

Survey of population structure and linkage disequilibrium pattern of Iranian native horse breeds by SNP markers

ABSTRACT

The aim of this study was to investigate the structure and genetic distances, inbreeding, and linkage disequilibrium (LD) in Iranian native horse breeds. Samples on 167 horses including Arabian Asil (24), Caspian (22), Dareshouri (15), Kurdish (66), and Turkmen (40) breeds were used. The samples were genotyped using Illumina 70k SNP equine bead chip. Data quality control was applied and loci with MAF < 2%, Mind < 5%, HWE ($<10^{-6}$), and genotype call rate < 1% were removed. Finally, 159 horses with 45,270 SNP loci were reminded for the genetic structure analyse. Clustering analyse showed that in K=5 Caspian and Kurdish breeds were placed in a common cluster, and Turkmen, Dareshouri, and Arabian Asil breeds were also located in distinct clusters. The LD decay among all breeds showed a significant decrease below 400 bp. Caspian and Dareshouri breeds experienced the lowest and highest LD and the highest and lowest effective population size, respectively. The inbreeding value based on the runs of homozygosity (ROH) were highest and lowest for Arabian Asil and Caspian breeds, respectively. The results showed that the Iranian native horse breeds were well identified by SNP markers, which can be used for stud base foundations of breeds with more accuracy.

Keywords: Iranian native horses, Genetic diversity, linkage disequilibrium, Runs of homozygosity, single nucleotide polymorphism markers.

بررسی ساختار جمعیتی و الگوی عدم تعادل پیوستگی اسب‌های بومی ایران با استفاده از نشانگرهای SNP

چکیده

هدف این پژوهش، بررسی ساختار و فواصل ژنتیکی، هم‌خونی و الگوی عدم تعادل پیوستگی (LD) و درک روابط بین نژادهای اسب بومی ایران بود. در مجموع از ۱۶۷ اسب که شامل ۲۴ اسب از نژاد عرب (اصیل)، ۲۴ اسب نژاد کاسپین، ۱۵ اسب نژاد دره‌شوری، ۶۶ اسب نژاد کرد و ۴۰ اسب ترکمن بودند، استفاده شد. نمونه‌ها با استفاده از تراشه (Illumina 70k SNP equine bead chip) ژنوتایپ شدند. در پالایش داده‌ها، جایگاه‌هایی با MAF کمتر از ۲ درصد، Mind کمتر از ۵ درصد، HWE کمتر از 10^{-6} و ژنوتایپ‌های کمتر از ۰/۰۱ حذف شدند و ۴۵۲۷۰ جایگاه SNP در ۱۵۹ اسب برای آنالیزهای بررسی ساختار ژنتیکی اسب‌های بومی ایران به کار گرفته شد. تجزیه ساختار ژنتیکی نشان داد که در K=5 نژادهای کاسپین و کرد در یک خوشه مشترک قرار گرفتند و نژادهای ترکمن، دره‌شوری و عرب (اصیل) نیز خوشه‌های متمایزی تشکیل دادند. فاز LD در بین تمام نژادها در فاصله زیر ۴۰۰bp کاهش معناداری داشت. نژاد کاسپین و دره‌شوری به ترتیب کم‌ترین و بیش‌ترین LD و بیش‌ترین و کم‌ترین اندازه موثر جمعیت را تجربه کردند. کم‌ترین و بیش‌ترین ضریب هم‌خونی براساس رشته‌های هموزیگوت ژنومی (ROH) نیز به ترتیب در نژادهای عرب اصیل و کاسپین دیده شد. نتایج نشان داد نژادهای بومی اسب ایران به خوبی با نشانگرهای (SNP) شناسایی شدند که این موضوع می‌تواند در تشکیل پایه‌های ژنومی نژادها (Data Base) و با صحت بالا مورد استفاده قرار گیرد.

واژه‌های کلیدی: اسب‌های بومی ایران، تنوع ژنتیکی، عدم تعادل پیوستگی، رشته‌های هموزیگوسیتی، نشانگرهای چندشکلی تک‌نوکلئوتیدی.

مقدمه

اسب (*Equus ferus caballus*) یکی از کهن‌ترین حیوانات اهلی شده توسط انسان است و از نظر گسترش جغرافیایی تقریباً در تمامی سرزمین‌هایی که انسان به آنها گام نهاده این حیوان نیز در کنار او حضور داشته است. بدین ترتیب می‌توان اسب را یکی از محبوب‌ترین حیوانات اهلی در جوامع بشری محسوب کرد. اسب یکی از حیوانات اهلی محسوب می‌شود که به واسطه کاربردش در کار، کشاورزی، حمل‌ونقل و منابع غذایی در زندگی بشر و تمدن جوامع انسانی در ۵۰۰۰ سال اخیر نقش بسزایی داشته است (Jun, 2009; Wade et al., 2013; Petersen et al., 2013; Mahrous et al., 2011; et al., 2014). بدون شک نژادهای اسب میراثی ارزشمند و بخش مهمی از سرمایه‌های ملی، منطقه‌ای و منابع ژنتیکی جهانی می‌باشند. با این حال، خوشبختانه امروزه عموم مردم به لطف فعالیت‌های جهانی و ملی در جهت

حفاظت از تنوع زیستی و منابع ژنتیکی از نیاز به حفظ نژادهای بومی خود آگاه هستند (Ivanković *et al.*, 2021). پس از اهلی شدن اسب در استپ‌های اوراسیا در ۵۵۰۰ سال گذشته (Warmuth *et al.*, 2011; Petersen *et al.*, 2013)، بر حسب موقعیت جغرافیایی، پیشینه فرهنگی و رویدادهای اجتماعی جوامع، نژادهای مختلفی از اسب به وجود آمد (Cosgrove *et al.*, 2020). پس از تشکیل پایه‌های ژنتیکی نژادها و از یاد نسل، تنوع ژنتیکی اسب‌ها دست‌خوش تغییراتی شدند که به سبب آن نژادهایی مختلف با خصوصیات مورفولوژیکی و رفتاری گوناگونی شکل گرفتند (McCue *et al.*, 2012). در حال حاضر، جمعیت‌های تشکیل دهنده اسب‌های بومی ایران شامل: اسب‌های ترکمن، کاسپین، کرد، دره شوری و عرب (اصیل) می‌باشد (Rafeie *et al.*, 2011). اسب ترکمن یکی از قدیمی‌ترین و با ارزش‌ترین نژادهای اسب دنیا محسوب می‌شود (Jiskrová *et al.*, 2016) و شامل زیرخانواده‌های آخال‌تکه، یموت، چناران و گولان است که امروزه در اکثر نقاط کشور پرورش داده می‌شود. این نژاد بومی منطقه ترکمن صحرا می‌باشد و سکونت‌گاه اصلی آن مناطق شمال شرقی ایران واقع در استان‌های گلستان و خراسان شمالی می‌باشند (Khalili, 2009; Khan Ahmadi *et al.*, 2017). طبق نتایج حاصل از مطالعات گذشته اسب ترکمن به همراه نژادهای خاورمیانه (عرب و کاسپین) در یک خوشه ژنتیکی قرار می‌گیرند که نشان دهنده منشأ ژنتیکی واحد و وجود اجداد مشترک بین آنهاست (Petersen *et al.*, 2013). اسب مینیاتوری کاسپین به عنوان یکی از قدیمی‌ترین اسب‌های موجود در خاورمیانه می‌باشند و طبق تحقیقات صورت گرفته این نژاد به همراه اسب ترکمن به عنوان منشأ و اجداد عمده اسب‌های شرقی هستند (Khalili, 2009). اسب دره شوری یکی از نژادهای اصلی ایران می‌باشد که به لحاظ ظاهری به نژاد عرب اصیل شباهت‌های زیادی دارد و منشأ آن نواحی مرکزی و جنوبی ایران است (Khalili, 2009). اسب‌های این نژاد به دلیل استفاده مردم عشایر در ایلات قشقای که مجبور به طی مسیرهای سخت و کوهستانی می‌باشند، از استقامت بدنی بالایی برخوردار هستند (<http://dareshurihorse.com>). اسب عرب اصیل از خالص‌ترین و مشهورترین نژادهای اسب در جهان است (Głazewska, 2010). زیستگاه اصلی این نژاد قسمت جنوب غرب ایران است و از لحاظ جغرافیایی در مجاورت با نژادهای کرد و دره‌شوری قرار دارد که در گذشته دست‌خوش نوعی آمیختگی نژادی یا رابطه اجدادی شده‌اند. به طوری که امروزه مطالعات ژنتیکی شباهت‌های ساختاری بین این جمعیت‌ها را تایید کرده است (Ovchinnikov *et al.*, 2018). اسب اصیل به زیرخانواده‌های کهیلان، هادیان، حمدانی، سگلاوی و عبیان تقسیم می‌شوند (Khalili, 2009). اسب دره‌شوری به لحاظ ظاهری شبیه اسب‌های عرب است و در منطقه "دره شور" پرورش می‌یابند. اسب‌های کرد که بومی مناطق غربی ایران هستند با شرایط سخت آب و هوایی سازگار شده‌اند و با داشتن سُم‌ها و بدنی قدرتمند توانایی حرکت در جاده‌های سنگلاخی و کوهستانی را دارند و زیرخانواده‌های جاف، افشاری و سنجایی را شامل می‌شوند (Khalili, 2009; Nazari *et al.*, 2022). در این بین، اسب‌های ترکمن و عرب بیش‌ترین تاثیر را در اصلاح نژاد و شکل‌گیری بسیاری از نژادها داشتند. بطوری که اغلب خط خونی نریان‌های نژادهای دیگر از این دو اسب نشأت گرفته‌اند (Wallner *et al.*, 2017). بنابراین، می‌توان نتیجه گرفت که براساس یافته‌های باستان‌شناسی، مراحل اولیه شکل‌گیری بیشتر نژادهای اسب اروپایی از خط خونی اسب‌های ایرانی نشأت گرفته و این نژادها در توسعه نژادهای اروپایی نقش بسزایی داشتند (Fages *et al.*, 2019).

با پیشرفت علم و فن آوری‌های نوین در عرصه ژنومیک، پژوهش‌های متنوع در زمینه ژنوم اسب‌ها صورت گرفت که در نهایت به توالی ژنوم اسب دست یافتند. بعد از توالی‌یابی ژنوم مادیان توایلایت^۱ (نژاد تاروبرد) به روش سانجر^۲ (Wade *et al.*, 2009)، تراشه‌های ژنومی (Illumina 50k SNP equine bead chip) توسعه یافت که قادر به نقشه برداری ژنوم و چندشکلی تک نوکلئوتیدی^۳ (SNP) است.

1. Twilight

2. Sanger

3. single-nucleotide polymorphism

با استفاده از این تراشه، ژنوم هفت اسب شامل: آندلوزین^۱، عرب، آخال تکه، ایسلندیک^۲، استانداردبرد^۳، تروبرد و کورتر هورس^۴ بررسی شدند (McCue *et al.*, 2012). مدت کوتاهی بعد از تراشه ۵۰K، تراشه ۷۰K (Illumina 70k SNP equine bead chip) توسعه یافت که قادر به تعیین ژنوتایپ ۷۴۰۵۶ جایگاه را دارد. با در دسترس بودن توالی مرجع ژنوم اسب، از این تراشه به طور گسترده‌ای در مطالعات سلامتی و عملکرد اسب استفاده می‌شود (wade *et al.*, 2009; McCoy *et al.*, 2014). اخیراً نیز تجزیه و تحلیل جامع تنوع نژادی اسب نشان داد که نشانگرهای SNP یکی از ابزارهای قدرتمند در مطالعات GWAS^۵ به شمار می‌رود (Petersen *et al.*, 2013). از نشانگرهای SNP به طور موفقیت‌آمیزی در مطالعات ژنتیکی مربوط به گاو (Qanbari *et al.*, 2010) و همچنین اسب برای آشکار ساختن پیشینه اهلی سازی اسب (wade *et al.*, 2009; Schubert *et al.*, 2014)، نشان‌های انتخاب (Zandi *et al.*, 2014; Maghsoodi *et al.*, 2017)، تخمین هم‌خونی (Khalt-Abadi Farahani; Moradi., 2018)، تنوع ژنتیکی (Sadeghi *et al.*, 2019; Cosgrove *et al.*, 2020) و بررسی ساختار جمعیتی (Petersen *et al.*, 2013; McCue *et al.*, 2012) استفاده شده است.

با مرور منابع علمی گذشته می‌توان دریافت که عمده هدف محققان به بررسی ساختار، میزان تنوع ژنتیکی، هم‌خونی و نشان‌های انتخاب در نژادهای بومی اسب ایران خلاصه شده است. به طور مثال زندی و همکاران (2014) با هدف پویس ژنومی، ساختار ژنتیکی و نشان‌های انتخاب دو نژاد کاسپین و ترکمن را بررسی کردند و نشان دادند که میزان عدم تعادل پیوستگی (LD) در نژاد کاسپین در بیشتر فواصل ژنتیکی بالاتر، هم‌خونی بیشتر و اندازه موثر جمعیت (Ne) کمتری نسبت به نژاد ترکمن داشت. در پژوهشی دیگر مقصودی و همکاران (2017) به بررسی ساختار جمعیت و شناسایی نواحی تحت انتخاب در ژنوم اسب‌های عرب و کرد پرداختند و نشان دادند که اسب‌های عرب ایرانی نسبت به اسب‌های کرد تنوع ژنتیکی بالاتری دارند. در تحقیق آنان، نشان‌های انتخاب در شش جایگاه ژنی شناسایی شد و نتیجه بررسی جایگاه‌های تحت انتخاب منجر به شناسایی ژن‌هایی شد که احتمالاً با عملکرد اسب در مسابقات ورزشی مرتبط می‌باشند. صادقی و همکاران (۲۰۱۹) با بررسی اسب‌های بومی ایران و تمرکز بر تنوع ژنتیکی و نشان‌های انتخاب نژاد عرب ایرانی پرداخت. یوسفی مشعوف و همکاران (۲۰۲۱) نیز تحقیقی با هدف مقایسه اسب‌های نژاد تارو برد، عرب ایرانی و کرد انجام دادند. در نتایج تحقیق آنان، الگوهای ساختار جمعیت معناداری بین نژادهای عرب و کرد مشاهده شد.

در ادامه تحقیقات پیشین و تکمیل اطلاعات ژنومی مبنی بر آگاهی از نحوه ارتباط ژنتیکی نژادهای اسب بومی ایران، پژوهش حاضر با استفاده از تراشه (Illumina 70k SNP Equine bead chip) با هدف بررسی و پویس ژنومی ساختار، تنوع ژنتیکی و الگوی عدم تعادل پیوستگی (LD) و برآورد میزان هم‌خونی پنج نژاد اسب ایران به دنبال آن است تا افزون بر تعیین فواصل و ساختار ژنتیکی مرتبط با جمعیت‌های مذکور، برخی راهکارهای حفاظتی را برای بررسی‌های بیشتر در این نژادها شناسایی کند.

روش‌شناسی پژوهش

تعیین ژنوتایپ نمونه‌ها و جمعیت‌های مورد بررسی

در این پژوهش از اطلاعات ژنومی ۱۶۷ راس اسب بومی کشور شامل: ۴۰ راس ترکمن، ۲۲ راس کاسپین، ۶۶ راس کرد، ۱۵ راس دره‌شوری و ۲۴ راس عرب استفاده شد. نمونه‌های مو اسب‌های ژنوتایپ شده توسط کارشناسان فدراسیون سوارکاری در طی سال‌های مختلف از استان‌هایی که خاستگاه اصلی این نژادها هستند، گرفته شد. همه نمونه‌ها علاوه بر تایید آزمون والدین فدراسیون سوارکاری، مراحل آزمایش انساب را گذرانده و از لحاظ خلوص نژادی مورد تایید و در کتاب انساب نژادی مربوطه قرار دارند. سعی شد تا برای آنالیز داده‌های

1. Andalusian

2. Icelandic

3. Standardbred

4. Quarter Horse

5- Genome-wide association study

ژنومی، هیچ یک از اسب‌های مورد مطالعه روابط خویشاوندی در شجره‌نامه نداشته باشند. نمونه‌های مو اخذ شده از طریق فدراسیون سوارکاری به آزمایشگاه Geneseek آمریکا ارسال و نمونه‌ها با استفاده از تراشه‌های ژنومی Illumina 70k SNP Equine bead chip و با بکارگیری دستورالعمل استاندارد شرکت Illumina تعیین ژنوتایپ شدند.

مراحل کنترل کیفیت و ویرایش داده‌ها

در پژوهش حاضر برای اطمینان از صحت و کیفیت داده‌های حاصل از تعیین ژنوتایپ در آنالیزهای نهایی مراحل فیلتراسیون روی داده‌های خام صورت گرفت و پس از تعیین ژنوتایپ، با عمل غربالگری، SNP‌های منتخب وارد مرحله دیگر آنالیز شدند (**Error! Reference source not found.**). برای بررسی دقیق ساختار ژنتیکی بین نژادهای بومی اسب ایران روش‌های مختلفی مورد استفاده قرار گرفت.

جدول ۱. مراحل مختلف ویرایش و کنترل کیفی داده‌های حاصل از تعیین ژنوتایپ نژادهای اسب بومی ایران

عرب	دره شوری	کرد	کاسپین	ترکمن	مراحل مختلف ویرایش
۲۴	۱۵	۶۶	۲۲	۴۰	تعداد کل حیوانات
۶۵۱۵۷	۶۵۱۵۷	۶۵۱۵۷	۶۵۱۵۷	۶۵۱۵۷	تعداد کل SNPs
۵۴۰۶	۷۹۳۶	۶۸۵۹	۵۸۷۵	۵۶۶۴	حذف SNPs با $Geno < 0.1$ در کل نمونه‌ها
۱۵۴۶۶	۱۵۳۳۲	۱۰۱۱۶	۹۰۱۹	۸۳۲۹	حذف SNPs با $MAF < 0.05$ در کل نمونه‌ها
۲	۱	۰	۳	۲	حذف SNPs با $Mind < 0.05$ در کل نمونه‌ها
۲	۰	۹۰۵	۰	۱۵۹۴	حذف SNPs که انحراف زیادی از (HWE) دارند (10^{-6})
۴۴۲۸۳	۴۱۸۹۹	۴۷۲۷۷	۵۰۲۶۳	۴۹۵۷۰	تعداد SNPs باقی مانده

آنالیزهای تنوع ژنتیکی و ساختار جمعیتی

آنالیز PCA روابط خویشاوندی ژنومی جهت تفکیک لایه‌بندی جمعیتی و شناسایی حیواناتی که خارج از گروه نژاد خود قرار گرفته‌اند با استفاده از نرم‌افزار PLINK بدست آمد. در ابتدا، اجرای آنالیز PCA بر روی داده‌های SNP با استفاده از تابع "Prcomp" در نرم‌افزار R V 4.2.1 (R Development Core Team, 2022) صورت گرفت. حالت یکسان^۱ (IBS) فواصل میان نمونه‌ها با استفاده از نرم‌افزار PLINK v 1.9 (Purcell *et al.*, 2007)، محاسبه و ترسیم آن با آنالیز مقیاس چندبعدی^۲ (MDS) انجام گرفت. برای بررسی ساختار کلی جمعیت‌های اسب ترکمن، کاسپین، کرد، دره‌شوری و عرب روی داده‌های ژنوتایپ، تحلیل مولفه‌های اصلی با استفاده از ماتریس IBS که توسط نرم‌افزاری Plink در بسته‌های نرم‌افزار R انجام شد. برای تحلیل مولفه‌های اصلی ضریب خویشاوندی ژنوتایپ، فاصله بین افراد از طریق مقیاس‌بندی چندبعدی کلاسیک (MDS) با استفاده از تابع cmdscale در بسته نرم‌افزاری stats نرم‌افزار R V 4.2.1 (R Development Core Team, 2022) محاسبه شد. نهایتاً با استفاده از اطلاعات بدست آمده در این مطالعه می‌توان نسبت به تفکیک نژادهای اسب بومی، بررسی ساختار جمعیتی اکوتیپ‌های حیوانات، آزمون انتساب و همچنین تعیین روابط خویشاوندی ژنومی بین افراد جهت تفکیک لایه‌بندی جمعیتی و شناسایی حیواناتی که خارج از گروه نژادی خود قرار گرفته‌اند نیز تصمیم‌گیری کرد، که انجام این مراحل با استفاده از بسته نرم‌افزاری ADEGENET 2.0.0 (Jombart *et al.*, 2018) در محیط R (R Development Core Team, 2022) به روش ترکیبی بیزین در

1. Identical By State

2. Multiple Dimensional Scaling

نرم افزار Structure v 2.3.4 (Pritchard *et al.*, 2000) انجام شد. در این نرم افزار، شناسایی بهینه خوشه‌ها با استفاده از افزایش تعداد K در مقایسه با معیار اطلاعات بیزی^۱ مورد استفاده قرار گرفت. برای تعیین خوشه‌ها نیاز است که از قبل جمعیت‌ها از پیش تعیین شده باشند. به همین منظور، از تابع find.cluster جهت تعیین تعداد خوشه و از الگوریتم k-means که روشی مستقل از مدل است برای خوشه‌بندی استفاده شد (Jombart *et al.*, 2018). از محتوای اجدادی از پیش تعریف شده (K)، به عنوان یک پیش فرض در تخمین سهم تعداد جمعیت‌های اجدادی برای هر یک از اسب‌های موجود استفاده شد (Alexander and Lange, 2011). برای تعیین مقدار بهینه K، مقدار K مربوط به حداکثر میانگین احتمال الگوریتم تخمینی داده‌ها در نظر گرفته شد (Pritchard *et al.*, 2000). تعداد تکرارها برای شبیه‌سازی جمعیت‌ها با ۱۰۰۰۰ دور قلق گیری و ۵۰۰۰۰ دور تعداد تکرار MCMC برای به دست آوردن $\ln Pr(X|K)$ انجام گرفت.

الگوی عدم تعادل پیوستگی (LD)

عدم تعادل پیوستگی به عنوان ارتباط غیر تصادفی میان آل‌های جایگاه‌های مختلف در داخل یک جمعیت تعریف شده است. عدم تعادل پیوستگی تحت تأثیر عامل‌های مختلفی از جمله تاریخچه و ساختار جمعیت مورد بررسی، اندازه جمعیت، نوع و تراکم نشانگری، روش برآورد هاپلوتیپ، نوترکیبی، دقت در ویرایش SNPها، تعادل هاردی-واینبرگ و استفاده از هاپلوتیپ‌های مادری و یا هر دو والد در برآورد LD قرار می‌گیرد (Bohmanova *et al.*, 2010). به دلیل اهمیت عدم تعادل پیوستگی بررسی‌های گسترده‌ای در این ارتباط در سطح ژنوم گونه‌های مختلفی از جمله اسب (Corbin *et al.*, 2010) صورت گرفته است. میانگین همبستگی بین جایگاهی بر مبنای اطلاعات جمعیت‌ها و آزمون تعادل هاردی - واینبرگ و عدم تعادل پیوستگی برای جمعیت منفرد و تخمین همبستگی بین جایگاهی میان و درون جمعیت‌ها برای جمعیت‌های چندگانه با استفاده از نرم افزار PLINK v 1.9 (Purcell *et al.*, 2007) محاسبه شد. مقادیر به دست آمده برای شاخص‌های الگوی تنوع ژنتیکی (هتروزیگوسیتی) بر اساس دستور (--het) و عدم تعادل پیوستگی (LD) بر اساس دستور (--r2) در نرم افزار PLINK v1.9 (Purcell *et al.*, 2007) به دست آمد. جهت بررسی الگوی ژنومی در جمعیت‌های مورد مطالعه، ارزش‌های تتا برای هر نشانگر SNP به روش برآوردگر ناریب تتا، (Weir and Cockerham, 1984) در محیط R محاسبه گردید.

برآورد ضریب هم‌خونی (F_{ROH})

هم‌خونی در جمعیت نژادهای اسب امری اجتناب‌ناپذیر است و از آنجایی که در سال‌های گذشته برخی از نژادهای اسب ایرانی به علت اندازه موثر کوچک و انتخاب مصنوعی به شدت تحت تأثیر قرار گرفته‌اند، احتمال اینکه اثرات منفی هم‌خونی را تجربه کرده باشند، بسیار زیاد است. از این رو، محاسبه و اندازه‌گیری میزان هم‌خونی یکی از بهترین راهکارها در جهت شناسایی افراد هم‌خون و تدوین استراتژی‌های لازم برای جلوگیری از خطای محاسبه هم‌خونی بر اساس اطلاعات شجره و تبارنامه است.

در این پژوهش با استفاده از نرم افزار PLINK، و با دستور (--homozyg) رشته‌های هموزیگوت^۲ (ROH) مورد نظر استخراج شدند. برای جلوگیری از تشکیل رشته‌های هموزیگوت حاصل از عدم تعادل لینکاژی با توالی کوتاه در توالی ژنوم، حداقل طول متوالی رشته‌های هموزیگوت ۵/۰ Mb در نظر گرفته شد. در ادامه، برای تعریف منطقه ژنومی به عنوان رشته‌های هموزیگوت، گزینه‌هایی با حداقل ۲۰ SNP در نواحی ژنومی و حداقل یک SNP با ژنوتیپ گمشده اعمال شد. همچنین، برای اینکه دو رشته هموزیگوت به صورت جداگانه قرار بگیرند حداقل فاصله بین SNPها برابر با یک میلیون جفت‌باز و حداقل تراکم SNP در داخل هر رشته برابر با یک SNP به ازای هر ۱۰

^۱. Bayesian Information Criterion

^۲. Runs of homozygosity

هزار جفت‌باز در نظر گرفته شد. در نهایت، برای محاسبه رشته‌های هموزیگوت و ضرایب هم‌خونی هر فرد در هر یک از نژادهای مورد بررسی از بسته detectRUNS (Biscarini *et al.*, 2018) در نرم‌افزار R (R Development Core Team, 2022) استفاده شد.

نتایج و بحث

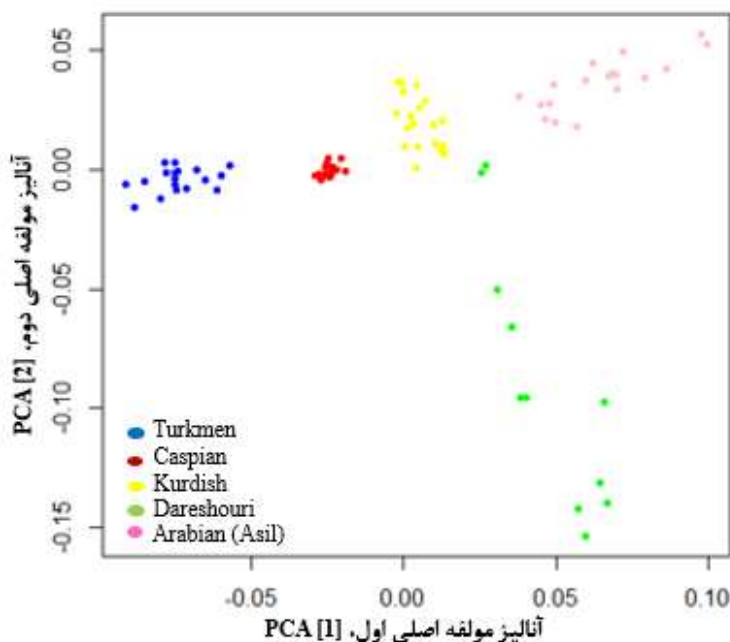
مراحل مختلف کنترل کیفیت داده‌ها بر روی ۶۵۱۵۷ جایگاه نشانگری SNP در ۱۶۷ اسب از نژادهای ذکر شده انجام شد. بعد از کنترل کیفیت، ۹۹۰۳ SNP به دلیل موقیعیته ناشناخته حذف شدند و در مراحل مختلف کنترل کیفیت روی SNPهای باقیمانده، تعداد هشت نمونه در جریان تعیین ژنوتایپ با کمتر از پنج درصد داده گم شده و ۷۹۴۲ SNP با MAF کمتر از ۰/۰۲ درصد حذف شدند. ۲۰۴۲ SNP نیز به دلیل خروج از تعادل هاردی-واینبرگ حذف شد. در نهایت، ۱۵۹ اسب با ۴۵۲۷۰ SNP وارد مرحله آنالیز نهایی شدند. همان‌گونه که در جدول ۱ مشخص است کم‌ترین و بیش‌ترین تعداد SNP باقیمانده به ترتیب در نژادهای دره‌شوری و کاسپین مشاهده می‌شود که SNPهای نژاد کاسپین از مطالعه نصیریپور و همکاران (۲۰۲۳) برای همین نژاد کمتر بود.

تجزیه و تحلیل مولفه‌های اصلی (PCA)

نتایج آنالیز PCA نشان داد که جمعیت‌های مورد مطالعه بر اساس اطلاعات PC1 و PC2 بر اساس منطقه جغرافیایی جدا شدند و کل نمونه‌ها به پنج مولفه اصلی تفکیک شدند. نتایج این آنالیز نشان داد که PCA1 باعث جدا شدن تمامی نژاد از یک دیگر شده است و به غیر از نژاد کاسپین، حیوانات دیگر بر روی PCA2 از گروه نژادی خود فاصله گرفته است. در این آنالیز نژاد کاسپین دارای همگن‌ترین افراد تشکیل دهنده نژادی بود، در صورتی که فواصل داخل نژادی مابقی جمعیت‌های مورد بررسی حاکی از وجود منشأ ژنتیکی مشابه و به طبع آن شباهت ظاهری آنهاست. همان‌طور که در شکل مشخص می‌گردد، تمایز واضحی بین جمعیت‌ها بسته به مناطق جغرافیایی آنها وجود دارد. در خروجی آنالیز خوشه‌بندی می‌توان مشاهده کرد که با تمام شباهت‌های فنوتیپی که جمعیت دره‌شوری با نژادهای عرب و کرد دارد اما به طور قابل توجهی از هر دو نژاد فاصله گرفته و برای خود گروهی مجزا از سایر نژادهای مورد بررسی تشکیل داده است. با توجه به محل قرارگیری نژاد دره‌شوری در آنالیز PCA می‌توان دید که این نژاد دارای مختصر آمیختگی ژنومی نیز با نژادهای عرب و کرد دارد که با در نظر گرفتن شرایط زیست بوم نژادهای مذکور این نتیجه دور از انتظار نیست. این گروه‌بندی می‌تواند با آمیزش اجدادی و وقوع جریان ژنی در مناطق جغرافیایی متفاوت مرتبط باشد (Eltannay *et al.*, 2015).

تجزیه و تحلیل SNP با تراکم بالا می‌تواند اطلاعات مربوط به گذشته اسب‌ها را در مطالعات ژنتیک جمعیت فراهم کند. با توجه به نتایج تحلیل مؤلفه‌های اصلی و نزدیکی مکانی محل زندگی دو نژاد عرب و کرد، ممکن است اسب‌های عرب و اسب‌های کرد جد مشترکی داشته باشند و یا این که احتمال دارد اسب‌های عرب در به وجود آمدن اسب‌های کرد نقش داشته باشند. در تحقیق بابایی و همکاران (۲۰۲۱) نیز تمام جمعیت‌های مورد بررسی (عرب، آخال‌تکه و کاسپین) به خوبی از یکدیگر متمایز شدند که با نتایج حاضر هم‌خوانی دارد. مشابه چنین گروه‌بندی با اندکی اختلاط نژادی در مطالعه یوسفی‌مشعوف و همکاران (۲۰۲۱) و مقصودی و همکاران (۲۰۱۷) نیز مشاهده شد. در مطالعه آنان نژادهای عرب اصیل و کرد هر دو یک گروه مجزا تشکیل دادند اما مقداری آمیختگی نژادی بین دو نژاد مذکور مشاهده شد. پیش‌تر نیز بررسی روی DNA میتوکندری اسب‌های عرب نشان داد که اسب‌های عرب با بسیاری از نژادهای اسب منشأ مشترک دارند (Głazewska, 2010). مشابه چنین نتایجی در مطالعات زندگی و همکاران (۲۰۱۴) و خلعت‌آبادی و مردای (۲۰۱۴) گزارش شد. به گونه‌ای که هر نژاد در PCA1 از یکدیگر و در PCA2 از گروه نژادی خود فاصله گرفتند. پترسون و همکاران (۲۰۱۳) به منظور تعیین تنوع ژنتیکی

۳۶ نژاد از سراسر جهان، بر اساس درخت هم‌جواری^۱، اسب‌های منطقه خاورمیانه (عرب، کاسپین و آخال‌تکه) را در یک گروه قرار داد، در صورتی که هر سه نژاد عرب اصیل، کاسپین و ترکمن در این پژوهش در گروه‌ها و مناطق جغرافیایی متفاوتی قرار گرفتند. در پژوهشی که در آن نژادهای منتخبی از آسیا، اروپا و آمریکا بررسی شدند، در نتایج PCA وجود مقداری تشابه نژادی در اسب‌های عرب، آخال‌تکه و کاسپین مشاهده شد (Jun et al., 2014). در مطالعه صادقی و همکاران (۲۰۱۹)، گراف PCA بدست آمده برای نژادهای کاسپین و ترکمن حاکی از هم‌پوشانی دو نژاد مذکور داشت که با مطالعه حاضر متفاوت بود. در تحقیق کاسگروو و همکاران (۲۰۲۰) نتایج حاصل از PCA چندین نژاد گوناگون به همراه اسب‌های ایرانی نشان داد که نژادهای عرب اصیل، کرد و دره‌شوری در کنار یکدیگر قرار گرفته و هم‌پوشانی قابل ملاحظه‌ای داشتند. نتایج حاصل از PCA در این پژوهش برای نژادهای اسب بومی ایران هم‌سو با مطالعه اسدالله‌پور عنایی و همکاران (۲۰۲۱) برای نژادهای آخال‌تکه و عرب و مطالعه موسوی و همکاران (۲۰۲۳) برای نژادهای شمالی (ترکمن و کاسپین) و نژادهای غربی/ جنوب غربی (کرد و عرب اصیل) و مطالعه نصیریپور و همکاران (۲۰۲۳) برای نژادهای کرد و کاسپین بود.



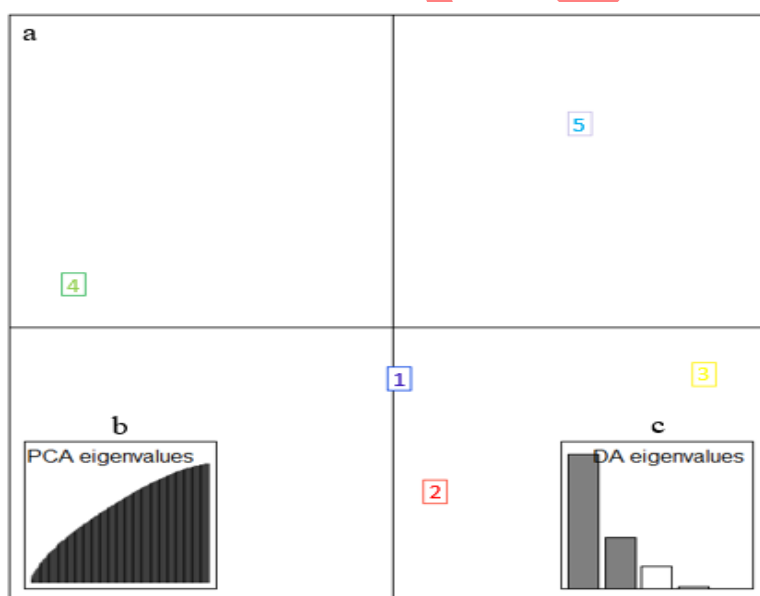
شکل ۱. خوشه‌بندی حیوانات براساس آنالیز PCA با استفاده از اطلاعات ژنوتایپ نمونه‌ها و بر اساس دو مولفه اصلی اول و دوم. نژادها: ترکمن (آبی)، کاسپین (قرمز)، کرد (زرد)، دره‌شوری (سبز) و عرب (صورتی)

نتایج بررسی ساختار جمعیت با استفاده از روش آنالیز چندمتغیره (Discriminant Analysis of Principal Components)

به منظور تشخیص تمایز و تنوع ژنتیکی در بین و داخل گروه‌ها آنالیز مشخص کننده مولفه‌های اصلی که یک روش غیر مبتنی بر مدل می‌باشد اجرا شد که این آنالیز توانایی تشخیص بهترین تمایز در بین نژادها را شناسایی و توصیف خوشه‌های ژنتیکی و همچنین احتمال عضویت هر فرد در جمعیت‌های مختلف مورد مطالعه را براساس توابع تفکیک شده، حفظ و فراهم می‌کند. همان‌گونه که در شکل ۲ مشاهده می‌شود روش DAPC به خوبی توانست تمام اسب‌های مورد مطالعه را در پنج گروه جداگانه خوشه‌بندی کند و تصویر دقیقی از نژادهای مورد بررسی در اختیار ما قرار دهد. نتایج روش DAPC نشان داد که نژادهای دره‌شوری و عرب در فاصله دورتری نسبت به سایر نژادها قرار گرفته است. دو نژاد کاسپین و ترکمن که از مناطق شمال و شمال شرق کشور بوده و در فاصله نزدیک‌تر قرار گرفته‌اند، که

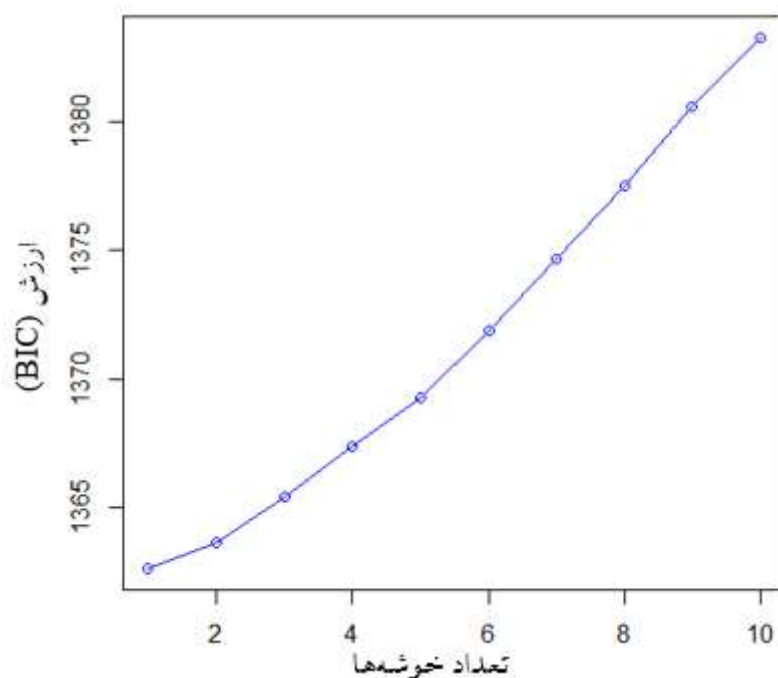
^۱. Neighbor Joining

احتمالاً این قرابت به دلیل پیشینه تاریخی و جد مشترک دو نژاد مذکور باشد. همچنین مشاهده می‌شود که نژاد کرد در فاصله نزدیک‌تری نسبت به سایر نژادها با نژادهای کاسپین قرار گرفته است، این مقدار نزدیکی می‌تواند به دلیل داشتن ژن‌های مشابه با نژادهای مذکور باشد که از تلاقی نژاد کرد و کاسپین در مناطق شمال غرب کشور به وجود آمده است. نکته قابل تامل در این شکل تشکیل گروه جداگانه نژاد دره‌شوری می‌باشد، در صورتی که این نژاد بر حسب منطقه جغرافیایی و زیستگاهی خود از نژادهای عرب و کرد به وجود آمده است، اما در شکل زیر با فاصله زیادی نسبت به نژادهای عرب و کرد قرار گرفته است. مقادیر ویژه مولفه‌های اصلی در تجزیه و تحلیل‌ها نمایش داده شده است که مقادیر ویژه مربوط به نسبت واریانس بین گروه‌ها بر روی واریانس درون گروه‌ها برای هر تابع قابل تفکیک می‌باشد که این موضوع نشان‌دهنده این است که اطلاعات می‌تواند در داخل دو تابع تفکیک قرار بگیرد (شکل ۲). مشابه تفکیک جمعیت‌ها به گروه‌های اختصاصی در تحقیق بابایی و همکاران (۲۰۲۱) نیز مشاهده شد که منطبق بر نتایج مطالعه حاضر بود. در تحقیق آنان، روش DAPC احتمال عضویت افراد جمعیت‌ها با صحت بالایی برآورد شد که در مقایسه با روش PCA، بهترین عملکرد را داشت. بنابراین، استفاده از روش‌های مبتنی بر مدل و پارامتری غیر مبتنی بر مدل DAPC در جمعیت‌های مورد بررسی، تداوم و عملکرد بالایی در تخمین نحوه ساختار جمعیت و استنباط احتمال عضویت هر یک از افراد به گروه‌ها را نشان می‌دهد (Pometti *et al.*, 2014). در این تحقیق، هر دو روش PCA و DAPC به درستی باعث تفکیک جمعیت نژادهای اسب بومی ایران شدند. که حاکی از کارآمدی این روش‌ها در شناسایی نژادها به گروه‌های اصلی خود است.



شکل ۲. تجزیه و تحلیل و اختصاص اسب‌ها در گروه‌های آنالیز مولفه‌های اصلی (۲a) به همراه اطلاعات کل مقادیر ویژه (۲b) و مقادیر ویژه مورد نیاز (۲c). نژادها: ترکمن (آبی)، کاسپین (قرمز)، کرد (زرد)، دره‌شوری (سبز) و عرب (صورتی).

در ادامه آنالیز نتایج بدست آمده از روش پارامتری DAPC، همانگونه که در شکل ۳ مشخص شد با افزایش تعداد K ، بهترین مقدار BIC در $K=5$ ترسیم شد که با ادامه روند افزایش K در نمودار زیر واضح است که بعد از $K=5$ مقدار BIC وجود واقعی خوشه‌ها را در همین مقدار مشخص کرده است.



شکل ۳. تعداد خوشه‌های بدست آمده از جمعیت‌های مورد مطالعه براساس معیار اطلاعات بیزین (BIC). محور X بیانگر تعداد خوشه و محور Y نشانگر ارزش معیار اطلاعات بیزی است.

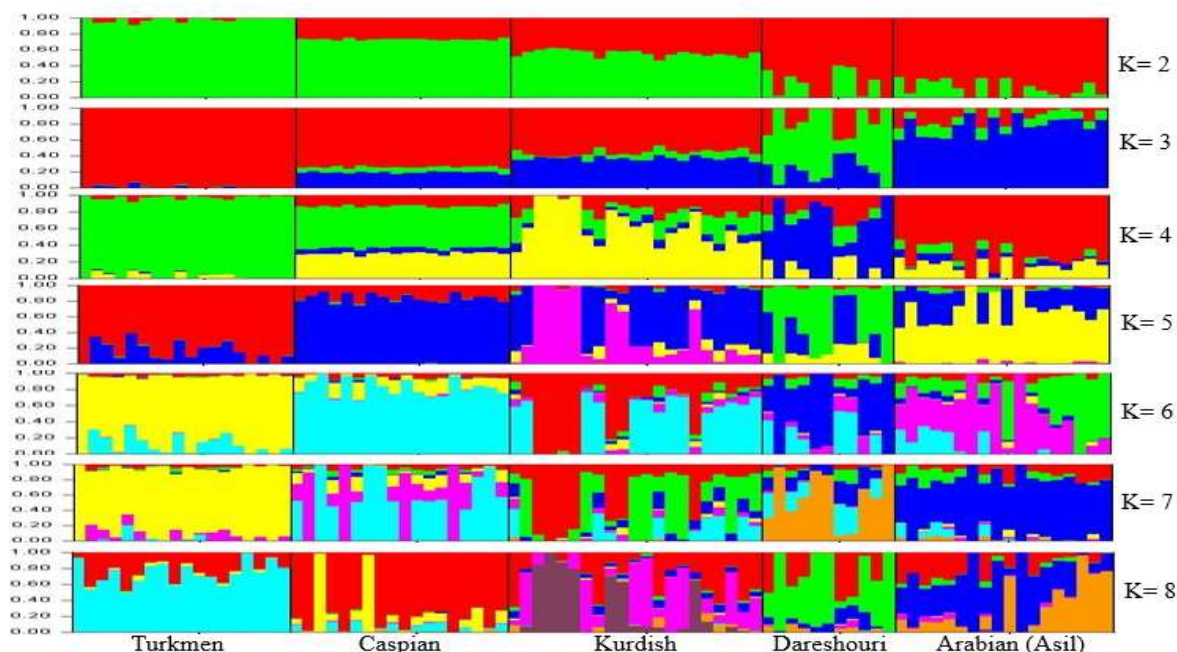
نتایج بررسی ساختار جمعیت با استفاده از روش مبتنی بر مدل

در مطالعه علم ژنتیک جمعیت، توانایی و نحوه تشخیص افراد متعلق به جمعیت‌های واقعی از اهمیت ویژه‌ای برخوردار است. بنابراین، برنامه Structure جهت بررسی ساختار ژنتیکی و آمیختگی در مجموعه داده‌های اسب‌های بومی شامل ۴۵۲۷۰ جایگاه SNP به کار گرفته شد. با توجه به اینکه دقیقاً تعداد نژادها براساس خلوص نژادی و همین‌طور زیرجمعیت‌ها برای نمونه‌های مورد بررسی معین نیست، لذا تعداد هشت پیمایش با افزوده شدن تعداد K از ۲ تا ۱۰ جهت بررسی ساختار جمعیتی در اسب‌های بومی ایران مورد استفاده قرار گرفت (نتایج $K=2$ تا $K=10$ در شکل ۴ ارایه گردیده است). در $K=2$ (فرض بر اینکه فقط دو جمعیت اسب تمامی تغییرات را در داده‌ها توجیه می‌کند) هر یک از نژادهای بومی در یک خوشه جداگانه قرار گرفتند، اما به لحاظ ساختاری نژادهای دره‌شوری، عرب با یکدیگر و نژادهای ترکمن، کاسپین و به نسبت کمتری کرد نیز با آنها دارای تشابه ژنتیکی هستند. در $K=3$ همچنان نژادهای دره‌شوری و عرب خوشه ژنتیکی متمایز اما مشابه خود را حفظ کردند، با افزایش تعداد جمعیت‌ها در $K=3$ به شباهت دو نژاد کاسپین و کرد افزوده می‌شود و نژاد ترکمن و کاسپین از هم متمایز می‌شوند. در $K=4$ نژادهای ترکمن، دره‌شوری و عرب همچنان در خوشه‌های مستقل خود ظاهر شده و متمایز از نژادهای دیگر بوده اما نژادهای کاسپین و کرد هر یک برای خود خوشه‌های جداگانه‌ای را تشکیل داده ولی از لحاظ ساختار ژنتیکی بیش‌ترین شباهت را با یکدیگر دارند. در $K=5$ رنگ غالب نژاد کاسپین با رنگ آبی مشخص شده است، همان‌گونه که از شکل ۳ قابل مشاهده است در $K=5$ نژادهای ترکمن، کرد، دره‌شوری و عرب دارای درصدی از خون نژاد کاسپین می‌باشند که نشان دهنده انشعاب نژادهای مذکور از اسب کاسپین است. در $K=6$ با وجود افزایش تعداد جمعیت فرض شده در نژادهای مورد بررسی، تغییر محسوسی در ساختار ژنتیکی نژادهای مورد بررسی مشاهده نشد. در مقادیر بالاتر از پنج نشانه‌های آمیختگی بیشتری در بین نژادهای کرد، دره‌شوری و عرب ظاهر شد، علت این تغییر را نیز شاید به توان با تشابه محل پرورش و شباهت‌های فوتوتیپی توجیه کرد. احتمالاً نسبت شباهت‌های ژنتیکی نژادهای کاسپین و

کرد به دلیل وجود نیای مشترک این دو نژاد است که ریشه هر دو نژاد به اسب نسایی باز می‌گردد (Abdoli et al., 2019). مشابه چنین نتایجی که در آن نژادهای عرب و دره‌شوری نسبت به دیگر نژادهای مورد بررسی، در یک گروه و مجاورت یکدیگر قرار گرفتند، در مطالعه صادقی و همکاران (۲۰۱۹) و کاسگروو و همکاران (۲۰۲۰) مشاهده شد. در مطالعه یوسفی مشعوف و همکاران (۲۰۲۱)، نژادهای عرب اصیل و کرد با شباهت بسیار بالای ژنتیکی نسبت به یکدیگر در مقابل اسب تروبرد در یک گروه مشترک قرار گرفتند شد ($K=2$)، اما با افزایش مقدار K ، تفاوت‌های نژادی دو نژاد ایرانی به شدت احساس شد ($K=3$). در مطالعه موسوی و همکاران (۲۰۲۳)، در $K=2$ و $K=3$ شباهت دو نژاد کاسپین و کرد هم‌راستا با مطالعه حاضر گزارش شد. در تحقیق دیگری که به بررسی شباهت ژنتیکی نژادهای مختلفی پرداخت، هر چند در $K=2$ نژادهای آسیایی (کاسپین و آخال‌تکه) آمیختگی ژنتیکی با نژاد عرب داشتند اما در یک خوشه جدا و مشابه قرار گرفتند (Jun et al., 2014). در مطالعه اسدالله پور نعایی و همکاران (۲۰۲۱) که ۱۶ نژاد از اسب‌های مختلف را مورد بررسی قرار داد، نژادهای آخال‌تکه و عرب در گروه‌های اختصاصی خود قرار گرفتند. نسبت انتساب هر یک از نژادها به خوشه‌های مختلف در $K=5$ در جدول ۲ نمایش داده شده است. نتایج حاصل از این مطالعه نشان دهنده مطابقت خوشه‌بندی جمعیت‌ها با فواصل جغرافیایی و محیط‌های پرورشی آنها است. نژادهای کاسپین و کرد که به ترتیب در مناطق شمالی و غربی کشور پرورش می‌یابند، در هر سطح از خوشه‌بندی در خوشه یکسانی طبقه‌بندی شده و هر یک از نژادهای ترکمن، دره‌شوری و عرب نیز به دلیل تفاوت در مناطق جغرافیایی خوشه‌های ژنتیکی جداگانه‌ای را تشکیل دادند.

جدول ۲. میزان عضویت هر یک از جمعیت‌ها در پنج خوشه استنباط شده ($K=5$)

تعداد افراد	خوشه‌های استنباط شده					نژادها
	۵	۴	۳	۲	۱	
۳۴	۰/۰۰۰	۰/۰۰۰	۰/۱۵۹	۰/۰۰۳	۰/۸۳۸	ترکمن
۱۸	۰/۰۰۰	۰/۰۰۱	۰/۸۳۷	۰/۰۰۴	۰/۱۵۷	کاسپین
۶۵	۰/۳۷۷	۰/۰۶۸	۰/۵۲۰	۰/۰۱۱	۰/۰۲۵	کرد
۱۴	۰/۰۰۰	۰/۱۲۹	۰/۲۵۷	۰/۵۹۵	۰/۰۱۸	دره‌شوری
۲۲	۰/۰۱۳	۰/۶۶۱	۰/۲۷۷	۰/۰۳۰	۰/۰۱۹	عرب (اصیل)



شکل ۴. میزان انتساب به خوشه‌های مختلف در ۱۰ مقدار بررسی شده از K (نمودارهای مربوط به برآوردهای مقادیر Q توسط نرم‌افزار Structure هر فرد با یک خط عمودی که به K قطعه رنگی تقسیم شده نشان داده شده که طول آن نشانگر سهم هر کدام از K خوشه است.

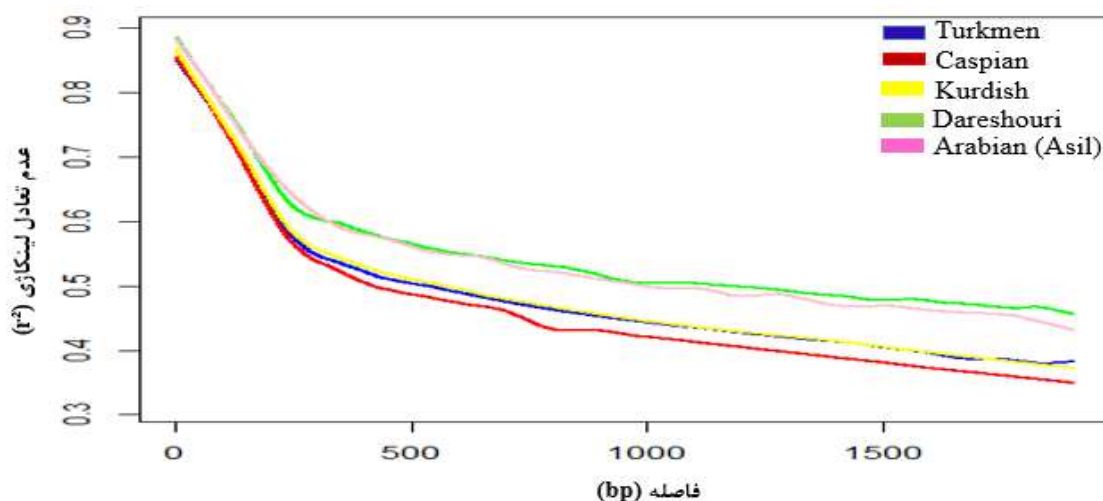
در مطالعه گسترده‌ای که توسط پترسون و همکاران (۲۰۱۳) با استفاده از SNP‌های اتوزومی برای بررسی ساختار ژنتیکی ۸۱۴ اسب از ۳۶ نژاد مختلف صورت گرفت، نتایج حاصل از محاسبات Parsimony، FST و فواصل ژنتیکی بین جمعیت‌های مورد بررسی انجام شد نشان داد که روابط میان نژادها عمدتاً منشا جغرافیایی و تاریخی شناخته شده‌ای از نژادها دارد. سطوح پایین اختلاف جمعیتی بین نژادهایی که روابط نزدیکی با فرآیند توسعه نژادی، و بین آنهایی که به دلیل اندازه موثر بزرگ جمعیت دارای سطح بالایی از تنوع ژنتیکی یا فنوتیپی درون نژادی داشتند مشاهده شد. جمعیت‌هایی که تحت فشار انتخاب، گردنه بطری قرار گرفته بودند و آنهایی که جمعیت بسته طولانی مدتی داشتند از تنوع ژنتیکی کم برخوردار بودند. در پژوهش المرزوک و همکاران (۲۰۱۷) با استفاده از Equine SNP70K BeadChip که برای بررسی تنوع ژنتیکی سه زیرشاخه اصلی اسب عرب (سگلاوی، کُهیلا و حَمدانی) انجام دادند، هتروزیگوسیتی مشاهده شده در محدوده‌ای بین ۰/۳۰ تا ۰/۳۲، هتروزیگوسیتی مورد انتظار بین ۰/۳۰ تا ۰/۳۱ گزارش شد. ضریب هم‌خونی (FIS) بین ۰/۰۲ تا ۰/۰۵ بود که نشان دهنده تنوع ژنتیکی بالا درون اسب‌های عرب سوریه است، همچنین تمایز ژنتیکی بین این سه تیره را بسیار پایین (FST ۰/۰۵) گزارش کردند. نتایج حاصل از خوشه‌بندی سلسله‌مراتبی نشان داد که تمایز آشکاری بین اسب‌های عرب و اسب پرزوالسکی وجود دارد.

الگوی عدم تعادل پیوستگی (LD) کل ژنوم

جهت نمایش میزان کاهش الگوی عدم تعادل پیوستگی، اطلاعات مربوط به فرآیند حاصل از نتیجه ژنوتایپ تمام نشانگرهای SNP در بین و داخل تمامی جمعیت‌های اسب بومی مورد بررسی قرار گرفت. همان‌گونه که از نمودار ۱ مشخص است فاز LD در بین نژادها بیشتر از داخل نژادها می‌باشد. فاز LD در بین تمام نژادها در فاصله زیر ۴۰۰bp با کاهش شدیدی را تجربه کرده و این میزان افت تا فاصله بین ۰/۴ تا ۰/۵ ادامه دارد. در داخل نژادها میزان r^2 در نژاد دره‌شوری و عرب در تمامی فواصل ژنتیکی بیش از نژادهای ترکمن، کاسپین و کرد می‌باشد. مقادیر بالای LD محاسبه شده برای جمعیت‌های دره‌شوری و عرب حاکی از سطح بالای هم‌خونی و کاهش اندازه موثر جمعیت

در این دو نژاد می‌باشد. با توجه به دو نژاد دره‌شوری و عرب که بیشترین شباهت فنوتیپی را در بین نژادهای بومی با یکدیگر دارند میزان LD دو نژاد مذکور با یکدیگر تفاوت چندانی نداشته و با رجوع به مقادیر برآورد شده حاصل از میزان مقدار FIS درون نژادها سطح بالای r^2 نژاد دره‌شوری قابل پیش‌بینی و دور از انتظار نبود. براساس گزارش وید و همکاران (۲۰۰۹) و کوربین و همکاران (۲۰۱۰) با افزایش فاصله بین نشانگرها تا ۰/۵ مگا جفت‌باز (Mb) در نژادهای مختلف اسب، مقدار r^2 از ۰/۶ به ۰/۲ کاهش می‌یابد. روند کاهش LD در جمعیت‌های مختلف با در نظر گرفتن سطح بهینه عدم تعادل پیوستگی تعیین‌کننده توان شناسایی جایگاه صفات کمی خواهد بود و به طور کلی تفاوت در میانگین LD کروموزوم‌ها به عوامل مختلفی هم‌چون: نسبت نوترکیبی، میزان هموزیگوسیتی، رانش ژنی و اثر ناشی از فشار انتخاب در هر کروموزوم بستگی دارد (Qanbari et al., 2010).

با توجه به شکل ۵ آشکار است که سطح پایین LD نژاد کاسپین حاکی از بیشتر بودن اندازه موثر این جمعیت از دیگر نژادهای مورد بررسی می‌باشد. نتایج بدست آمده از این تحقیق مشابه نتایج مطالعه پترسون و همکاران (۲۰۱۳) بود، در مطالعه آنها اندازه موثر جمعیت کاسپین ۳۵۱ و عرب ۳۴۶ بود. با آگاهی از میزان LD به دست آمده برای نژادهای بومی مورد بررسی می‌توان برای شناسایی نقاط مرتبط با بیماری‌های ژنتیکی و بهبود صفات دخیل در عملکردهای ورزشی و اقتصادی استفاده برد. عکس چنین نتایجی در تحقیق زندگی و همکاران (۲۰۱۴) گزارش شد، نتایج حاصل از پژوهش آنان نشان داد که سطح (LD) و همخونی نژاد کاسپین در مقایسه با نژاد ترکمن بالاتر و در نتیجه اندازه موثر کمتری نسبت به نژاد ترکمن داشت. در پژوهش خلعت‌آبادی و مرادی (۲۰۱۸) که با استفاده از رشته‌های هموزیگوت ۱ ژنومی (ROH) برای تخمین میزان هم‌خونی و اندازه موثر جمعیت نژادهای آخال‌تکه، عرب، کاسپین، مغول و تارو برد صورت گرفت، میانگین هتروزیگوسیتی مورد انتظار و مشاهده‌شده در نژادهای مذکور به ترتیب در دامنه ۰/۲۷۸-۰/۳۰۸ و ۰/۲۷۸-۰/۲۹۴ بود. محاسبه ضریب هم‌خونی نشان داد که مقادیر این شاخص بین ۰/۰۰۳ (نژاد مغولی) و ۰/۱۶۶ (نژاد تارو برد) در نوسان می‌باشد. در نتایج تحقیق آنان مشخص شد که در بین اسب‌های نژاد آسیایی، نژاد کاسپین با وجود تنوع ژنتیکی مناسب و ضریب هم‌خونی پایین، اندازه موثر آن نسبت به سایر نژادهای مورد مطالعه در تحقیق مذکور کمتر می‌باشد. در پژوهش اسدالله‌پور نعنایی و همکاران (۲۰۲۱) به منظور شناسایی هم‌خونی ژنومی گزارش شد که میانگین تعداد رشته‌های هموزیگوت (ROH) در نژاد آخال‌تکه نسبت به برخی نژادها کمتر بود. همچنین میانگین تعداد رشته‌های هموزیگوت کمتر از ۱۰۰۰kb که نشان‌دهنده هم‌خونی است، در نسل‌های جدید نژاد آخال‌تکه نسبت به نژادهای دیگر و عرب بیشتر بود. در کل، در نتایج آنان گزارش شد که بیش‌ترین میزان عدم تعادل پیوستگی در تمام فواصل ژنوم مربوط به ترتیب متعلق به نژادهای ترو برد، استاندارد برد و آخال‌تکه بود. در مطالعه نصیرپور و همکاران (۲۰۲۳) نیز بالاترین مقدار عدم تعادل پیوستگی برای نژاد کرد در کروموزوم‌های ۳، ۲۸، ۲۵، ۳۰ و ۸ و برای نژاد کاسپین در کروموزوم‌های ۳، ۴، ۳۰، ۲۳، ۱۷، ۸ و ۲۸ گزارش شد.



شکل ۵. مقایسه میزان عدم تعادل پیوستگی در بین نژادهای بومی اسب ایران

تجزیه و تحلیل رشته‌های هموزیگوت (ROH)

با انجام آنالیزهای ژنومی مشخص شد که میانگین بیش‌ترین ضریب هم‌خونی فردی برای هر نژاد به ترتیب مربوط به اسب عرب اصیل، دره‌شوری، کرد، ترکمن و کاسپین می‌باشد (شکل ۵). نمودار ویولون دو نقطه پرت و خارج از گروه را در نژادهای کرد و کاسپین شناسایی کرد. این نقاط با استفاده از محاسبات نرم‌افزار PLINK مربوط به هم‌خونی، به عنوان افراد ناخالص و خارج نژاد شناسایی شدند. با این حال، با توجه به اطلاعات محدود شجره و کتاب تبارنامه موجود در دو نژاد کرد و کاسپین، امکان تعیین اجداد مشترک افراد در دو نژاد مذکور مسیر نبود. همان‌گونه که از شکل شماره ۵ مشخص است در بین نمونه‌های مورد بررسی، تنوع اسب دره‌شوری نسبت به دیگر نژادها دارای بیش‌ترین مقدار بود. در تفسیر نتیجه حاصل از مشاهده هم‌خونی در نژادهای مورد بررسی باید خاطر نشان کرد مقادیر بدست آمده برای نژادها تا حدود زیادی قابل انتظار و طبیعی می‌باشد. زیرا بالا بودن میزان هم‌خونی در نژاد عرب اصیل به علت بسته شدن کتاب تبارنامه در سال ۱۳۵۵، انسجام و مدیریت صحیح آمیزش در بین اسب‌ها با وجود تیره‌های مختلف است. از طرفی نیز با مشاهده نمودار زیر می‌توان به میزان فشار انتخاب که طی قرون متمادی بر نژاد عرب اصیل صورت گرفته پی برد. همان‌گونه که مشخص است مقدار میانگین FIS در جمعیت کاسپین کمتر از بقیه نژادها بود و حاکی از بالا بودن میزان هتروزیگوسیتی و پلی‌مورفیسم در این نژاد است. همچنین، نحوه ساختار نمودار نژادهای کرد و دره‌شوری بیانگر ورود خون خارج نژادی به این جمعیت‌ها است. در تکمیل نتایج حاصل از تخمین هم‌خونی نژادهای ایرانی باید خاطر نشان کرد با وجود اینکه نژاد کاسپین به همراه اسب عرب اصیل جزو اولین اسب‌ها با تبارنامه مشخص و معتبر است، اما در سالیان گذشته به دلیل جمعیت کم آن اسب‌های ناخالصی به خزانه ژنی این نژاد ورود پیدا کرده‌اند. در پژوهشی که توسط پترسون و همکاران (۲۰۱۳) بر روی ۸۱۴ اسب از ۳۶ نژاد مختلف با استفاده از نشانگرهای SNP صورت گرفت، میزان هم‌خونی نژادهای عرب، کاسپین و آخال تکه را به ترتیب (۰/۰۳۳)، (-۰/۰۲۲) و (۰/۰۱۵) گزارش شد که میزان هم‌خونی نژاد کاسپین با نتیجه این تحقیق برای همین نژاد تفاوت معناداری دارد. در مطالعه یوسفی و همکاران (۲۰۲۱) جمعیت اسب کرد ROH کمتری معادل ۶Mb نسبت به اسب عرب اصیل و ترابرد داشت. در آن مطالعه میانگین کم‌ترین مقدار ضریب هم‌خونی به ترتیب برای اسب عرب اصیل، کرد و ترابرد گزارش شد که به طور معنی‌داری با نتایج حاصل از پژوهش حاضر متفاوت بود. با توجه به آنالیزهای آماری و تفسیر نتایج به دست آمده مشخص شد که علی‌رغم خزانه ژنی کوچک و مقدار هم‌خونی گزارش شده، اسب نژاد کاسپین از تنوع ژنتیکی بالایی برخوردار است و می‌توان اظهار کرد به استثنای اثرات رانش ژنتیکی، تمامی نژادهای مشتق‌شده از مناطق اولیه اهلی‌سازی دارای بالاترین ضریب هتروزیگوسیتی در بین

نژادهای حیوانی هستند (Warmuth et al., 2011). در کل نتایج حاصل از مطالعات پیشین تایید می‌کند که صرف نظر از منشأ ژنتیکی حیوانات و شناخت الگوهای آمیزشی، تنوع موجود در جمعیت نژادهای اسب تحت تاثیر جنسیت آنهاست که احتمالاً به دلیل وجود جریانات ژنی بین جمعیت‌ها در طی سال‌های طولانی فرآیند اهلی‌سازی و استفاده از آنها به عنوان وسیله جابجایی بوده است. (Petersen et al., 2013).



شکل ۶. نمودار ویولن میانگین، چارک‌ها و فراوانی (عرض نمودار) ضریب هم‌خونی مبتنی بر ROH (FROH) برای هر گروه نژادی

نتیجه گیری کلی

مطالعه انجام گرفته نشان دهنده یک تجزیه و تحلیل در ژنوم اسب‌های بومی ایران می‌باشد، که به طور کلی تصویری را از روابط بین نژادهای ایرانی فراهم می‌کند. توالی‌یابی و تکثیر ژنوم اسب یکی از دستاوردهای مهم است که کاربرد گسترده‌ای در بهبود عملکرد و سلامت حیوان و نیز درک بیشتر تفاوت‌های تکاملی و مولکولی با دیگر پستانداران دارد. این تحقیق با بررسی تفاوت فراوانی SNPها در سطح ژنوم اسب با استفاده از تراشه Equine SNP 70k و سپس شناسایی ساختار دقیق نژادهای بومی صورت گرفت. در این پژوهش با توجه به این که از روش‌های مختلفی برای بررسی ساختار جمعیت به کار برده شد و کلیه روش‌های مذکور توانست ساختار جمعیت‌های بومی را نشان دهد، نتایج به دست آمده نشان داد که با وجود افراد خالص در این جمعیت‌ها نژادهای ایرانی از اکوتیپ‌های مختلف و نژادهای متفاوتی می‌باشند که در مقایسه با یکدیگر اشتراک ژنتیکی زیادی دارند. اما به هر حال در بررسی نژادهای بومی به صورت مجزا مشخص شد که این نژادها نیز با یکدیگر متفاوت بوده و در خوشه‌های جداگانه‌ای قرار می‌گیرند.

تشکر و قدردانی

از پرسنل و کارشناسان محترم فدارسیون سوارکاری ایران که در تهیه نمونه‌های مو و داده‌های ژنومی مربوط به این تحقیق ما را یاری کردند، تشکر و قدردانی می‌گردد.

- Abdoli, M. (2019). Genetic structure survey of Iranian horse breeds by SSR and SNP markers. Master Thesis. Department of Animal Science, University of Zanjan, Iran. pp.197. (In Persian).
- Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC bioinformatics*, 12, 1-6.
- Almarzook, S., Reissmann, M., Arends, D., & Brockmann, G. A. (2017). Genetic diversity of Syrian Arabian horses. *Animal genetics*, 48(4), 486-489.
- Asadollahpour Nanaei, H., Nosrati, M. and Mohammadabadi, M.R. (2021). Genetic structure analysis of Akhal-Teke horse population and comparison with other horse breeds by using whole genome sequencing data. *Modern Genetics Journal*. 16(4), pp.299-307. (In Persian).
- Baba ʔyi, N., Rafat, S.A., Moradi, M.H. and Feizi Derakhshi, M.R., 2021. Comparison of principal component analysis (PCA) and discriminant analysis of principal component (DAPC) methods for analysis of population structure in Akhal-Teke, Arabian, and Caspian horse breeds using genomic data. *Iranian Journal of Animal Science Research*, 13(3), pp.453-462. (In Persian).
- Biscarini, F., Cozzi, P., Gaspa, G., & Marras, G. (2018). detectRUNS: Detect runs of homozygosity and runs of heterozygosity in diploid genomes.
- Bohmanova, J., Sargolzaei, M., & Schenkel, F. S. (2010). Characteristics of linkage disequilibrium in North American Holsteins. *BMC genomics*, 11(1), 1-11.
- Corbin, L. J., Blott, S. C., Swinburne, J. E., Vaudin, M., Bishop, S. C., & Woolliams, J. A. (2010). Linkage disequilibrium and historical effective population size in the Thoroughbred horse. *Animal Genetics*, 41, 8-15.
- Cosgrove, E. J., Sadeghi, R., Schlamp, F., Holl, H. M., Moradi-Shahrbabak, M., Miraei-Ashtiani, S. R., ... & Brooks, S. A. (2020). Genome diversity and the origin of the Arabian horse. *Scientific reports*, 10(1), 9702.
- Eltanany, M., Elfaroug Sidahmed, O., & Distl, O. (2015). Assessment of genetic diversity and differentiation of two major camel ecotypes (*Camelus dromedarius*) in Sudan using microsatellite markers. *Archives Animal Breeding*, 58(2), 269-275.
- Fages, A., Hanghøj, K., Khan, N., Gaunitz, C., Seguin-Orlando, A., Leonardi, M., ... & Orlando, L. (2019). Tracking five millennia of horse management with extensive ancient genome time series. *Cell*, 177(6), 1419-1435.
- Głazewska, I. (2010). Speculations on the origin of the Arabian horse breed. *Livestock Science*, 129(1-3), 49-55.
- Jombart, T., Kamvar, Z. N., Collins, C., Lustrik, R., Beugin, M. P., Knaus, B. J., & Jombart, M. T. (2018). Package 'adegenet'. Github repository:< <https://github.com/thibautjombart/adegenet>.
- Jun, J., Cho, Y. S., Hu, H., Kim, H. M., Jho, S., Gadhvi, P., ... & Bhak, J. (2014). Whole genome sequence and analysis of the Marwari horse breed and its genetic origin. *BMC genomics*, 15(9), 1-10.
- Ivanković, A., Bittante, G., Konjačić, M., Kelava Ugarković, N., Pećina, M., & Ramljak, J. (2021). Evaluation of the Conservation Status of the Croatian Posavina Horse Breed Based on Pedigree and Microsatellite Data. *Animals*, 11(7), 2130.
- Khalili, M. (2009). Horse and my expertise. Tehran, Nashr-e Zare Publication. pp. 694. (In Persian).
- Khalt-Abadi Farahani, A. H., & Moradi, M. H. (2018). Estimation of inbreeding values using genomic run of homozygosity and study of an evolutionary trend for effective population size in some Asian horse breeds. *Iranian Journal of Animal Science*, 49(3), 381-392. (In Persian).
- Khan Ahmadi, A. (2017). Genomic exploration to track selection patterns in Turkmen and Kurdish horses using high-density SNP markers. Ph.D. Thesis, Sari University of Agricultural Sciences and Natural Resources, Sari, Iran, pp. 140. (In Persian).
- Mahrous, K. F., Hassanane, M., Mordy, M. A., Shafey, H. I. and Hassan, N. (2011). Genetic variations in horses using microsatellite markers. *Journal of Genetic Engineering and Biotechnology*. 9:103-9.
- Maghsoodi, S. M., Mehrabani Yeganheh, H., Nejati Javaremi, A. and Yousefi Mashouf, N. 2017. Investigating population structure and identifying signatures of selection in Iranian Kurdish and

- Arabian horses. *Iranian Journal of Animal Science*, 48(3), 429-438. (In Persian).
- McCoy, A. M., & McCue, M. E. (2014). Validation of imputation between equine genotyping arrays. *Animal genetics*, 45(1), 153.
- McCue, M. E., Bannasch, D. L., Petersen, J. L., Gurr, J., Bailey, E., Binns, M. M., ... & Mickelson, J. R. (2012). A high density SNP array for the domestic horse and extant Perissodactyla: utility for association mapping, genetic diversity, and phylogeny studies. *PLoS genetics*, 8(1), e1002451.
- Metzger, J., Karwath, M., Tonda, R., Beltran, S., Águeda, L., Gut, M., ... & Distl, O. (2015). Runs of homozygosity reveal signatures of positive selection for reproduction traits in breed and non-breed horses. *BMC genomics*, 16, 1-14.
- Mousavi, S. F., Razmkabir, M., Rostamzadeh, J., & Seyedabadi, H. R. (2023). Evaluating of population diversity and detecting of genomic footprints of selection in four main Iranian horse breeds. *Agricultural Biotechnology Journal*, 15(2), 19-44. (In Persian).
- Saghi, D., & Mobaraki, A. (2018). Estimation of inbreeding and survey of the pedigree structure of Iranian Turkmen horses population. *Research On Animal Production (Scientific and Research)*, 9(22), 131-137. (In Persian).
- Nasirpour, M., & Moradi Shahr Babak, H. (2024). Analysis of SNP Chip 70k Genomic Data to Identify Loci Related to Differentiation of Caspian and Kurdish Horses using Selection Sweep. *Research On Animal Production*, 14(39), 154-162. (In Persian).
- Ovchinnikov, I. V., Dahms, T., Herauf, B., McCann, B., Juras, R., Castaneda, C., & Cothran, E. G. (2018). Genetic diversity and origin of the feral horses in Theodore Roosevelt National Park. *Plos one*, 13(8), e0200795.
- Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, J. Z. (2012). Genomic patterns of homozygosity in worldwide human populations. *The American Journal of Human Genetics*, 91(2), 275-292.
- Petersen, J. L., Mickelson, J. R., Cothran, E. G., Andersson, L. S., Axelsson, J., Bailey, E., ... & McCue, M. E. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PloS one*, 8(1), e54997.
- Pometti, C. L., Bessega, C. F., Saidman, B. O., & Vilardi, J. C. (2014). Analysis of genetic population structure in *Acacia caven* (Leguminosae, Mimosoideae), comparing one exploratory and two Bayesian-model-based methods. *Genetics and molecular biology*, 37, 64-72.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3), 559-575.
- Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC genetics*, 13, 1-11.
- Qanbari, S., Pimentel, E. C. G., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A. R., & Simianer, H. (2010). The pattern of linkage disequilibrium in German Holstein cattle. *Animal genetics*, 41(4), 346-356.
- R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rafeie, F., Amirinia, C., Javaremi, A. N., Mirhoseini, S. Z., & Amirmozafari, N. (2011). A study of patrilineal genetic diversity in Iranian indigenous horse breeds. *African Journal of Biotechnology*, 10(75), 17347-17352.
- Sadeghi, R., Moradi-Shahrbabak, M., Miraei Ashtiani, S. R., Schlamp, F., Cosgrove, E. J., & Antczak, D. F. (2019). Genetic diversity of Persian Arabian horses and their relationship to other native Iranian horse breeds. *Journal of Heredity*, 110(2), 173-182.
- Schubert, M., Jónsson, H., Chang, D., Der Sarkissian, C., Ermini, L., Ginolhac, A., ... & Orlando, L. (2014). Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *Proceedings of the National Academy of Sciences*, 111(52), E5661-E5669.
- Wade, C. M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Imsland, F., ... & Lindblad-Toh, K. (2009). Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, 326(5954), 865-867.

- Warmuth, V., Eriksson, A., Bower, M. A., Cañon, J., Cothran, G., Distl, O., ... & Manica, A. (2011). European domestic horses originated in two Holocene refugia. *PLoS one*, 6(3), e18194.
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *evolution*, 1358-1370.
- Yousefi-Mashouf, N., Mehrabani-Yeganeh, H., Nejati-Javaremi, A., Bailey, E., & Petersen, J. L. (2021). Genomic comparisons of Persian Kurdish, Persian Arabian and American thoroughbred horse populations. *PLoS One*, 16(2), e0247123.
- Zandi, M. B. 2014. A study of patrilineal genetic diversity in Iranian indigenous horse breeds. Ph.D. Thesis, University of Tehran, Iran. (*In Persian*).

Survey of population structure and linkage disequilibrium pattern of Iranian native horse breeds by SNP markers

Introduction: The horse (*Equus ferus caballus*) has been one of the oldest animals domesticated by humans. In terms of geographical distribution, this animal is also present in all the lands, and horses are included as the most popular domesticated animals in human societies (Jun et al., 2014; Mahrous et al., 2011 and Petersen et al., 2013). At present, the populations of Iranian native horse breeds include Turkmen, Caspian, Kurdish, Dareshouri, and Arabian (Asil) (Rafeie et al., 2011). From the five official registered horse breeds of Iran, Turkmen and Asil horses had the greatest impact on the breeding program and foundation of many breeds. So, most bloodlines of stallions from other breeds originate from these two horse breeds (Wallner et al. 2017). Therefore, based on archaeological findings, the early stages of the establishment of most European horse breeds originated from the bloodline of Iranian horses and these breeds played a significant role in the development of horse breeds in globe (Fages et al., 2019). So, this research was carried out to investigate and genomic scan of the structure, genetic diversity, linkage disequilibrium (LD), and estimate the level of inbreeding by Illumina 70k SNP Equine bead chip to determine the distances and genetic structure related to the mentioned populations, to identify some conservation strategies for further investigations.

Material and methods: In this study, the genotype dataset of 167 horses including Turkmen (40), Caspian (22), Kurdish (66), Dareshouri (15), and Arabian Asil (24) from the Equestrian federation of Iran databases were used. It was tried to ensure that none of the horses have family relationships in the pedigree for the analysis of genomic data. The hair samples were sent to the Geneseek laboratory in the United States and were genotyped using Illumina 70k SNP Equine bead chip. Data quality control was applied by PLINK v 1.9 (Purcell et al., 2007) software, and loci with MAF < 2%, Mind < 5%, HWE (<10⁻⁶), and genotype call rate < 1% were removed. Finally, 159 horses with 45,270 SNP loci were reminded for further analysis. Principal component analysis (PCA) was performed on SNP data using the "Prcomp" function in R v 4.2.1 (R Development Core Team, 2022). Finally, to examine the population structure, the assignment test, also determines the genomic kinship relationships between individuals in order to distinguish populations and identify outliers, we used adegenet v 2.0.0 package (Jombart et al., 2018) in R. Assigning of individual into their genetic groups was examined using the Structure v 2.3.4 (Pritchard et al., 2000) assuming K=2 to 8. The Structure algorithm included the admixture model and correlated allele frequencies.

Results and discussion: Data quality control was applied and loci with MAF < 2%, Mind < 5%, HWE (<10⁻⁶), and genotype call rate < 1% were removed. Finally, 159 horses with 45,270 SNP loci were reminded. The results of PCA analysis showed that the populations were separated based on PC1 and based on PC2 geographic regions of all samples were separated into five main components. According to the PCA analysis, it can be seen that Dareshouri breed has a slight genomic admixture with Arabian Asil and Kurdish horses, while Arabian Asil, Turkmen, and Caspian were well differentiated from each other, which is close with results of Babayi et al., (2021). Similar to such grouping with a little admixture was also observed in Yousefi Mashouf et al., (2021) and Maghsoudi et al., (2017). In this research, both PCA and DAPC methods correctly separated the population of Iranian native horse breeds. The analysis of the genetic structure within the whole sample of Iranian horses with the admixture algorithm confirmed lack of differentiation between Caspian and Kurdish, Arabian Asil, and Dareshouri breeds at K=2, then we select different K from 2 to 5 as the optimal structure.

The further results for K from 5 to 8 did not add information about genetic differentiation between populations. The LD decay among all breeds showed a significant decrease below 400 bp. Caspian and Dareshouri breeds experienced the lowest and highest LD and the highest and lowest effective population size, respectively. The low level of LD of the Caspian breed indicates that the effective size of this population is greater than that of other breeds. The inbreeding value based on runs of homozygosity (ROH) were highest and lowest for Arabian Asil and Caspian breeds, respectively. The results of this study showed that the Iranian native horse breeds were well defined by SNP markers, which can be used for stud base foundations of breeds with more accuracy.

Keywords: *Iranian native horses, Genetic diversity, linkage disequilibrium, Runs of homozygosity, single nucleotide polymorphism markers.*

فیلد استنادی
نقشه