

## پویش ژنومی نشانه‌های انتخاب در گاو‌میش‌های خوزستانی و مازندرانی

مهدی مخبر<sup>۱</sup>، محمد مرادی شهربابک<sup>۲\*</sup>، مصطفی صادقی<sup>۳</sup>، حسین مرادی شهربابک<sup>۴</sup> و جان ویلیامز<sup>۵</sup>

۱، ۲، ۳ و ۴. دانشجوی دکتری ژنتیک و اصلاح نژاد، استاد و استادیاران گروه مهندسی علوم دامی،

پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج

۵. استاد و مدیر بخش ژنومیک مرکز تحقیقات Padano، ایتالیا

(تاریخ دریافت: ۱۳۹۳/۱۱/۲۶ - تاریخ تصویب: ۱۳۹۴/۴/۲)

## چکیده

شناسایی مناطق ژنومی که هدف انتخاب بوده‌اند، از بحث‌برانگیزترین مباحث در حوزه تحقیقات ژنتیک حیوانی، به‌خصوص در حیوانات اهلی است. در این مطالعه، پویش کل ژنوم برای شناسایی مناطقی از ژنوم که در گاو‌میش‌های خوزستانی و مازندرانی هدف انتخاب‌های طبیعی یا مصنوعی قرار گرفته‌اند، اجرا شده است. بدین منظور ۱۴۸ رأس گاو‌میش رودخانه‌ای شامل جمعیت‌های خوزستانی (۱۲۱ رأس) و مازندرانی (۲۷ رأس)، به وسیله آرایه‌های ژنومیک Axiom® Buffalo Genotyping 90K تعیین ژنوتیپ شدند. جهت جستجوی نشانه‌های انتخاب از برآوردگر ناریب  $F_{ST}$  ( $\theta$ ) استفاده شد. در مجموع ۲۳ منطقه که نشانگرهای SNP آن‌ها بالاتر از ۰/۱ درصد حد بالای توزیع تجربی  $F_{ST}$  بود، به‌عنوان نشانه‌های انتخاب شناسایی شده و مورد بررسی‌های بیشتر قرار گرفتند. بعد از انطباق مناطق ژنومی انتخاب‌شده با مناطق ژنومی متناظر آن روی ژنوم گاو (UMD3.1 *Bos Taurus* Genome)، ۶۴ ژن و ۲۷ QTL شناسایی شد. تعدادی از این ژن‌ها در مطالعات قبلی نیز به‌عنوان نشانه‌های انتخاب گزارش شده‌اند. برخی از این ژن‌ها احتمالاً در مسیرهای بیولوژیکی که با اهلی شدن حیوانات (توسعه مغز و عملکردهای رفتاری، پیگماسیون و آداپته شدن به محیط زندگی و شرایط جغرافیایی) و تولید شیر مرتبط می‌باشند، درگیرند. بررسی‌ها همچنین نشان داد که QTL‌های شناسایی شده در این تحقیق با صفات مهم اقتصادی از جمله صفات مرتبط با راندمان تبدیل غذایی، وزن بدن، چربی زیرپوستی، افزایش وزن روزانه، تیپ، تردی گوشت، ترکیبات شیر، اتصالات پستانی، اندازه گوساله و گوساله‌زایی آسان ارتباط دارند. به هر حال جهت شناسایی نقش دقیق این ژن‌ها و QTL‌ها لازم است مطالعات پیوستگی و عملکردی بیشتری انجام گیرد.

**واژه‌های کلیدی:** پویش ژنوم، شاخص تمایز جمعیتی، گاو‌میش خوزستانی و مازندرانی، نشانه‌های انتخاب.

## مقدمه

قبیل اندازه مؤثر جمعیت و تمایز جمعیت‌ها با استفاده از روش‌های آنالیز ژنومیک جمعیت استنباط گردید. یکی از جنبه‌های مهم و کاربردی آنالیزهای مربوط به ژنومیک جمعیت شناسایی نشانه‌های انتخاب است. نشانه‌های انتخاب در واقع الگوهایی روی ژنوم هستند

عبارت ژنومیک جمعیت اولین بار در Gulcher & Stefansson (1998) به کار رفت و برای آنالیز تنوع جمعیتی درون و بین جمعیت‌ها بر اساس داده‌های کل ژنوم استفاده شد و خصوصیات خاص جمعیت‌ها از

دسته‌بندی‌های مختلفی برای روش شناسایی وجود دارد. در حالت کلی می‌توان این روش‌ها را در ۵ دسته مجزا شامل آزمون‌های آماری بر پایه تمایز جمعیت‌ها (Akey *et al.*, 2002)، آزمون‌های آماری مبتنی بر عدم تعادل لینکاژی و طول هاپلوتایپی (Sabeti *et al.*, 2002)، آزمون‌های آماری مبتنی بر توزیع فراوانی آللی (Tajima, 1989)، آزمون‌های آماری مبتنی بر کاهش مکانی تنوع ژنتیکی (Rubin *et al.*, 2010) و آزمون‌های آماری مبتنی بر جهش‌های تغییردهنده عملکرد (Nielsen & Yang, 1988) قرار داد. همچنین آزمون‌هایی ارائه شده‌اند که ترکیبی از روش‌های جستجوی نشانه‌های انتخاب هستند. نتایج بررسی‌ها نشان‌دهنده توانایی و حساسیت بیشتر روش‌های ترکیبی در شناسایی نشانه‌های انتخاب است (Grossman *et al.*, 2010). علاوه بر روش‌های ترکیبی، روش‌های دیگری نیز جهت افزایش دقت جستجوی نشانه‌های انتخاب در یک‌سری از مطالعات استفاده شده است که از این دسته می‌توان به ترکیب اطلاعات جستجوی نشانه‌های انتخاب با مطالعات پیوستگی ژنوم (GWAS)، پویش تکراری جمعیت (Scanning replicated population) و آنالیز عملکردی (Functional analysis) اشاره کرد (Oleksyk *et al.*, 2010).

در دهه اول قرن ۲۱ با تکمیل پروژه‌های توالی‌یابی DNA انسان و بسیاری از گونه‌های دیگر، شناسایی نشانه‌های انتخاب در سطح ژنوم نیز وارد مرحله جدیدی شده و تحقیقات زیادی در این زمینه در گونه‌های مختلف طراحی شده است و در برخی موارد ژن‌های موجود در مناطق مختلف ژنومی شناسایی و حتی عملکرد بیولوژیکی آن‌ها نیز تعیین شده است. در بررسی نشانه‌های انتخاب در سطح ژنوم، تمام ژنوم و ژن‌های موجود در آن‌ها، همزمان بررسی می‌شود. بنابراین دیگر نیازی به اطلاعات قبلی از عملکرد تک ژن‌های مورد بررسی نیست (Sabeti *et al.*, 2002; Akey, 2002). همچنین در این روش می‌توان تا حدود زیادی آریب ناشی از عوامل دموگرافیک مثل دریافت ژنتیکی، مهاجرت و غیره را کنترل کرد. بنابراین بررسی همزمان جایگاه‌های مختلف با استفاده از

که در اثر نیروهای انتخاب روی نواحی خاصی از ژنوم تشکیل می‌شوند (Simianer *et al.*, 2014). در سال‌های اخیر با توجه به دسترسی وسیع به داده‌های چند شکلی مولکولی، شناسایی الگوهای درون ژنومی اهمیت یافته است، زیرا این الگوها نشان‌دهنده انتخاب مثبت درون جمعیت‌هاست. دیدگاه کلیدی در این خصوص پدیده انتقال همراه (Genetic Hitchhiking) است (Maynard-Smith & Haigh, 1974). در نتیجه این پدیده، الگوی تنوع ژنتیکی و عدم تعادل لینکاژی در اطراف جایگاه مورد انتخاب تغییر خواهد کرد و هرچه به آلل مورد انتخاب نزدیک‌تر شویم، تنوع ژنتیکی در اطراف آن کاهش و عدم تعادل لینکاژی افزایش می‌یابد. در نهایت ممکن است هاپلوتایپ خاص که حاوی چندین ژن است تنها هاپلوتایپ یا مهم‌ترین هاپلوتایپ موجود در جمعیت باشد. به کاهش یا حذف تنوع در جایگاه‌های مجاور یک آلل جهش‌یافته سودمند طی انتخاب مثبت Selective Sweep گفته می‌شود (Fan *et al.*, 2010; Akey *et al.*, 2009). بنابراین هنگامی که یک آلل سودمند در زمان‌های مختلف هدف انتخاب مثبت قرار می‌گیرد، باعث ایجاد نشانه‌هایی در سطح ژنوم می‌گردد (Biswas & Akey, 2006). این نشانه‌ها شامل کاهش مکانی تنوع ژنتیکی، انحراف از توزیع فراوانی آللی (Site Frequency Spectrum) و افزایش عدم تعادل لینکاژی (Linkage Disequilibrium) و توسعه ساختار هاپلوتایپی (Haplotype Structure) هستند (Qanbari *et al.*, 2014). در حال حاضر ابزارها و روش‌های مختلفی برای شناسایی این نشانه‌های انتخاب در سطح ژنوم جمعیت‌های مختلف، به‌طور موفقیت‌آمیز توسعه یافته و استفاده می‌گردد (Li *et al.*, 2014). با اینکه روش‌های موجود برای شناسایی نشانه‌های انتخاب روی داده‌های ژنومیک در فرض‌ها و اجرا متفاوت هستند و انتظار می‌رود رابطه ضعیفی بین نتایج آن‌ها وجود داشته باشد، مناطقی که با روش‌های مختلف به‌عنوان نشانه‌های انتخاب شناسایی می‌شوند، بر اساس شواهد قطعی نشان‌دهنده وجود انتخاب در این نواحی هستند (Utsunomiya *et al.*, 2013). بر اساس مکانیسم عملکرد شناسایی نشانه‌های انتخاب،

آرایه‌های SNP امکان شناسایی مناطقی از ژنوم را که تنها تحت تأثیر انتخاب قرار گرفته‌اند، فراهم می‌آورد (Akey, 2009; MacEachern *et al.*, 2009).

با توجه به اینکه پویش ژنوم در حد وسیع در انسان انجام گرفته، پژوهش‌های فراوان دیگری در این خصوص روی انسان اجرا شده و گزارش‌های آن‌ها در دسترس است. انواع روش‌ها و فناوری‌هایی که همزمان با مراحل تکمیل پروژه ژنوم انسان و بعد از آن توسعه پیدا کرده‌اند، به‌طور کامل وارد حوزه مطالعات ژنومیک حیوانات اهلی شده است (Fan *et al.*, 2010; Goddard, 2009). از آن رو که حیوانات اهلی علاوه بر انتخاب طبیعی، به‌خصوص در سال‌های اخیر در معرض انتخاب‌های مصنوعی شدیدی نیز قرار گرفته‌اند، نمونه‌های مناسب‌تری برای شناسایی سیگنال‌های انتخاب در سطح ژنوم هستند و از آن‌ها برای شناسایی ژن‌های کنترل‌کننده فنوتیپ‌های مختلف استفاده می‌گردد (Biswas & Akey, 2006).

اولین گزارش در خصوص استفاده از آرایه‌های ژنومیکی روی گاو میش، به مطالعه و پویش ژنوم گاو میش با استفاده از آرایه‌های ژنومیکی گاو (Illumina Bovine SNP50K BeadChip) برمی‌گردد که توسط Michelizzi *et al.* (2011) انجام گرفته است. این مطالعه به منظور بررسی کارایی استفاده از آرایه‌های طراحی‌شده برای گاو در پویش ژنوم گاو میش صورت گرفت. این آرایه دارای ۵۴۰۰۱ SNP است و ۱۰ گاو میش آبی با این آرایه‌ها تعیین ژنوتیپ شدند. نتایج آن تحقیق نشان داد که ۴۳۶۸۷ جایگاه از مجموع کل SNP‌های روی آرایه، با جایگاه‌های گاو میش هم‌خوانی داشت و در روی ژنوم گاو میش وجود داشت ولی از این تعداد فقط ۱۱۵۹ جایگاه حالت چند شکلی نشان داد. این نتیجه گویای این مطلب است که با وجود اینکه این جایگاه‌ها چند شکل نیستند ولی در بین این دو گونه حفظ شده‌اند. اولین گزارش مربوط به مطالعات پویش ژنوم گاو میش با آرایه‌های اختصاصی طراحی‌شده برای گاو میش توسط *Iamartino et al.* (2013) ارائه شد. این مطالعه در واقع گزارش اولیه‌ای از کیفیت ژنوتایپینگ با استفاده از آرایه‌های ژنومیکی مربوط به گاو میش (Axiom®)

همچنین مطالعات پیوستگی کل ژنوم روی ۹۰۱ گاو میش از نژادهای مدیترانه‌ای (Mediterranean) ایتالیا و مورای (Murrah) برزیل انجام گرفت. در این مطالعه کارایی آرایه‌های ژنومیکی توسط این محققان مثبت ارزیابی شد و همچنین مطالعات پیوستگی کل ژنوم وجود ۹ جایگاه را که با صفات مربوط به تولید شیر ارتباط داشتند، تأیید کرد (Iamartino *et al.*, 2013).

گاو میش به خانواده گاوها (*Bovinae*) متعلق است و شباهت خیلی زیادی با گاو دارد (Khan *et al.*, 2012). این دام از نظر تولید شیر و گوشت دارای خصوصیات ممتازی است. گاو میش در مناطقی که گاو قادر به زندگی در آن نیست، به خوبی سازش یافته و به تولید خود ادامه می‌دهد (Borghese, 2011). گاو میش رودخانه‌ای از نظر اقتصادی گونه‌ای اهلی با اهمیتی در بیشتر کشورهای آسیایی و مدیترانه‌ای است. بهبود ژنتیکی آن به‌ویژه در عملکرد تولیدمثلی و کمیت تولید گوشت و شیر، به رتبه‌های بالایی در میان تحقیقات وابسته به کشاورزی در این کشورها نیاز دارد (Othman, 2006).

هدف از مطالعه حاضر استفاده از آرایه‌های ژنومیکی گاو میش (Axiom® Buffalo Genotyping 90K Array) جهت تعیین ژنوتیپ نژادهای گاو میش خوزستانی و مازندرانی و شناسایی مناطقی از ژنوم بود که در سال‌های متمادی تحت تأثیر نیروهای انتخاب طبیعی یا مصنوعی قرار گرفته‌اند. این مناطق ژنومی می‌توانند جزو مناطق ژنومی متمایزکننده نژادهای مذکور باشند. نتایج این بررسی می‌تواند در تحلیل تفاوت‌های بین دو نژاد و برنامه‌ریزی‌های اصلاح نژادی مفید باشد. همچنین مطالعه حاضر اولین مطالعه انجام‌گرفته روی ژنوم گاو میش جهت شناسایی نشانه‌های انتخابی است و اهمیت علمی فراوان دارد.

### مواد و روش‌ها

در مطالعه حاضر از ۱۴۸ نمونه (۱۲۱ نمونه مو مربوط به جمعیت گاو میش خوزستانی از شهرستان‌های اهواز، شوش، شوشتر، سوسنگرد و قصر شیرین و ۲۷ نمونه خون مربوط به جمعیت گاو میش مازندرانی از میانکاله

کل ۱۴۶ حیوان از ۱۴۸ حیوان مراحل کنترل کیفیت از Affypipe را گذراندند و ۲ حیوان به دلیل DQC کمتر از ۰/۸۲ و CR کمتر از ۰/۹۷ حذف شدند. در مجموع از بین ۸۹۸۶۴ جایگاه بررسی شده، ۷۳۹۳۵ جایگاه که ژنوتیپ آن‌ها به لحاظ کیفی مناسب بوده و در گروه‌بندی‌های ژنوتیپی تعریف شده در نرم‌افزار Affypipe قرار گرفته بودند، انتخاب شدند و خروجی آن در فرمت ورودی نرم‌افزار Plink ذخیره شد. برای اطمینان از بالابودن کیفیت داده‌های مورد استفاده در آنالیزهای بعدی، با اعمال فیلترهایی داده‌های با کیفیت پایین و داده‌های دارای اطلاعات کم حذف شدند. این بخش با نرم‌افزار Plink انجام گرفت (Purcell et al., 2007). داده‌های با اطلاعات کم شامل ژنوتیپ‌های دارای حداقل فراوانی آلی (MAF) کمتر از ۰/۰۲ و جایگاه‌های با بیش از ۰/۰۵ ژنوتیپ از دست‌رفته (CR<sub>SNP</sub>) هستند. سپس SNP‌هایی که p-value تعادل هاردی-واینبرگ (HWE) برای آن‌ها کمتر از سطح معناداری بود، کنار گذاشته شدند. سطح معناداری به‌عنوان معیاری از خطای تعیین ژنوتیپ (Teo et al., 2007) در نظر گرفته می‌شود. سطح معناداری در این تحقیق با استفاده از تصحیح بنفرونی ( $\beta = \alpha/n$ )،  $6/8 \times 10^{-7}$  تعیین شد. در این رابطه، n تعداد آزمون است که در اینجا برابر تعداد SNP (n=۷۳۹۳۵) و  $\alpha$  خطای آزمایش است که در این بررسی ۰/۰۵ در نظر گرفته شده بود. کنترل‌های کیفی داده‌های ژنوتیپی جداگانه روی جمعیت‌های مورد مقایسه اعمال شد (جدول ۱). سپس جهت داشتن دیدگاهی کلی در خصوص ساختار جمعیتی حیوانات و نژادهای مورد مطالعه و شناسایی حیواناتی که خارج از گروه نژادی خود قرار گرفته‌اند، آنالیز مؤلفه‌های اصلی (Principal Component Analysis) با استفاده از دستور prcomp در محیط R صورت گرفت. در انتها و قبل از آنالیزهای شناسایی نشانه‌های انتخاب، جمعیت‌های مورد مطالعه برای خویشاوندی‌های نزدیک با تعیین IBS (Identical by State) یا به عبارت دیگر ضرایب خویشاوندی ژنومیکی که توسط Leutenegger et al. (2003) ارائه شده، غربال شدند. این بخش در محیط R و با استفاده از پکیج GenABEL صورت گرفت و

شهرستان بهشهر) از گله‌هایی که تحت سیستم ثبت شجره و رکوردگیری مرکز اصلاح نژاد کشور قرار داشتند، جمع‌آوری شد. در مجموع ۸ رأس از نمونه‌ها شامل ۷ رأس خوزستانی و ۱ رأس مازندرانی، نر بودند. نمونه‌های مو به همراه ریشه مو از بخش‌هایی از بدن دام که آلودگی کمتری داشت، به دست آمد (Campbell et al., 1997) و نمونه‌های خون از وداچ زیر دمی گاو میش تهیه گردید. استخراج DNA از مو طبق روش Alberts et al. (2010) با اندکی تغییر انجام گرفت. DNA نیز با استفاده از روش بهینه نمکی (Grimberg et al., 1989) از خون استخراج و در ۷۰ میکرولیتر بافر TE حل شد. تعیین کمیت و کیفیت DNA استخراجی به وسیله دستگاه نانودراپ (Thermo 2000C) و ژل آگارز ۱/۲ درصد صورت گرفت. میانگین مقدار و خلوص DNA استخراج شده از مو به ترتیب ۲۷/۶۷ میکروگرم و ۱/۸۷ (۲/۰۷-۱/۶۵) بود، این مقادیر برای خون ۱۷/۸۶ میکروگرم و ۱/۸۳ (۱/۶۳-۱/۹۸) گزارش شد. سپس نمونه‌ها برای مراحل بعدی توالی‌یابی به آزمایشگاه ژنومیک مرکز تحقیقات پادانو (Parco Tecnologico Padano) در کشور ایتالیا منتقل شدند. در آزمایشگاه مرکز تحقیقات پادانو، قبل از تعیین ژنوتیپ، دوباره کمیت نمونه‌ها بررسی شد. این قسمت از کار با استفاده از Quant-it dsDNA Assay Kit شرکت Life Technology صورت گرفت. در این روش کمیت DNA دو رشته‌ای با مقدار نور فلورسنت ساطع شده از نمونه‌ها تعیین شد. برای خواندن کمیت فلورسنت از دستگاه Spectrophotometer-fluorometer Tecan Genios Reader و نرم‌افزار Magellan استفاده شد. بعد از تعیین کمیت DNA، نمونه‌ها در غلظت ۵۰ (ng/ul) رقیق‌سازی شدند. تعیین ژنوتیپ نمونه‌ها با استفاده از آرایه‌های Axiom® Buffalo Genotyping 90K مربوط به شرکت افی‌متریکس (Affymetrix) صورت گرفت. کیفیت و فیلتراسیون اولیه داده‌های خام توسط نرم‌افزار Affypipe تعیین شد (Nicolazzi et al., 2014). اساس فیلتراسیون در این مرحله، اطلاعات به‌دست‌آمده از کیفیت داده‌های خام بوده و حیوانات بر اساس معیارهای DQC (Dish Quality Control) و CR (Call rate) انتخاب شدند. در

فقط مناطقی که بیشترین تمایز جمعیتی را بین دو نژاد نشان دادند، انتخاب شدند. جهت بررسی بصری نشانه‌های انتخاب، گراف منهن پلات به وسیله بسته نرم‌افزاری QQman در محیط R رسم شد (شکل ۲). در مرحله بعد اطلاعات مربوط به این نواحی (شامل کروموزوم و موقعیت SNPهای انتخابی بر روی کروموزوم) جهت شناسایی ژن‌هایی که احتمالاً با نواحی انتخابی مرتبط هستند، تعیین شدند. موقعیت کروموزومی این جایگاه‌ها با موقعیت ژن‌های لیست‌شده برای ژنوم گاو Ensemble Biomart Tool در (UMD3.1 *Bos Taurus*) (EntrezGene <http://www.ensembl.org/biomart>) تعیین شدند. همچنین این جایگاه‌ها با موقعیت QTLهای موجود برای ژنوم گاو (UMD3.1 *Bos Taurus*) در پایگاه اطلاعاتی آنلاین QTLهای گاو (<http://www.animalgenome.org/cgi-bin/QTLdb>) (BT/browse) مقایسه شدند. در نهایت ژن‌ها و QTLهای مرتبط با مناطق انتخابی به دست آمد و مورد بررسی‌های بیشتر قرار گرفت.

### نتایج و بحث

پس از مراحل مختلف غربالگری بر اساس اطلاعات کنترل کیفی داده‌های خام و خویشاوندی حیوانات، در نهایت ۶۱۱۹۳ نشانگر SNP مربوط به ۱۳۴ حیوان (۱۱۱ گاومیش خوزستانی و ۲۳ گاومیش مازندرانی) برای آنالیزهای بعدی انتخاب شدند (جدول ۱).

۱۰ رأس حیوان به دلیل داشتن IBS بالای ۰/۵ از آنالیزهای بعدی حذف شدند (جدول ۱). البته در شرایطی که بر طبق گفته P'erez O'Brien *et al.* (2014)، حد آستانه بیشتر از ۰/۸ در نظر گرفته شود، هیچ حیوانی حذف نمی‌شود.

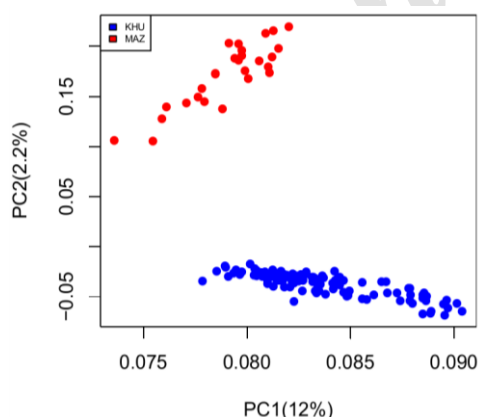
جهت بررسی الگوی ژنومیکی انتخاب مثبت در دو جمعیت گاومیش ایرانی، ارزش‌های  $F_{ST}$  برای هر SNP به روش ناریب تتا و در محیط R محاسبه گردید (Weir & Cockerham, 1984). مزیت این روش به روش پایه‌ای  $F_{ST}$  که توسط رایت ارائه شده، این است که در روش تتا اندازه نمونه‌ها در فرمول لحاظ گردیده و در واقع خطای نمونه‌گیری در نظر گرفته شده است (Weir & Cockerham, 1984). با توجه به اینکه این روش ناریب است، احتمال به‌دست‌آمدن ارزش‌های منفی نیز وجود دارد (Akey *et al.*, 2002). ارزش‌های  $F_{ST}$  (تتا) برای هر SNP و جداگانه برای کروموزوم‌های اتوزومی و کروموزوم X محاسبه شد. جهت شناسایی بهتر سیگنال‌های انتخاب در سطح ژنوم به جای در نظر گرفتن ارزش عددی هر SNP از میانگین‌گیری ارزش‌های عددی SNPهای مجاور به روش Creeping Window (CW) با طول ۳۰۰ کیلو جفت باز استفاده شد (Qanbari *et al.*, 2012) و تنها ۰/۱ درصد مناطقی از ژنوم که در آن تمام نشانگرهای مجاور ارزش بالایی داشتند، به‌عنوان نشانه‌های انتخاب شناسایی و تعیین شدند. حد آستانه در نظر گرفته‌شده سخت‌گیرانه بود و

جدول ۱. مراحل مختلف کنترل کیفیت داده‌های حاصل از تعیین ژنوتیپ گاومیش خوزستانی و مازندرانی

تعداد کل حیوانات برای هر نژاد	نژاد مازندرانی	نژاد خوزستانی	مجموع بین دو نژاد
۲۷	۱۱۹	۱۴۶	
۷۳۹۳۵	۷۳۹۳۵	۷۳۹۳۵	
۰/۹۹۷۳	۰/۹۹۵۴	-	
۲	۸	۱۰	
۲۳	۱۱۱	۱۳۴	
۲	۰	۲	
۱۸	۱۷	-	
۱۱۳۵۳	۹۴۲۸	-	
۲	۱۱۹	-	
۶۳۳	۲۴۳	-	
۶۱۹۶۱	۶۴۱۳۹	۶۱۱۹۳	

۰/۰۵۹ و در کروموزوم جنسی X، ۰/۰۲۹ با انحراف استاندارد ۰/۰۴۱ بود. اینکه این میانگین‌ها در کروموزوم‌های اتوزومی و جنسی به هم نزدیک هستند و نبود تفاوت معنادار بین آن‌ها، می‌تواند به سبب تعداد کم نرها در این بررسی باشد. همچنین این تفاوت اندک می‌تواند به علت استفاده کمتر از SNP‌هایی باشد که در بخش‌های غیرهمولوگ کروموزوم جنسی Y است.

جهت شناسایی بهتر سیگنال‌های انتخاب در سطح ژنوم به جای در نظر گرفتن ارزش عددی هر SNP، میانگین‌گیری ارزش‌های عددی SNP‌های مجاور به روش CW با طول ۱۰ SNP صورت گرفت و تنها ۰/۱ درصد مناطقی از ژنوم که در آن تمام نشانگرهای مجاور دارای ارزش بالایی باشند، به‌عنوان نشانه‌های انتخاب شناسایی و تعیین شدند (شکل ۲). حد آستانه در نظر گرفته‌شده سخت‌گیرانه بود و فقط مناطقی که بیشترین تمایز جمعیتی را در بین دو نژاد نشان می‌دهند، انتخاب شدند. این حد آستانه در مطالعات مختلف از ۰/۰۱ درصد (Moradi et al., 2012) تا ۲/۵ درصد (Qanbari et al., 2012) متغیر بوده است. در نهایت ۲۳ ناحیه روی ژنوم برای بررسی‌های بعدی انتخاب شدند که حد آستانه را رد کرده بودند.



شکل ۱. آنالیز PCA مربوط به دو نژاد گاومیش خوزستانی (KHU) و مازندرانی (MAZ) با استفاده از اطلاعات نشانگری SNP از سطح ژنوم

از آنجا که حاشیه‌نویسی (Annotated) ژنوم گاومیش هنوز کامل نشده و توالی‌اش با توالی ژنوم گاو

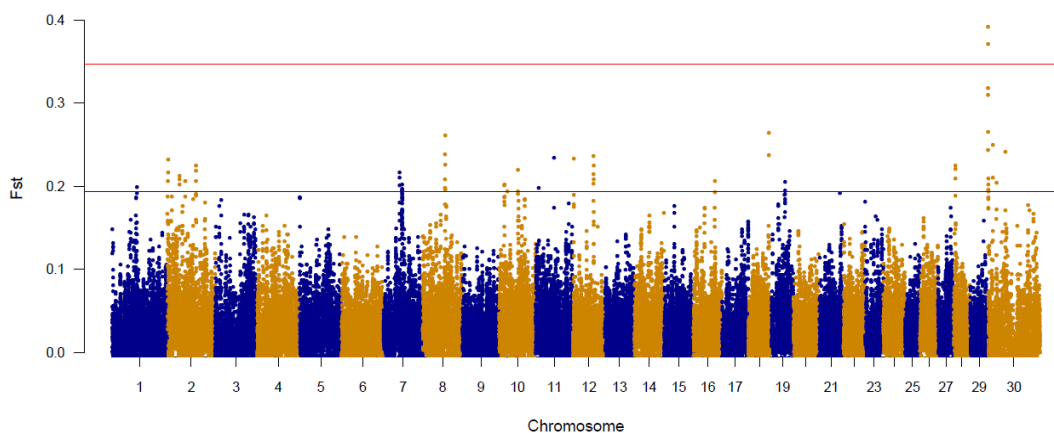
با این تعداد نشانگر SNP در مجموع Mbp ۲۶۴۹/۴۱ از ژنوم با میانگین فاصله ۴۳/۳ Kbp تحت پوشش قرار گرفت. در این مطالعه، ساختار جمعیتی نژادهای گاومیش خوزستانی و مازندرانی کشور بررسی شد. نتایج آنالیز PCA بر اساس PC1 و PC2 نشان داد که این دو نژاد در گروه‌های کاملاً مجزا قرار دارند و در هیچ نقطه‌ای هم‌پوشانی ندارند. این دو PCA در مجموع حدود ۱۴/۲ درصد واریانس (۱۲ درصد برای ویژه بردار یک و ۲/۲ درصد برای ویژه بردار دو) را توجیه می‌کنند (شکل ۱). در گونه‌های اهلی دیگر از قبیل گاو و گوسفند و خوک نیز، حیوانات فقط بر اساس ویژه بردارهای یک و دو بر اساس نژاد و منطقه جغرافیایی دسته‌بندی شده‌اند (Yang et al., 2014; The Bovine HapMap Consortium et al., 2009; Kijas et al., 2012). با اینکه پایین بودن مقدار واریانس توجیهی نشان‌دهنده ساختارهای جمعیتی کمتر تمایز یافته است ولی این جمعیت‌ها همان‌طور که گراف نشان می‌دهد، کاملاً از هم تفکیک شده‌اند. بخشی از واریانس پایین می‌تواند به خاطر تنوع بالای نمونه‌گیری (۵۱ محل نمونه‌برداری) باشد. در مطالعه‌ای که توسط Kijas et al. (2012) روی جمعیت‌های گوسفند دنیا صورت پذیرفت، ۲۰ تا PC اول در مجموع فقط ۱۶ درصد واریانس را توجیه کردند که بزرگ‌ترین مقدار مربوط به PC1 به مقدار ۲/۹۸ درصد بود و PC2 و PC3 به ترتیب ۱/۴۴ و ۱/۱۹ درصد از واریانس توجیه کردند. در مطالعه Barendse et al. (2009) واریانس توجیهی برای PC1 و PC2 به ترتیب برابر ۶/۱ و ۳/۷ درصد بود. با همین مقدار تفرق جمعیتی می‌توانیم امیدوار به یافتن نقاطی از ژنوم باشیم که به‌طور معنادار انتخاب شده‌اند.

با توجه به اینکه SNP‌های موجود روی کروموزوم X در حیوانات نر هموزیگوت است، می‌تواند باعث ایجاد تفاوت در ارزش عددی میانگین  $F_{ST}$  شود. ولی در این بررسی ارزش‌های عددی میانگین‌های  $F_{ST}$  تا در کروموزوم‌های اتوزومی و جنسی به هم نزدیک بود و تفاوت معناداری بین آن‌ها مشاهده نشد. میانگین ارزش‌های  $F_{ST}$  برای هر SNP به روش ناریب تا در کروموزوم‌های اتوزومی ۰/۰۳۲ با انحراف استاندارد

مرتبط با مناطق انتخابی به وسیله Biomart ([www.ensembl.org/biomart/martview](http://www.ensembl.org/biomart/martview)) بررسی شد. از ۲۳ منطقه مورد بررسی تنها ۱۳ منطقه حاوی مناطق کدکننده ژن‌ها بودند. در مجموع ۵۶ ژن که در ارتباط با منطقه انتخابی بودند از این طریق تعیین شدند. اطلاعات مربوط به ۵۰ مورد از ژن‌ها در جدول ۲ آورده شده است. غیر از این ۵۰ ژن ۱۴ پروتئین نامشخص (Uncharacterized protein) دیگر نیز در منطقه انتخابی روی کروموزوم ۷ (۵۳۶۱۳۷۴۷-۵۴۱۳۳۳۷۲ جفت بازی) شناسایی شدند که همه این موارد جزو گروه پروتئینی پروتوکادهرین هستند (جدول ۲). در این بخش برخی از ژن‌ها که احتمالاً تحت انتخاب بوده‌اند مورد بحث و تحلیل قرار می‌گیرند. تعدادی از این ژن‌ها که در این بررسی به‌عنوان نشانه‌های انتخاب گزارش شده‌اند، با مطالعات قبلی هم‌خوانی داشت (Bernard *et al.*, 2007; Ogorevc *et al.*, 2009; Utsunomiya *et al.*, 2013). برخی از این ژن‌ها احتمالاً در مسیرهای بیولوژیکی دخالت دارند که با اهلی‌شدن حیوانات و سازگار شدن با محیط زندگی و شرایط جغرافیایی مرتبط هستند. از مسیرهای بیولوژیکی‌ای که با اهلی‌شدن حیوانات در ارتباط بوده است، می‌توان به توسعه مغز و عملکردهای رفتاری، ادراکات حسی، پیگماسیون (تجمع رنگدانه‌ها در بافت‌ها)، سیستم ایمنی و سیستم انعقاد خون، اشاره کرد (Qanbari *et al.*, 2014).

هم ردیف شده است، بنابراین مشخصات توالی ژنومیکی گاو میش بر اساس اطلاعات توالی ژنوم گاو ارائه شده که از جمله مشخصات جایگاه نشانگر SNP مربوط به گاو میش بر اساس موقعیت‌های ژنومیکی کروموزوم‌های گاو است و نتایج به‌دست‌آمده بر اساس ترتیب ژنوم گاو بود. به دلایل عنوان‌شده استخراج اطلاعات ژن‌هایی که در مناطق انتخابی قرار دارند، آسان‌تر شده است. این هم‌ردیف‌سازی و حاشیه‌نویسی ژنوم گاو میش بر اساس اطلاعات ژنوم گاو به این دلیل بود که ژنوم گاو و گاو میش برخلاف تفاوت در تعداد کروموزوم‌ها (۳۰ جفت کروموزوم گاو در مقایسه با ۲۵ جفت کروموزوم گاو میش آبی)، باهم همولوگ هستند. همچنین ترتیب ژن‌ها در هر دو گونه با هم منطبق بوده و به لحاظ سیتوژنتیکی کل ژنوم دو گونه با در نظر گرفتن بازوهای کروموزومی به جای کروموزوم کامل، در کنارهم قرار می‌گیرد (Amaral *et al.*, 2007). ۲۳ منطقه که  $F_{ST}$  ارزش عددی بالاتر از ۰/۱۹ را داشتند به ترتیب روی کروموزوم‌های ۱، ۲ (منطقه ۶)، ۷ (منطقه ۳)، ۸، ۱۰ (منطقه ۳)، ۱۱، ۱۲ (منطقه ۲)، ۱۶، ۱۸، ۱۹، ۲۱، ۲۸ و X قرار دارند (شکل ۲). جهت شناسایی ژن‌های مرتبط با بخش‌های انتخابی ژنوم، مناطق ارتولوگوس با مناطق انتخابی و ۲۰۰ کیلو جفت باز قبل و بعد از نقاط شروع و پایان توالی (۱۲/۳۸ مگا جفت باز)

KHMZ Manhattan plot of Fst



شکل ۲. تصویر منتهن پلات مربوط به ارزش‌های میانگین گیری شده  $F_{ST}$  تنا با طول پنجره ۳۰۰ کیلو جفت بازی مقایسه بین دو نژاد گاو میش خوزستانی و مازندرانی را نشان می‌دهد. خط آبی حد آستانه نشانه‌های انتخاب روی کروموزوم‌های بدنی و خط قرمز حد آستانه نشانه‌های انتخاب روی کروموزوم جنسی X (۳۰) را نشان می‌دهد. حد آستانه در این مطالعه بر اساس توزیع تجربی  $F_{ST}$  تعیین شده است.

جدول ۲. فهرست ژن‌های شناسایی شده مربوط به نواحی مورد انتخاب مربوط به دو نژاد گاومیش خوزستانی و مازندرانی

کروموزوم	شروع	انتها	ژن‌های شناسایی شده	صفت مرتبط با	QTL_ID	رفرنس QTL
۱	۷۱۰۳۹۹۷۰	۷۱۴۸۷۹۶۵	KIAA0226 - MUC20 SLC51A	میانگین افزایش وزن روزانه	۲۲۷۷۷	(Peters <i>et al.</i> , 2012)
۲	۱۲۶۸۷۹۳	۱۶۶۸۷۹۳	PTPN18 - AMER3 ARHGEF4	-	-	-
۲	۳۴۷۸۳۰۱۱	۳۵۱۸۳۰۱۱	-	چربی زیر پوستی	۴۵۱۷	(Takasuga <i>et al.</i> , 2007)
۲	۵۰۶۹۹۵۱۵	۵۱۰۹۹۵۱۵	-	باقیمانده غذای مصرفی	۵۰۹۶	(Sherman <i>et al.</i> , 2008)
۲	۵۰۶۹۹۵۱۵	۵۱۰۹۹۵۱۵	-	راندمن تبدیل غذایی	۵۰۹۸	(Sherman <i>et al.</i> , 2008)
۷	۴۵۶۶۰۰۸۲	۴۶۰۶۰۰۸۲	ONEC3 SCAM4 CSNK2 SOWA	وزن کشتار	۴۵۱۸	(Takasuga <i>et al.</i> , 2007)
۷	۴۵۶۶۰۰۸۲	۴۶۰۶۰۰۸۲	ONEC3 SCAM4 CSNK2 SOWA	زاویه پا	۱۵۷۱	(Ashwell <i>et al.</i> , 2005)
۷	۴۵۶۶۰۰۸۲	۴۶۰۶۰۰۸۲	ONEC3 SCAM4 CSNK2 SOWA	عمق بدن	۱۵۷۰	(Ashwell <i>et al.</i> , 2005)
۷	۴۵۶۶۰۰۸۲	۴۶۰۶۰۰۸۲	ONEC3 SCAM4 CSNK2 SOWA	طول بدن	۱۶۸۲	(Schnabel <i>et al.</i> , 2005)
۷	۵۲۶۵۰۴۷۶	۵۲۹۹۸۰۲۰	PSD2- PURA- IGIP CYSTM1-SLC4A9 EIF4EBP3- SLC35A4	راندمن تبدیل غذایی	۴۴۱۶	(Nkrumah <i>et al.</i> , 2007)
۷	۵۳۰۴۷۸۱۰	۵۴۱۳۳۳۷۲	TMCO6 - IK -WDR55 HARS2 - ZMAT2 PCDHA3 - PCDHA5 PCDHA11 - PCDHA13 DNAJB1 - PCDHB11 PCDHGA2-PCDHGA3 PCDHGB2-PCDHGA5	تولید پروتئین شیر	۶۱۹۰	(Daetwyler <i>et al.</i> , 2008)
۷	۵۳۰۴۷۸۱۰	۵۴۱۳۳۳۷۲	TMCO6 - IK -WDR55 HARS2 - ZMAT2 PCDHA3 - PCDHA5 PCDHA11 - PCDHA13 DNAJB1 - PCDHB11 PCDHGA2-PCDHGA3 PCDHGB2-PCDHGA5	باقیمانده غذای مصرفی	۱۱۴۵۰	(Marquez <i>et al.</i> , 2010)
۸	۶۳۵۰۹۷۲۰	۶۳۶۶۹۷۲۲	FOXE1 - NANS	-	-	-
۱۰	۱۴۶۸۴۴۹۵	۱۵۰۸۴۴۹۵	MAP2K5 - SKOR1 PIAS1 - FEM1B	درصد آلفا کازئین شیر	۲۱۵۴۱	(Schopen <i>et al.</i> , 2009)
۱۰	۱۴۶۸۴۴۹۵	۱۵۰۸۴۴۹۵	MAP2K5 - SKOR1 PIAS1 - FEM1B	طول بدن	۱۵۸۰	(Ashwell <i>et al.</i> , 2005)
۱۰	۱۴۶۸۴۴۹۵	۱۵۰۸۴۴۹۵	MAP2K5 - SKOR1 PIAS1 - FEM1B	عمق بدن	۱۵۷۸	(Ashwell <i>et al.</i> , 2005)
۱۰	۲۲۸۹۹۸۰۱	۲۳۲۹۹۸۰۱	TRAC	محل استقرار سر پستان	۴۶۳۴	(Schrooten <i>et al.</i> , 2000)
۱۰	۲۲۸۹۹۸۰۱	۲۳۲۹۹۸۰۱	TRAC	اتصالات پستانی	۴۶۳۳	(Schrooten <i>et al.</i> , 2000)
۱۱	۸۰۵۳۴۱۲	۸۴۵۳۴۱۲	-	وزن بدن	۴۵۴۶	(Takasuga <i>et al.</i> , 2007)
۱۱	۸۰۵۳۴۱۲	۸۴۵۳۴۱۲	-	میانگین افزایش وزن روزانه	۴۳۵۸	(Nkrumah <i>et al.</i> , 2007)
۱۱	۸۰۵۳۴۱۲	۸۴۵۳۴۱۲	-	مصرف ماده خشک	۴۳۵۹	(Nkrumah <i>et al.</i> , 2007)
۱۲	۲۰۶۰۸۶۸	۲۴۶۰۸۶۸	DIAPH3	-	-	-
۱۶	۶۰۸۹۷۷۸۶	۶۱۲۹۷۷۸۶	RASAL2	-	-	-
۱۸	۵۸۹۰۰۰۱۵	۵۹۷۵۲۴۸۵	-	اندازه گوساله	۳۰۵۴۵	(Höglund <i>et al.</i> , 2012)
۱۸	۵۸۹۰۰۰۱۵	۵۹۷۵۲۴۸۵	-	گوساله‌زایی آسان	۳۰۵۴۰	(Höglund <i>et al.</i> , 2012)
۱۸	۵۸۹۰۰۰۱۵	۵۹۷۵۲۴۸۵	-	شاخص تولد	۳۰۵۳۶	(Höglund <i>et al.</i> , 2012)
۱۹	۳۹۶۰۳۰۰۱	۴۰۱۰۷۹۸۶	MRPL45 - SOCS7 ARHGAP23- MLLT6 CISD3 - PSMB3 LASP1	درصد اسید کاپریلیک شیر	۱۰۰۲۳	(Morris <i>et al.</i> , 2007)
۱۹	۳۹۶۰۳۰۰۱	۴۰۱۰۷۹۸۶	MRPL45 - SOCS7 ARHGAP23- MLLT6 CISD3 - PSMB3 LASP1	درصد اسید کاپریک شیر	۱۰۰۲۴	(Morris <i>et al.</i> , 2007)
۱۹	۳۹۶۰۳۰۰۱	۴۰۱۰۷۹۸۶	MRPL45 - SOCS7 ARHGAP23- MLLT6 CISD3 - PSMB3 LASP1	درصد اسید لوریک شیر	۱۰۰۲۵	(Morris <i>et al.</i> , 2007)
۲۱	۶۰۴۲۴۲۶۸	۶۰۸۲۲۴۶۸	ERICH1	کیفیت و تردی گوشت	۲۰۸۱۴	(McClure <i>et al.</i> , 2012)
۲۸	۴۴۴۳۸۴۴	۴۸۴۳۸۴۴	FMN2	کیفیت و تردی گوشت	۲۰۸۲۸	(McClure <i>et al.</i> , 2012)

شروع و انتهای توالی جهت خلاصه کردن ژن‌های مجاور هم در یک ردیف آورده شده است. این نتایج بر اساس توالی ژنوم گاو (UMD3.1 Bos Taurus) است.



2004). به دلیل نقش ضد آپوپتوسیزی این ترکیب می‌تواند نقش مهمی در تردی گوشت در گاوهای گوشتی داشته باشد (Bernard *et al.*, 2007). در مطالعه دیگری Marcos-Carcavilla *et al.* (2010) نشان دادند که جهش در بخش‌های خاصی از ژن HSP90AA1، با سازگار شدن گوسفندان در شرایط محیطی مختلف ارتباط دارد. این محققان نتیجه گرفتند در گوسفندانی که در مناطقی با دامنه دمایی کمتر، میانگین دمای سالانه بیشتر و طول و عرض جغرافیایی کمتر زندگی می‌کنند، ژنوتیپ‌های خاصی از این ژن بیشترین فراوانی را دارد.

ژن‌های SLC51A، SLC35A4 و SLC4A9 از خانواده ژنی SLC نیز در ناحیه بین ۵۳/۲۴ و ۵۳/۹۵ مگا جفت باز روی کروموزوم ۷ گاوی قرار گرفته‌اند که در این مطالعه نشانه‌های انتخاب را نشان دادند. خانواده ژنی SLC که بیشتر در جمعیت‌های انسانی مورد تمایز و انتخاب قرار گرفته‌اند، نقش‌های متنوعی در مکانیسم‌های زیستی دارند. در مطالعات مختلف ارتباط این خانواده ژنی با رنگدانه‌های پوست (Lamason *et al.*, 2005; Pardridge, 2005; Voight *et al.*, 2006)، فرایندهای از قبیل بالا بردن (SLC27A4) و اکسیداسیون (SLC25A20) اسیدهای چرب غذا (Voight *et al.*, 2006) نشان داده شده است. این خانواده ژنی نقش‌های بیولوژیکی بسیار گسترده‌ای از قبیل سیگنال‌دهی پرولاکتین، ترشح انسولین، جذب طیف وسیعی از مواد مغذی، سنتز هورمون تیروئید و مسیرهای متابولیکی مختلف دیگر دارند (EntrezGene). به علاوه در مطالعات دیگری ارتباط این خانواده ژنی با صفات تولید و ترکیبات شیر از قبیل چربی و پروتئین گزارش شده است. از این میان می‌توان به ارتباط ژن‌های SLC39A8، SLC34A2، SLC1A4، SLC30A4، SLC7A5، SLC6A2، SLCA3 و SLC35B1 با صفات مربوط با تولید و کیفیت شیر اشاره کرد (Ogorevc *et al.*, 2009). ولی با توجه به طیف فعالیت‌های وسیع ژن‌های انتخابی این خانواده ژنی، اختصاص دقیق ژن‌های انتخابی این خانواده ژنی به مسیر بیولوژیکی خاصی امکان‌پذیر نیست.

از جایگاه‌های ژنی مهم دیگر که در این بررسی به‌عنوان نشانه‌های انتخاب شناسایی شدند، ژن‌های

ژن‌های PCDHA3 (Protocadherin Alpha 3)، PCDHA5، PCDHA11، PCDHA13، PCDHB11، PCDHGA2، PCDHGA3، PCDHGB2 و WDR55، و چندین ژن دیگر مرتبط با خانواده ژنی پروتوکادهرین در ناحیه بین ۵۳ و ۵۴/۱۳ مگا جفت باز روی کروموزوم ۷ گاوی قرار گرفته‌اند که در این مطالعه نشانه‌های انتخاب را نشان دادند. این ژن‌ها عضو گروه ژنی پروتوکادهرین آلفا (Protocadherin Alpha) هستند. گروه ژنی آلفا یکی از سه دسته ژنی مرتبط به هم است. این گروه ژنی مرکب از ۱۵ بالا خانواده است که ۱۳ تا از آن‌ها به لحاظ توالی به هم خیلی شبیه‌اند و دو بالا خانواده دیگر فاصله بیشتری با توجه به توالی‌هایشان دارند. این گروه پروتئین‌های عصبی (neural cadherin-like cell proteins adhesion) جزو پروتئین‌های غشای پلاسمایی هستند که احتمالاً نقش اساسی در ساختار و عملکرد ارتباطات عصبی بین سلولی در مغز دارند. به علاوه این پروتئین‌ها در توسعه سیستم عصبی، تجمع سیناپسی و انتقال سیناپسی دخالت دارند (EntrezGene). ژن‌های PCDHB4، PCDHB6، PCDHB7، PCDHB13 و چند ژن دیگر از گروه ژنی پروتوکادهرین بتا به‌عنوان نشانه‌های انتخاب در بین نژادهای گاو انتخاب شدند (Qanbari *et al.*, 2014). در مطالعه Utsunomiya *et al.* (2013) ژن PCDHA5 به‌عنوان نشانه انتخاب شناسایی شده و احتمالاً تحت انتخاب مثبت قرار داشتند. همچنین ژن WDR55 یک پروتئین مولتی دوماین انتقال غشایی است که به‌طور قابل توجهی در مغز بیان می‌شود (EntrezGene).

ژن DnajB1 (ENSBTAG00000045858) DnaJ homolog, subfamily B, member 1 ((Hsp40)) در ناحیه بین ۵۳/۸۵ و ۵۳/۸۶ مگا جفت‌باز روی کروموزوم ۷ گاوی قرار گرفته است و ژن پروتئین بیندینگ استرس گرمایی را کد می‌کند. این ژن کو-چاپرون Hsp70 (پروتئین با وزن مولکولی ۷۰ کیلو دالتونی، از مهم‌ترین پروتئین‌های استرس گرمایی) است که می‌تواند آثار ترکیبی روی عملکردهای بیولوژیکی داشته باشد. به طور مثال یک ژن از همین خانواده با Hsp70 (ترکیب DNaJA1/Hsp70) به‌طور مستقیم از آپوپتوسیز (مرگ‌ومیر سلول) جلوگیری می‌کند (Gotoh *et al.*, )

منتشر نشده است) از این دو نژاد در این بررسی از لحاظ صفات تیپ و تولید شیر تفاوت‌های معناداری داشتند و همچنین به دلیل وجود تفاوت‌های اقلیمی محل پرورش این دو نژاد، وجود این تفاوت‌ها محتمل بود و شناسایی این محل‌های ژنتیکی از مقایسه ژنومیک بین این دو نژاد، وجود این تفاوت‌ها را تأیید می‌کند.

#### نتیجه‌گیری

در این تحقیق پویش ژنومی جهت شناسایی نشانه‌های انتخاب در نژادهای گاو میش ایرانی خوزستانی و مازندرانی انجام گرفت و نشانه‌های انتخاب در بخش‌هایی از ژنوم شناسایی شد که در مطالعات قبلی در انسان، گاو و دام‌های اهلی دیگر نیز تأیید شده بودند. همان‌طور که در بخش نتایج و بحث نیز آورده شد، ژن‌های مربوط به نواحی انتخابی احتمالاً با مسیرهای بیولوژیکی که با اهلی شدن حیوانات از جمله توسعه مغز و عملکردهای رفتاری، پیگمسیون و آداپته‌شدن به محیط زندگی و شرایط جغرافیایی و نیز یکسری از صفات تولید شیر نقش دارند، در ارتباط هستند. البته یکسری از این مسیرهای بیولوژیکی در مطالعات قبلی مشخص شده‌اند و تأثیر ژن‌های انتخابی نیز اثبات شده است. همین نتایج گویای این است که احتمالاً انتخاب شدیدی روی صفات تولیدی انجام نگرفته و فقط تعداد اندکی ژن مرتبط با صفات تولید شیر در این بررسی نشانه‌های انتخاب را نشان دادند. همچنین QTL‌های شناسایی عمدتاً با صفات رشد و تیپ ارتباط داشتند. در هر صورت با توجه به اینکه این دو نژاد به شرایط محیطی مختلف آداپته شده‌اند و همچنین به لحاظ عملکردی و تولید شیر با هم متفاوتند، وجود این تفاوت‌ها در بین دو نژاد منطقی به نظر می‌رسد. با وجود این ژن‌هایی در مجاورت مناطق و SNP‌های انتخابی شناخته شده‌اند که نقش بیولوژیکی تعدادی از آن‌ها به‌طور کامل شناخته نشده است و همچنین برای برخی از این ژن‌ها احتمال وجود تأثیرات متقابل نیز وجود دارد که شناخته‌شده نیستند. به‌رحال جهت شناسایی نقش دقیق این ژن‌ها لازم است مطالعات پیوستگی و عملکردی بیشتری همراه با مطالعات سیستم بیولوژیک انجام گیرد.

NANS (در ناحیه بین ۶۳/۵-۶۳/۶ مگا جفت‌بازی کروموزوم ۸ گاوی) و SOCS7 (در ناحیه بین ۳۹/۶-۴۰/۱ مگا جفت‌بازی کروموزوم ۱۹ گاوی) هستند. در مطالعات قبلی ارتباط این ژن‌ها با صفات تولیدی شیر گزارش شده است (Ogorevc *et al.*, 2009). ژن NANS در مسیر بیوسنتز اسید سیالیک و لیپوئیک اسید دخالت داشته و در نهایت در تعداد سلول‌های سوماتیکی شیر نقش دارد (Ogorevc *et al.*, 2009). واریانتهای ژن SOCS7 در مقاومت به انسولین و متابولیسم لیپیدها و صفات مرتبط با چربی نقش دارد (Tellechea *et al.*, 2013). این ژن همچنین ممانعت‌کننده پرولاکتین، هورمون رشد و لپتین است و این کار را با جلوگیری از فعال‌شدن STAT3 و STAT5 و جداکردن آن‌ها در سیتوپلاسم و کاهش قدرت اتصال آن‌ها به DNA انجام می‌دهد (GeneCards) و با توجه به اینکه دو هورمون پرولاکتین و رشد در تولید شیر نقش عمده دارند، نقش این ژن در شیردهی نمایان می‌شود (Biswas *et al.*, 2006).

همچنین ژن‌های FOXE1 و MAP2K5 که در این بررسی به عنوان نشانه انتخاب تعیین شدند، احتمالاً با مکانیسم‌های مرتبط با تولید شیر ارتباط دارند. ژن‌های FOXE1 و MAP2K5 به ترتیب به عنوان سیگنال‌های مسیر تولید اکسی‌توسین و هورمون تیروئید نقش دارند. در ضمن ارتباط ژن‌هایی از همین خانواده‌ها مثل FOXH1، FOXF1، FOXB1 (Mitogen Activated Protein Kinase 7) و MAP3K8 با ورم پستان گزارش شده است (Zheng *et al.*, 2006).

علاوه بر ژن‌های شناسایی شده در مناطق تحت انتخاب، یکسری QTL در مناطق تحت انتخاب شناسایی شدند (جدول ۲). این QTL‌ها بیشتر با صفات مرتبط با راندمان تبدیل غذایی، وزن بدن، افزایش وزن روزانه، تیپ، تردی گوشت، ترکیبات شیر، اتصالات پستانی، اندازه گوساله و گوساله‌زایی آسان ارتباط دارند. وجود این QTL‌ها و نیز ژن‌های مرتبط با آداپتاسیون و تولید در این دو نژاد، نشانگر وجود تفاوت و تنوع در صفات بین دو نژاد مورد بررسی است. با توجه به اینکه نمونه‌های مورد استفاده (داده‌ها

## سپاسگزاری

از مؤسسه تحقیقاتی Padano به‌خاطر کمک‌های ارزنده‌شان در خصوص تعیین ژنوتیپ و از جناب آقای دکتر صابر قنبری و دکتر محمدحسین مرادی به‌جهت راهنمایی‌های ارزنده ایشان در خصوص آنالیز داده‌ها و از جناب آقای مهندس جواد رحمانی برای مشارکت در مراحل نمونه‌برداری و کارهای آزمایشگاهی، قدرانی می‌گردد.

نویسندگان مقاله از مرکز اصلاح نژاد و بهبود تولیدات دامی کشور و مراکز امور دام استان‌های خوزستان، کرمانشاه و مازندران و شرکت دانشگاهی و دانش‌بنیان نواندیش‌البرز که حمایت مالی این پروژه را بر عهده داشتند و در نمونه‌برداری و ثبت رکورد صفات تیپ و تولید همکاری کردند، کمال تشکر را دارند. همچنین

## REFERENCES

1. Akey, J. M. (2009). Constructing genomic maps of positive selection in humans: Where do we go from here?. *Genome Res*, 19, 711-722.
2. Akey, J. M., Zhang, G., Zhang, K., Jin, L. & Shriver, M. D. (2002). Interrogating a high-density SNP map for signatures of natural selection. *Genome Res*, 12(12), 1805-1814.
3. Alberts, C. C., Ribeiro-Paes, J. T., Aranda-Selverio, G., Cursino-Santos, J. R., Moreno-Cotuli, V. R., Oliveir, A. L. D., Santos Departamento, W.F. & Souza, E. B. (2010). DNA extraction from hair shafts of wild Brazilian felids and canids. *Gene and Mol Res*, 9 (4), 2429-2435.
4. Amaral, M. E., Owens, K. E., Elliott, J. S., Fickey, C. & Schaffer, A. A. (2007). Construction of a river buffalo (*Bubalus bubalis*) whole-genome radiation hybrid panel and preliminary RH mapping of chromosomes 3 and 10. *Anim Genet*, 38, 311-314
5. Ashwell, M. S., Heyen, D. W., Weller, J. I., Ron, M., Sonstegard, T. S., Van-Tassell, C. P. & Lewin, H. A. (2005). Detection of quantitative trait loci influencing conformation traits and calving ease in Holstein-Friesian cattle. *J Dairy Sci*, 88(11), 4111-9.
6. Bernard, C., Cassar-Malek, I., Le Cunff, M., Dubroeuq, H., Renand, G. & Hocquette, J. F. (2007). New indicators of beef sensory quality revealed by expression of specific genes. *J Agric Food Chem*, 55, 5229-5237.
7. Biswas, S. & Akey, J. M. (2006). Genomic insights into positive Selection. *Trends in Genetics*, 22(8), 437-436.
8. Borghese, A. (2011). Situation and perspectives of buffalo in the world, Europe and Macedonia. *Macedonian J Anim Sci*, 1(2), 281-296.
9. Campbell, A. M., Williamson, J., Padula, D. & Sundby, S. (1997). Use PCR & Single Hair to produce a "DNA Fingerprint". *The American Biol Teach*, 59(3), 172-178.
10. Daetwyler, H. D., Schenkel, F. S., Sargolzaei, M. & Robinson, J. A. B. (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. *J Dairy Sci*, 91 (8), 3225-36.
11. Ensembl BioMart: Ensembl online genome data base BioMart Tool. <http://www.ensembl.org/biomart/martview/>.
12. EntrezGene: NCBI Resources EntrezGene. <http://www.ncbi.nlm.nih.gov/>.
13. Fan, B., Du, Z. Q., Gorbach, D. M. & Rothschild, M. F. (2010). Development and Application of High-density SNP Arrays in Genomic Studies of Domestic Animals. *Asian-Aust. J Anim Sci*, 23(7), 833-847.
14. GeneCards. <http://www.genecards.org/cgi-bin/carddisp.pl?gene=STAT>
15. Goddard, M. (2009). Genomic selection: prediction of accuracy and maximization of long term response. *Genetics*, 136, 245-257.
16. Gotoh, T., Terada, K., Oyadomari, S. & Mori, M. (2004). Hsp70 DnaJ chaperone pair prevents nitric oxide and CHOP induced apoptosis by inhibiting translocation of Bax to mitochondria. *Cell Death Differ*, 11, 390-402.
17. Grimberg, J., Nawoscihik, S., Belluscio, L., McKee, R., Turk, A. & Eisenberg, A. (1989). A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. *Nucl Acids Res*, 17, 8390.
18. Grossman, S. R., Shylakhter, I., Karlsson, E. K., Byrne, E. H. & Morales, S. *et al.* (2010). A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*, 327, 883-886.
19. Gulcher, J. & Stefansson, K. (1998). Population Genomics: Laying the Groundwork for Genetic Disease Modeling and Targeting. *Clin Chem Lab Med*, 36, 523-527.

20. Höglund, J. K., Gulbrandsen, B., Lund, M. S. & Sahana, G. (2012). Analyzes of genome-wide association follow-up study for calving traits in dairy cattle. *BMC Genet*, 13, 71.
21. Khan, M. S., Rehman, M. S. & Hassan, F. (2012). Breeding Buffaloes in Genomics era –Issues of recording and evaluation. *J. Anim Plant Sci*, 22(3), 174-180.
22. Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S. & Porto Neto, L. R., *et al.* (2012). Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. *PLoS Biol*, 10(2), e1001258. doi:10.1371/journal.pbio.1001258.
23. Lamason, R. L., Mohideen, M., Mest, J. R., Wong, A. C. & Norton, H. L., *et al.* (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*, 310, 1782–1786.
24. Leutenegger, A. L., Prum, B., Génin, E., Verny, C., Lemaître, A., Clerget-Darpoux, F. & Thompson, E. A. (2003). Estimation of the inbreeding coefficient through use of genomic data. *Am J Hum Genet*, 73, 516-523.
25. Li, Z., Xie, J., Li, W., Tang, A., Li, X., Jiang, Z., Han, Y., Ye, J., Jing, J., Gui, Y. & Cai, Z. (2011). Identification and characterization of human PCDH10 gene promoter. *Gene*, 475, 49-50.
26. MacEachern, S., Hayes, B., McEwan, J. & Goddard, M. (2009). An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (*Bos taurus*) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic cattle. *BMC Genomics*, 10, 181.
27. Marcos-Carcavilla *et al.* (2010). Polymorphisms in the HSP90AA1 5' flanking region are associated with scrapie incubation period in sheep. *Cell Stress & Chaperones*, 15(4), 343-349.
28. Marquez, G. C., Enns, R. M., Grosz, M. D., Alexander, L. J. & Macneil, M. D. (2010). Quantitative trait loci with effects on feed efficiency traits in Hereford x composite double backcross populations. *Anim Genet*, 40(6), 986-8.
29. Maynard-Smith, J. & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genet Res*, 23, 23-35.
30. McClure, M. C., Ramey, H. R., Rolf, M. M., McKay, S. D., Decker, J. E. & Chapple, R. H., *et al.* (2012). Genome-wide association analysis for quantitative trait loci influencing Warner-Bratzler shear force in five taurine cattle breeds. *Anim Genet*, 43 (6), 662-73.
31. Michelizzi, V. N., Wu, X., Dodson, M., Michal, J. J., Zambrano-Varon, J., McLean, D. J. & Jiang, Z. (2011). A Global View of 54,001 Single Nucleotide Polymorphisms (SNPs) on the Illumina BovineSNP50 BeadChip and Their Transferability to Water Buffalo. *Int. J. Biol. Sci*, 7(1), 18-27.
32. Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K. G. & McEwan, J. C. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics*, 13, 10.
33. Morris, C. A., Cullen, N. G., Glass, B. C., Hyndman, D. L., Manley, T. R., Hickey, S. M., McEwan, J. C., Pitchford, W. S., Bottema, C. D. K. & Lee, M. A. H. (2007). Fatty acid synthase effects on bovine adipose fat and milk fat. *Mammal Genome*, 18 (1), 64-74.
34. Nicolazzi, E. L., Iamartino, D. & Williams, J. L. (2014). AffyPipe: an open-source pipeline for Affymetrix Axiom genotyping workflow. *Bioinformatics*, 30(21), 3118-9. doi: 10.1093/bioinformatics/btu486. Epub 2014 Jul 15.
35. Nielsen, R. & Yang, Z. (1988). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*, 148, 929-936.
36. Nkrumah, J. D., Sherman, E. L.; Li, C., Marques, E., Crews, D. H., Bartusiak, R., Murdoch, B., Wang, Z., Basarab, J. A. & Moore, S. S. (2007). Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. *J Anim Sci*, 85(12), 3170-81.
37. Ogorevc, J., Kunej, T., Razpet, A. & Dovc, P. (2009). Database of cattle candidate genes and genetic markers for milk production and mastitis. *Anim Genet*, 40, 832-851.
38. Oleksyk, T. K., Smith, M. W. & O'Brien, S. J. (2010). Genome-wide scans for footprints of natural selection. *Philosophical Transactions of the Royal Society B*, 365, 185-205.
39. Othman, O. E. (2006). Restriction fragment length polymorphism and gene mapping of two genes associated with composition in Egyptian river buffalo. *J Dairy Sci*, 1(1), 84-92.
40. Pardridge, W. M. (2005). The blood–brain barrier: bottleneck in brain drug development. *NeuroRx: J Am Soc Exp NeuroTherapeutics*, 2(1), 3-14.
41. Pérez O'Brien, A. M., Utsunomiya, Y. T., Gábor Mészáros, V. B., Bickhart, D. M., Liu, G. E., Van Tassell, C. P., Sonstegard T. S., Silva, M. D., Garcia, J. F. & Sölkner, J. (2014). Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genet Sel Evol*, 46, 19.
42. Peters, S. O., Kizilkaya, K., Garrick, D. J., Fernando, R. L., Reecy, J. M., Weaber, R. L., Silver, G. A. & Thomas, M. G. (2012). Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. *J Anim Sci*, 90 (10), 3398-409.

43. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J. & Sham, P. C. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet*, 81, 559-575.
44. Qanbari, S., Pausch, H., Jansen, S., Somel, M. & Strom, T. M., *et al.* (2014). Classic Selective Sweeps Revealed by Massive Sequencing in Cattle. *PLoS Genet*, 10(2), e1004148. doi:10.1371/journal.pgen.1004148
45. Qanbari, S., Strom, T. M., Haberer, G., Weigend, S. & Gheyas, A. A., *et al.* (2012) A High Resolution Genome-Wide Scan for Significant Selective Sweeps: An Application to Pooled Sequence Data in Laying Chickens. *PLoS ONE*, 7(11), e49525. doi:10.1371/journal.pone.0049525.
46. Rubin, C. J., Zody, M. C., Eriksson, J., Meadows, J. R. S., Sherwood, E., *et al.* (2010). Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, 464, 587-591.
47. Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z. P. & Richter, D. J., Schaffner, S. F., Gabriel, S. B., Platko, J. V., Patterson, N. J., McDonald, G. J., *et al.* (2002). Detecting recent positive selection in the human genome from Haplotype structure. *Nature*, 419, 832-837.
48. Schnabel, R. D., Sonstegard, T. S., Taylor, J. F. & Ashwell, M. S. (2005). Whole-genome scan to detect QTL for milk production, conformation, fertility and functional traits in two US Holstein families. *Anim Genet*, 36(5), 408-16.
49. Schopen, G.C.B., Koks, P.D., van-Arendonk, J.A.M., Bovenhuis, H. & Visker, M.H.P.W. (2009). Whole genome scan to detect quantitative trait loci for bovine milk protein composition. *Anim Genet*, 40 (4), 524-37.
50. Schrooten, C., Bovenhuis, H., Coppieters, W. & Van Arendonk, J. A. (2000). Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. *J Dairy Sci*, 83(4), 795-806.
51. Sherman, E.L., Nkrumah, J.D., Murdoch, B.M. & Moore, S.S. (2008). Identification of polymorphisms influencing feed intake and efficiency in beef cattle. *Anim Genet*, 39 (3), 225-31.
52. Simianer, H., Ma, Y. & Qanbari, S. (2014). Statistical problems in livestock population genomics. Proceedings of 10<sup>th</sup> Congress of Genetics Applied to Livestock Production, 17-22 August, Vancouver, BC, Canada.
53. Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585-595.
54. Takasuga, A., Watanabe, T., Mizoguchi, Y., Hirano, T., Ihara, N., Takano, A. & Yokouchi, K. F. (2007). Identification of bovine QTL for growth and carcass traits in Japanese Black cattle by replication and identical-by-descent mapping. *Mammal Genome*, 18 (2), 125-36.
55. Tellechea, M. L., Steinhart, A. P., Rodriguez, G., Taverna, M. J., Poskus, E. & Frechtel, G. (2013). Common variants in SOCS7 gene predict obesity, disturbances in lipid metabolism and insulin resistance. *Nutr Metab Cardiovasc*, 23(5), 424-431.
56. Teo, Y. Y., Fry, A. E., Clark, T. G., Tai, E. S. & Seielstad, M. (2007). On the usage of HWE for identifying genotyping errors. *Annal of Hum Genet*, 71, 701-703.
57. The Bovine HapMap Consortium, Gibbs, R. A., Taylor, J. F., Van Tassell, C. P. & Barendse, W., *et al.* (2009) Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. *Science*, 324, 528-532. doi:10.1126/science.1167936.
58. The R Project for Statistical Computing: Free software environment for statistical computing and graphics. [http:// www.r-project.org/](http://www.r-project.org/).
59. Utsunomiya, Y.T., Perez O'Brien, A.M., Sonstegard, T.S., Van Tassell, C.P. & do Carmo, A.S., *et al.* (2013). Detecting Loci under Recent Positive Selection in Dairy and Beef Cattle by Combining Different Genome-Wide Scan Methods. *PLoS ONE*, 8(5), e64280. doi:10.1371/journal.pone.0064280.
60. Voight, B.F., Kudaravalli, S., Wen, X. & Pritchard, J.K. (2006). A Map of Recent Positive Selection in the Human Genome. *PLoS Biol*, 4(3), e72.
61. Weir, B. S. & Cockerham, C. C. (1984). Estimating F-Statistics for the analysis of population structure. *Evol*, 38(6), 1358-1370.
62. Zheng, J., Watson, A.D. & Kerr, D.E. (2006). Genome-wide expression analysis of lipopolysaccharide-induced mastitis in a mouse model. *Infect and Immunity*, 74, 1907-15.

## Genome-wide survey of signature of positive selection in Khuzestani and Mazandrani buffalo breeds

Mahdi Mokhber<sup>1</sup>, Mohammad Moradi Shahrabak<sup>2\*</sup>, Mostafa Sadeghi<sup>3</sup>,  
Hossein Moradi Shahrabak<sup>4</sup> and John Williams<sup>5</sup>

1, 2, 3, 4. Ph.D. Candidate, Professor, and Assistance Professors, Department of Animal Science,  
Faculty of Agricultural Science and Engineering, University College of Agriculture and Natural  
Resources (UTCAN), University of Tehran, Karaj, Iran

5. Parco Tecnologico Padano, Lodi, Italy,

(Received: Apr. 15, 2014 - Accepted: Jun. 23, 2015)

### ABSTRACT

Identification of selection targeted genomic regions is one of the most challenging areas of research in animal genetics, particularly in livestock. We carried out a genome-wide scan for signatures of positive selection to identify genomic regions that had been under selection in Iranian Khuzestani and Mazandrani buffalo breeds. A total of 148 water buffalo from Khuzestani (N=121) and Mazandrani (N=27) buffalo breeds were genotyped using Axiom® Buffalo Genotyping 90K Array. Unbiased method of population differentiation index ( $F_{ST}$ ) was applied to detect signatures of selection. In total, 23 regions exceeding the 0.1 percent threshold of the empirical posterior distribution were identified as extremely differentiated. These selected genomic regions were surveyed to find encoding putative candidate genes and 64 genes and 27 QTL were extracted from the corresponding areas in UMD3.1 Bos Taurus Genome Assembly. Some of these genes have previously reported as signature of positive selection in the last studies. Some of these genes were also found to be involved in milk production traits and domestication-related changes include sensory perceptions, brain and neural system development, pigmentation, and geographic adaptation. Also, survey on extracted QTLs was shown that these QTLs involved in some economical important traits in buffalo such as feed conversion ratio, subcutaneous fat, body weight, average daily gain, type, Meat tenderness, milk production content, udder attachment, calf size and calving ease traits. However, it will be necessary to carry out more association and functional studies to demonstrate the implication of these genes.

**Keywords:** genome-wide survey, Khuzestani and Mazandrani buffalo breeds, population differentiation index, signatures of selection.