تحقیقات تولیدات دامی
سال چهارم/شماره سوم/پاییز ۱۳۹۴ (۱۰۴-۹۳)

آنالیز ساختار ژنتیکی جمعیت گاوهای بومی ایران با استفاده از نشانگرهای متراکم چندشکل تک‌نوکلئوتیدی

کریم کریمی[۱،۴]، علی اسماعیلی زاده کشکوئیه[۲*]، مسعود اسدی فوزی[۳]

۱- دانشجوی دکترای ژنتیک و اصلاح نژاد دام، گروه علوم دامی، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان
۲- استاد گروه علوم دامی، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان
۳- دانشیار گروه علوم دامی، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان
۴- عضو انجمن پژوهشگران جوان، دانشگاه شهید باهنر کرمان

(تاریخ دریافت: ۹۳/۹/۱ - تاریخ پذیرش: ۹۴/۶/۲۱)

چکیده

این تحقیق با هدف بررسی ساختار ژنتیکی گاوهای بومی بر اساس استفاده از نشانگرهای چندشکل تک‌نوکلئوتیدی انجام شد. بدین منظور تعداد ۹۰ راس گاو از نژادهای سرابی (۲۰)، کردی (۱۰)، مازندرانی (۱۰)، تالشی (۱۰)، سیستانی (۱۰)، کرمانی (۱۰)، نجدی (۱۰) و توده بومی فارس (۱۰) مورد نمونه‌برداری قرار گرفتند. تعیین ژنوتیپ نمونه‌ها با استفاده از چیپ Illumina High density Bovine BeadChip با تراکم ۷۷۷۹۶۲ جایگاه انجام شد. پس از حذف جایگاه‌های دارای عدم تعادل پیوستگی (جفت جایگاه‌های دارای r^2 بالاتر از ۰/۲)، ۶۴۳۳۳ جایگاه SNP جهت آنالیزهای بعدی باقی ماندند. برنامه STRUCTURE جهت بررسی ساختار ژنتیکی گاوهای بومی کشور در این مجموعه داده به کار گرفته شد. تعداد جمعیت‌های فرض شده در هر اجرا از دو تا هشت جمعیت در نظر گرفته شد. نتایج حاصل از این تحقیق نشان داد که در اولین سطح خوشه‌بندی (K=۲)، جمعیت‌های سرابی (۹۷/۶٪) و کردی (۹۴/۳٪) در خوشه مشترکی قرار گرفته‌اند و نژاد سیستانی (۹۲٪) نیز خوشه متمایز دیگری را به خود اختصاص داده است، در حالی‌که سایر نژادها آمیخته‌ای از این دو خوشه بودند. تعداد سه خوشه (K=۳) به بهترین شکل تعداد گروه‌های ژنتیکی موجود در مجموعه داده‌ها را توجیه کرد. وجود سه خوشه (K=۳) در داده‌ها موجب قرار گرفتن نژاد کردی در یک خوشه مستقل شد. در این مطالعه تفاوت‌های ژنتیکی گاوهای بومی ایران به خوبی به وسیله داده‌های متراکم SNP شناسایی شدند. همچنین شباهت‌های مربوط به توزیع جغرافیایی و خصوصیات تولیدی گاوها نیز با روابط ژنتیکی یافت شده در این مطالعه انطباق داشتند.

واژه‌های کلیدی: استنباط بیزی، چندشکلی تک‌نوکلئوتیدی، ساختار ژنتیکی، گاوهای بومی ایران

* نویسندهٔ مسئول: aliesmaili@uk.ac.ir

مقدمه

همزمان با وقوع تغییرات گسترده در شرایط آب و هوایی زمین، لزوم تغییر نگرش در برنامه‌های اصلاح نژادی دام‌های اهلی بیشتر احساس می‌شود. کاهش تنوع ژنتیکی، افت ناشی از همخونی و بروز بیماری‌ها از جمله نگرانی‌های صنعت دام محسوب می‌شوند. این در حالی است که سازگاری حاصل از انتخاب طبیعی در طی سالیان متمادی موجب شده است که دام‌های بومی در مقایسه با نژادهای تجاری از مقاومت ژنتیکی و قدرت سازگاری بالاتری برخوردار باشند (Hoffmann, 2010). وجود این تنوع ژنتیکی لازمه‌ی اجرای اهداف اصلاح نژادی پیش‌رو و سازگاری سریع‌تر نسبت به تغییرات محیطی خواهد بود. به همین جهت حفاظت از منابع ژنتیکی دام‌های بومی دارای اهمیت بوده و تغییر نگرش جهانی در جهت حفاظت از دام-های بومی مناطق مختلف ضروری به نظر می‌رسد (Herrero-Medrano *et al.*, 2013). مطالعه وضعیت اکولوژیکی و تکاملی جمعیت‌ها به دلیل توزیع جغرافیایی غیرتصادفی آنها نیازمند شناخت صحیح ساختار ژنتیکی جمعیت‌ها است. همچنین شناسایی ساختار و تنوع ژنتیکی جمعیت‌ها گامی مهم در جهت توسعه برنامه‌های حفاظتی و مدیریت پایدار منابع ژنتیکی محسوب می‌شود (Groeneveld *et al.*, 2010).

خوشه‌بندی به گروه‌بندی مجموعه‌ای از افراد بر اساس یک یا چند ویژگی مشترک میان آنها گفته می‌شود. روش‌های خوشه‌بندی بر اساس داده‌های مولکولی را می‌توان به دو گروه عمده روش‌های مبتنی بر فاصله ژنتیکی و روش‌های مبتنی بر مدل تقسیم‌بندی کرد. در روش مبتنی بر فاصله ابتدا یک ماتریس فاصله ژنتیکی تشکیل شده و در مرحله بعد به کمک الگوریتم‌های مختلف، درخت فیلوژنتیک این ماتریس رسم می‌شود. این روش تا حد زیادی به معیار فاصله و الگوریتم ترسیم فواصل وابسته است و ارزیابی میزان اطمینان حاصل از خوشه‌بندی مشکل است (Franzen, 2008). در روش‌های مبتنی بر مدل، مشاهدات هر خوشه نمونه‌های تصادفی از یک مدل پارامتریک[1] محسوب می‌شوند و استنتاج پارامترهای هر خوشه با استفاده از روش‌های آماری استاندارد انجام می‌شود. روش آماری بیزی یکی از روش‌های بکار گرفته شده در خوشه‌بندی مبتنی بر مدل برای استنتاج ساختار جمعیت‌ها است. این روش چارچوب منسجمی را جهت در نظر گرفتن نامعینی ذاتی در برآورد پارامترها و ارزیابی قدرت خوشه‌بندی فراهم آورده است (Corander *et al.*, 2003). در این روش اطلاعات پیشین مربوط به جمعیت‌ها و فراوانی‌های آللی با اطلاعات موجود از ژنوتیپ افراد تلفیق شده و از آنها جهت بدست آوردن توزیع پسین جمعیت‌ها و فراوانی‌های آللی استفاده می‌شود. بنابراین امکان در نظر گرفتن اطلاعات اولیه در مورد جمعیت‌ها نظیر مکان‌های جغرافیایی آنها نیز فراهم است (Holsinger *et al.*, 2004). استنباط بیزی اولین بار به وسیله Symons (1981) جهت خوشه‌بندی مبتنی بر مدل مورد استفاده قرار گرفت و در حال حاضر به روشی استاندارد در تعیین ساختار ژنتیکی جمعیت‌ها تبدیل شده است.

الگوهای حاصل از داده‌های مولکولی می‌توانند جهت شناسایی ارتباط میان جمعیت‌های یک گونه بکار گرفته شوند. در سال‌های اخیر، نشانگرهای چندشکل تک-نوکلئوتیدی (SNP) در مطالعات ژنوم دام‌های اهلی به میزان زیادی مورد استفاده قرار گرفته‌اند. کارآیی این نشانگرها در پویش‌های سرتاسری ژنوم (Bolormaa *et al.*, 2010)، انتخاب ژنومی (Goddard *et al.*, 2007) و شناسایی نقشه-های ژنومی (McKay *et al.*, 2007) به اثبات رسیده است. SNPها فراوان‌ترین منبع تغییرات ژنتیکی داخل ژنوم محسوب می‌شوند و میزان جهش در آنها در مقایسه با سایر نشانگرها پایین است. همچنین قابلیت پوشش سرتاسری ژنوم به وسیله این نشانگرها به واسطه فناوری‌های ریزآرایه فراهم شده است (Gray *et al.*, 2000 and Helyar *et al.*, 2011). دسترسی به آرایه‌های SNP با تراکم بالا فرصت مناسبی را جهت بررسی تنوع و ساختار ژنتیکی جمعیت‌های کوچک فاقد شجره فراهم نموده است (Engelsma *et al.*, 2012). نشانگرهای چندشکل تک‌نوکلئوتیدی در بسیاری از حیوانات اهلی از جمله گاو (McKay *et al.*, 2007;

1. Parametric model

(Pertoldi *et al.*, 2010)، سگ (VonHoldt *et al.*, 2010) و اسب (Petersen *et al.*, 2013) به منظور شناسایی ساختار ژنتیکی جمعیت‌ها بر اساس روش بیزی بکار گرفته شده‌اند و به عنوان ابزاری استاندارد در این زمینه شناخته می‌شوند.

پرورش و اهلی کردن گاو در سرزمین ایران دارای منشا تاریخی است. گاوهای بومی ایران در مناطق جغرافیایی مختلف پراکنده شده‌اند و از تنوع بسیاری در صفات ظاهری خود برخوردار هستند. در حالی‌که نواحی سرد و کوهستانی شمال غرب و غرب کشور محل نگهداری نژادهایی چون سرابی و کردی است، مناطق معتدل سواحل دریای خزر نژادهای مازندرانی و تالشی را در خود جای داده‌اند و مناطق گرم جنوب کشور بستر پرورش توده‌هایی چون سیستانی، نجدی و.... می‌باشند (توکلیان، ۱۳۷۸). همزمان با ورود نژادهای پرتولید اروپایی نظیر هلشتاین به کشور، تلاقی گاوهای بومی با این نژادها جهت افزایش تولید مورد ترویج قرار گرفت. نبود برنامه مناسب جهت انجام این تلاقی‌ها، خلوص نژادهای داخلی را به شدت تحت تاثیر قرار داده است. اقتصادی نبودن میزان تولید، به طور خاص تولید شیر، در گاوهای بومی در مقایسه با نژادهای اروپایی موجب کاهش تمایل روستائیان به نگهداری این توده‌ها شده است و بروز خشکسالی‌های مکرر نیز شدت این امر را دوچندان کرده است. در حال حاضر منابع ژنتیکی گاوهای بومی کشور در معرض خطر جدی قرار دارند و لزوم توجه به حفاظت از گاوهای خالص باقی‌مانده امری ضروری به نظر می‌رسد. شناخت ساختار ژنتیکی و روابط موجود میان این نژادها نه تنها ارزش حفاظتی آنها را روشن‌تر می‌سازد بلکه می‌تواند جهت اولویت‌بندی برنامه‌های حفاظت ژنتیکی نیز بکار گرفته شود (FAO, 2007). از آنجا که تاکنون تحقیق جامعی در خصوص میزان خلوص و ساختار ژنتیکی جمعیت گاوهای بومی کشور صورت نگرفته است، این تحقیق با هدف بررسی ساختار ژنتیکی گاوهای بومی بر اساس استفاده از نشانگرهای چندشکل تک‌نوکلئوتیدی انجام شد. امید است نتایج این مطالعه در جهت حفاظت از توده‌های گاو بومی کشور به کار گرفته شوند.

مواد و روش‌ها

به منظور بررسی ساختار جمعیت گاوهای بومی کشور، تعداد ۹۰ راس گاو از نژادهای سرابی (۲۰)، کردی (۱۰)، مازندرانی (۱۰)، تالشی (۱۰)، سیستانی (۱۰)، کرمانی (۱۰)، نجدی (۱۰) و توده بومی فارس (۱۰) مورد نمونه‌برداری قرار گرفتند. بر اساس داده‌های شجره‌ای و یا با پرسش از صاحبان دام‌ها، نمونه‌هایی مورد انتخاب قرار گرفتند که دارای رابطه خویشاوندی نبوده و یا روابط خویشاوندی اندکی داشتند. نمونه‌های مو از ناحیه دم گاوها جدا شدند و در بسته‌های جداگانه به آزمایشگاه منتقل شدند. استخراج دی.ان.ای از ریشه‌های مو انجام شد. نمونه‌های دی.ان.ای با استفاده از Illumina High-density Bovine BeadChip (Illumina, Inc, San Diego, CA, USA) شامل ۷۷۷۹۶۲ جایگاه SNP تعیین ژنوتیپ شدند. یک نمونه مربوط به نژاد نجدی به دلیل عدم کیفیت کافی تعیین ژنوتیپ از مجموعه داده‌ها حذف شد و در مجموع تعداد ۸۹ نمونه از گاوهای بومی ایران در ۷۷۷۹۶۲ جایگاه SNP مورد استفاده قرار گرفت. به منظور انجام کنترل کیفیت روی داده‌ها از نرم-افزار PLINK 1.07 (Purcell *et al.*, 2007) استفاده شد. کنترل داده‌ها بر اساس کیفیت نمونه‌ها، جایگاه‌ها و مکان-های فیزیکی انجام شد. تنها جایگاه‌های واقع روی کروموزوم‌های اتوزوم مورد انتخاب قرار داده شدند. خلاصه‌ای از معیارهای مورد استفاده جهت اجرای کنترل کیفی روی داده‌ها و تعداد جایگاه‌های حذف شده به ازای هر یک از آستانه‌های اعمال شده در جدول ۱ ارائه شده است. پس از انجام کنترل کیفیت روی مجموعه داده‌ها، تعداد ۲۸۱۳۳۳ جایگاه SNP باقی ماند و در آنالیزهای بعدی مورد استفاده قرار گرفت.

اختصاص مدل مناسب برای مشاهدات هر خوشه (جمعیت) اولین چالش در روش‌های خوشه‌بندی مبتنی بر مدل است. در اینجا فرض شده است که هر خوشه به وسیله یک مجموعه مشخص از فراوانی‌های آللی مدل‌بندی شده است. فرضیات اساسی مدل عبارتند از: وجود تعادل هاردی-وینبرگ داخل جمعیت‌ها و تعادل پیوستگی کامل میان جایگاه‌ها. تحت این فرضیات هر آلل در هر جایگاه و در هر ژنوتیپ یک نمونه مستقل از یک توزیع فراوانی خاص در نظر

گرفته می‌شود. اگر X بردار ژنوتیپ افراد نمونه‌گیری شده باشد و بردار جمعیت‌های منشا افراد (ناشناخته) و بردار فراوانی‌های آللی جمعیت‌ها (ناشناخته) نیز به‌ترتیب با Z و P نشان داده شود، به کمک روش بیز می‌توان توزیع پسین P و Z را به شکل زیر بدست آورد:

$$pr(Z,P|X) \propto pr(Z)\,pr(P)pr(X|Z,P)$$

در این معادله pr(Z) توزیع پیشین اطلاعات جمعیت‌ها (به‌طور پیش‌فرض می‌توان سهم مساوی برای هر یک از جمعیت‌ها در نظر گرفت) و pr(P) توزیع پیشین فراوانی‌های آللی (به‌طور پیش‌فرض می‌توان از توزیع دیریخله[1] استفاده کرد) را نشان می‌دهند و pr(X|Z,P) اطلاعات موجود از ژنوتیپ افراد را نمایش می‌دهد. در مدل آمیختگی می‌توان اطلاعات پیشین نسبت آمیختگی هر فرد را نیز در محاسبات وارد کرد. به منظور بدست آوردن تقریبی از توزیع پسین، از زنجیره‌های مونت کارلو مارکوف (MCMC) جهت گرفتن نمونه‌هایی از توزیع پسین استفاده می‌شود و استنباط‌های P و Z بر اساس خلاصه آماری حاصل از این نمونه‌ها ارائه می‌شود (Pritchard *et al.*, 2000; Sethuraman, 2013). به منظور حذف SNPهایی که در حالت عدم تعادل پیوستگی با یکدیگر قرار داشتند، تراکم مجموعه داده‌های حاصل از کنترل کیفیت با استفاده از نرم‌افزار PLINK 1.07 کاهش داده شد. بدین منظور در پنجره‌هایی شامل ۵۰ SNP و با حرکت ۱۰ SNP رو به جلو در هر مرحله، SNPهای دارای r^2 (معیار عدم تعادل پیوستگی) بیش از ۰/۲ با یکدیگر از مجموعه داده‌ها حذف شدند. پس از حذف جایگاه‌های دارای عدم تعادل پیوستگی، تعداد ۶۴۳۳۳ SNP جهت آنالیز ساختار ژنتیکی مورد استفاده قرار گرفتند. نرم‌افزار STRUCTURE 2.3.4 (Pritchard *et al.*, 2000) جهت خوشه‌بندی جمعیت‌ها به گروه‌های ژنتیکی مورد استفاده قرار گرفت. مدل آمیختگی و فراوانی‌های آللی تصحیح شده در نرم‌افزار STRUCTURE جهت آنالیز ساختار داده‌ها مورد انتخاب قرار گرفتند. استنتاج تعداد خوشه‌ها به کمک اجرای مستقل برنامه به تعداد ۴ بار انجام شد. در هر بار اجرای برنامه تعداد ۳۵۰۰۰ تکرار به همراه ۱۵۰۰۰ دور سوختن برای مقادیر مختلف K (تعداد جمعیت‌های فرض شده) بکار گرفته شد. ارزیابی رسیدن به نقطه همگرایی در نرم‌افزار به کمک بررسی مقادیر آلفا و مقادیر درست‌نمایی (که به وسیله نرم‌افزار در خروجی‌های هر تکرار ارائه می‌شوند) انجام شد. مناسب‌ترین تعداد جمعیت فرض شده جهت تشریح گروه‌های ژنتیکی موجود در داده‌ها به کمک روش Evanno (Evanno *et al.*, 2005) انتخاب شد. بر اساس این روش از لگاریتم درست‌نمایی داده‌ها جهت تعیین مقدار حقیقی K (تعداد جمعیت‌های فرض شده) استفاده می‌شود. در ابتدا از درست‌نمایی داده‌ها (L(K)) (که در خروجی نرم‌افزار داده شده است) در اجراهای مختلف میانگین گرفته شد. در مرحله دوم میزان تغییر تابع درست‌نمایی نسبت به K بدست آورده شد:

$$L'(K) = |L(K) \quad L(K \quad 1)|$$

و پس از این، مرتبه دوم تغییر L(K) نسبت به K نیز محاسبه شد:

$$\acute{L}(K) = |\acute{L}(K+1) - \acute{L}(K)|$$

در نهایت مقدارΔK با تقسیم میانگین مقادیر مرتبه دوم تغییرL(K) نسبت به K بر انحراف استاندارد مقادیر درست‌نمایی بدست آورده شد:

$$\Delta K = \frac{m|\acute{L}(K)|}{s(L(K)}$$

مقداری از K که به ازای آن مقدار ΔK حداکثر شد به عنوان بهترین مقدار K برگزیده شد.

1. Dirichelt distribution

جدول ۱- معیارهای مورد استفاده در کنترل کیفیت داده‌ها و تعداد جایگاه‌های حذف شده به ازای هر معیار

Table 1. Measures of genotype quality control and number of SNPs removed per measure

Quality control thresholds	Number of SNPs removed per threshold
SNPs located on the X chromosome	39367
SNPs located on the Y chromosome	1224
SNPs located on the Mitochondrial DNA	343
Call rates per marker < 90%	445091
Minor allele frequency (MAF) < 0.01	33152
Markers deviated from H-W[a] equilibrium (P-value<10^{-7})	2159
Remained SNPs	281333

[a] Hardy-Weinberg

نتایج و بحث

مقادیر هتروزیگوسیتی مشاهده شده، هتروزیگوسیتی مورد انتظار و سطوح چندشکلی در هر یک از جمعیت‌های مورد مطالعه در جدول ۲ ارائه شده است. بالاترین سطح چندشکلی در بین نژادهای مورد بررسی در نژاد سرابی مشاهده شد، در حالی که کمترین میزان آن مربوط به نژاد سیستانی بود. کمترین مقدار هتروزیگوسیتی مورد انتظار (۰/۲۴۶) نیز در نژاد سیستانی مشاهده شد. از سوی دیگر، بالاترین مقادیر هتروزیگوسیتی مورد انتظار در نژادهای کردی (۰/۳۳۹) و سرابی (۰/۳۳۸) بدست آمد. این نتایج نشان‌دهنده تطبیق مناسب میان مقادیر هتروریگوسیتی با سطوح چندشکلی مشاهده شده در جمعیت‌ها بود. برنامه STRUCTURE (مبتنی بر استنباط بیزی) جهت بررسی ساختار ژنتیکی و آمیختگی در مجموعه داده‌های گاوهای بومی ایران شامل ۶۴۳۳۳ جایگاه بکار گرفته شد. نتایج حاصل از خوشه‌بندی برای هشت مقدار مختلف K در هشت نژاد گاو بومی در شکل ۱ نمایش داده شده است. در K=۲ (فرض شده است که وجود تنها دو جمعیت تمامی تغییرات موجود در داده‌ها را توجیه می‌کند) نژادهای سرابی و کردی به ترتیب به میزان ۹۷/۶٪ و ۹۴/۳٪ در یک خوشه سهیم بودند. در حالی‌که نژاد سیستانی (۹۲٪) در یک خوشه متمایز قرار گرفته است و سایر نژادها آمیخته‌ای از این دو خوشه متمایز بوده‌اند. در K=۳ ، نژاد کردی در یک خوشه مستقل ظاهر شده است. نژادهای سرابی و سیستانی نیز همچنان در خوشه‌هایی متمایز جای دارند و سایر نژادها درصدهای آمیختگی متفاوتی از این سه خوشه دارا می‌باشند. در مقادیر K بالاتر از ۳، نشانه‌های آمیختگی در نژاد کردی ظاهر شده است در حالیکه نژادهای سیستانی و سرابی توانستند همچنان خلوص خود را حفظ کنند. تفاوت میان نژادهای تالشی و مازندرانی با سایر نژادها برای اولین بار در K= ۴ ظهور یافته است به طوری‌که به‌ترتیب ۷۹/۶٪ و ۶۸/۹٪ از افراد نژادهای تالشی و مازندرانی در K= ۴ به یک خوشه یکسان نسبت داده شده‌اند. تغییرات مقادیرΔK به ازای مقادیر مختلف K در شکل ۲ نمایش داده شده است. بیشترین مقدار ΔK در K=۳ بدست آمد. این امر نشان می‌دهد که تعداد ۳ خوشه به بهترین شکل تعداد گروه‌های ژنتیکی موجود در مجموعه داده‌ها را توجیه می‌کند. نسبت انتساب هر یک از نژادها به خوشه‌های مختلف در K=۳ در جدول ۳ نمایش داده شده است. هممچنین مقادیر مختلف F_{st} میان جمعیت‌های فرض شده در K=۳ در جدول ۴ نشان داده شده است. داده‌های این جدول نشان می‌دهند که فاصله ژنتیکی میان اولین و دومین خوشه استنباط شده (به‌ترتیب مربوط به نژادهای کردی و سرابی) در مقایسه با فاصله ژنتیکی خوشه‌های اول و دوم با خوشه سوم (نژاد سیستانی) کمتر بوده است. همچنین فاصله ژنتیکی نژاد کردی با نژاد سیستانی در مقایسه با فاصله نژاد سرابی با نژاد سیستانی بیشتر بوده است.

جدول ۲- سطح چندشکلی، هتروزیگوسیتی مشاهده شده و هتروزیگوسیتی مورد انتظار در هر یک از نژادهای گاو بومی ایران

Table 2. Level of polymorphism, observed heterozygosity and expected heterozygosity for each Iranian native cattle breed

Breed	Number of samples	% SNPs with MAF > 0.01	Observed heterozygosity	Expected heterozygosity
Sarabi	20	9.1	0.322	0.338
Taleshi	10	15.8	0.306	0.317
Mazandarani	10	12.8	0.320	0.326
Pars	10	16	0.291	0.314
Kermani	10	13.1	0.283	0.312
Najdi	9	16.8	0.295	0.307
Sistani	10	25.5	0.229	0.246
Kurdi	10	12.4	0.330	0.339
Total	89	15.2	0.297	0.312

تایورین فاقد کوهان (*Bos taurus*) و ایندیکاین کوهان‌دار (*Bos indicus*) دو گروه اصلی گاوهای امروزی را در جهان تشکیل داده‌اند. بر اساس خصوصیت ظاهری داشتن و یا نداشتن کوهان، فرض اولیه این تحقیق این بود که نژادهای سیستانی، مازندرانی، تالشی و نجدی در گروه بوس ایندیکوس قابل طبقه‌بندی هستند، در حالی‌که نژادهای کردی، سرابی، کرمانی و فارس بیشتر به گروه بوس تایروس شباهت داشتند. خوشه‌بندی به روش بیزی جهت بررسی این فرض بکار گرفته شد. نتایج حاصل از این تحقیق نشان داد که در K=۲، جمعیت‌های سرابی و کردی در خوشه مشترکی قرار گرفته‌اند و نژاد سیستانی نیز خوشه متمایز دیگری را به خود اختصاص داده است. سایر نژادها آمیخته‌ای از این دو خوشه بودند. نتایج حاصل از اولین سطح خوشه‌بندی (K=۲) منعکس‌کننده تقسیم‌بندی کلی گاوها به دو گروه تایورین و ایندیکاین است. در مجموع با توجه به سایر ویژگی‌ها، بر اساس نتایج این تحقیق می‌توان نژادهای سرابی و کردی را به عنوان نژادهای بوس تایروس و نژاد سیستانی را به عنوان نژاد بوس ایندیکوس معرفی کرد. سایر جمعیت‌ها نشانه‌هایی از آمیختگی حاصل از تلاقی تایورین و ایندیکاین را نمایان ساختند. ساختار جمعیتی هشت نژاد گاو از گروه‌های تایورین و ایندیکاین در مطالعه‌ای به کمک داده‌های چند شکل تک نوکلئوتیدی مورد بررسی قرار گرفت (McKay *et al.*, 2008). بالاترین سطح تفاوت‌های ژنتیکی مشاهده شده در این تحقیق میان نژادهای بوس تایروس و بوس ایندیکوس قابل شناسایی بود. به طور مشابه، در مطالعه‌ای دیگر از اطلاعات ۴۳۰۴۳ نشانگر چندشکل تک‌نوکلئوتیدی جهت بررسی الگوی آمیختگی و تفاوت‌های ژنتیکی ۱۳۴ نژاد مختلف گاو بهره گرفته شد (Decker *et al.*, 2014). نتایج این تحقیق نیز به خوبی موجب تمایز سه گروه اصلی شامل ایندیکاین آسیایی، تایوراین اروپایی و تایوراین آفریقایی در میان نژادهای مورد بررسی شد. همچنین در مطالعه‌ای که به وسیله Bovine (2009) HapMap انجام شد، نژادهای تایوراین و ایندکاین در اولین سطوح خوشه‌بندی به خوبی قابل تشخیص بودند. به‌طور جالب توجهی مشاهده شد که نژاد سرابی نقش چندانی در آمیختگی سایر جمعیت‌های گاو بومی کشور نداشته است و به‌صورت خوشه‌ای متمایز در سطوح مختلف خوشه‌بندی مشاهده شد. به نظر می‌رسد این نژاد به لحاظ منشا تاریخی با سایر نژادهای کشور دارای تفاوت باشد. هر چند که نژاد کردی در سطح K=۳ به عنوان نژادی خالص در یک خوشه متمایز قرار گرفته است، اما با افزایش تعداد جمعیت‌های فرض شده، درصد بالایی از آمیختگی در این نژاد مشاهده شده است.

نتایج حاصل از این مطالعه نشان‌دهنده مطابقت خوشه‌بندی ژنتیکی جمعیت‌ها با فواصل جغرافیایی و محیط‌های پرورش آنهاست. نژادهای کردی و سرابی که در مناطق سرد و کوهستانی ناحیه غرب و شمال غرب کشور نگهداری می‌شوند در اولین سطح خوشه بندی (K=۲) در خوشه یکسانی طبقه‌بندی شده‌اند و به لحاظ ژنتیکی نیز فاصله کمتری را در سطوح بالاتر خوشه‌بندی نشان داده‌اند. نتایج این مطالعه

نشان داد که در K=۴، ۷۹/۶ درصد از افراد نژاد تالشی و ۶۸/۹ درصد از افراد نژاد مازندرانی در انطباق با فاصله جغرافیایی کمتر میان آنها، در خوشه یکسانی طبقه‌بندی شده‌اند. هر دو نژاد در نواحی شمالی کشور در مناطق جنگلی شمال رشته کوه البرز نگهداری می‌شوند. همچنین توده‌های کرمانی، فارس و نجدی نیز که در نواحی گرم جنوب کشور نگهداری می‌شوند به میزان زیادی از آمیختگی با نژاد سیستانی متاثر شده‌اند و الگوی آمیختگی مشابهی را در سطوح مختلف خوشه‌بندی نشان داده‌اند. در بین این نژادها، کرمانی بیشترین تاثیر را از نژاد سیستانی داشته است که علت آن را می‌توان به فاصله جغرافیایی کمتر نواحی نگهداری این دو نژاد نسبت داد (شکل ۱). بر اساس ویژگی-های تولیدی نیز می‌توان انطباق مناسبی را بین خوشه‌بندی ایجاد شده در سطح اول خوشه‌بندی با خصوصیات عمده تولیدی نژادها مشاهده کرد. نژاد سیستانی که به عنوان تنها نژاد گوشتی کشور شناخته می‌شود در خوشه‌ای متمایز قرار گرفته است در حالی‌که در خوشه مقابل نژاد سرابی با ویژگی‌های برجسته تولید شیر قرار گرفته است.

در این مطالعه تعداد زیادی از جایگاه‌های SNP به دلیل نداشتن کیفیت لازم از مجموعه داده‌ها کنار گذاشته شده‌اند. به نظر می‌رسد بتوان این امر را بیشتر ناشی از اثر اریب ناشی از شیوه انتخاب[1] SNPها در تولید تراشه تجاری دانست. گاوهای بومی ایران در پروژه‌های تعیین توالی ژنوم گاو حضور نداشته‌اند. از سوی دیگر ژنوم مرجع مورد استفاده در این پروژه‌ها، ژنوم یک نژاد اروپایی بوده است و ژنوم گاوهای اروپایی بوس تایروس نسبت به سایر نژادها بیشتر توالی یابی شده است. بنابراین همان گونه که انتظار می‌رود فراوانی آلل نادر درSNPهای شناسایی شده در نژادهای اروپایی در مقایسه با نژادهای بوس ایندیکوس بالاتر است. سطح چند شکلی مشاهده شده در ژنوم گاوهای بوس تایروس نسبت به گاوهای بوس ایندیکوس در بسیاری از مطالعات مبتنی بر داده‌های SNP بیشتر گزارش شده است (Mustafa *et al.*, 2014; Lin *et al.*, 2010; Edea *et al.*, 2013). در میان نژادهای ایرانی پایین‌ترین سطح چندشکلی در نژاد سیستانی مشاهده شد که بخشی از آن را می‌توان به اثرات عمده بوس ایندیکوس در این نژاد نسبت داد. همچنین سطح بالاتر چندشکلی مشاهده شده در نژادهای سرابی و کردی نیز با ارتباط نزدیک‌تر آنها با نژادهای بوس تایروس همخوانی دارد. در این مطالعه سطح بالایی از آمیختگی در نژادهای ایرانی مشاهده شد. در حالی‌که نژادهای سرابی و کردی را می‌توان به گروه بوس تایروس و گاو سیستانی را به گروه بوس ایندیکوس نسبت داد، سایر نژادها درجات متفاوتی از آمیختگی را نشان می‌دهند. منشا آمیختگی گاوهای بومی ایران در این مطالعه به خوبی قابل تفسیر نیست. نزدیک بودن مناطق جغرافیایی پرورش گاوهای بومی و نگهداری این دام‌ها بر اساس سیستم چرا می‌تواند دلیلی بر منشا تاریخی آمیختگی در میان نژادهای گاو بومی کشور باشد. از سوی دیگر در طی پنجاه سال اخیر، نژادهای خارجی (عمدتاً هلشتاین) به میزان زیادی وارد کشور شده‌اند و با نژادهای بومی تلاقی داده شده‌اند. بنابراین بررسی نقش نژادهای خارجی نظیر هلشتاین و جرسی در آمیختگی گاوهای بومی کشور نیازمند مطالعه بیشتر است.

در این مطالعه تفاوت‌های ژنتیکی گاوهای بومی ایران به خوبی به وسیله داده‌های ژنوتیپی پرتراکم SNP شناسایی شد. شباهت‌های مربوط به توزیع جغرافیایی و خصوصیات تولیدی گاوها نیز به خوبی در روابط ژنتیکی یافت شده انعکاس یافتند. انطباق مناسب میان گروه‌های ژنتیکی یافت شده در این مطالعه با فواصل جغرافیایی و خصوصیات تولیدی، نشان‌دهنده خلوص ژنتیکی است که طی سالیان دراز در اثر سازگاری با محیط زندگی در گاوهای بومی ایران ایجاد شده است و لزوم توجه به حفاظت از این منابع ژنتیکی را روشن‌تر می‌سازد.

1. Ascertainment Bias

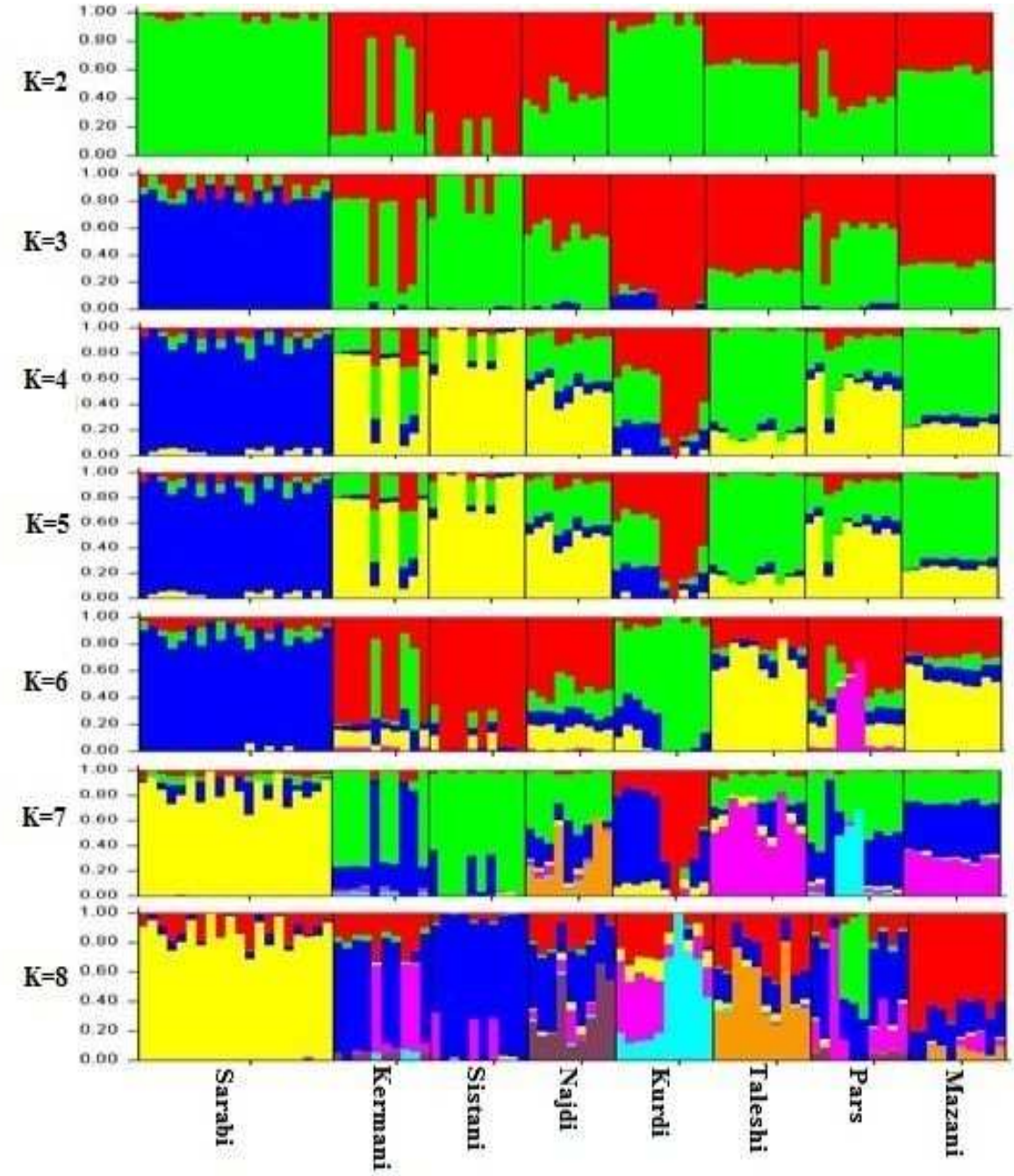

Fig. 1. Cluster assignment for eight values of K investigated. Each individual is represented by one vertical line with the proportion of assignment to each cluster shown on the y axis and colored by cluster

شکل ۱- میزان انتساب به خوشه‌های مختلف در هشت مقدار مختلف بررسی شده از K. هر فرد با یک خط عمودی نمایش داده شده است و میزان انتساب این فرد به هر خوشه در محور y با رنگ خاص هر خوشه نمایش داده شده است

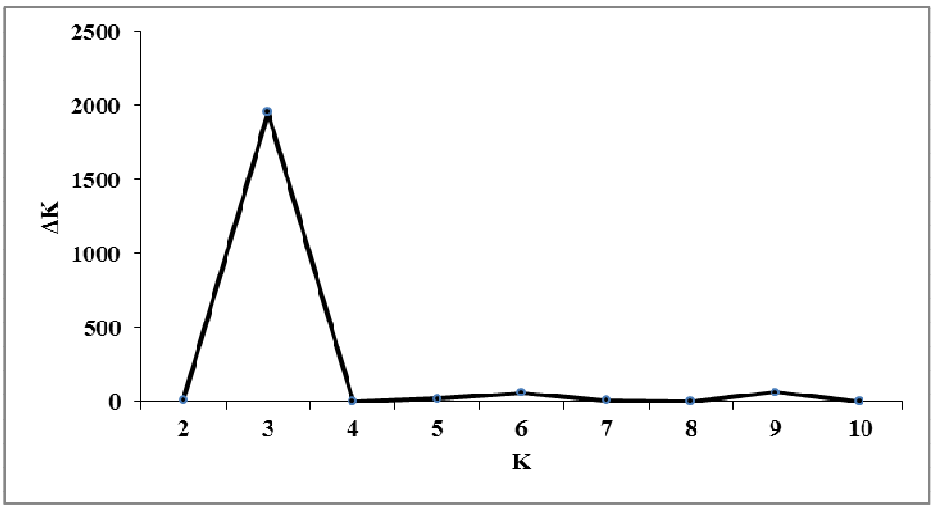

Fig. 2. Representation of changes in ΔK values per different K (number of assumed populations) values. Maximum point of ΔK depicts the optimal K value

شکل ۲- نمایش تغییرات مقادیر ΔK به ازای مقادیر مختلف K (تعداد جمعیت‌های فرض شده). حداکثر مقدار ΔK نشان‌دهنده مناسب‌ترین مقدار K است

جدول ۳- میزان عضویت هر یک از جمعیت‌ها در سه خوشه استنباط شده (K=۳)

Table 3. Proportion of membership of each population in each of the three Inferred clusters (K= 3)

Population	Inferred clusters			Number of individuals
	1	2	3	
Sarabi	0.106	0.093	0.800	20
Kermani	0.616	0.375	0.009	10
Sistani	0.901	0.089	0.010	10
Najdi	0.549	0.429	0.023	9
Kurdi	0.023	0.914	0.063	10
Taleshi	0.290	0.710	0.001	10
Fars	0.567	0.408	0.025	10
Mazandarani	0.344	0.655	0.000	10

جدول ۴- مقادیر F_{st} برآورد شده میان هر یک از جفت خوشه‌های استنباط شده در K=۳

Table 4. Estimated pairwise F_{st} between Inferred clusters in K=3

Cluster number	Cluster 1	Cluster 2	Cluster 3
Cluster 1	0.000	0.0107	0.2927
Cluster 2	0.0107	0.000	0.1618
Cluster 3	0.2927	0.1618	0.000

سپاسگزاری

بدین وسیله مراتب سپاسگزاری خود را از کلیه دامداران و کارکنان ایستگاه‌های پرورش گاو بومی کشور به پاس همکاری ایشان در تهیه نمونه از دام‌های مورد مطالعه اعلام می‌نماییم.

فهرست منابع

توکلیان ج. ۱۳۷۸. نگرشی بر ذخایر ژنتیکی دام و طیور بومی ایران. موسسه تحقیقات علوم دامی کشور، کرج، ص. ۴۵-۳.

Bolormaa S., Pryce J. E., Hayes B. J. and Goddard M. E. 2010. Multivariate analysis of a genome-wide association study in dairy cattle. Journal of Dairy Science, 93: 3818-3833.

Bovine HapMap C., Gibbs R. A., Taylor J. F., Van Tassell C. P., Barendse W., Eversole K. A., Gill C. A., Green R. D., Hamernik D. L., Kappes S. M., Lien S., Matukumalli L. K., McEwan J. C., Nazareth L. V., Schnabel R. D., Weinstock G. M., Wheeler D. A., Ajmone-Marsan P., Boettcher P. J., Caetano A. R., Garcia J. F., Hanotte O., Mariani P., Skow L. C., Sonstegard T. S., Williams J. L., Diallo B., Hailemariam L., Martinez M.L., Morris C. A., Silva L. O., Spelman R. J., Mulatu W., Zhao K., Abbey C. A., Agaba M., Araujo F. R., Bunch R. J., Burton J., Gorni C., Olivier H., Harrison B. E., Luff B., Machado M. A., Mwakaya J., Plastow G., Sim W., Smith T., Thomas M. B., Valentini A., Williams P., Womack J., Woolliams J. A., Liu Y., Qin X., Worley K. C., Gao C., Jiang H., Moore S. S., Ren Y, Song X. Z., Bustamante C. D., Hernandez R. D., Muzny D. M., Patil S., San Lucas A., Fu Q., Kent M. P., Vega R., Matukumalli A., McWilliam S., Sclep G., Bryc K., Choi J., Gao H., Grefenstette J. J., Murdoch B., Stella A., Villa-Angulo R., Wright M., Aerts J., Jann O., Negrini R., Goddard M. E., Hayes B. J., Bradley D. G., Barbosa da Silva M., Lau L. P., Liu G.E., Lynn D. J., Panzitta F. and Dodds K. G. 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science, 324: 528-532.

Corander J., Waldmann P. and Sillanpäa M. J. 2003. Bayesian analysis of genetic differentiation between populations. Genetics, 163: 367–374.

Decker J. E., McKay S. D., Rolf M. M., Kim J., Molina Alcala A., Sonstegard T. S., Hanotte O., Gotherstrom A., Seabury C. M., Praharani L., Babar M. E., Correia de Almeida Regitano L., Yildiz M. A., Heaton M. P., Liu W. S., Lei C. Z., Reecy J. M., Saif-Ur-Rehman M., Schnabel R. D. and Taylor J. F. 2014. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. PLoS Genetics, 10: e1004254.

Edea Z., Dadi H., Kim S. W., Dessie T., Lee T., Kim H., Kim J. J. and Kim K. S. 2013. Genetic diversity, population structure and relationships in indigenous cattle populations of Ethiopia and Korean Hanwoo breeds using SNP markers. Frontiers in Genetics, 4: 1-9.

Engelsma K. A., Veerkamp R. F., Calus M. P., Bijma P. and Windig J. J. 2012. Pedigree and marker-based methods in the estimation of genetic diversity in small groups of Holstein cattle. Journal of Animal Breeding and Genetics, 129: 195-205.

Evanno G., Regnaut S. and Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14: 2611-2620.

FAO. 2007. The state of the world's animal genetic resources for food and agriculture. In: Rischkowsky B, Pilling D (eds) United Nations Food and Agriculture Organization (FAO), Rome, Italy.

Franzen J. 2008. Bayesian cluster analysis: some extensions to non-standard situations. Doctoral dissertation, Stockholm university, Stockholm.
Goddard M. E. and Hayes B. J. 2007.Genomic selection. Journal of Animal Breeding and Genetics, 124: 323–330.
Groeneveld L. F., Lenstra J. A., Eding H., Toro M. A. and Scherf B. 2010. Genetic diversity in farm animals-a review. Animal Genetics, 41: 6-31.
Gray I. C., Campbell D. A. and Nigel K. 2000. Single nucleotide polymorphisms as tools in human genetics. Human Molecular Genetics, 16: 2403–2408.
Helyar S. J., Hemmer-Hansen J., Bekkevold D., Taylor M. I., Ogden R., Limborg M. T., Cariani A., Maes G. E., Diopere E., Carvalho G. R. and Nielsen E. E. 2011. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources, 11: 123-136.
Herrero-Medrano J. M., Megens H., Groenen M. A. M., Ramis G., Bosse M., Pérez-Enciso M. and Crooijmans R. P. M. A. 2013 . Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. BMC Genetics, 14:106-120.
Hoffmann I. 2010. Climate change and the characterization, breeding and conservation of animal genetic resources. Animal Genetics, 41: 32-46.
Holsinger K. E. and Wallace L. E. 2004. Bayesian approaches for the analysis of population structure: an example from Platanthera leucophaea (Orchidaceae). Molecular Ecology, 13: 887-894.
Lin B. Z., Sasazaki S. and Mannen H. 2010. Genetic diversity and structure in Bos taurus and Bos indicus populations analyzed by SNP markers. Animal Science Journal, 81: 281–289.
McKay S. D., Schnabel R. D., Murdoch B. M., Matukumalli L. K., Aerts J., Coppieters W., Crews D., Dias Neto E., Gill C. A., Gao C., Mannen H., Stothard P., Wang Z., Van Tassell C. P., Williams J. L., Taylor J. F. and Moore S. S. 2007. Whole genome linkage disequilibrium maps in cattle. BMC Genetics, 8: 74.
McKay S. D., Schnabel R. D., Murdoch B. M., Matukumalli L. K., Aerts J., Coppieters W., Crews D., Dias Neto E., Gill C. A., Gao C., Mannen H., Wang Z., Van Tassell C. P., Williams J. L., Taylor J. F. and Moore S. S. 2008. An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. BMC Genetics, 9: 37.
Mustafa H., Heather H. J., EuiSoo K., Ahmad N., Ali A., Ahmad Khan W., Naseer Pasha T., Farooq M. Z., Javed K., Ajmal A. and Sonstegard T. S. 2014. Comparative analysis of genome wide difference in Red Sindhi and Holstein cattle breeds using dense SNP marker. International Journal of Advanced Research, 2: 300-304.
Pertoldi C., Tokarska M., Wójcik J. M., Kawałko A., Randi E., Kristensen T. N., Loeschcke V., Coltman D., Wilson G. A., Gregersen V. R. and Bendixen C. 2010. Phylogenetic relationships among the European and American bison and seven cattle breeds reconstructed using the BovineSNP50 Illumina Genotyping BeadChip. Acta Theriologica, 55: 97-108.
Petersen J. L., Mickelson J. R., Cothran E. G., Andersson L. S., Axelsson J., Bailey E. and Bannasch D. 2013. Genetic diversity in the modern horse illustrated from genome-wide SNP data. Plos One, 8: 1-15.
Pritchard J. K., Stephens M. and Donnelly P. 2000. Inference of population structure using multi locus genotype data. Genetics, 155: 945–959.
Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A., Bender D., Maller J., Sklar P., de Bakker P. I., Daly M. J. and Sham P. C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics, 81: 559-575.
Sethuraman A. 2013. On inferring and interpreting genetic population structure - applications to conservation, and the estimation of pairwise genetic relatedness. Ph.D. dissertation, Iowa State University, Iowa State.
Symons M. J. 1981. Clustering criteria and multivariate normal mixtures. Biometrics, 37 :35-43.
Vonholdt B. M., Pollinger J. P., Lohmueller K. E., Han E., Parker H. G., Quignon P., Degenhardt J. D., Boyko A. R., Earl D. A., Auton A., Reynolds A., Bryc K., Brisbin A., Knowles J. C., Mosher D. S., Spady T. C., Elkahloun A., Geffen E., Pilot M., Jedrzejewski W., Greco C., Randi E., Bannasch D., Wilton A., Shearman J., Musiani M., Cargill M., Jones P. G., Qian Z., Huang W., Ding Z. L., Zhang Y. P., Bustamante C. D., Ostrander E. A., Novembre J. and Wayne R. K. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature, 464: 898-902.

Animal Production Research
Vol. 4, No. 3, 2015 (93-104)

University of Guilan

Analysis of genetic structure of Iranian indigenous cattle populations using dense single nucleotide polymorphism markers

K. Karimi[1,4], A. Esmailizadeh Koshkoiyeh[2*], M. Asadi Fuzi[3]

1. Ph.D. Candidate of Animal Breeding and Genetics, Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Iran
2. Professor, Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran
3. Associate Professor, Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran
4. Young Researchers Society, Shahid Bahonar University of Kerman, Kerman, Iran

(Received: 22-11-2014 – Accepted: 12-9-2015)

Abstract

The objective of this study was to detect genetic structure of Iranian indigenous cattle using dense single nucleotide polymorphism (SNP) markers. In total, 90 cattle individuals comprising Sarabi (n=20), Kurdi (n=10), Mazandarani (n=10), Taleshi (n=10), Sistani (n=10), Kermani (n=10), Najdi (n=10) and Fars (n=10) breeds were sampled. Genotyping was performed using Illumina High-density Bovine BeadChip designed to genotype 777,962 SNPs. After removing sites being in linkage disequilibrium (pairwise sites having r^2 greater than 0.2), 64333 SNPs remained for further analyses. STRUCTURE software was used to detect genetic structure of native cattle in this data set. Number of assumed populations per run was selected between two to eight. Results of this study indicated that in the first level of clustering (K=2), Sarabi (97.6%) and Kurdi (94.3%) populations shared the same cluster and Sistani (92%) had another separate cluster while other breeds were admixed from these two clusters. Kurdi population appeared as an independent cluster at K=3. Most suitable number of genetic groups in data set was explained by three clusters (K=3). Genetic differentiation of Iranian indigenous cattle was well detected in this study. Also, similarities of geographical distribution and production characteristics were in good accordance with founded genetic relationships in this study.

Keywords: Bayesian inference, Single nucleotide polymorphism, Genetic structure, Iranian indigenous cattle

*Corresponding author: aliesmaili@uk.ac.ir