

بررسی سطوح عدم تعادل پیوستگی در ژنوم گاوهای بومی استان فارس با استفاده از داده‌های متراکم نشانگرهای چند شکل تک نوکلئوتیدی

کریم کریمی^{۱*}، علی اسماعیلی زاده کشکوئی^۲ و مسعود اسدی فوزی^۲

تاریخ دریافت: ۱۳۹۴/۱۰/۶ تاریخ پذیرش: ۱۳۹۵/۳/۳

^۱ دانش آموخته دکتری گروه علوم دامی دانشکده کشاورزی دانشگاه شهید باهنر کرمان و عضو انجمن پژوهشگران جوان دانشگاه شهید

باهنر کرمان

^۲ به ترتیب استاد و دانشیار گروه علوم دامی دانشکده کشاورزی دانشگاه شهید باهنر کرمان

* مسئول مکاتبه: Email: Karim.karimi81@gmail.com

چکیده

زمینه مطالعاتی: شناسایی سطوح عدم تعادل پیوستگی در میان جمعیت‌ها ابزاری مفید در مطالعه تاریخچه جمعیت‌ها و شناسایی نواحی ژنومی مرتبط با صفات مهم اقتصادی است. هدف: در این تحقیق، به منظور فراهم آوردن اطلاعات پایه مورد نیاز در طراحی مطالعات ارتباطی کل ژنوم و بررسی تغییرات اندازه موثر جمعیت در گاوهای بومی فارس، سطوح عدم تعادل پیوستگی در ژنوم افراد این جمعیت مورد مطالعه قرار گرفت. روش کار: بدین منظور تعداد ۱۰ راس از گاوهای بومی فارس به طور تصادفی نمونه برداری شدند و به کمک تراشه Illumina Bovine HD در ۷۷۷۹۶۲ جایگاه SNP تعیین ژنوتیپ شدند. کنترل کیفی داده‌ها بر اساس نرخ فراخوانی تعیین ژنوتیپ، انحراف از تعادل هادری-وینبرگ و فراوانی آلل‌های نادر در جایگاه‌های مختلف انجام شد و ۵۵۷۱۸ جایگاه برای انجام آنالیزهای بعدی انتخاب شدند. کلیه جفت مقایسه‌های بین SNPها در سه دسته فاصله‌ای صفر تا ۱۰، ۱۰ تا ۱۰۰ و ۱۰۰ تا ۱۰۰۰ کیلوبازی تقسیم بندی شدند و متوسط آماره r^2 برای تمامی کروموزوم‌های غیرجنسی در این فواصل محاسبه شد. نتایج: متوسط r^2 بالاتر از ۰/۳ در فواصل کمتر از ۹ Kb و متوسط r^2 بالاتر از ۰/۲ در فواصل کمتر از ۶۰ Kb مشاهده گردید. مقادیر r^2 بدست آمده در فواصل بین نشانگری کوتاه‌تر (کمتر از ۱۰۰ Kb) از اطمینان بیشتری برخوردار بودند. مقادیر بالای r^2 یافت شده در این مطالعه نشان‌دهنده سطح بالای بروز همخونی و کاهش شدید جمعیت موثر در گاوهای بومی فارس است. نتیجه‌گیری نهایی: نتایج این مطالعه می‌تواند در تعیین تراکم نشانگری مورد نیاز جهت کسب دقت کافی در مطالعات ژنومی احتمالی در گاوهای بومی فارس به کار گرفته شود.

واژگان کلیدی: ژنوم گاو، عدم تعادل پیوستگی، گاو بومی فارس، نشانگرهای چند شکل تک نوکلئوتیدی

مقدمه

عدم تعادل پیوستگی (LD) به عنوان ارتباط غیر تصادفی میان آلل‌های جایگاه‌های مختلف در داخل یک جمعیت تعریف شده است. سطح LD در یک جمعیت توسط برخی از عوامل ژنتیکی نظیر انتخاب، جهش، مهاجرت، ساختار جمعیتی و نرخ نوترکیبی تحت تاثیر قرار می‌گیرد (زو و همکاران ۲۰۱۳). شناسایی نشانه‌های انتخاب در ژنوم، بررسی اساس ژنتیکی صفات اقتصادی مهم و تنوع ژنتیکی بین نژادهای مختلف می‌تواند به کمک تعیین سطوح LD در یک جمعیت مورد بررسی قرار داده شوند (مک کی و همکاران ۲۰۰۷؛ گویا و همکاران ۲۰۱۴). انتخاب ژنومی در حال حاضر با موفقیت در صنعت گاو شیری مورد استفاده قرار می‌گیرد و استفاده از آن در حال گسترش به سایر دام‌های اهلی نیز می‌باشد. انتخاب ژنومی بر پایه عدم تعادل پیوستگی میان ژن‌های مسبب ایجاد تغییرات فنوتیپی و نشانگرهای مولکولی بنا نهاده شده است. به همین دلیل، میزان دقت و کارایی این روش تا حدودی وابسته به میزان عدم تعادل پیوستگی در ژنوم حیوانات مورد انتخاب است (میوسن و همکاران ۲۰۰۱؛ هیز و همکاران ۲۰۰۹). علاوه بر این، کارایی مطالعاتی نظیر انتخاب به کمک نشانگرها، مطالعه ارتباطی کل ژنوم، نقشه یابی جایگاه‌های موثر بر صفات کمی و آزمون انساب نیز به سطح LD در جمعیت وابسته است (سلنر و همکاران ۲۰۰۷).

سطوح عدم تعادل پیوستگی در بین مطالعات مختلف بر حسب جمعیت مورد مطالعه و آماره مورد استفاده تفاوت زیادی داشته است. در مقایسه با مطالعات انجام شده بر روی ژنوم انسان (کونارد و همکاران ۲۰۰۶)، گسترش عدم تعادل پیوستگی به میزان بیشتری در ژنوم دام‌های اهلی مشاهده شده است که می‌توان علت آن را ناشی از اندازه موثر کوچک‌تر جمعیت و فشار انتخاب شدیدتر در این جمعیت‌ها دانست (میدوز و همکاران ۲۰۰۸). مطالعات اولیه

صورت گرفته بر ژنوم گاو اغلب به کمک نشانگرهای ریزماهواره انجام شده‌اند (فارنیر و همکاران ۲۰۰۰). با این حال، در حال حاضر با ظهور تکنولوژی های پیشرفته توالی یابی ژنوم، نشانگرهای چند شکل تک نوکلئوتیدی به میزان زیادی جهت بررسی سطوح LD ژنوم گاو مورد استفاده قرار گرفته‌اند (سالمون-تورس و همکاران ۲۰۱۴). میزان متفاوت عدم تعادل پیوستگی و تفاوت های مشاهده شده بین کروموزوم‌های مختلف به لحاظ سطح LD در میان نژادهای مختلف نشان داده است که این تفاوت‌ها می‌تواند جهت شناسایی خصوصیات یک نژاد مورد استفاده باشند. علاوه بر این، از آنجا که کاهش میزان LD در سطح ژنوم تابعی از تعداد نسل‌ها می‌باشد، داده‌های عدم تعادل پیوستگی به میزان زیادی در برآورد اندازه موثر جمعیت‌ها مورد استفاده قرار گرفته‌اند (فلوری و همکاران ۲۰۱۰).

استان فارس به عنوان یکی از مهدهای تمدن ایران زمین از قدیمی‌ترین مراکز اهلی سازی و پرورش دام در کشور محسوب می‌شود. گاوهای بومی فارس به لحاظ جثه کوچک بوده و اغلب به رنگ سیاه دیده می‌شوند، با این حال گاوهای طلایی، قهوه‌ای روشن و خاکستری نیز در این توده یافت می‌شوند. تاکنون مطالعه‌ای در خصوص تعیین ویژگی‌های تولیدی و فنوتیپی این گاوها به عمل نیامده است و اطلاعات چندانی از ویژگی‌های این گاوها در دسترس نیست. از جمله ویژگی‌های بارز این جمعیت بومی می‌توان به سازگاری با شرایط چرا، استفاده از خوراک‌های با کیفیت پایین، مقاومت به انگل‌ها و بیماری‌های منطقه اشاره نمود. در سال‌های اخیر تعداد گاوهای بومی به شدت رو به کاهش نهاده است و تلاقی‌گری های بدون برنامه، یافتن گاوهای خالص را مشکل کرده است. هدف از این مطالعه تعیین میزان عدم تعادل پیوستگی در ژنوم گاوهای بومی فارس بوده است. نتایج این تحقیق می‌تواند به بررسی روند تغییرات اندازه این جمعیت بومی

¹- Linkage disequilibrium (LD)

آل نادر (MAF) برای کلیه جایگاه‌ها محاسبه شد و کلیه جایگاه‌های با MAF کمتر از ۰/۰۱ از محاسبات کنار گذاشته شدند. همچنین، جایگاه‌های ژنی بر اساس مقادیر MAF در ۶ دسته شامل مقادیر کمتر از ۰/۰۵، ۰/۰۵-۰/۱، ۰/۱-۰/۲، ۰/۲-۰/۳، ۰/۳-۰/۴ و ۰/۴-۰/۵ تقسیم بندی شدند و نمودار توزیع فراوانی آنها رسم گردید.

برآورد عدم تعادل پیوستگی

در این مطالعه از آماره r^2 (هیل ۱۹۷۴) به منظور برآورد سطوح عدم تعادل پیوستگی در داده‌ها بهره گرفته شد. r^2 که به عنوان ضریب همبستگی بین جفت SNPها تعریف شده است، بر اساس معادله زیر محاسبه گردید (هیل و ویر ۱۹۹۴):

$$r^2 = \frac{(freqAB * freqab - freqAb * freqaB)^2}{(freq A * freq a * freq B * freq b)} \quad [1]$$

در معادله بالا $freq A$ ، $freq a$ ، $freq B$ و $freq b$ به ترتیب نشان‌دهنده فراوانی آلل‌های A ، a ، B و b بوده و $freq AB$ ، $freq ab$ ، $freq Ab$ و $freq aB$ نیز به ترتیب فراوانی هاپلوتایپ‌های AB ، ab ، Ab و aB را نشان می‌دهند. آماره r^2 به کمک نرم افزار SnpIddHD (سرگلزایی و همکاران ۲۰۰۹) برای تمامی جفت نشانگرهای موجود بر روی هر یک از کروموزوم‌ها محاسبه گردید. محاسبه r^2 تنها برای جفت SNPهایی انجام شد که در فاصله حداکثر ۱۵ Mb از یکدیگر قرار داشتند. به کمک معادله زیر کلیه r^2 های محاسبه شده برای تعداد کم نمونه، مورد تصحیح قرار داده شدند (اسود ۱۹۷۱):

$$r^2_{corrected} = \frac{r^2_{computed} - \frac{1}{n}}{1 - \frac{1}{n}} \quad [2]$$

در این معادله n نشان دهنده تعداد هاپلوتایپ‌های موجود در نمونه می‌باشد (تعداد نمونه ضربدر دو).

حداکثر فاصله بین SNPها به سه دسته تقسیم بندی شد: فواصل کمتر از ۱۰ Kb، فواصل بیشتر از ۱۰ Kb و کمتر از ۱۰۰Kb و فواصل کمتر از ۱۰۰۰ Kb و بیشتر از ۱۰۰Kb سپس فواصل موجود در هر یک از دسته های

در نسل‌های گذشته کمک کرده و به عنوان بخشی از خصوصیات ژنتیکی این جمعیت مورد توجه باشد. همچنین میزان کارایی و تراکم نشانگری مورد نیاز برای اجرای روش‌هایی از قبیل انتخاب ژنومیک و مطالعات ارتباطی کل ژنوم در این جمعیت نیز به کمک داده‌های LD قابل بررسی است.

مواد و روش‌ها

جمع‌آوری نمونه، تعیین ژنوتیپ و کنترل کیفی داده‌ها
تعداد ۱۰ راس گاو به صورت تصادفی و کنترل شده از جمعیت گاوهای بومی استان فارس مورد نمونه‌برداری قرار گرفتند. نمونه‌های مو از ناحیه انتهایی دم گاو‌ها جدا شدند و در قالب بسته‌های جداگانه به آزمایشگاه انتقال داده شدند. مراحل استخراج دی.ان.ای و تعیین ژنوتیپ نمونه‌ها در آزمایشگاه تحقیقات ژنتیک حیوانی دانشگاه کوئینزلند (بریسبن، استرالیا) انجام گرفت. استخراج دی.ان.ای به کمک کیت Gentra Puregene Tissue (California, USA Qiagen Inc., Valencia, 117) از ریشه‌های مو انجام شد و تعیین ژنوتیپ نمونه‌ها با استفاده از Illumina BovineHD SNP Beadchip انجام گرفت. کنترل کیفی داده‌ها با استفاده از نرم افزار PLINK 1.07 (پارسل و همکاران ۲۰۰۷) انجام شد. تنها SNPهای واقع شده بر روی کروموزوم‌های غیرجنسی مورد استفاده قرار گرفتند. کلیه جایگاه‌های ژنی دارای نرخ فراخوانی تعیین ژنوتیپ کمتر از ۹۰٪ از مجموعه داده‌ها کنار گذاشته شدند. از آنجا که انحراف بسیار معنی‌دار از تعادل هاردی-وینبرگ در برخی از جایگاه‌های ژنی می‌تواند دلیلی بر خطای تعیین ژنوتیپ در این جایگاه‌ها باشد (سن و بورمیستر، ۲۰۰۸)، کلیه جایگاه‌هایی که آماره (کای اسکویر) آزمون تعادل هاردی وینبرگ در آنها دارای P -value کمتر از 10^{-7} بود نیز از داده‌ها حذف شدند. فراوانی

کوتاه‌ترین طول در کروموزوم ۲۵ (۴۳/۰۸ Mb) مشاهده گردید. متوسط فاصله بین دو SNP مجاور برابر با ۴۸/۱ کیلوباز بدست آورده شد و میانگین کل MAF برابر با ۰/۲۲۳ بود. در مجموع تعداد ۵۱۸۱۶۷ جفت مقایسه برای بدست آوردن r^2 مورد استفاده قرار گرفت. شکل ۱ نمودار توزیع MAF را در دسته‌های مختلف نشان داده است. ۴۴/۵ درصد از کل جایگاه‌ها دارای MAF کمتر از ۰/۲ بودند و بالاترین نسبت SNPها در دسته با MAF ما بین ۰/۴ تا ۰/۵ دیده شدند. MAF به عنوان یکی از عوامل اثرگذار در برآورد r^2 معرفی شده است، به طوری که به ویژه در فواصل بین نشانگری کوتاه‌تر (فواصل کمتر از ۱۰ کیلوباز) افزایش مقادیر MAF موجب افزایش میانگین مقادیر r^2 شده است (ختکار و همکاران ۲۰۰۸؛ زو و همکاران ۲۰۱۳). توزیع نسبتاً یکنواخت جایگاه‌ها در دسته‌های با MAF متفاوت (شکل ۱)، نشان می‌دهد که این عامل نقش چندان مهمی در اختلاف برآوردهای r^2 در این تحقیق نداشته است. متوسط r^2 بین SNPهای مجاور به ازای هر یک از کروموزوم‌ها بدست آمد. میانگین کل LD بین SNPهای مجاور برابر با ۰/۲۹ بود و بالاترین متوسط r^2 بین SNPهای مجاور در کروموزوم ۲۲ (۰/۳۰۴) و کمترین آن در کروموزوم ۱۵ (۰/۲۰۸) مشاهده شد. سطح LD در میان کروموزوم‌های مختلف به شکل قابل توجهی متفاوت بود. میانگین LD در فواصل مختلف ژنومی مورد محاسبه قرار گرفت. جفت مقایسه‌های بین SNPها در سه دسته فاصله‌ای شامل ۰ Kb تا ۱۰ Kb (با پنجره‌های ۱Kb)، ۱۰ Kb تا ۱۰۰ Kb (با پنجره‌های ۱۰ Kb) و ۱۰۰ Kb تا ۱۰۰۰ Kb (با پنجره‌های ۱۰۰ Kb) تقسیم بندی شدند. جدول ۱ اطلاعات آماری مربوط به متوسط r^2 حاصل شده در فواصل مختلف ژنومی را نمایش داده است. بررسی روند تغییرات میزان r^2 در فواصل مختلف نشان داد که در فواصل بین نشانگری کوتاه‌تر، سطح LD نوسان بیشتری داشته است. همچنین روند کاهش LD در فواصل بیش از ۴۰۰ کیلوبازی بسیار کند شده است و به سطح نسبتاً ثابتی رسیده است. متوسط r^2 همراه با افزایش

اول تا سوم به ترتیب به کمک پنجره‌های ۱، ۱۰ و ۱۰۰ کیلوبازی تقسیم‌بندی شدند. میانگین r^2 در هر یک از پنجره‌ها به ازای هر یک از کروموزوم‌ها محاسبه شد و نمودار تغییرات میانگین r^2 بر حسب این فواصل رسم گردید.

تغییرات اندازه موثر جمعیت

شدت عدم تعادل پیوستگی در فواصل مختلف روی ژنوم می‌تواند جهت برآورد اندازه موثر جمعیت در دوره‌های گذشته مورد استفاده باشد. اندازه موثر جمعیت (Ne) را می‌توان به کمک معادله زیر برآورد نمود (سلنر و همکاران ۲۰۰۷):

$$E(r^2) = \frac{1}{1+4Nec} \quad [۳]$$

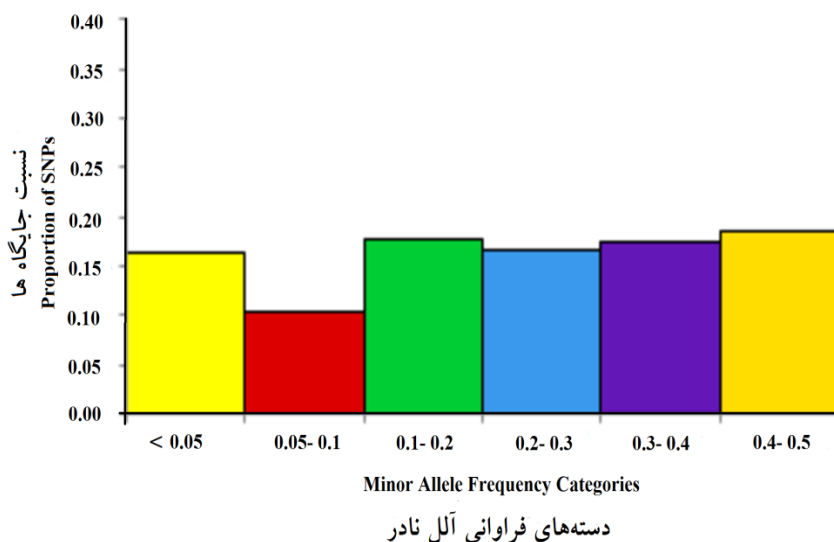
در این معادله Ne اندازه موثر جمعیت بوده و c نشان دهنده فاصله ژنتیکی (بر حسب مورگان) می‌باشد. با فرض اینکه cM ۱ برابر با ۱ Mb باشد، فواصل فیزیکی ۲۵ Kb، ۵۰ Kb، ۱۰۰ Kb، ۵۰۰ Kb، ۱ Mb، ۵ Mb و ۱۰ Mb جهت برآورد اندازه موثر جمعیت از ۲۰۰۰ نسل قبل تا ۵ نسل پیش مورد استفاده قرار داده شدند.

نتایج و بحث

تعداد ۱۰ راس گاو بومی در ۷۷۷۹۶۲ جایگاه تعیین ژنوتیپ شدند. تعداد SNPهای واقع شده بر روی کروموزوم‌های X، Y، میتوکندری و جایگاه‌های نامشخص به ترتیب برابر با ۳۹۳۶۷، ۱۲۲۴، ۳۴۳ و ۱۷۳۵ بود که همگی آنها از محاسبات بعدی کنار گذاشته شدند. همچنین، اجرای کنترل کیفی بر روی داده‌ها بر اساس نرخ فراخوانی تعیین ژنوتیپ (کمتر از ۹۰٪)، انحراف از تعادل هاردی-وینبرگ (P-value کمتر از 10^{-7}) و آستانه MAF (کمتر از ۰/۰۱) نیز به ترتیب موجب حذف ۶۱۲۳۰۷، ۲۱۵۹ و ۱۲۰۸۲۷ جایگاه از داده‌ها شد. پس از انجام کنترل کیفی بر روی مجموعه داده‌ها تعداد ۵۵۷۱۸ SNP جهت آنالیزهای بعدی مورد انتخاب قرار گرفتند. طول کل ژنوم اتوزومی برابر با ۲۱۶۲/۸۲ Mb بود و بزرگترین طول کروموزوم مربوط به کروموزوم یک (۱۶۰/۸۸ Mb) و

کل r^2 در تمامی کروموزوم‌های مورد بررسی برابر با $0/118$ بود و بالاترین سطح r^2 در کروموزوم ۱۷ ($r^2=0/153$) و کمترین حد آن در کروموزوم ۲۶ ($r^2=0/093$) دیده شد (شکل ۲).

فاصله فیزیکی بین SNP ها کاهش یافت، با این حال درجه تغییرات آن در بین کروموزوم‌های مختلف یکسان نبود. این امر را می‌توان به نرخ وقوع نوترکیبی و فشار انتخاب متفاوت در بین کروموزوم‌های مختلف نسبت داد. میانگین



شکل ۱- توزیع فراوانی آلل‌های نادر (MAF) در گاوهای بومی فارس

Figure 1- Distribution of minor allele frequencies (MAF) in Fars native cattle

در بررسی چندین نژاد بوس تایروس و بوس ایندیگوس، محدوده $0/55$ تا $0/75$ در این فاصله گزارش شده است (پرز ابرین و همکاران ۲۰۱۴). از سوی دیگر، در بررسی هشت نژاد گاو سطح LD در فواصل کمتر از ۱۰ Kb در محدوده‌ای ما بین $0/25$ (هرفورد) تا $0/49$ (برهمن) قرار داشته است (پورتو نتر و همکاران ۲۰۱۴). همچنین مقدار گزارش شده برای متوسط r^2 در همین فاصله در نژادهای آمیخته برزیلی در محدوده $0/25$ تا $0/4$ قرار گرفته است (موکری و همکاران ۲۰۱۴). کاهش r^2 در فواصل بین ۱۰ تا ۱۰۰ کیلوبازی نیز مورد بررسی قرار گرفت. نتایج نشان می‌دهند که متوسط r^2 در این فاصله از محدوده $0/365$ - $0/294$ (در پنجره ۱۰ تا ۲۰ کیلوبازی) تا محدوده $0/159$ - $0/239$ (در پنجره ۹۰ تا ۱۰۰ کیلوبازی) در تغییر بوده است (جدول ۱).

جدول ۲ اطلاعات مربوط به تغییرات اندازه موثر جمعیت را بر مبنای میزان عدم تعادل پیوستگی در فواصل مختلف ژنومی نمایش داده است. روند تغییرات اندازه موثر جمعیت در گاوهای بومی فارس نشان می‌دهد که در فاصله زمانی ۲۰۰۰ نسل قبل تا ۵ نسل قبل، اندازه موثر جمعیت از ۳۰۷۶ راس به ۱۴ راس کاهش یافته است. نقشه-های LD می‌توانند جهت بررسی پیشینه تاریخی نژادها و تنوع بین آنها مورد استفاده باشند. در این مطالعه متوسط r^2 در فواصل کمتر از ۱ Kb برابر با $0/46$ بدست آمد. همچنین متوسط r^2 بین SNP‌هایی که در فاصله کمتر از ۱۰ Kb از یکدیگر قرار داشتند، برابر با $0/35$ بود. سطح LD در فواصل کمتر از ۱۰ Kb در مطالعات انجام شده بر روی نژادهای نلور (اسپیلجولان و همکاران ۲۰۱۳) و فریزین-هلشتاین استرالیایی (ختکار و همکاران ۲۰۰۸) به ترتیب برابر با $0/34$ و $0/76$ گزارش شده است. همچنین

جدول ۱- اطلاعات آماری مربوط به میانگین r^2 بر حسب فاصله بین جفت SNP ها در فواصل مختلف بین ۰ تا ۱ Mb در ژنوم گاوهای بومی فارس

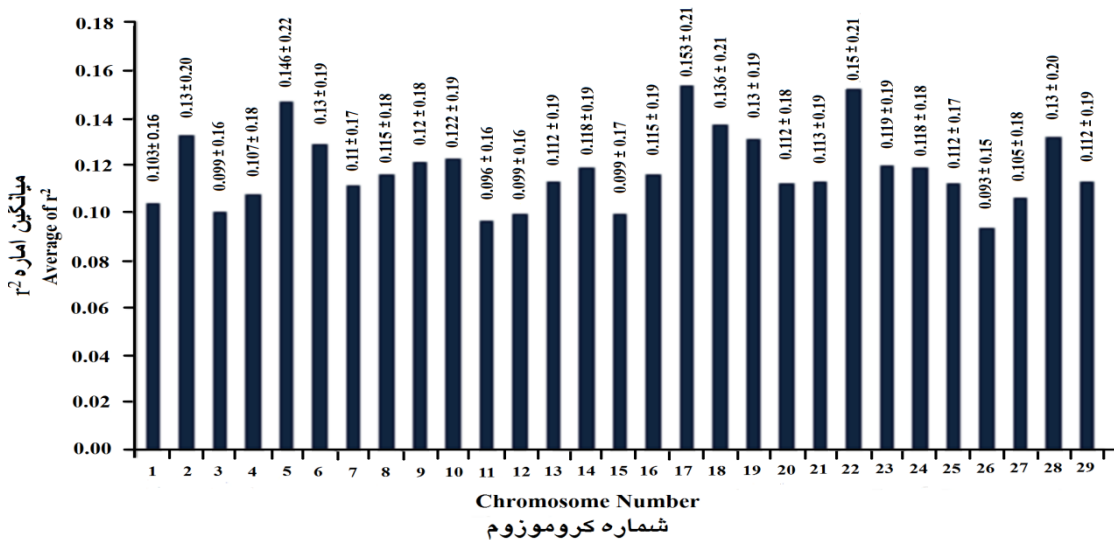
Table 1- Statistical information of average r^2 based on SNP pair distances for various intervals from 0 to 1 Mb in Fars native cattle genome

فاصله (کیلوبازی) Distance (Kb)	میانگین تعداد جفت مقایسه‌ها Average number of pairwise comparisons	میانگین r^2 Average of r^2	انحراف معیار r^2 Standard deviation of r^2	درصد r^2 های بالاتر از ۰/۳ Percent of $r^2 > 0.3$
0-1	25	0.495	0.374	57.43
1-2	47	0.442	0.379	52.06
2-3	48	0.390	0.373	49.20
3-4	53	0.355	0.365	44.95
4-5	50	0.348	0.363	43.88
5-6	49	0.318	0.354	40.92
6-7	48	0.320	0.352	42.18
7-8	49	0.324	0.346	41.83
8-9	47	0.301	0.340	38.83
9-10	47	0.290	0.337	38.17
10-20	460	0.277	0.333	36.94
20-30	441	0.245	0.320	32.46
30-40	428	0.224	0.307	30.72
40-50	438	0.206	0.295	28.25
50-60	427	0.201	0.293	27.48
60-70	423	0.192	0.286	26.39
70-80	419	0.189	0.280	26.12
80-90	419	0.187	0.281	25.85
90-100	409	0.176	0.271	24.78
100-200	4097	0.162	0.248	22.60
200-300	4026	0.146	0.248	20.39
300-400	3950	0.139	0.241	19.63
400-500	3938	0.134	0.236	19.00
500-600	3904	0.130	0.234	18.47
600-700	3864	0.129	0.232	18.30
700-800	3876	0.129	0.232	18.25
800-900	3848	0.128	0.230	18.25
900-1000	3847	0.127	0.230	18.00

جدول ۲- نمایش تعداد نسل‌های پیشین و اندازه موثر جمعیت بر اساس فواصل فیزیکی بین نشانگرها در گاوهای بومی فارس

Table 2- Representation of ancient generations number and effective population sizes based on physical distances between markers in Fars native cattle

	فاصله بین جفت SNPها Distance between pairs of SNPs						
	25 Kb	50 Kb	100 Kb	500 Kb	1 Mb	5 Mb	10 Mb
فاصله ژنتیکی (مورگان) Genetic distance (Morgan)	0.00025	0.0005	0.001	0.005	0.01	0.05	0.1
تعداد نسل‌های پیشین Number of ancient generations	2000	1000	500	100	50	10	5
اندازه موثر جمعیت Effective population size	3076	1950	1228	327	171	26	14



شکل ۲- مقایسه میزان عدم تعادل پیوستگی در بین کروموزوم‌های مختلف در جمعیت گاو بومی فارس

Figure 2- Comparison of linkage disequilibrium levels among different chromosomes in Fars native cattle population

(موکری و همکاران ۲۰۱۴) و سمینتال چینی (۰/۰۵-۰/۰۸) (زو و همکاران ۲۰۱۳)، چنین به نظر می‌رسد که اعداد بدست آمده در این محدوده بالاتر از حدود معمول برآورد شده‌اند. هر چند مطالعات شامل اندازه نمونه‌های کوچک در معرض اریب و خطای برآورد بیشتری قرار دارند، اما میزان اریب بر حسب فاصله بین نشانگرها متغیر است (ختکار و همکاران ۲۰۰۸). به طور قطع، برآوردهای بدست آمده در این مطالعه در فواصل بالاتر از ۱۰۰ کیلوباز تحت تاثیر تعداد کم نمونه، بالاتر برآورد شده‌اند. با این حال، اعداد بدست آمده در فواصل بین نشانگری کوتاه‌تر (کمتر از ۱۰۰ Kb) از اطمینان بیشتری برخوردار هستند. به طور کلی، مقایسه I^2 در بین مطالعات مختلف امری مشکل است، زیرا میزان I^2 گزارش شده در هر یک از این مطالعات تا حدودی به اندازه نمونه، آستانه MAF، تراکم نشانگری مورد استفاده و فاصله بین نشانگرها وابسته است. مقادیر I^2 بدست آمده در این تحقیق، علیرغم اثر اریب ناشی از تعداد کم نمونه، با سطح بالای همخوانی و کاهش جمعیت شدید مشاهده شده در گاوهای بومی فارس در انطباق هستند. بنابراین، برآورد مقادیر بالا I^2 را تا حدودی نیز می‌توان به کوچک بودن اندازه موثر این جمعیت نسبت داد. گاوهای بومی فارس در طی سالیان اخیر تحت تاثیر

در بررسی ۱۹ نژاد گاو نشان داده شد که در محدوده بین ۹۵ تا ۱۰۰ کیلوبازی، کمترین سطح LD در نژادهای پیدمانتس (۰/۰۸۵)، شیکو (۰/۱۰۴) و شاروله (۰/۱۰۵) وجود داشته است، در حالی که بالاترین سطوح LD در نژادهای هرفورد (۰/۲۲۲)، جرسی (۰/۲۰۱) و براون سوئیس (۰/۱۷۷) دیده شد (سالمون-تورس و همکاران ۲۰۱۴).

علاوه بر این، در همین فاصله سطوح LD در محدوده‌های ۰/۰-۱۱/۲۳ (نلور) (اسپیلجولان و همکاران ۲۰۱۳)، ۰/۰-۱۵۳/۴۰۲ (هلشتاین-فریزین استرالیا) (ختکار و همکاران ۲۰۰۸)، ۰/۱۳-۰/۲۷ (کانشیم) (پرز ابرین و همکاران ۲۰۱۴) و ۰/۱۶-۰/۳ (سمینتال چینی) (زو و همکاران ۲۰۱۳) گزارش شده است. مقایسه این نتایج نشان می‌دهد که اعداد بدست آمده در این مطالعه در دو دسته فاصله‌ای نخست در محدوده منطقی برای I^2 قرار گرفته‌اند. با این حال، متوسط I^2 در سومین محدوده فاصله‌ای مورد بررسی (۱۰۰ تا ۱۰۰۰ کیلوبازی) برابر با ۰/۱۶۲ تا ۰/۱۲۹ بوده است. در مقایسه با نتایج حاصله در نژاد فریزین-هلشتاین استرالیایی (۰/۰۵۷-۱/۰۸) (ختکار و همکاران ۲۰۰۸)، گاوهای بومی سوئد (۰/۰-۰/۶۱۴) (فلوری و همکاران ۲۰۱۰)، نژاد کانشیم (۰/۰-۰/۷/۱)

کیلوبازی و کمتر از ۶۰ کیلوبازی دیده شدند. از سوی دیگر، متوسط فاصله نشانگرهای مورد استفاده برابر با ۴۸/۱ کیلوباز بود. بنابراین، بر اساس این تراکم از نشانگرها، دستیابی به فواصل به طور متوسط کمتر از ۶۰ کیلوباز بین نشاگرها امکان پذیر است، اما فواصل بین نشانگرها به طور متوسط بیش از ۹ کیلوباز خواهند بود. بنابراین، بر اساس استدلال‌های برگرفته از مطالعات مورد اشاره، تراکم نشانگری به کار برده شده در این مطالعه (۵۵۷۱۸ نشانگر) برای بدست آوردن دقت کافی در مطالعات انتخاب ژنومی در گاوهای بومی فارس کافی است، اما قادر به کسب دقت کافی در مطالعات ارتباطی کل ژنوم (GWAS) نخواهد بود.

نتیجه‌گیری کلی

نتایج این تحقیق نشان داد که روند تغییرات سطوح LD در فواصل بین نشانگری کوتاه‌تر دارای نوسان بیشتری بوده است. با وجود اثر اریب ناشی از تعداد کم نمونه، مقادیر r^2 مشاهده شده در این تحقیق حاکی از بروز همخونی و کاهش شدید اندازه جمعیت در گاوهای بومی فارس هستند. به همین جهت، تدوین برنامه‌هایی جهت حفاظت از این توده ژنتیکی ضروری به نظر می‌رسد.

خشکسالی، عدم قدرت رقابت اقتصادی با نژادهای خارجی و آمیخته‌گری‌های کنترل نشده دچار افت جمعیتی محسوسی شده‌اند. کند شدن روند کاهش LD در نسل‌های متاخرتر نشان می‌دهد که این توده کاهش جمعیت شدیدی را در طی نسل‌های اخیر تجربه کرده است. به روشنی می‌توان گفت در حال حاضر برنامه‌ریزی جهت جلوگیری از انقراض گاوهای بومی اندک باقی‌مانده ضروری است. فرض وجود عدم تعادل پیوستگی میان نشانگرها و جایگاه‌های ژنی مرتبط با صفات کمی اساس شکل‌گیری مطالعاتی نظیر انتخاب ژنومی، مطالعات ارتباطی کل ژنوم (GWAS)، نقشه‌یابی جایگاه‌های صفات کمی و... بوده است. تراکم نشانگرهای مورد نیاز جهت اجرای این گونه مطالعات وابسته به سطوح LD در جمعیت‌های مورد بررسی است. در همین راستا، مطالعاتی جهت تعیین میزان r^2 مورد نیاز جهت کسب دقت کافی در این گونه مطالعات انجام شده است. میوسن و همکاران (۲۰۰۱) سطوح LD بالاتر از ۰/۲ را جهت دستیابی به حداقل دقت ۰/۸۵ در برآورد ارزش‌های اصلاحی ژنومی لازم دانسته‌اند. همچنین LD مفید جهت دستیابی به قدرت آماری کافی در مطالعات GWAS بالاتر از ۰/۳ ذکر شده است (آردلی و همکاران ۲۰۰۱). در مطالعه حاضر، مقادیر LD بالاتر از ۰/۳ و ۰/۲ به ترتیب در فواصل نشانگری کمتر از ۹

منابع مورد استفاده

- Ardlie KG, Kruglyak L and Seielstad M, 2002. Patterns of linkage disequilibrium in the human genome. *Nature Reviews Genetics* 3: 299-309.
- Conrad DF, Coop MG, Wen X, Wall JD, Rosenberg NA and Pritchard JK, 2006. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nature Genetics* 38: 1251-1260.
- Espigolan R, Baldi F, Boligon AA, Souza FR, Gordo DG, Tonussi RL, Oliveira HN, Tonhati H, Sargolzaei M, Schenkel FS, Carvalheiro R, Ferro JA and Albuquerque LG, 2013. Study of whole genome linkage disequilibrium in Nellore cattle. *BMC Genomics* 14: 305.
- Farnir F, Coppieters W, Arranz J, Berzi P, Cambisano N, Grisart B, Karim L, Marcq F, Moreau L, Mni M, Nezer C, Simon P, Vanmanshoven P, Wagenaar W and Georges M, 2000. Extensive genome-wide linkage disequilibrium in Cattle. *Genome Research* 10: 220-227.
- Flury C, Tapio M, Sonstegard T, Drogemuller C, Leeb T, Simianer H, Hanotte O and Rieder S, 2010. Effective population size of an indigenous Swiss cattle breed estimated from linkage disequilibrium. *Journal of Animal Breeding and Genetics* 127(5): 339-347.

- Gouveia JJS, Da Silva MGBV, Paiva SR and de Oliveira SMP, 2014. Identification of selection signatures in livestock species. *Genetics and Molecular Biology* 37(2): 330-342.
- Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME, 2009. Invited review: Genomic selection in dairy cattle: progress and challenges. *Journal of Dairy Science* 92(2): 433-443.
- Hill WG, 1974. Estimation of linkage disequilibrium in randomly mating populations. *Heredity* 33: 229-239.
- Hill WG and Weir BS, 1994. Maximum likelihood estimation of gene location by linkage disequilibrium. *American Journal of Human Genetics* 54: 705-714.
- Khatkar MS, Nicholas FW, Collins AR, Zenger KR, Cavanagh JA, Barris Schnabel W, Taylor RDJF and Raadsma HW, 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC Genomics* 9: 187.
- McKay SD, Schnabel RD, Murdoch BM, Matukumalli LK, Aerts J, Coppeters W, Crews D, Dias Neto E, Gill CA, Gao C, Mannen H, Stothard P, Wang Z, Van Tassell CP, Williams JL, Taylor JF and Moore SS, 2007. Whole genome linkage disequilibrium maps in cattle. *BMC Genetics* 8: 74.
- Meadows JR, Chan EK and Kijas JW, 2008. Linkage disequilibrium compared between five populations of domestic sheep. *BMC Genetics* 9: 61.
- Meuwissen TH, Hayes BJ and Goddard ME, 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157(4): 1819-1829.
- Mokry F, Buzanskas M, de Alvarenga Mudadu M, do Amaral Grossi D, Higa R, Ventura R, de Lima A, Sargolzaei M, Conceição Meirelles S, Schenkel F, da Silva M, Méo Niciura S, de Alencar M, Munar D and de Almeida Regitano L, 2014. Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC Genomics* 15(Suppl 7): S6.
- Pérez O'Brien AM, Mészáros G, Utsunomiya YT, Sonstegard TS, Garcia JF, Van Tassell CP, Carneiro R, da Silva MVB and Sölkner J, 2014. Linkage disequilibrium levels in *Bos indicus* and *Bos taurus* cattle using medium and high density SNP chip data and different minor allele frequency distributions. *Livestock Science* 166: 121-132.
- Porto-Neto LR, Kijas JW and Reverter A, 2014. The extent of linkage disequilibrium in beef cattle breeds using high-density SNP genotypes. *Genetics Selection Evolution* 46: 22 .
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ and Sham PC, 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81(3): 559-575.
- Salomon-Torres R, Matukumalli LK, Van Tassell CP, Villa-Angulo C, Gonzalez-Vizcarra VM and Villa-Angulo R, 2014. High density LD-based structural variations analysis in cattle genome. *PLoS One* 9(7): e103046.
- Sargolzaei M, Schenkel FS and VanRaden P, 2009. *gebv*: Genomic breeding value estimator for livestock. Technical Report to Dairy Cattle Breed Genetic Community, University of Guelph, Canada, 1-7.
- Sellner EM, Kim JW, McClure MC, Taylor KH, Schnabel RD and Taylor JF, 2007. Board-invited review: Applications of genomic information in livestock. *Journal of Animal Science* 85(12): 3148-3158.
- Sen S and Burmeister M, 2008. Hardy-Weinberg analysis of a large set of published association studies reveals genotyping error and a deficit of heterozygotes across multiple loci. *Human Genomics* 3: 36-52.
- Sved JA, 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theoretical Population Biology* 2: 125-141.
- Villa-Angulo R, Matukumalli LK, Gill CA, Choi J, Van Tassell CP and Grefenstette JJ, 2009. High-resolution haplotype block structure in the cattle genome. *BMC Genetics* 10: 19.
- Zhu M, Zhu B, Wang YH, Wu Y, Xu L, Guo LP, Yuan ZR, Zhang LP, Gao X, Gao HJ, Xu SZ and Li JY, 2013. Linkage disequilibrium estimation of Chinese beef Simmental cattle using high-density SNP panels. *Asian-Australasian Journal of Animal Science* 26(6): 772-779.

Linkage disequilibrium levels in Fars province native cattle population using high-density SNP data

K Karimi^{1,4*}, A Esmaili Zadeh Koshkuieh² and M Asadi Fuzi³

Received: December 27, 2015 Accepted: May 23, 2016

¹PhD Graduate, Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman and Member of Young Researchers Society, Shahid Bahonar University of Kerman, Kerman, Iran

²Professor and Associate Professor, respectively, Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran

*Corresponding author: Karim.Karimi81@gmail.com

Introduction: Knowledge on the extent of linkage disequilibrium (LD) among populations is a useful tool to study the evolutionary history of populations and to detect the genomic regions associated with economically important traits. In recent years, high dense SNP data has been widely applying as the standard tools in LD analysis of livestock populations. In the current study, LD levels in the genome of Fars native cattle were investigated in order to provide basic information required to design genome-wide association studies and to survey on probable changes in the effective size of this population. Moreover, the applicability of the modern genomic technologies such as genomic selection can be evaluated using LD data.

Material and methods: Ten individuals were randomly sampled from Fars native cattle population and were genotyped using Illumina Bovine HD beadchip for 777962 SNP markers. Data quality control was performed based on genotyping call rate, deviation from Hardy-Weinberg equilibrium and minor allele frequency. Thereby, a total of 55,718 SNPs were remained for further analyses. MAF was calculated using PLINK v1.07 for all autosomal SNPs and the distribution of the allelic frequencies was graphed as the proportion of SNPs represented in 6 different categories of MAF: <0.05 , ≥ 0.05 to <0.1 , ≥ 0.1 to <0.2 , ≥ 0.2 to <0.3 , ≥ 0.3 to <0.4 and ≥ 0.4 to ≤ 0.5 . The measure of LD (r^2) was calculated for all marker pairs of each chromosome (syntenic SNPs) using the SnppldHD software. SNP pairwise comparisons were grouped in the three distance sets including zero to 10 Kb, 10 Kb to 100 Kb and 100 Kb to 1000 Kb intervals and average r^2 was computed for all of the autosomes at these distances.

Results and discussion: Number of SNPs located on X, Y, mitochondrial chromosomes and unknown sites were respectively equal to 39367, 1224, 343 and 1735 and all of those were excluded from further analysis. Average distance between adjacent SNPs was equal to 48.1 Kb and the average of MAF was equal to 0.223. Also, 44.5% of sites had MAF < 0.2 and the highest proportion of SNPs was observed at category with MAF 0.4 to 0.5. In this study, average r^2 at distances less than 1 Kb was equal to 0.46. Moreover, average r^2 between SNPs with distances less than 10 Kb was equal to 0.35. Average r^2 values above 0.3 and 0.2 were observed at distances less than 9 Kb and 60 Kb, respectively. Survey on LD decay in 10 to 100 Kb distances indicated that based on 10 to 20 Kb SNP bins, average r^2 changed from 0.294 to 0.365. The average r^2 values obtained at shorter distances (less than 100 Kb) had higher reliabilities than those obtained at longer distances. The pattern of LD was significantly different among various chromosomes. LD decay in distances more than 400 Kb was very slow and reached to a relatively stable level. Average r^2 decreased with increase in physical distance between SNPs. However, change levels were not the same for different chromosomes. This can be attributed to different recombination rates and selection intensities between chromosomes. Average r^2 for all chromosomes was equal to 0.118 and the highest level of r^2 was observed in chromosome 17 ($r^2=0.153$) and the lowest one was in chromosome 26 ($r^2=0.093$). Change trend in effective population size of Fars native cattle indicated that during 2000 to 5 previous generations, effective population size was decreased from 3076 to 14 in this population. Slower LD decay in more

recent generations indicated that this population has experienced a severe decline in recent generations. Clearly, at this situation, design a conservation program is essential to avoid extinction of Fars native cattle. Markers density applied in this study (55,718) will be appropriate to obtain enough accuracy in genomic selection studies in Fars native cattle. However, this marker density will not be able to get enough accuracy in Genome Wide Association Studies (GWAS).

Conclusion: Results from this study can be applied to determine the optimum marker density required to achieve enough accuracy in the possible genomic studies on the Fars native cattle. Due to recent intense inbreeding and considerable reduction in effective population size of Fars native cattle population, the high r^2 values found in this study were as expected.

Key words: Bovine genome, Fars native cattle, Linkage disequilibrium, Single nucleotide polymorphism