررسی به عنوان نشانه‌های انتخاب گزارش شده‌اند، که با بررسی‌های پیشین همخوانی داشت (Kijas *et al.*, 2012; Fariello *et al.*, 2014; Mancini *et al.*, 2014; Jarvis *et al.*, 2012; Li & Kim, 2015; Yang *et al.*, 2016). ژن‌های FBP1 و FBP2 که در این تحقیق با منطقه تحت انتخاب مثبت روی کروموزوم ۲ همپوشانی دارند، نقش کلیدی در چرخه پنتوز فسفات و چرخه کربس داشته و به عنوان کاتالیز آنزیم تبدیل فروکتوز ۱-۶-فسفات به فروکتوز ۶-فسفات می‌باشند. علاوه بر این در هموستازی گلوکز، تنظیم گلوکونئوزنز و ایزوزیم عضله نقش دارند و نقص در این آنزیم‌ها باعث محدود شدن افزایش وزن می‌شود (GeneCards).

مجموع کل SNPها ویرایش داده‌ها انجام شد. در این تحقیق، برای شناسایی بهتر مناطق ژنومی تحت انتخاب در کل ژنوم گراف منهن ترسیم شد. در این گراف بخش‌هایی از ژنوم که ارزش‌های XP-EHH بالایی دارند، نشان دهنده تمایز جمعیت‌های با وزن از شیرگیری بالا و وزن از شیرگیری پایین در گوسفند نژاد زندی در آن مناطق ژنومی در اثر انتخاب مثبت هستند که به دنبال انتخاب‌های طبیعی و مصنوعی در طی نسل‌های مختلف برای جایگاه‌های مورد نظر به وجود آمده است. نتایج به دست آمده نشان داد که در چندین منطقه ژنومی SNPهای مجاور تفرق جمعیتی بالایی داشتند، که در این تحقیق ۱۰ منطقه ژنومی روی کروموزوم‌های ۱، ۲، ۳، ۵، ۶، ۹، ۱۲، ۱۳، ۲۱ و ۲۲ شناسایی شد (شکل ۱). ارزش‌های XP-EHH بالاتر از ۳/۵ در مقایسه بین دو وزن متفاوت در صدک ۹۹/۹ کل ارزش‌های XP-EHH قرار داشت و بسیار معنی‌دار بود. در این مقایسه نشانه‌های انتخاب برای جمعیت‌های با وزن بالای از شیرگیری دارای XP-EHH کمتر از ۳/۵۰- به دست آمد. در این بخش برخی از ژن‌هایی که تحت انتخاب بودند، بحث می‌شود. ژن‌های شناسایی شده در نواحی مختلف ژنومی می‌توانند به عنوان کاندیدهای تحت انتخاب مطرح باشند. در این بررسی جمعیت با وزن از شیرگیری بالا و جمعیت با وزن از شیرگیری پایین را بررسی کردیم. جایگاه‌های مرتبط با رشد،



شکل ۱- توزیع ارزش‌های آماره XP-EHH در سطح ژنوم جمعیت گوسفندان نژاد زندی با موقعیت کروموزومی SNPها بر روی محور X و ارزش‌های XP-EHH استاندارد شده آنها بر روی محور Y نمایش داده شده است. خط ترسیم شده نشان دهنده ۹۹/۹ صدک کروموزوم‌های اتوزوم و X می‌باشد.

**Figure 1- Distribution genomes of XP-EHH values in Zandi sheep breed population by chromosome: SNP position in the genome is shown on the X-axis, and XP-EHH standard values are plotted on the Y-axis. The values above the lines are in 99.9 percentile of autosomal and chromosome X SNPs.**

میانگین خوراک مصرفی روزانه، طول لاشه، وزن لاشه و ضخامت چربی پشت گزارش شده است (Guo *et al.*, 2007).

ژن ABOP که بر روی کروموزوم ۳ واقع شده است، نقش‌های کلیدی در ارتباط با تنظیم میزان خوراک مصرفی، متابولیسم چربی، ذخیره پروتئین و همچنین جذب و هضم ویتامین‌ها در بدن دارد (GeneCards). ژن APOB در متابولیسم لیپید از طریق جذب و ترشح لیپیدها شامل کلسترول و تری گلیسریدها دخالت دارد.

این ژن‌ها در فرآیندهای بیولوژیکی در متابولیسم قندهای فروکتوز و بیوستنز ساکارز نقش دارند. ژن FBP1 فرم کبدی آنزیم او۶ فسفاتاز و ژن FBP2 فرم عضله‌ای آنزیم او۶ فسفاتاز است (UniProtKB Gene).

در مطالعه Weikard *et al.* (2012) نقش این دو ژن در ارتباط با عضله‌های اسکلتی و غدد پستانی معنی دار گزارش شده است. همچنین در مطالعه‌ای روی ۴ نژاد گاو گوشتی (آنگوس، سمیتال، آمیخته‌های سمیتال و هرورد) ارتباط معنی‌داری بین چند شکلی در ژن FBP1 با

جدول ۳- ژن‌های شناسایی شده مرتبط با مناطق تحت انتخاب در گوسفند نژاد زندی.

Table 3- The identified genes under selection in Zandi sheep breed.

موقعیت ژنومی گوسفند (کروموزومی) Location on Ovine genome (Chromosome)	موقعیت ژنومی انسان (کروموزومی) Location on Human genome (Chromosome)	ژن (ها) Gene(s)	صفات مرتبط Related Traits	منابع References
(2) 31653188-31684127	(9) 94603133-94640249	FBP1	میانگین خوراک مصرفی روزانه، طول لاشه Average Daily Feed Intake, Carcass Length	Guo <i>et al.</i> , 2007
(2) 31709019-31765595	(9) 94558720-94593793	FBP2	رشد ماهیچه‌های اسکلتی، تولید شیر Growth of skeletal muscle; Milk Production	Weikard <i>et al.</i> , 2012
(3) 28671825-28710732	(2) 21044073-21001429	APOB	وزن بدن در سنین مختلف، چاقی body weight in different age; obesity	Zhang <i>et al.</i> , 2006
(5) 5056451-5057662	(19) 17816512-17821574	INSL3	کیفیت اسپرم، باروری Sperm quality, Fertility	Mishra <i>et al.</i> 2013
(6) 36514210-36556824	(4) 88257620-88284769	ABCG2	کارایی مصرف خوراک، چربی داخل ماهیچه ای (ماربیلینگ) Feed efficiency; Intramuscular fat	Mishra <i>et al.</i> , 2014
(6) 37138667-37257756	(4) 17629306-17781512	FAM184 B	صفات رشد، افزایش وزن روزانه Growth traits; Average daily gain	Jin <i>et al.</i> , 2015
(6) 36837847-36848482	(4) 87799581-87812435	IBSP	اندازه گوساله، سقط جنین Calf size, Dystocia	Olsen <i>et al.</i> , 2010
(6) 37092323-37118578	(4) 17577192-17607972	LAP3	رشد جنین و اندازه لاشه لاشه Foetal growth and Carcass size; Carcass traits	Bongiorni <i>et al.</i> , 2012; Saatchi <i>et al.</i> , 2014
(6) 36804948-36811949	(4) 87799581-87846817	MEPE	اندازه گوساله، سقط جنین Dystocia	Olsen <i>et al.</i> , 2010
(6) 36566367-36630153	(4) 88007668-88077777	PKD2	وزن لاشه گرم و میزان چربی داخل ماهیچه ای (ماربیلینگ) Hot carcass weight; Intramuscular fat	Abo-Ismael <i>et al.</i> , 2014
(6) 36404281-36418026	(4) 88257620-88284769	PPM1K	وزن لاشه گرم، باقیمانده خوراک مصرفی Hot carcass weight; Residual feed intake	Abo-Ismael <i>et al.</i> , 2014
(6) 36651734-36658288	(4) 87975650-87983426	SPP1	کشتار، افزایش وزن روزانه وزن بدن قبل از Yearling body Weight, Body weight before Slaughter, Average Daily Gain	White <i>et al.</i> , 2007
(9) 37556752-37589709	(8) 58552924-58582860	SDCBP	وزن لاشه گرم Hot carcass weight	Lee <i>et al.</i> , 2013
(12) 25191495-25241603	(1) 223701593- 223776018	CAPN2	کیفیت گوشت Meat quality	Tizioto <i>et al.</i> , 2013
(12) 25081358-25156543	(1) 223538007- 223665734	CAPN8	بازدهی خوراک مصرفی، افزایش وزن روزانه Feed efficiency; Average Daily Gain	Olivieri <i>et al.</i> , 2016
(13) 41742022-41775868	(20) 25006230-25058980	ACSS1	سنتز اسیدهای چرب اشباع saturated fatty acid	Toral <i>et al.</i> , 2016
(21) 46224390-46229122	(11) 69698232-69704642	FGF19	چاقی Obesity	Kharitononkov, 2009

2009). این کلاستر ژنی در ۷۴ نژاد گوسفند در سراسر جهان (Kijas *et al.*, 2012) و نیز گوسفندان جنوب غربی آسیا و آفریقا (Fariello *et al.*, 2014) تحت انتخاب بوده است. ژن SPP1 که نام دیگر آن استئوپتین (OPN) می‌باشد و در فسفریله کردن گلیکوپروتئین‌ها و تشکیل استخوان در انسان و موش از طریق تفرق سلوهای استئوبلاست نقش کلیدی دارد (Denninger *et al.*, 2015). ژن‌های ABCG2، LAP3 و PPM1K که بر روی کروموزوم ۶ واقع شده‌اند دارای نقش‌های کلیدی در رشد می‌باشند که در اینجا بحث می‌شود. ارتباط معنی‌داری بین چند شکلی در ژن ABCG2 با صفات مرتبط با کیفیت گوشت گزارش شده است (Abo-Ismael *et al.*, 2014). نقش اصلی ژن کاندیدای ABCG2 در سلول‌های اپیتلیال پستانی در تولید و ترکیبات چربی و پروتئین شیر می‌باشد به طوریکه موجب افزایش تولید چربی و درصد چربی و پروتئین و کاهش تولید شیر می‌شود. ژن LAP3 ارتباط معنی‌داری با صفات مرتبط با رشد جنین و راحت گوساله‌زایی در نژاد گوشتی Piedmontese ایتالیایی گزارش شده است. ژن PPM1K در فرآیند متابولیک فسفر و دفسفریلاسیون اسیدهای آمینه نقش دارد. در مطالعه Abo-Ismael *et al.* (2014) ارتباط ژن PPM1K با باقیمانده مصرف خوراکی و افزایش وزن لاشه گرم گزارش کردند. ژن SDCBP، که

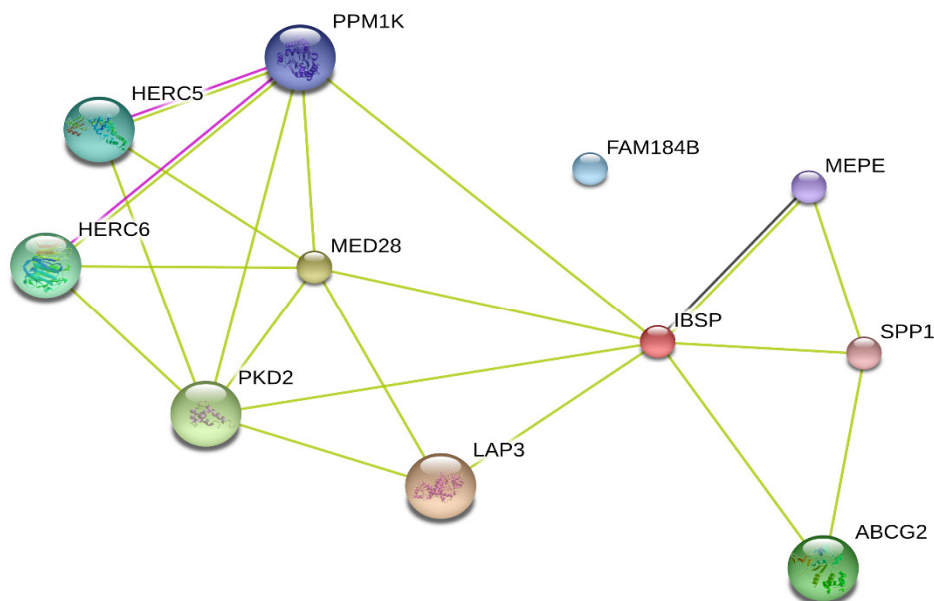
در جوجه‌های گوشتی ارتباط معنی‌داری بین چندشکلی در این ژن با وزن بدن و ذخیره چربی گزارش شده است (Zhang *et al.*, 2006). ژن INSL3 که RLF هم نامیده می‌شود روی کروموزوم ۵ بوده و از سلول‌های لایدیگ بیضه در دوران جنینی و بلوغ ترشح شده و نقش کلیدی در تفرق بیضه در نرها در دوران جنینی و در دوران بلوغ بر تولید و کیفیت اسپرم دارد (Mishra *et al.*, 2013). ژن INSL3 جزو خانواده ژنی هورمون‌های شبه انسولین می‌باشد که بیشتر در اندام‌های جنسی تولید می‌شود و در توسعه اندام‌های تناسلی در حیوانات ماده و باروری نقش دارد (Pitia *et al.*, 2016). ژن PKD2 بر روی کروموزوم ۶ مکان یابی شده و در انتقال یون کلسیم و سیگنالینگ کلسیم در سلول‌های اپیتلیال و هوموستازی کلسیم و همچنین در پاسخ سلولی به استرس‌های اسموتیک نقش دارد (GeneCards). ژن‌های SPP1، IBSP و MEPE که بر روی کروموزوم ۶ واقع شده‌اند، یک کلاستر ژنی موجود در استخوان و غضروف می‌باشند که جزو فسفوگلیکوپروتئین‌های ماتریکس خارج سلولی (ECM) هستند و نقش اساسی در تشکیل فرم طبیعی استخوان در جنین در سن ۱۸۰ روزگی دارند که به موثری بر اندازه گوساله نقش دارند و به طور غیرمستقیم بر روی سخت‌زایی و یا مرده‌زایی معنی گزارش شده است (Olsen *et al.*, )

جنینی، بازیابی بافت، تقسیمات سلولی و متابولیسم انرژی دخالت دارند ( UniProtKB Gene). بررسی QTL های شناسایی شده در مناطق ژنومی مورد نظر در گوسفند نشان داد که بر روی کروموزوم ۲، ۶ و ۹ QTL های گزارش شده مرتبط با صفات مهم اقتصادی از جمله وزن بدن در هنگام کشتار، وزن لاشه گرم و میزان گوشت در لاشه بود (Cavanagh et al., 2010). همچنین QTL گزارش شده بر روی کروموزوم ۱۲ با تراکم حجم استخوان که همبستگی بالایی با وزن بدن، ماهیچه و چربی داشت، گزارش شده بود (Campbell et al., 2003).

برای درک بهتر عملکردهای مولکولی ژن های تحت انتخاب مثبت، هستی شناسی ژن ها مورد بررسی قرار گرفت. آنالیز هستی شناسی ژن در جدول ۴ در برخی از مسیرهای مهم و معنی دار ارائه شده است. همچنین آنالیز مسیر در جدول ۵ ارائه شده است. برای پیش بینی ارتباط متقابل پروتئین های شناسایی شده در مناطق اورتولوگوس انسان از پایگاه STRING برای ترسیم شبکه ژنی استفاده شد. نتایج حاصل از شبکه پروتئینی نشان داد که ژن های به دست آمده به طور مستقیم و غیر مستقیم با توسعه سیستم اسکلتی و رشد در ارتباط هستند (شکل ۲).

بر روی کروموزوم ۹ واقع شده است در مورفولوژی و توسعه بافت نقش دارد (GeneCards) به طوریکه در فرآیندهای بیولوژیکی نقش تنظیمی مثبت در رشد سلول دارد (UniProtKB Gene). در مطالعه Lee et al. (2013) ارتباط معنی داری بین چند شکلی در ژن SDCBP با وزن لاشه گرم در گاوهای گوشتی هانفو گزارش شده است. ژن های کاندیدای CAPN2 و CAPN8 بر روی کروموزوم ۱۲ واقع شده اند از زیر واحدهای آنزیم کالپین می باشند (GeneCards). تاثیر این دو ژن بر روی کارایی مصرف خوراک در جمعیت گاوهای گوشتی نلور شامل میزان ماده خشک مصرفی و کاهش باقیمانده خوراک مصرفی و میانگین افزایش وزن روزانه معنی دار گزارش شده است (Olivieri et al., 2016).

ژن ACSS1 که بر روی کروموزوم ۱۳ واقع شده است نقش پیشرو در سنتز اسیدهای چرب دارد (GeneCards). ژن ACSS1 در فرآیندهای بیولوژیکی در تنظیم بیوسنتز کلسترول بوسیله پروتئین متصل شونده به عناصر کنترلی استرول (SREBP) نقش دارد (UniProtKB Gene). ژن FGF19 بر روی کروموزوم ۲۱ واقع شده است و از خانواده ژنی فاکتور رشد فیبروبلاستی یا (FGF) شامل ۲۲ ژن می باشد (GeneCards). این خانواده ژنی در بسیاری از فرآیندهای سلولی نظیر تمایز و رشد سلول های



شکل ۲- شبکه ژنی و پروتئینی ژن‌های تحت انتخاب مثبت. دایره‌های با رنگ‌های مختلف نشان دهنده ژن یا پروتئین‌ها می‌باشد و خطوط بین آنها تعاملات بین آنهاست بطوریکه خطوط صورتی برای پس از رونویسی، سیاه، واکنشی و آبی رنگ نشان دهنده تعاملات پروتئینی با یکدیگر است.

**Figure 2- Gene and Protein network of genes under positive selection. Nodes are genes or proteins; lines indicate interactions between proteins with: pink lines for post-translational, yellow lines for expression, black lines for reaction, and blue lines for binding.**

جدول ۴- آنالیز ژن آنتولوژی در مناطق تحت انتخاب مثبت.

**Table 4- Analysis of Gene ontology (GO) terms in regions under positive selection.**

دسته Category	ژن آنتولوژی GO term	توصیف ژن آنتولوژی Description GO	ارزش P P- Value	نام ژن Gene Names
GOTERM_BP_FAT	GO:0006006	glucose metabolic process	4.5E-2	FBP1 FBP2
GOTERM_BP_FAT	GO:0031214	ossification	1.5E-2	IBSP MEPE SPP1
GOTERM_MF_FAT	GO:0004197	cysteine-type endopeptidase activity	2.8E-2	CAPN2 CAPN8
GOTERM_MF_FAT	GO:0008234	cysteine-type peptidase activity	6.3E-2	CAPN2 CAPN8



جدول ۵- آنالیز غنی‌سازی مسیر در مناطق تحت انتخاب مثبت.

**Table 5- Analysis of Enriched pathway terms in regions under positive selection.**

مسیر	توصیف	تعداد ژن‌ها	ارزش P
Pathway term	Description	Number of Genes	P-value
hsa00010	Glycolysis / Gluconeogenesis	2	3.8E-2
hsa04910	Insulin signaling pathway	2	7.8E-2
hsa04512	ECM-receptor interaction	2	4.9E-2
hsa04152	AMPK signaling pathway	2	8.5E-2

### نتیجه‌گیری

بررسی ژن‌ها و QTL‌های موجود در نواحی تحت انتخاب مثبت نشان داد که جمعیت گوسفندان نژاد زندگی به صورت متمایز هدف انتخاب‌های طبیعی و یا مصنوعی قرار گرفته‌اند. بررسی ژن‌هایی که تا کنون در این مناطق گزارش شده است مرتبط با صفات رشد، متابولیسم چربی، متابولیسم قند و انرژی، لاشه و تولیدمثل بودند، اما بررسی‌های پیوستگی و عملکردی بیشتری جهت شناسایی عملکرد ژن‌ها و QTL‌های مرتبط نیاز می‌باشد. در مجموع، نتایج این تحقیق می‌تواند منبع اطلاعاتی با ارزشی در

جهت شناسایی ژن‌های متمایز کننده و شناسایی نواحی ژنومی کاندید برای بسیاری از صفات مهم اقتصادی موجود در گوسفندان نژاد زندگی فراهم آورد.

### سپاسگزاری

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.

### منابع

- Abo-Ismael MK, Voort GV, Squires JJ, Swanson KC, Mandell IB, Liao X, Stothard P, Moore S, Plastow G, Miller SP (2014). Single nucleotide polymorphisms for feed efficiency and performance in crossbred beef cattle. *BMC Genetics* 15: 1471-1480.
- Al-Mamun HA, Clark S, Kwan P, Gondro C (2015). Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. *Genetics Selection Evolution* 47: 1-11.

- Barrett JC, Fry B, Maller J, Daly MJ (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265.
- Bongiorni S, Mancini G, Chillemi G, Pariset L, Valentini A (2012). Identification of a short region on chromosome 6 affecting direct calving ease in piedmontese cattle breed. *PLoS One* 7: 137-145.
- Chen M, Pan D, Ren H, Fu J, Li J, Su G, Wang A, Jiang L, Zhang Q, Liu JF (2016). Identification of selective sweeps reveals divergent selection between Chinese Holstein and Simmental cattle populations. *Genetics Selection Evaluation* 48: 76-87.
- 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.
- Campbell AW, Bain WE, McRae AF, Broad TE, Johnstone PD, Dodds KG, Veenvliet BA, Greer GJ, Glass BC, Beattie AE, Jopson NB, McEwan JC (2003). Bone density in sheep: genetic variation and quantitative trait loci localisation. *Bone* 33: 540–548.
- Denninger KC, Litman T, Marstrand T, Moller K, Svensson L, Labuda T, Andersson A (2015). Kinetics of gene expression and bone remodelling in the clinical phase of collagen induced arthritis. *Arthritis Research Theoretical* 17: 43-55.
- De Roos AP, Hayes BJ, Spelman RJ, Goddard ME (2008). Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. *Genetics* 179: 1503–1512.
- Espigolan R, Baldi F, Boligon AA, Souza FRP, Gordo DGM, Tonussi RL, Cardoso DF, Oliveira HN, Tonhati H, Sargolzaei M, Schenkel FS, Carneiro R, Ferro JA, Albuquerque LG (2013). Study of whole genome linkage disequilibrium in Nellore cattle. *BMC Genomics* 14: 305-314.
- Ensembl BioMart: Ensembl online genome data base BioMart Tool. <http://www.ensembl.org/biomart/martview/>.
- Fariello MI, Servin B, Tosser-Klopp G, Rupp R, Moreno C, International Sheep Genomics Consortium, Cristobal MS, Boitard S (2014). Selection Signatures in Worldwide Sheep Populations. *PLoS One* 9: 813-82.
- Gholizadeh M, Rahimi Mianji G, Nejati Javaremi A (2014). Linkage disequilibrium estimation and haplotype based genome-wide association to detect QTLs affecting twinning rate in Baluchi sheep. *Research on Animal Production* 10: 166-178.
- Guo H, Liu WS, Takasuga A, Eyer K, Landrito E, Xu SZ, Gao X, Ren HY (2007). Characterization and Mapping of the Bovine FBP1 Gene. *Asian-Australian Journal of Animal Science* 9: 1319–1326.
- GeneCards. <http://www.genecards.org/cgi-bin/carddisp.pl?gene=STAT>.
- Iso-Touru T, Tapio M, Vilkki J, Kiseleva T, Ammosov I, Ivanova Z, Popov R, Ozerov M, Kantanen J (2016). Genetic diversity and genomic signatures of selection among cattle breeds from Siberia, eastern and northern Europe. *Animal Genetics* 6: 1-12.
- Jarvis JP, Scheinfeldt LB, Soi S, Lambert C, Omberg L, Ferwerda B, Froment A, Bodo JM, Beggs W, Hoffman G, Mezey J, Tishkoff SA (2012). Patterns of Ancestry, Signatures of

- Natural Selection, and Genetic Association with Stature in Western African Pygmies. *PLoS Genetics* 8: 641-653.
- Javanmard A, Mohammadabadi MR, Zarrigabayi GE, Gharahedaghi AA, Nassiry MR, Javadmash A, Asadzadeh N (2008). Polymorphism within the intron region of the bovine leptin gene in Iranian Sarabi cattle (Iranian *Bos taurus*). *Russian Journal of Genetics* 44: 495-497.
- Jin CF, Chen YJ, Yang ZQ, Shi K, Chen CK (2015). A genome-wide association study of growth trait-related single nucleotide polymorphisms in Chinese Yancheng chickens. *Genetics and Molecular Research* 14: 15783-15792.
- Kim H, Song KD, Kim HJ, Park W, Kim J, Lee T, Shin DH, Kwak W, Kwon YJ, Sung S Moon S, Lee KT, Kim K, Hong JK, Eo KU, Lee WK, Kim DK, Oh JD, Kim ES, Cho ES, Lee HK, Kim TU, Kim H (2015). Exploring the Genetic Signature of Body Size in Yucatan Miniature Pig. *PLoS One* 4: 1732-1742.
- Kijas JW, Johannes A, Ben Hayes L, Boitard S, PortoNeto 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.
- Kharitononkov A (2009). FGFs and metabolism. *Current Opinion in Pharmacology* 9: 805–810.
- Khodabakhshzadeh R, Mohammadabadi MR, Esmailzadeh Koshkoieh A, Moradi-Shahrehabak H, Ansari Namin S (2015a). Study of mutations available in first-half exon 2 of GDF9 gene in crossbred sheep born from crossing of Romanov rams with Kermani ewes. *Iranian Journal of Animal Science Research* 6: 395-403 (In Farsi).
- Khodabakhshzadeh R, Mohammadabadi MR, Moradi H, Esmailzadeh AK, Ansari NS (2015b). Identify of G→A point mutation at positions 477 and 721 in exon 2 of GDF9 gene in Kermani sheep. *Modern Genetics* 10: 261-268 (In Farsi).
- Khodabakhshzadeh R, Mohammadabadi MR, Moradi-Shahrehabak H, Esmailzadeh Koshkoieh A (2016a). Identification of available mutations in the first-half (from 5' end) of exon 2 of GDF9 gene in crossbred sheep from crossing of Romanov and Lori-Bakhtiari breeds. *Animal Production Research* 4: 15-26 (In Farsi).
- Khodabakhshzadeh R, Mohammadabadi MR, Esmailzadeh Koshkoieh A, Moradi-Shahrehabak H, Bordbar F, Ansari Namin S (2016b). Identification of point mutations in exon 2 of GDF9 gene in Kermani sheep. *Polish Journal of Veterinary Sciences* 19: 281–289.
- Lee SH, Choi BH, Lim D, Gondro C, Cho YM, Dang CG, Sharma A, Jang GW, Lee KT, Yoon D, Lee HK, Yeon H, Yang BS, Kang HS, Hong SK (2013). Genome-Wide Association Study Identifies Major Loci for Carcass Weight on BTA14 in Hanwoo (Korean Cattle). *PLoS One* 8: 677-688.
- Li Y, Kim JJ (2014). Effective Population Size and Signatures of Selection Using Bovine 50K SNP Chips in Korean Native Cattle (Hanwoo). *Evolutionary Bioinformatics* 11: 143-153.
- Liu S, He S, Chen L, Li W, Di J, Liu M (2017). Estimates of linkage disequilibrium and effective population sizes in Chinese Merino (Xinjiang type) sheep by genome-wide SNPs. *Genes Genome* 17: 1-13.
- Mancini G, Gargani M, Chillemi G, Nicolazzi EL, Marsan PA, Valentini A, Pariset, L. (2014). Signatures of selection in five Italian cattle breeds detected by a 54 K SNP panel. *Molecular Biological Reports* 41: 957–965.

- Meadows JR, Chan EK, Kijas JW (2008). Linkage disequilibrium compared between five populations of domestic sheep. *BMC Genetics* 9: 61-73.
- Meuwissen THE, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome wide dense marker maps. *Genetics* 157: 1819-1829.
- Mishra C, Palai TK, Sarangi LN, Prusty BR, Maharana BR (2013). Candidate gene markers for sperm quality and fertility in bulls. *Veterinary World* 6: 905-910.
- Mohammadabadi MR, Sattayimokhtari R (2013). Estimation of (co) variance components of ewe productivity traits in Kermani sheep. *Slovak Journal of Animal Science* 46: 45-51.
- Mohammadi H, Moradi shahrehabak M, Sadeghi M (2011). Study and comparative of polymorphism (T487C) in exon 17 of the DGAT1 gene and its relationship with carcass weight and dressing percentage in the Lori-Bakhtiari and Zel sheep breed. *Iranian Journal of Biotechnology* 1: 77-89 (In Farsi).
- Mousavizadeh A, Mohammadabadi MR, Torabi A, Nassiry MR, Esmailizadeh AK (2009). Genetic polymorphism at the growth hormone locus in Iranian Talli goats by polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP). *Iranian Journal of Biotechnology* 7: 51-53 (In Farsi).
- Niu H, Zhu B, Guo P, Zhang W, Xue J, Chen Y, Li J (2016). Estimation of linkage disequilibrium levels and haplotype block structure in Chinese Simmental and Wagyu beef cattle using high-density genotypes. *Livestock Science* 190: 1-9.
- Olsen HG, Hayes BJ, Kent MP, Nome T, Svendsen M, Lien S (2010). A genome wide association study for QTL affecting direct and maternal effects of stillbirth and dystocia in cattle. *Animal Genetics* 41: 273-280.
- Olivieri BF, Mercadante MEZ, Cyrillo JNDSG, Branco RH, Bonilha SFM, deAlbuquerque LG, Rafael MS, Baldi R (2016). Genomic Regions Associated with Feed Efficiency Indicator Traits in an Experimental Nellore Cattle Population. *PLoS One* 9: 11-23.
- Pritchard JK, Przeworski M (2001). Linkage disequilibrium in humans: models and data. *American Journal of Human genetics* 69: 1-14.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Chang A (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics* 81: 559-575.
- Pitia AM, uchiyama K, Sano H, Kinukawa M, Minato Y, Sasada H, Kohsaka T (2016). Functional insulin-like factor 3 (INSL3) hormone-receptor system in the testes and spermatozoa of domestic ruminants and its potential as a predictor of sire fertility. *Animal Science Journal* 10: 1269-1278.
- Porto-Neto LR, Kijas JW, Reverter A (2014). The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genetics Selection Evaluation* 46: 22-35.
- 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: 318-328.
- Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R, Lander ES (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419: 832-837.
- 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.

- Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* 449: 913–918.
- Saatchi M, Schnabel RD, Taylor JF, Garrick DJ (2014). Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics* 15: 442–449.
- Scheet P, Stephens M (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American Journal of Human Genetics* 78: 629–644.
- Soufy B, Mohammadabadi MR, Shojaeyan K, Baghizadeh A, Ferasaty S, Askari N, Dayani O (2009). Evaluation of Myostatin gene polymorphism in Sanjabi sheep by PCR-RFLP method. *Animal Science Researches* 19: 81–89.
- Teo YY, Fry AE, Clark TG, Tai ES, Seielstad M (2007). On the usage of HWE for identifying genotyping errors. *Annals of Human Genetics* 71: 701–703.
- Tizioto PC, Decker JE, Taylor JF, Schnabel RD, Mudadu, MA, Silva FL, Mourao GB, Coutinho LL, Tholon P, Sonstegard TS, Rosa AN, Alencar MM, Tullio RR, Medeiros SR, Nassu RT, Feijó GL, Silva LO, Torres RA, Siqueira F, Higa RH, Regitano LC (2013). Genome scan for meat quality traits in Nelore beef cattle. *Physiological Genomics* 45: 1012–1020.
- Toral PG, Hervás G, Suarez-Vega A, Arranz JJ, Frutos P (2016). Isolation of RNA from milk somatic cells as an alternative to biopsies of mammary tissue for nutrigenomic studies in dairy ewes. *Journal of Dairy Science* 99: 8461–8471.
- Utsunomiya YT, Perez O'Brien AM, Sonstegard TS, Solkner J, Garcia JF (2015). Genomic data as the “hitchhiker’s guide” to cattle adaptation: tracking the milestones of past selection in the bovine genome. *Frontiers in Genetics* 6: 36–47.
- UniProtKB Gene. [http://www.uniprot.org/help/gene\\_name](http://www.uniprot.org/help/gene_name).
- Wall JD, Pritchard JK (2003). Haplotype blocks and linkage disequilibrium in the human genome. *Nature Reviews Genetics* 4: 587–597.
- Weikard R, Goldammer T, Brunner RM, Kuehn C (2012). Tissue specific mRNA expression patterns reveal a coordinated metabolic response associated with genetic selection for milk production in cows. *Physiological Genomics* 44: 728–739.
- Yang W, Qi Y, Fu J (2016). Genetic signals of high-altitude adaptation in amphibians: a comparative transcriptome analysis. *BMC Genetics* 17: 134–140.
- Zamani P, Akhondi M, Mohammadabadi MR (2015). Associations of Inter-Simple Sequence Repeat loci with predicted breeding values of body weight in sheep. *Small Ruminant Research* 132: 123–127.
- Zhao FP, Wang GK, Zeng T, Wei CH, Zhang L, Wang HH, Zhang SZ, Liu RZ, Liu Z, Du LX (2014). Estimations of genomic linkage disequilibrium and effective population sizes in three sheep populations. *Livestock Science* 170: 22–29.
- Zhao FP, Wei CH, Zhang L, Liu JS, Wang GK, Zeng T, Du LX (2016). A genome scan of recent positive selection signatures in three sheep populations. *Journal of Integrative Agriculture* 15: 162–174.
- Zhang S, Li H, Shi H (2006) Single marker and haplotype analysis of the chicken apolipoprotein B gene T123g and D9500D9- polymorphism reveals association with body growth and obesity. *Poultry Science* 85: 178–184.

## Estimation of linkage disequilibrium and whole-genome scan for detection of loci under selection associated with body weight in Zandi sheep breed

Mohammadi H.<sup>1</sup>, Rafat S.A.<sup>2</sup>, Moradi Shahrabak H.\*<sup>3</sup>, Shodja J.<sup>4</sup>, Moradi M.H.<sup>5</sup>

<sup>1</sup> Ph.D Graduated Animal Science Department, Faculty of Agriculture, Tabriz University, Tabriz, Iran.

<sup>2</sup> Professor Animal Science Department, Faculty of Agriculture, Tabriz University, Tabriz, Iran.

<sup>3</sup> Assistant Professor Animal Science Department, University College of Agriculture and Natural Resources, Tehran University, Tehran, Iran.

<sup>4</sup> Professor Animal Science Department, Faculty of Agriculture, Tabriz University, Tabriz, Iran.

<sup>5</sup> Assistant Professor Animal Science Department, Faculty of Agriculture, Arak University, Arak, Iran.

### Abstract

Body weight is the most economically important trait in sheep industry. In genome-wide association study and genomic selection, determining of extent and level of linkage disequilibrium (LD) is critical in sample size and marker density. Moreover, selection for increasing of frequency in new mutations that are advantageous only in a subset of populations leaves some signatures in the genome. Locations of selection signatures are often correlated with genes and QTLs affecting economically important traits. Therefore, the objectives of this research were to study LD pattern and identify the genomic regions that have been under artificial and natural selection in Zandi sheep breeds with different body weight. For this purpose, the blood samples were collected from 73 Zandi sheep and 54241 markers were genotyped by using Illumina Ovine SNP50K BeadChip array based on latest assembly (Oar\_4.0) sheep genome. After quality control, 40879 SNPs belonging to 71 animals were used in the final analysis. LD was calculated between all pairs were calculated with  $r^2$  by PLINK software. To detect the selection sweep, due to linkage disequilibrium, associated with these signatures we carried out the cross-population Extended Haplotype Homozygosity (XP-EHH) test. In this study, the extent of LD in this study was 40 kb with  $r^2=0.2$ . The results revealed ten genomic regions on 1, 2, 3, 5, 6, 9, 12, 13, 21 and 22 chromosomes. Bioinformatics analysis demonstrated that some of these genomic regions overlapped with reported genes included in the growth traits, fat metabolism, development of the skeletal system, energy metabolism and meat quality traits such as FBP1, FAM184B, PKD2, FGF19, SPP1, MEPE CAPN2 genes.

**Keywords:** *Linkage Disequilibrium, Selection Signatures; XP-EHH Test; Zandi sheep; Body Weight.*

\* Corresponding author: Moradi Shahrabak H.

Tel:09133915306

Email: hmoradis@ut.ac.ir