

Evaluating of population diversity and detecting of genomic footprints of selection in four main Iranian horse breeds

Seyedeh Fatemeh Mousavi 

Ph.D. Student, Department of Animal Science, Faculty of Agriculture Engineering, University of Kurdistan, Sanandaj, Iran. E-mail address: mousavisfateme@gmail.com

Mohammad Razmkabir 

*Corresponding author. Associate Professor, Department of Animal Science, Faculty of Agriculture Engineering, University of Kurdistan, Sanandaj, Iran. E-mail address: m.razmkabir@uok.ac.ir

Jalal Rostamzadeh 

Assistant Professor, Department of Animal Science, Faculty of Agriculture Engineering, University of Kurdistan, Sanandaj, Iran. E-mail address: j.rostamzadeh@uok.ac.ir

Hamid-Reza Seyedabadi 

Associate Professor, Animal Science Research Institute of Iran, Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran. E-mail address: hseyedabady@gmail.com

Abstract

Objective

Natural and artificial selections change the allele frequencies among populations and consequently leave detectable patterns on the genome during domestication. A little information about the differences of genetic structures of Iranian native breeds that result in these selective pressures are available. Therefore, in this study, we examined genetic diversity using four native horse breeds including Turkmen (29 samples), Caspian (21 samples), Kurdish (67 samples) and Persian Arabian (52 samples).

Materials and methods

For this purpose, three methods including Principal Component Analysis (PCA), neighbor joining (NJ) and admixture were investigated. In addition, with the integrating of three different selection of signature methods, including TajimaD, nucleotide divergency (π) and integrated haplotype

homozygosity score could identify the selection traces in these breeds. To integrate the different results of these three methods, De-correlated composite of multiple signals framework was used.

Results

All the methods used to examine the genetic structure well separated these four breeds and showed a pattern related to geographical origin for their genetic diversity. By combining different methods of identifying the selection signatures, many genes affected by the selection pressure were identified in all four breeds. These genes were associated with specific GO terms and QTL, in which 16 shared candidate genes among all four breeds were identified as suggested candidates for selection. Additionally, 11 different QTL types was identified in all four studied breeds, which was divided into the category of traits related to adaptation, fitness through genetic disorders or morphological and behavioral traits.

Conclusions

Overall, the results of this study can help better understand the process of natural and artificial selections in Iranian horses. In addition, this research has helped to improve our understanding about the genetic differences between studied Iranian horse breeds, which can help in finding the best solution to preserve and improve genetic diversity.

Keywords: selection signatures, Asil horse, Caspian horse, Turkmen horse, Kurdish horse

Paper Type: Research Paper.

Citation: Mousavi SF, Razmkabir M, Rostamzadeh J, Seyedabadi HR (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.

Agricultural Biotechnology Journal 15 (2), 19-44.

DOI: 10.22103/jab.2023.20153.1426

Received: February 04, 2023.

Received in revised form: March 22, 2023.

Accepted: March 23, 2023.


Published online: June 10, 2023.

Publisher: Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society.




© the authors

بررسی تنوع ژنتیکی و شناسایی نشانه های ردپای انتخاب در چهار نژاد اصلی اسب ایرانی

سیده فاطمه موسوی 


دانشجوی دکتری، گروه علوم دامی، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران. رایانامه:

mousavisfateme@gmail.com


محمد رزم کبیر 

*نویسنده مسئول: دانشیار، گروه علوم دامی، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران. رایانامه:

m.razmkabir@uok.ac.ir

جلال رستمزاده 

استادیار، گروه علوم دامی، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران. رایانامه: j.rostamzadeh@uok.ac.ir

حمیدرضا سیدآبادی 

دانشیار، موسسه تحقیقات علوم دامی کشور، کرج، ایران. رایانامه: hseyedabady@gmail.com

تاریخ دریافت: ۱۴۰۱/۱۱/۱۵ تاریخ دریافت فایل اصلاح شده نهایی: ۱۴۰۲/۰۱/۰۲ تاریخ پذیرش: ۱۴۰۲/۰۱/۰۳

چکیده

هدف: انتخاب های طبیعی و مصنوعی باعث تغییر در فراوانی آلی در بین جمعیت ها شده و به صورت مداوم الگوهای قابل شناسایی را بر روی سطح ژنوم طی فرآیند اهلی سازی برجای می گذارند. اطلاعات کمی در خصوص تفاوت های ساختارهای ژنتیکی در نژادهای اسب بومی ایران که منتج از فشارهای انتخابی هستند وجود دارد. بنابراین، در این مطالعه با استفاده از اطلاعات چهار نژاد اسب بومی شامل ترکمن (۲۹ نمونه)، کاسپین (۲۱ نمونه)، کرد (۶۷ نمونه) و اصیل ایرانی (۵۲ نمونه) به بررسی تنوع ژنتیکی پرداخته شد.

مواد و روش ها: برای این منظور سه روش تجزیه مولفه های اصلی (PCA)، پیوستگی مجاور (NJ) و اختلاط جمعیتی مورد بررسی قرار گرفت. در ادامه با کمک ادغام سه روش مختلف شناسایی نشانه های انتخاب شامل TajimaD، تنوع نوکلئوتیدی (pi) و درجه هاپلوتایپی یکپارچه (iHS) اقدام به شناسایی ردپاهای انتخاب در این نژادها گردید. برای ادغام نتایج مختلف این سه روش، از چهارچوب ترکیب همبسته سیگنال های چند گانه استفاده گردید.

نتایج: تمام روش های مورد استفاده برای بررسی ساختار ژنتیکی، به خوبی این چهار نژاد را از هم جدا نموده و یک الگوی مرتبط با منشاء جغرافیایی را برای تنوع ژنتیکی آن ها نشان داد. با ترکیب روش های مختلف شناسایی نشانه های انتخاب، تعداد زیادی از ژن های تحت تاثیر فشار انتخاب در هر چهار نژاد شناسایی گردید. این ژن ها با مسیرهای خاص GO و نواحی مرتبط با QTL در ارتباط بودند. تعداد ۱۶ ژن مشترک در بین چهار نژاد به عنوان ژن های کاندید انتخاب شناسایی گردید. بعلاوه، در هر چهار نژاد تحت بررسی تعداد ۱۱ نوع QTL مشترک شناسایی گردید که به دسته صفات مرتبط با سازگاری، شایستگی از طریق نقص های ژنتیکی و یا رفتاری و ظاهری قابل تقسیم بودند.

نتیجه گیری: در مجموع نتایج بدست آمده در این تحقیق می تواند در درک بهتر فرآیند انتخاب طبیعی و مصنوعی در اسب های ایران کمک نماید. همچنین این تحقیق به درک بهتر تفاوت های ژنتیکی بین نژادهای اسب بومی ایران کمک نموده که می تواند در یافتن بهترین راه حل برای حفظ و بهبود تنوع ژنتیکی کمک شایانی نماید.

کلیدواژه ها: نشانه های انتخاب، اسب اصیل، اسب کاسپین، اسب ترکمن، اسب کرد.

نوع مقاله: پژوهشی.

استناد: موسوی سیده فاطمه، رزم کبیر محمد، رستم زاده جلال، سیدآبادی حمیدرضا (۱۴۰۲) بررسی تنوع ژنتیکی و شناسایی نشانه های ردپای انتخاب در چهار نژاد اصلی اسب ایرانی. *مجله بیوتکنولوژی کشاورزی*، ۱۵(۱)، ۱۹-۴۴.

Publisher: Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society.



© the authors

مقدمه

مطالعات نشان می دهد که اسب ها برای اولین بار در حدود ۳۵۰۰ سال پیش در استپ های قزاقستان و روسیه اهلی شدند (Anthony & Brown 2011). پس از اهلی سازی، افزایش مکانیزاسیون سبب کاهش نیاز به اسب و در نتیجه کاهش جمعیت آن ها در سرتاسر دنیا از جمله ایران گردید. وجود تفاوت های زیاد در شرایط زراعی-اکولوژیکی^۱ در نقاط مختلف ایران، موجب پیدایش و تکامل نژادهای مختلف اسب با خصوصیات منحصر به فرد گردید. در حال حاضر پنج نژاد شاخص با قدمت دیرینه ی کشاورزی، اهمیت فرهنگی و تاریخی در ایران وجود دارند که شامل: نژادهای ترکمن، دره شوری، کاسپین، کرد و اصیل ایرانی (عرب ایرانی) می باشند (Fotovati 2000). برخی از این نژادها در خطر انقراض قرار دارند و بایستی برنامه های اصلاحی با قابلیت

¹ Agro-ecological conditions

افزایش تنوع ژنتیکی جهت جلوگیری از انقراض آن‌ها طراحی گردد. دانستن تفاوت‌های ژنتیکی می‌تواند در یافتن بهترین راه حل برای حفظ و بهبود تنوع ژنتیکی کمک شایانی نماید. اسب اصیل ایرانی از نژادهای بسیار مهم می‌باشد که در سطح جهانی با جمعیتی حدود ۱ میلیون راس توزیع شده است (World Arabian Horse Organization (WAHO)). این نژاد دارای جمعیت‌های اصلاحی در بسیاری از کشورها از جمله کشورهای خاورمیانه مانند عربستان سعودی، سوریه و ایران با خصوصیات ظاهری بسیار متفاوت می‌باشد (Forbis 1976). اسب ترکمن به عنوان یکی از قدیمی‌ترین نژادهای دنیا شناخته شده که دارای چندین منبع ژنتیکی منحصر به فرد با تنوع ظاهری، ساختاری و کاربردی متفاوت می‌باشد (Jiskrová et al. 2016). اسب کاسپین به عنوان اسب باستانی با اندازه کوچک شناخته می‌شود که منشا آن شمال ایران در حاشیه دریای خزر می‌باشد. این نژاد به همراه اسب ترکمن به عنوان منشا تمام اسب‌های شرقی پیشنهاد می‌شوند (Firouz 1998). اسب کرد که منشا آن غرب ایران و استان کرمانشاه می‌باشد، با شرایط کوهستانی و دمای نسبتاً سرد سازگاری بسیار خوبی دارد. شرایط محیطی بسیار سخت و خشن کرمانشاه منجر به تکامل اسب کرد با خصوصیات منحصر به فردی شده است. استراتژی انتخاب و آمیزش‌های کنترل شده می‌تواند برای حفظ این تنوع ژنتیکی بسیار مفید باشد (Nazari et al. 2022). با وجود مطالعات فراوان انجام شده در نژادهای بومی اسب ایران برای بررسی تنوع ژنتیکی، آزمون انساب، ساختار ژنتیکی جمعیت‌ها و یا شناسایی نشانه‌های انتخاب با استفاده از اطلاعات ریزماهورها^۲ و چندشکلی تک نوکلئوتیدی^۳ (SNP) بر روی تنها یک جمعیت (Seyedabadi et al. 2006; Gharahveysi & Irani 2011; Rahimi-Mianji et al. 2015; Ala-Amjadi et al. 2017; Nazari et al. 2022) و یا چند نژاد (Moridi et al. 2013; Sadeghi et al. 2019; Salek Ardestani et al. 2022)، در حال حاضر شواهد برای درک بهتر دلایل ژنتیکی تنوع جمعیت‌های اسب ایران بسیار اندک می‌باشد.

انتخاب‌های طبیعی و مصنوعی طی مراحل اهلی‌سازی باعث تغییر در فراوانی آللی بین جمعیت‌ها شده که موجب ایجاد الگوهای تنوع قابل شناسایی در سطح ژنوم می‌شوند (Saravanan et al. 2020). شناسایی این نواحی ژنتیکی که تحت تاثیر انتخاب قرار داشته‌اند به دلیل ارائه دید جدید از تاریخچه تکاملی جمعیت‌ها و کمک به شناسایی ریشه اصلی تنوع فنوتیپی بین نژادها یکی از مباحث جذاب جهت مطالعات ژنتیک دامی می‌باشد (Ceballos et al. 2018). از سویی دیگر، استفاده از تکنیک‌های مولکولی در سال‌های اخیر جهت مطالعه موجودات بومی و حفاظت شده، کاربرد گسترده‌ای یافته است (Askari et al. 2010; Mohammadabadi 2017). میزان اطلاعات به‌دست‌آمده از این تکنیک‌های ژنتیکی، یکی از پارامترهای قابل ارزیابی برای مطالعه جمعیت‌های مختلف و درک تفاوت‌های ژنتیکی بین جمعیت‌هاست (Mohammadifar et al. 2014; Mohammadifar and Mohammadabadi 2018). همچنین، مطالعه نژادهای مختلف با استفاده از تکنیک‌های مولکولی بسیار مهم و برای طبقه‌بندی آن‌ها مفید است (Mohammadabadi et al. 2017). حفاظت باید بر اساس دانش عمیقی از

² Microsatellite

³ Single nucleotide polymorphism

منابع ژنتیکی نژادهای خاص باشد، لذا تلاش برای شناسایی و تعیین خصوصیات ژنتیکی نژادهای مختلف بسیار اهمیت دارد (Ghasemi et al. 2010; Mohammadabadi et al. 2021) تنوع ژنتیکی یک عنصر اساسی برای پیشرفت ژنتیکی، حفظ جمعیت‌ها، تکامل و سازگاری به شرایط محیطی متغییر و مختلف می‌باشد (Askari et al. 2008; Mohammadifar and Mohammadabadi 2011). بررسی کل ژنوم برای شناسایی نواحی تحت تاثیر انتخاب به دلیل ارائه روش‌های پیشرفته ژنوتیپ کردن با قیمت‌های بسیار پایین امکان‌پذیر شده است. با وجودی که آزمون‌های آماری متنوعی برای شناسایی نشانه‌های انتخاب پیشنهاد شده است، استفاده از یک روش آماری دارای محدودیتی است که یک جایگاه تحت انتخاب ممکن است بر روی مکان‌های مجاور تأثیر بگذارد و وضوح جایگاه‌های تحت انتخاب را کاهش دهد (Lotterhos et al. 2017). علاوه بر این، مطالعات نشان دادند که روش‌های آماری مختلف به دلایل مختلف مانند حساسیت به اندازه جمعیت، ریشه تنوع ژنتیکی، و تاریخچه جمعیتی نتایج متفاوتی ارائه می‌دهند (Lotterhos & Whitlock 2015; Schlamp et al. 2016; Vatsiou et al. 2016). برای حل این مشکلات، استراتژی‌های متفاوتی مانند روش‌های برپایه ترکیب مقادیر-p (p-values) از روش‌های آماری مختلف (معیارهای ترکیبی انتخاب) استفاده می‌شوند (Randhawa et al. 2014; Lotterhos & Whitlock 2015). در این مطالعه به بررسی ساختارهای جمعیتی چهار نژاد مهم اسب بومی ایران شامل ترکمن، کاسپین، کرد و اصیل ایرانی (عرب ایرانی) پرداخته خواهد شد.

مواد و روش‌ها

جمع‌آوری داده‌ها: تمامی داده‌ها از چهار منبع اطلاعاتی مختلف جمع‌آوری شده‌اند. تعداد نمونه در هر نژاد و منابع اطلاعاتی در اطلاعات تکمیلی، جدول تکمیلی شماره ۱ گزارش شده است. اولین منبع اطلاعاتی مربوط به موسسه تحقیقات علوم دامی کشور^۴ (ASRI) با تعداد ۸۶ نمونه می‌باشد که شامل اطلاعات ژنومیک ۸۱ (۵۱ راس نر) اسب نژاد کرد، ۱ اسب اصیل نر و ۲ (۱ راس نر) اسب ترکمن است. این اسب‌ها بر اساس اطلاعات شجره‌ای انتخاب شده و هیچ تلاقی با دیگر نژادها در شجره شان مشاهده نشده است. همچنین تست انساب با استفاده از روش میکروستلایت برای آن‌ها انجام شده است. این نمونه‌ها از چهار استان کرمانشاه، کردستان، اصفهان و کرمان جمع‌آوری شده‌اند. منبع اطلاعاتی بعدی مربوط به مرکز اصلاح نژاد و بهبود تولیدات دامی کشور^۵ (ABC) می‌باشد. در این منبع اطلاعاتی، ۸۶ نمونه از سه نژاد ترکمن (۳۴ راس که ۳۰ راس نر بوده)، کاسپین (۲۲ راس که ۷ راس نر بوده) و کرد (۳۰ راس که ۲۸ راس نر بوده) در شش استان (کرمانشاه، کردستان، اصفهان، گلستان، مازندران و تهران) وجود داشت. این داده‌ها بین سال‌های ۲۰۱۴ و ۲۰۱۵ جمع‌آوری شده‌اند. برای هر نژاد، اسب‌ها بر اساس شجره و خصوصیات نژادیشان نمونه‌گیری می‌شوند. در این دو منبع اطلاعاتی، در مجموع اطلاعات مربوط به ۶۵۱۵۷ جایگاه SNP جهت مطالعه در

⁴ Animal Science Research Institute of IRAN

⁵ National Animal Breeding Center

دسترس قرار داشت. برای افزایش تعداد نمونه‌های مورد مطالعه و نیز اضافه نمودن نمونه‌های نژاد اسب اصیل ایرانی، داده‌های جمع آوری شده با اطلاعات دو منبع دیگر ادغام گردید که شامل ۹۶ نمونه (نژادهای اصیل ۷۱ راس (۲۳ راس نر)، ترکمن ۱۱ راس (۵ راس نر)، کاسپین ۷ راس (۴ راس نر) و کرد ۷ راس (۴ راس نر)) از مطالعه (Sadeghi et al. 2019) و همچنین تعداد ۹ (۲ راس نر) نمونه اسب اصیل از مطالعه Cosgrove et al. 2020 می‌باشد.

ژنوتیپ کردن و کنترل کیفیت داده‌ها: DNA نمونه‌های ASRI و ABC از تعداد ۵۰ عدد ریشه مو از یال اسب‌ها با استفاده از کیت خون و بافت DNeasy (کمپانی Qiagen آلمان) استخراج گردید. سپس این DNA استخراج شده با استفاده از آرایه Equine 70K SNP BeadChip در آزمایشگاه GeneSeek Inc. واقع در آمریکا ژنوتایپ گردید (Neogen GeneSeek, Lincoln, NE). دو منبع اطلاعاتی دیگر با استفاده از چیپ K Equine SNP۶۷۰ ژنوتایپ شد که در مقالات قبلی به جزئیات آن اشاره شده است (Sadeghi et al. 2019; Cosgrove et al. 2020). جایگاه SNP‌های مورد استفاده با توجه به نقشه فیزیکی ژنوم اسب EquCab3.0 تصحیح گردید (Kalbfleisch et al. 2018). ابتدا با استفاده از نرم افزار PLINK داده‌های مختلف که با دو تکنولوژی Affymetrix و GeneSeek ژنوتایپ شده بودند با یکدیگر ادغام شدند (Purcell et al. 2007). با توجه به اینکه مطالعه بر روی SNP‌های قرار گرفته بر روی کروموزوم‌های غیرجنسی بودند، SNP‌های قرار گرفته بر روی کروموزوم‌های جنسی حذف گردید. بعد از ادغام داده‌ها تعداد ۲۷۵ نمونه و ۴۰۸۱۳ جایگاه SNP برای کنترل کیفیت باقی ماند. در این مرحله SNP‌ها بر اساس نرخ ژنوتیپ کردن کمتر از ۹۰ درصد، فراوانی آلل کوچک^۶ (MAF) کمتر از ۱ درصد، یا خارج از تعادل هاردی واینبرگ^۷ با p-value کوچکتر از 10^{-6} حذف شدند. با توجه به احتمال حضور نمونه‌های با خویشاوندی بسیار نزدیک در هر نژاد که می‌توانند موجب ایجاد اربیی در آنالیزها گردند، افراد برای همسانی با منشأ تباری^۸ (IBD) کوچکتر از 0.25 مورد آزمون قرار گرفته و در جفت‌های با خویشاوندی نزدیک، یکی حذف گردید. برای انجام این بخش، از دستور genome- در نرم افزار PLINK استفاده گردید. قبل از حذف افراد خویشاوند، با استفاده از نرم افزار Beagle V5.2 ژنوتایپ‌های گم شده در داخل هر نژاد بصورت جداگانه ایمپوت گردید (Browning & Browning 2007). با توجه به اینکه هرس بر پایه LD^۹ می‌تواند باعث کاهش اثرات اربیی تشخیصی^{۱۰} در برآوردهای تنوعی گردد (Malomane et al. 2018)، برای بررسی ساختارهای جمعیتی با روش‌های تجزیه مولفه‌های اصلی^{۱۱} (PCA)، پیوستگی مجاور^{۱۲} (NJ) و اختلاط^{۱۳} جمعیتی، یک زیرمجموعه از SNP‌ها بر پایه هرس LD با استفاده از PLINK انتخاب گردید. برای آنالیز PCA، با استفاده از دستور

⁶ Minor allele frequency

⁷ Hardy-weinberg

⁸ Identity-by-descent

⁹ LD based pruning

¹⁰ Ascertainment biases

¹¹ Principal component analysis

¹² Neighbor-joining

¹³ Admixture

indep-pairwise با در نظر گرفتن شرایط پنجره‌های ۵۰ SNP و جابجایی ۵ SNP برای هر پنجره، و نیز همبستگی بین SNP (r^2) بیشتر از 0.5، SNP‌های اضافی حذف شدند. هرس مشابه برای آنالیزهای اختلاط و NJ به جز برای r^2 (بیشتر از 0.3) در نظر گرفته شد.

بررسی ساختار جمعیت: از دستور prcomp (با scale = TRUE) در نرم افزار آماری R version 4.0.2 جهت

انجام آنالیز PCA برای بررسی ارتباط بین جمعیت‌ها استفاده گردید (Team 2013). با توجه به اینکه سه PC اول، هر چهار جمعیت را به خوبی جدا کرده، از این سه معیار برای شناسایی نمونه‌های پرت استفاده گردید. دلیل حذف نمونه‌های پرت احتمال وجود خطا در شجره‌ها و تلاقی بین نژادی ثبت نشده در گذشته بوده که در نژادهای ایرانی وجود دارد. از الگوریتم فاکتور پرت محلی^{۱۴} (LOF) برای امتیازدهی برای هر نمونه در یک فرآیند چند بعدی استفاده گردید. LOF یک روش تشخیص ناهنجاری بدون نظارت بوده که انحراف چگالی محلی را با در نظر گرفتن تعداد محدودی از نمونه‌های مجاور در هر بعد برای هر نمونه محاسبه می‌نماید (Breunig et al. 2000). با فرض اینکه نمونه‌های پرت دارای یک شکاف فاصله‌ای با سایر نمونه‌ها می‌باشد، تعداد ۱۹ نمونه از داده‌ها حذف گردید (اطلاعات تکمیلی، شکل تکمیلی ۱).

ارتباط بین جمعیت‌های مورد مطالعه با استفاده از محاسبه فواصل ژنتیکی افراد با کمک NJ مورد بررسی قرار گرفته است. درخت NJ با استفاده از VCF-kit (Cook & Andersen 2017) ایجاد و در FigTree نمایش داده شد (<http://tree.bio.ed.ac.uk/software/figtree>). از نرم افزار پرکاربرد ADMIXTURE برای استنباط منشا نژادی و میزان اختلاط بین جمعیت‌ها استفاده گردید (Alexander et al. 2009). برای هر فرد، نسبت اجدادی با در نظر گرفتن اجزاء جدی از پیش تعیین شده (K) به عنوان تعداد جمعیت‌های اجدادی محاسبه گردید (Alexander & Lange 2011). یک روش اعتبارسنجی 5-fold برای هر K (از ۱ تا ۷) برای محاسبه خطای استاندارد جهت مقایسات استفاده شده است به طوری که K با کوچکترین خطا به عنوان محتمل‌ترین تعداد جمعیت‌های اجدادی در نظر گرفته شد.

شناسایی نواحی تحت انتخاب: در این مطالعه از سه روش تک جمعیتی شامل: (۱) TajimaD (Tajima 1989)، (۲)

تنوع نوکلئوتیدی (pi) (Nei & Li 1979)، و (۳) درجه هاپلوپولیمورفیسم یکپارچه^{۱۵} (Voight et al. 2006) استفاده گردید. برای این منظور دو روش TajimaD و pi با کمک نرم افزار vcftools آنالیز گردید (Danecek et al. 2011). برای محاسبه آماره‌های TajimaD در هر نژاد و کروموزوم به صورت جداگانه، از دستور TajimaD- بدون در نظر گرفتن همپوشانی و اندازه پنجره‌های 300 Mb استفاده گردید. مقادیر D محاسبه شده برای هر پنجره 300-Mb به همان SNP‌های قرار گرفته در داخل آن پنجره متناسب شده و مقادیر محاسبه نشده به صفر تبدیل گردید. آماره‌های pi نیز با استفاده از دستور site-pi- همانند روش قبلی برای

¹⁴ Local outlier factor

¹⁵ Integrated haplotype score

هر نژاد و کروموزوم به صورت جداگانه محاسبه گردید. برای کاهش اختلال (نویز)، خروجی‌های این روش برای هر کروموزوم با استفاده از دستور runmed در R با پنجره‌های ۳۱ SNPs ("constant" = endrule) هموار گردید.

روش iHS بر پایه این اصل می‌باشد که جهش‌ها و نوترکیبی‌ها تمایل به ایجاد ساختارهای هموزیگوت کوتاه‌تری دارند، بنابراین آلل اجدادی (قدیمی‌تر) باید هاپلوتایپ‌های هموزیگوت کوتاه‌تری از آلل‌های مشتق شده (جدیدتر) داشته باشند (Voight et al. 2006). برای محاسبه آماره‌های iHS از پکیج rehh در R استفاده گردید (Gautier & Vitalis 2012).

ترکیب همبسته سیگنال‌های چندگانه (DCMS): ترکیب آماره‌های تست‌های مختلف نشانه‌های انتخاب موجب بهبود نرخ اختلال (نویز) شده و در نتیجه باعث افزایش وضوح شناسایی نواحی تحت انتخاب می‌شود (Ma et al. 2015; Lotterhos et al. 2017). برای این منظور آماره‌های سه روش شناسایی نواحی تحت انتخاب به مقادیر DCMS تبدیل گردید. این روش همبستگی بین آماره‌های مختلف را در نظر گرفته که با ترکیب مقادیر p محاسبه شده برای هر جایگاه یک مقدار DCMS را ارائه می‌دهد. برای محاسبه DCMS در جایگاه l از فرمول زیر استفاده می‌شود:

$$DCMS_l = \sum_{t=1}^m \frac{\log\left(\frac{1-p_{lt}}{p_{lt}}\right)}{\sum_{i=1}^m |r_{it}|}$$

که در اینجا p_{lt} مقادیر p-value در جایگاه l برای روش t بوده ($t = 1, \dots, m$) به عنوان همبستگی در سطح گستره ژنومی بین تست‌های آماری t ام و t ام می‌باشد. مخرج کسر در این معادله نشان‌دهنده عامل وزنی برای جلوگیری از تاثیر همبستگی زیاد بین آماره‌ها بوده که بسته به میزان همبستگی گستره کل ژنومی آنها از مقادیر $1/m$ تا ۱ متغیر می‌باشد.

در این مطالعه مقادیر DCMS برای هر جایگاه در هر نژاد با ترکیب مقادیر آماره روش‌های ذکر شده (iHS، TajimaD و pi) با توجه به مطالعه (Yurchenko et al. 2018) انجام شد. محاسبه مقادیر DCMS برای هر جایگاه SNP طی چند مرحله صورت گرفت که شامل: (۱) مقادیر p-value محاسبه شده در هر روش ابتدا توسط پکیج MINOTAUR در R با آزمون یک طرفه راست^{۱۶} برای iHS و یک طرفه چپ^{۱۷} برای pi و TajimaD با استفاده از دستور stat_to_pvalue محاسبه شد (Verity et al. 2017). (۲) ماتریکس کواریانس بین آماره‌های مختلف با استفاده از دستور CovNAMcd در پکیج R به نام rrcovNA با آلفا 0.75 و ۳۰۰۰۰ نمونه تصادفی از SNPs برآورد شد (Todorov et al. 2011). (۳) این ماتریکس کواریانس محاسبه شده به عنوان ورودی برای دستور DCMS در MINOTAUR برای محاسبه مقادیر DCMS برای هر SNP استفاده گردید. (۴) خروجی دستور DCMS در یک توزیع نرمال با استفاده از مدل خطی در دستور rml در پکیج MASS محاسبه شد (Venables & Ripley 2013) (model = rlm (dcms ~ 1)). دستور rlm مقادیر میانگین و انحراف استاندارد محاسبه کرده که به عنوان ورودی در دستور pnorm با lower.tail برابر FALSE استفاده گردید.

¹⁶ Right-tailed test

¹⁷ Left-tailed test

شناسایی ژن‌های کاندیدا و جایگاه‌های صفات کمی^{۱۸} (QTLs)، و آنالیز کارکردی^{۱۹}: اطلاعات جایگاه ژن‌های اسب و هستی‌شناسی^{۲۰} (GO) از ژنوم EquCab3.0 از پایگاه داده Biomart جمع‌آوری گردید (Kasprzyk 2011). برای شناسایی ژن‌های تحت تاثیر انتخاب، فاصله حداکثر ۴۰ kbp برای تمام ۱ درصد SNPs با کمترین مقادیر p-value در نظر گرفته شده و ژن‌های قرار گرفته در این فواصل جدا شد. آنالیزهای غنی‌سازی مسیر^{۲۱} GO با استفاده از دستور enricher از پکیج clusterProfiler در R انجام شد (Wu et al. 2021). یک لیست از QTLs‌هایی که در نزدیک این SNP‌ها قرار گرفته بودند نیز از کتابخانه داده animalQTLdb استخراج گردید (http://www.animalgenome.org/QTLdb, Release (46, December 27, 2021)). این منبع اطلاعاتی شامل ۲۶۰۵ QTLs از ۱۰۴ مطالعه انجام شده در اسب برای ۶۴ صفت مختلف می‌باشد.

نتایج

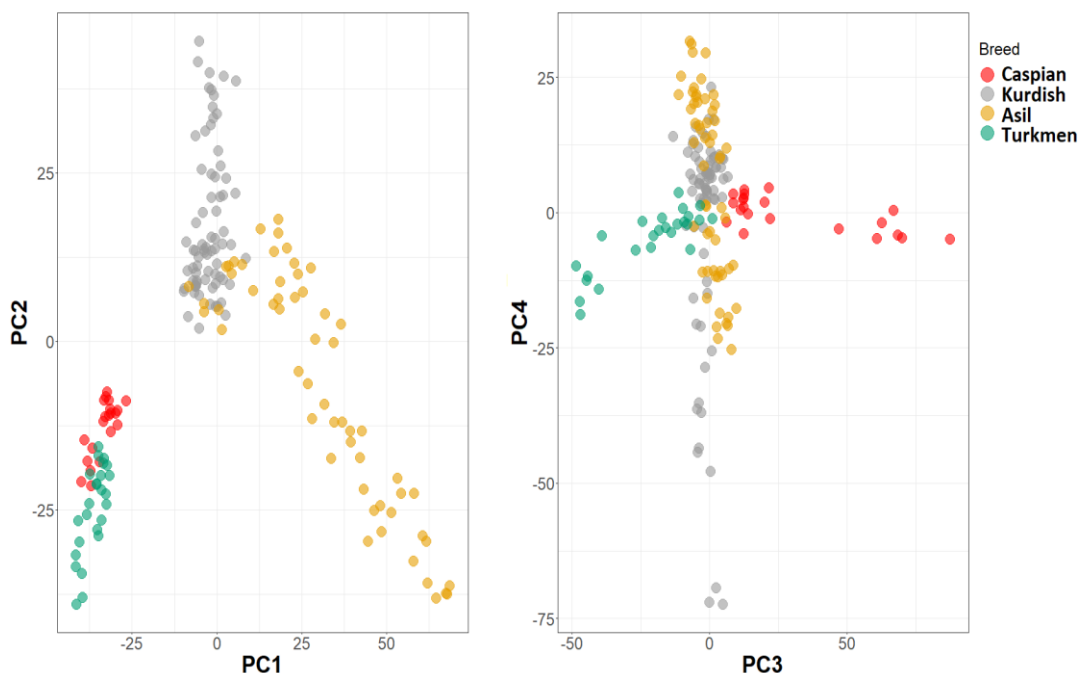
ساختار ژنتیکی جمعیت‌ها: برای بررسی ارتباط بین چهار جمعیت اسب ایرانی شامل ترکمن، کاسپین، کرد و اصیل از PCA استفاده گردید (شکل ۱). این شکل چهار PCs اول را نشان می‌دهد که سه PCs اول به خوبی این چهار نژاد را از هم تفکیک نموده است. در مجموع چهار PCs (6.43% PC1, PC2, PC3, PC4) به ترتیب 2.84، 1.43، ۱.۱۴ و 1.02 درصد را توجیح نمودند) از کل واریانس ژنتیکی را توجیح نمود. این پلات نشان داد که رابطه نزدیکی بین نژادهای ترکمن و کاسپین وجود داشته است و همچنین بین نژادهای کرد و اصیل ارتباط نزدیک‌تر مشاهده گردید. این ارتباط ژنتیکی با منشا ژنتیکی آنها هم خوانی نشان داد. PC اول (PC1) نژادها را به دو دسته با منشا مختلف جغرافیایی شامل شمال (کاسپین و ترکمن) و جنوب غرب یا غرب (کرد و اصیل) تقسیم نمود. همچنین PC1 تا حدود زیادی دو نژاد اصیل و کرد را از هم جدا نموده است. در ادامه PC2 باعث جدا شدن بیشتر دو نژاد کرد و اصیل شده و همچنین در این PC نژاد ترکمن و کاسپین تا حدودی از هم تفکیک شدند. در آخر PC3 به خوبی دو نژاد ترکمن و کاسپین را از هم جدا نمود. ارتباط ژنتیکی که در روش NJ مشاهده شد از نتایج PCA پشتیبانی می‌نماید (شکل ۲). شاخه‌های کاسپین و ترکمن نسبتاً به همدیگر نزدیک‌تر بوده اما این دو نژاد از دو نژاد دیگر (کرد و اصیل) به طور نسبی فاصله نشان دادند. همان طور که قبلاً اشاره شد این پدیده می‌تواند به دلیل منشاء جغرافیایی آنها باشد. در این آنالیز جمعیت کرد به دو شاخه مجزا تقسیم شده که جمعیت اصیل بین آنها قرار گرفت. وجود شاخه‌های ادغام شده در جمعیت‌های کرد و اصیل نشان از رابطه ژنتیکی نزدیک این دو نژاد و یا وجود اجداد مشترک ثبت نشده شجره‌ای در گذشته می‌باشد.

¹⁸ Quantitative trait locus

¹⁹ Functional analysis

²⁰ Gene ontology

²¹ Pathway enrichment



شکل ۱. نمودار آنالیز برای PCs ۱ و ۲ (سمت چپ) و ۳ و ۴ (سمت راست) برای چهار نژاد ترکمن (سبز)، کاسپین (قرمز)، کرد (خاکستری) و اصیل (زرد)

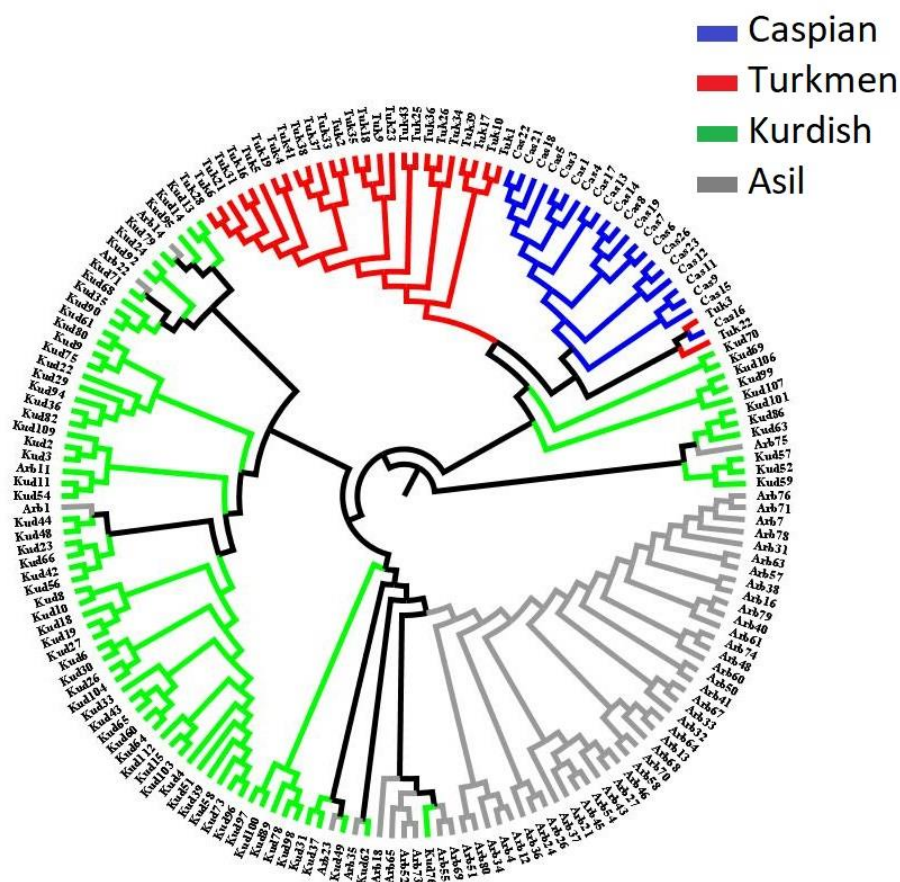
Figure 1. Plot for Principal component analysis based on 1 and 2 (left), and 3 and 4 (right) PCs for Turkmen (green), Caspian (red), Kurdish (gray) and Asil (yellow) breeds

آنالیز اختلاط جمعیتی با فرض وجود حداکثر ۷ جمعیت اجدادی ($K = 7$) برای یافتن محتمل‌ترین تعداد جمعیت اجدادی با استفاده از روش CV انجام شد (شکل ۳-A). با توجه به نتایج CV، کمترین خطای مدل زمانی حاصل گردید که مقادیر K برابر ۲ و ۳ بود (شکل ۳-B). در مقادیر K برابر 1، نمونه‌های مورد مطالعه به سه دسته شامل ۱) کاسپین و ترکمن، ۲) اصیل و ۳) کرد (حد واسط دو دسته قبلی)، قابل تقسیم هستند. در K برابر ۳، دسته بندی تقریباً مشابهی مشاهده گردید. در این آنالیز مشابه دو روش قبلی، دو نژاد کرد و اصیل و دو جمعیت کاسپین و ترکمن نیز غرابت نزدیک‌تری بهم داشتند، که با منشاء جغرافیایی آنها هم‌خوانی داشت. شاید بتوان گفت که یکی از دلایل غرابت بیشتر در بین نژادهای نزدیک از نظر جغرافیایی، وجود تلاقی‌های بین این نژادها در گذشته باشد که موجب ایجاد اشتراک‌های ژنتیکی بیشتر بین آنها شده است. این امر شاید در دو نژاد کرد و اصیل بیشتر نمود داشته باشد، به این دلیل که این دو نژاد در شرایط محیطی بسیار متفاوت از هم قرار گرفته‌اند.

نشانه‌های انتخاب و شناسایی GO و QTLsهای مرتبط با این نواحی انتخابی: بعد از محاسبه مقادیر DCMS

در هر یک از نژادها به صورت جدا گانه، مقادیر p-value این آماره برای ۴۰۱۲۰ جایگاه SNP در یک توزیع نرمال برآزش گردید. در هر نژاد ۱ درصد از SNPs با قوی‌ترین سیگنال (۴۰۰ SNPs) به عنوان نواحی کاندیدا در نظر گرفته شد (شکل ۴). انتخاب ۱

درصد SNPs با قوی‌ترین سیگنال در مطالعات شناسایی نشانه‌های انتخاب در حیوان اهلی بسیار معمول می‌باشد (Eydivandi et al. 2021a; Eydivandi et al. 2021b).



شکل ۲. نمودار درختی برای نتایج NJ برای چهار نژاد مورد مطالعه شامل کاسپین (آبی)، ترکمن (قرمز)، کرد (سبز) و اصیل (خاکستری)

Figure 2. Plot of a neighbor-joining phylogenetic tree for four studied breeds including, Caspian (blue), Turkmen (red), Kurdish (green) and Asil (gray)

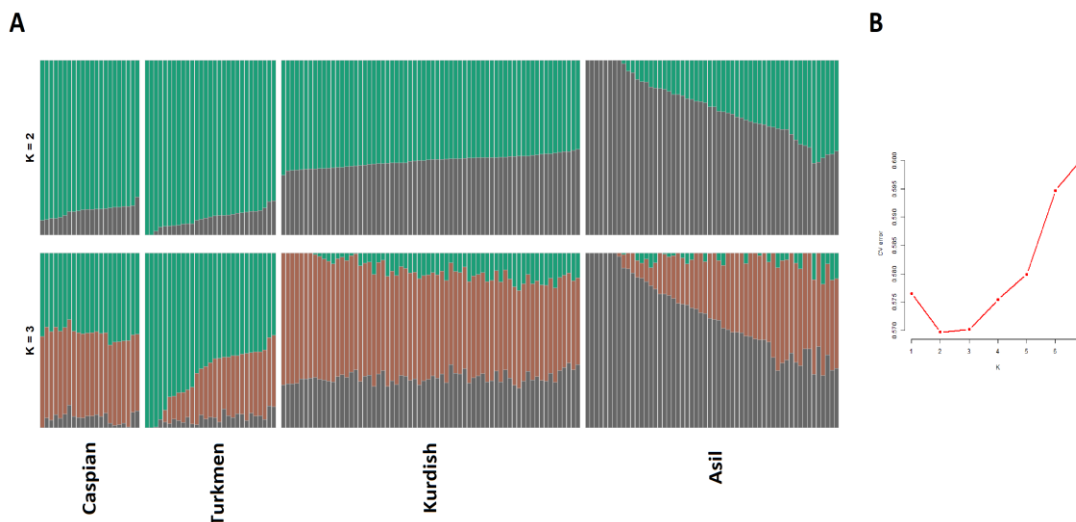
کاسپین: با در نظر گرفتن فاصله ۴۰ kbp به عنوان معیاری از قرارگیری در محدوده SNPهای کاندیدا، تعداد ۴۸۰ ژن در نزدیکی ۴۰۰ SNP برتر شناسایی گردید (فایل تکمیلی ۱). با دسته‌بندی این ژن‌ها براساس نوع، به ترتیب ۳۷۰، ۸۹، ۷، ۴، ۳ و ۱ ژن در دسته‌های ژن‌های کد کنند پروتئین^{۲۲}، RNAهای غیر کد کننده بلند^{۲۳} (lncRNA)، RNAهای هسته‌ای کوچک^{۲۴}

²² Protein coding

²³ Long non-coding RNAs

²⁴ Small nuclear RNAs

(snRNAs)، میکرو RNA^{۲۵} (miRNA)، RNA های هستکی کوچک^{۲۶} (snoRNAs)، ژن های کاذب^{۲۷} و RNA های مختص بدن کزال کوچک^{۲۸} (scaRNAs) قرار می گیرند. معنی دارترین SNP بر روی کروموزوم شماره ۱۱ در جایگاه ۵۸,۷ Mbp قرار داشته که نزدیک ژن آنتی ژن اسپرم با همولوژی کالپونین و دامنه سیم پیچی مانند^{۲۹} (SPECC1) می باشد.



شکل ۳. نمودار نتایج آنالیز اختلاط جمعیتی (Admixture) برای K برابر ۲ و ۳ و همچنین سهم هر جمعیت اجدادی در تشکیل جمعیت های مورد مطالعه (A). میزان خطای محاسبه شده با روش CV برای مقادیر K مساوی ۱ تا ۷ نشان داده شده است (B)

Figure 3. Plot of the results of admixture analysis for k = 2 and 3, and the contribution of each ancestral population in forming the studied populations. The error calculated by the CV method for K values equal to 1 to 7 is shown (B)

مطالعات نشان دادند که این ژن در بافت بیضه و بافت های سرطانی به شدت بیان شده و در فرآیندهای مختلف سلولی مانند بیان tRNA، میتوز، رشد سلولی و مرگ برنامه ریزی شده سلول^{۳۰} در سلول های سرطانی نقش دارد. همچنین بیان این ژن در مراحل بسیار حساس رشد جنینی یعنی تمام مراحل نمو پیش از لانه گزینی جنین^{۳۱} موش مشاهده گردید (Lee & Choi 2021). بعلاوه، دومین SNP معنی دار روی کروموزوم ۵ و فاصله 46.5 Mbp در کنار دو ژن مهم به نام فسفولیسیرید دهیدروژناز^{۳۲} (PHGDH) و ۳-هیدروکسی-۳-متیل گاوتاریل-CoA سنتتیز^{۳۳} (HMGCS2) قرار گرفته است. مطالعات نشان دادند که ژن

²⁵ MicroRNA

²⁶ Small nucleolar RNAs

²⁷ Pseudogenes

²⁸ Small Cajal body-specific RNAs

²⁹ Sperm antigen with calponin homology and coiled-coil domains 1

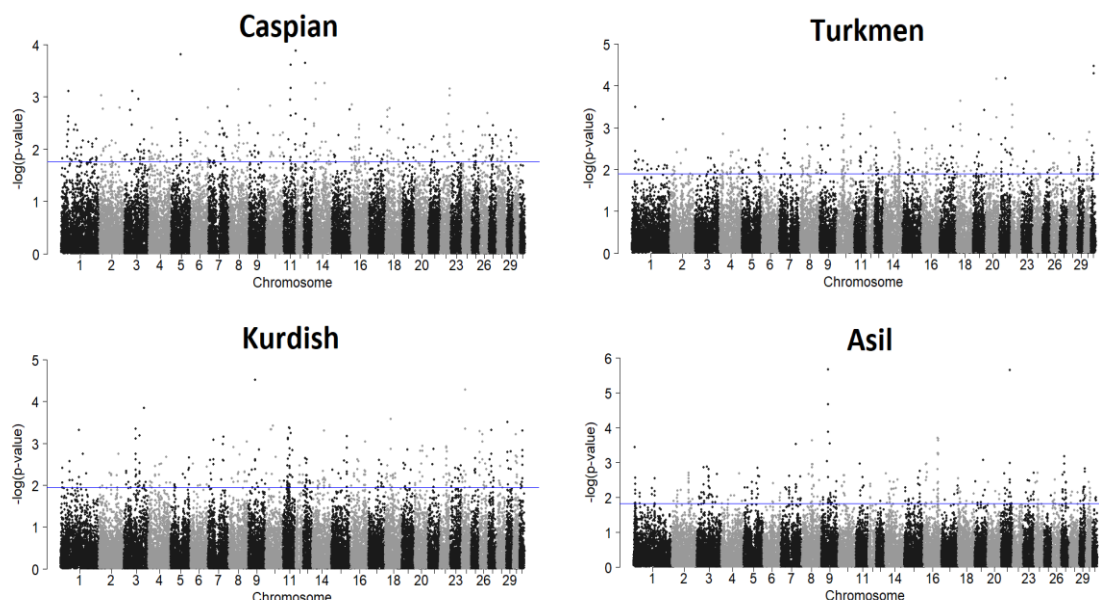
³⁰ Apoptosis

³¹ Preimplantation development

³² Phosphoglycerate Dehydrogenase

³³ 3-Hydroxy-3-Methylglutaryl-CoA Synthase 2

HMGCS2 به عنوان عامل مرکزی در بسیاری از شبکه‌ها و مسیرهای بیولوژیکی می‌باشد (Shukla et al. 2017). همچنین PHGDH به عنوان یک ژن مهم در ایمنی هومورال^{۳۴} معرفی می‌گردد (D'Avola et al. 2021).



شکل ۴. نمودار منتهن برای مقادیر $-\log_{10}$ p-values محاسبه شده در قالب DCMS برای جایگاه‌های SNP در چهار نژاد مختلف تحت بررسی

Figure 5. Manhattan plot of the $-\log_{10}$ (P-values) calculated by the DCMS framework at each SNP position in the four studied breeds

تجزیه و تحلیل غنی سازی برای این ۴۸۰ ژن برای شناسایی GOs معنی دار با در نظر گرفتن تصحیح برای خطای چندین آزمون غیر مستقل صورت گرفت. تعداد ۳ GOs شامل فعالیت اپوکسیژناز اسید آراشیدونیک^{۳۵} (GO:0008392 در دسته MF)، مسیر اپوکسیژناز P450^{۳۶} (GO:0019373 در دسته BP) و فرآیند متابولیک زئوبیوتیک^{۳۷} در این آنالیز با در نظر گرفتن تصحیح بونفرونی^{۳۸} (GO:0006805 در دسته BP) کمتر از 0.05 معنی دار گردید. بررسی QTL در نواحی پیشنهادی برای انتخاب در نژاد کاسپین موجب شناسایی ۷۸ QTL مختلف در ۱۸ دسته مختلف گردید (فایل تکمیلی ۱).

ترکمن: در مجموع تعداد ۳۸۲ ژن در نزدیکی SNPهای کاندیدا تحت انتخاب برای نژاد ترکمن قرار گرفته بودند (فایل تکمیلی ۲). مانند کاسپین، بیشترین نوع ژن مربوط به ژن های کد کننده پروتئین با ۲۷۹ ژن، در رتبه های بعدی lncRNAs

³⁴ Humoral immunity
³⁵ Arachidonic acid epoxygenase activity
³⁶ Epoxygenase P450 pathway
³⁷ Xenobiotic metabolic process
³⁸ Bonferroni correction

snRNAs، miRNAs و snoRNAs به ترتیب با تعداد ۹۳، ۵، ۳ و ۲ ژن بودند. اولین و دومین SNP با بیشترین معنی‌داری در کنار یکدیگر و روی کروموزوم ۳۱ در فاصله 14.2 Mbp قرار داشته که ژنی در نزدیکی این دو SNP قرار نداشت. سومین SNP در کنار ژن ENSECAG00000040940 قرار گرفته که تاکنون نقشی برای آن معرفی نشده است. در مقابل SNPs شماره چهار و پنج با بیشترین معنی‌داری (کروموزوم ۲۰؛ ۵۳،۳ Mbp) برای این نژاد در کنار ژن خانواده با تشابه توالی ۸۳، عضو B³⁹ (FAM83B) قرار گرفته بودند. مطالعات نشان از اهمیت این ژن در فرآیندهای سلولی مختلف مانند تبادل بین غشائی، نگهداری سیستم اسکلتی، ترمیم DNA و سیگنال Wnt می‌دهند (Cirello et al. 2022).

با بررسی بیشتر بر روی ژن‌های کاندید نزدیک به SNP‌های برتر، تعداد ۲ GOs معنی‌دار شناسایی گردید. اولین GO مربوط به فعالیت کانال کلسیمی (GO:0005262 در دسته MF) بوده که در سطح p تصحیح شده 0.05 معنی‌دار بود. دومین GO معنی‌دار، GO:0008392 بوده که در کاسپین نیز معنی‌دار گردید. در ادامه ۶۶ QTLs مختلف در نزدیک SNPs کاندید شناسایی گردید که در ۱۵ دسته مختلف قابل تقسیم‌بندی بودند (فایل تکمیلی ۲).

نژاد کرد: در این نژاد تعداد ۴۳۴ ژن مختلف در نزدیکی ۱ درصد SNPs‌های برتر شناسایی گردید (فایل تکمیلی ۳). از بین این ژن‌ها، تعداد ۳۳۳ مربوط به کدکننده‌های پروتئینی بوده و ۷۸، ۸، ۵، ۵ و ۵ عدد از آنها به ترتیب مربوط به lncRNAs، miRNAs، ژن‌های کاذب، snRNAs و snoRNAs بودند. اولین SNP (chr9: 32.4 Mbp) در کنار ژن پروکسیداسین مانند^{۴۰} (PXDNL) قرار گرفته که پارالوگ ژن پروکسیداسین (PXDN) می‌باشد. بیان ژن PXDN در بسیاری از سلول‌ها مانند سلوهای اندوتلیال، اپیتلیال و فیبروبلاست‌ها مشاهده شده، و اختلال در بیان این ژن باعث بروز مشکلات جدی مانند سرطان‌های مختلف از جمله سرطان در سیستم تخمدانی خواهد شد (Zheng & Liang 2018). دومین SNP (chr24: 18.4 Mbp) در فاصله کمتر از ۴۰ Kbp از ژن کدکننده خانواده پروتئینی D4 به نام D4 دارای روی و دو انگشت PHD از خانواده^{۴۱} (DPF3) قرار گرفته است. این ژن دارای یک بیان مختص بافتی در قلب و همچنین مراحل نمو گونه‌های مختلف می‌باشد. مطالعات نشان دادند که این ژن می‌تواند در نمو صحیح قلب و ماهیچه‌های بدنی نقش مهمی داشته باشد (Lange et al. 2008).

با مطالعه تجزیه و تحلیل غنی‌سازی برای ژن‌های کاندید، GO معنی‌داری برای این نژاد مشاهده نگردید. در مجموع تعداد ۷۰ جایگاه مختلف QTL با فاصله کمتر از ۴۰ Kbp از SNPs کاندید قرار گرفته بودند که در ۱۹ نوع QTL قابل دسته‌بندی بودند (فایل تکمیلی ۳).

اصیل ایرانی: تعداد ژن شناسایی شده در نزدیکی SNP‌های کاندید برای انتخاب ۳۸۴ عدد بودند (فایل تکمیلی ۴). همانند سایر نژادهای مورد مطالعه، بیشترین ژن‌های شناسایی شده مربوط به دسته ژن‌های کدکننده پروتئین با تعداد ۲۸۴ ژن بوده و پس

³⁹ Family with sequence similarity 83, member B

⁴⁰ Peroxidasin Like

⁴¹ D4, zinc and double PHD fingers, family 3

از آن به ترتیب مربوط به snRNAs، lncRNAs و ژن‌های کاذب با تعداد ۸۵، ۱۰ و ۳ ژن بودند. برای دسته‌های snoRNAs و miRNAs فقط یک ژن شناسایی گردید. اولین SNP معنی‌دار در نزدیکی ژنی نبوده ولی دومین SNP (chr9: 33.7 Mbp) در نزدیکی ژن مهم تنظیم‌کننده رتیکولوفاژی^{۴۲} (RETREG1) قرار گرفته است. مطالعات انجام شده بر روی این ژن نشان از اهمیت این ژن در بسیاری از بیماری‌ها و سرطان‌ها می‌دهد (Islam et al. 2018)، که دال بر اهمیت این ژن در بسیاری از فعالیت‌های حیاتی بدن است. سومین و چهارمین SNP معنی‌دار (chr9: 33.4 Mbp) در نزدیکی ژن سینتروفین گامای^{۴۳} (SNTG1) قرار گرفته بودند. مطالعات انجام شده نشان دادند که این ژن می‌تواند به عنوان فاکتور مهمی برای عمر و اندازه بدن در گاو باشد (Mészáros et al. 2014; An et al. 2019).

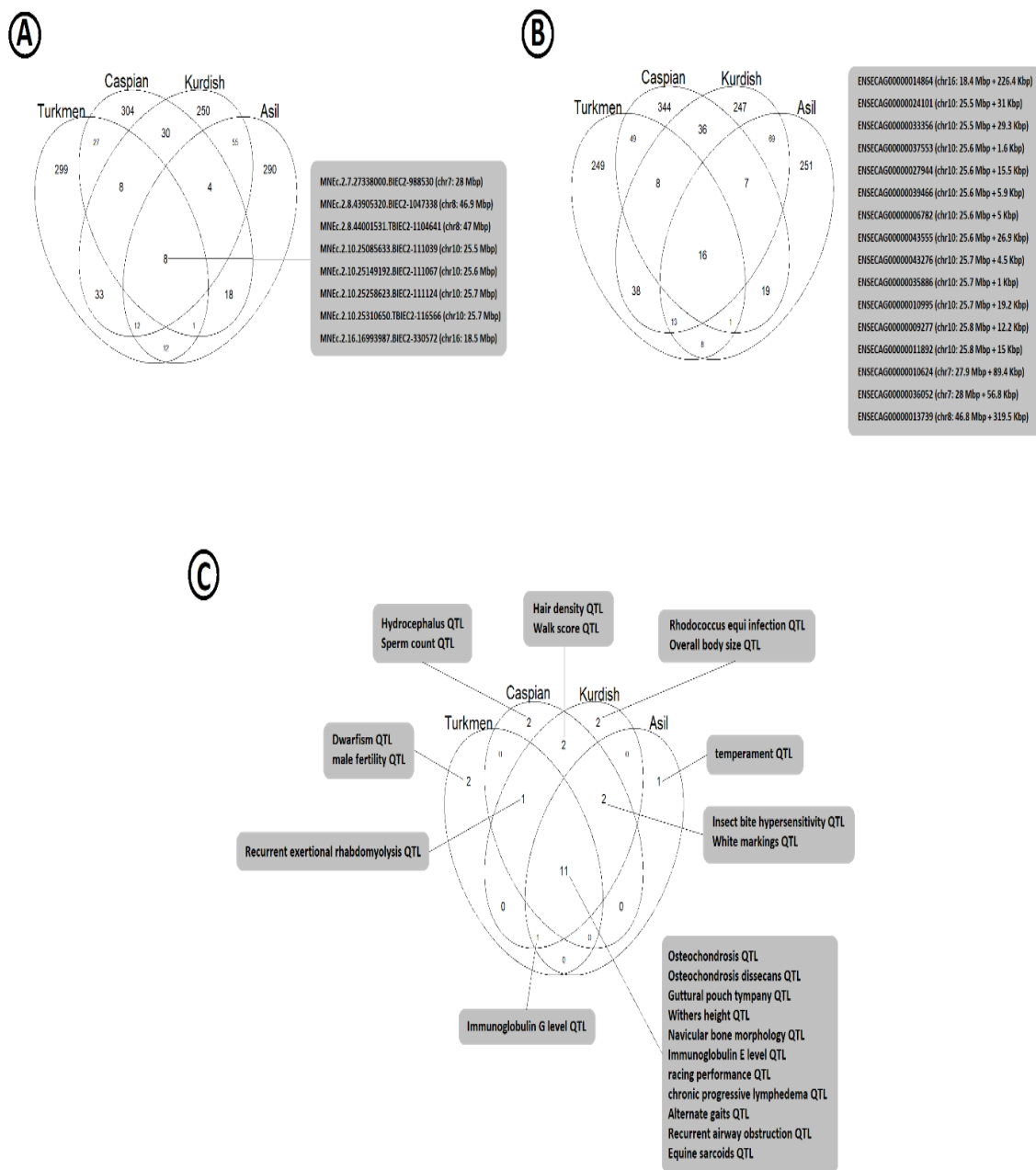
با مطالعه تجزیه و تحلیل غنی‌سازی برای ژن‌های کاندید، تعداد ۳ GOs معنی‌دار برای این نژاد شناسایی گردید. این ۳ GOs شامل بخش اساسی سلول (GO:0045178 در دسته CC)، چسبندگی سلول‌های هموفیل از طریق مولکول‌های چسبنده غشای پلازما (GO:0007156 در دسته BP) و چسبندگی سلولی (GO:0007155 در دسته BP) بودند. با بررسی QTLs نزدیک به نواحی کاندید، تعداد ۷۷ QTLs مختلف در نزدیک آن‌ها شناسایی گردید که در ۱۵ دسته مختلف قابل تقسیم‌بندی بودند (فایل تکمیلی ۴).

شباهت‌های الگو نشانه‌های انتخاب در چهار نژاد تحت مطالعه: شباهت در نتایج بررسی نشانه‌های انتخاب برای

چهار نژاد مورد مطالعه شامل نژادهای کاسپین، ترکمن، کرد و اصیل در ۳ سطح SNP، ژن و QTL در شکل ۵ نشان داده شده است. برای سطح GO تنها بین جمعیت‌های کاسپین و ترکمن یک GO مشابه (GO:0008392) مشاهده گردید. در سطح SNP، تعداد ۸ جایگاه SNP در بین تمام نژادهای مشترک بود (شکل ۵-A). چهار SNP مشترک مشاهده شده مربوط به کروموزوم ۱۰ در فاصله 25.5 تا 25.7 Mbp قرار گرفته بودند. همچنین ۲ SNP بر روی کروموزوم ۸ و ۱ SNP بر روی کروموزوم‌های ۷ و ۱۶ مشاهده گردید. در سطح ژن، تعداد ۱۲ ژن مشابه بر روی کروموزوم ۱۰ مشاهده گردید (شکل ۵-B). همچنین ژن‌های مشابه در کروموزوم‌های ۷ (n = 2)، ۸ (n = 1) و ۱۶ (n = 1) مشاهده گردید. این ژن‌های مشترک که در هر چهار نژاد به عنوان ژن‌های کاندید انتخاب معرفی گردیدند، تحت آنالیز هستی‌شناسی قرار گرفته تا درک بهتری از این شباهت‌ها بدست آید. نتایج غنی‌سازی برای این ۱۶ ژن وجود تعداد ۱۱ GOs را در سطح معنی‌داری بر اساس p بونفرونی 0.05 را نشان داد (جدول ۱). در نهایت تشابهات موجود بین جمعیت‌های مختلف تحت مطالعه در سطح QTLs مورد بررسی قرار گرفته شد (شکل ۵-C). در این سطح تعداد ۱۱ QTLs مشترک شناسایی گردید. همچنین برای ترکمن (n = 2)، کاسپین (n = 2)، کرد (n = 2) و اصیل (n = 1) نیز QTLs اختصاصی هر نژاد مشاهده گردید.

⁴² Reticulophagy Regulator 1

⁴³ Syntrophin Gamma 1



شکل ۵. نمودار ون برای تشابهات بدست آمده در مطالعه شناسایی نشانه‌های انتخاب در چهار نژاد اسب بومی ایران شامل: ترکمن، کاسپین، کرد و اسیل. تعداد تشابهات بدست آمده برای SNPs (A)، ژن (B) و QTLs (C) نشان داده شده است

Figure 6. Vann diagrams for number of overlapped candidate, SNPs (A), genes (B) and QTL types (C) under selection among the four studied Iranian horse breeds including: Turkmen, Caspian, Kurdish and Asil

جدول ۱. نتایج GOs معنی‌دار در سطح p بونفرونی ۰,۰۵ برای ۱۶ ژن مشترک کاندید تحت انتخاب در چهار نژاد کاسپین، ترکمن، کرد و اصیل

Table 1. Results of the significant GOs based on Bonferroni adjustment at 0.05 for 16 common genes that were candidate for signature of selection in four Iranian horse breeds including Caspian, Turkmen, Kurdish and Asil

شناسه GO (GO term)	توضیح شناسه GO (GO term definition)	P تصحیح شده (Adjusted P)	نوع GO ¹ (GO type)
GO:0023052	سیگنالینگ (signaling)	0.029	BP
GO:0007528	نمو اتصال عصبی عضلانی (neuromuscular junction) (development)	0.029	BP
GO:0007157	چسبندگی سلول-سلول هتروفیل از طریق مولکول‌های چسبنده سلول غشای پلازما (heterophilic cell-cell adhesion via plasma) (membrane cell adhesion molecules)	0.029	BP
GO:0061844	پاسخ ایمنی هومورال ضد میکروبی با واسطه پپتید ضد میکروبی (antimicrobial humoral immune response mediated by) (antimicrobial peptide)	0.029	BP
GO:0050829	پاسخ دفاعی به باکتری گرم منفی (Gram-) (negative bacterium)	0.035	BP
GO:0007218	مسیر سیگنالینگ نوروپپتیدی (neuropeptide signaling pathway)	0.047	BP
GO:0045202	سیناپس (synapse)	0.025	CC
GO:0031594	اتصال عصبی عضلانی (neuromuscular junction)	0.041	CC
GO:0005912	اتصال چسبیده (adherens junction)	0.048	CC
GO:0014069	چگالی پس سیناپسی (postsynaptic density)	0.049	CC
GO:0005911	اتصال سلول به سلول (cell-cell junction)	0.049	CC

¹ BP: فرآیند بیولوژیکی؛ CC: ساختار سلولی

بحث

دانستن مبنای ژنتیکی تنوع فنوتیپی درون و بین گونه‌ها یکی از موضوعات چالش برانگیز در زیست‌شناسی مدرن می‌باشد. انتخاب مصنوعی در کنار انتخاب طبیعی یکی از مهم‌ترین منابع تنوع زیستی برای فرآیند اهلی‌سازی بوده که به دلیل هزاران سال تلاش برای اصلاح انتخابی حیوانات اهلی ایجاد شده‌اند (Andersson & Georges 2004). جمعیت اسب‌های ایرانی تحت فرآیند انتخاب برای اهداف اصلاحی مختلف با منشاء جغرافیایی متفاوت و تنوع اجزاء محیطی مانند دما، رطوبت و ارتفاع بوده است.

در این مطالعه با استفاده از تکنولوژی ژنوتیپ کردن SNP، یک مجموعه نسبتاً بزرگی از اسب‌های ایرانی متعلق به چهار نژاد اصلی (کاسپین، ترکمن، کرد و اصیل) برای بررسی تنوع ژنتیکی، ساختار جمعیتی و نواحی تحت انتخاب با استفاده از ترکیب اطلاعات دو نوع آرایه متفاوت (Equine SNP670k و GGP Equine 70K) مطالعه گردید. به منظور بررسی دقیق افتراق ژنتیکی، یک مطالعه در سطح بین‌المللی در ۳۶ نژاد اسب برای بررسی ارتباط ژنتیکی اسب‌های سرتاسر دنیا انجام شد (Petersen et al. 2013). در این مطالعه یک افتراق ژنتیکی واضح با منشاء جغرافیایی بین نژادهای اسب مشاهده گردید. نتایج مشابه برای دیگر نژادها مانند گوسفند (Eydivandi et al. 2021a; Eydivandi et al. 2021b)، بز (Colli et al. 2018) و گاو (Karimi et al. 2016) مشاهده شد. مشاهده تنوع ژنتیکی با منشاء جغرافیایی در مقیاس جهانی می‌تواند موجب این پرسش شود که آیا این عامل می‌تواند در مقیاس‌های کوچک‌تر (مانند تنوع مشاهده شده در درون یک کشور) باعث ایجاد تنوع ژنتیکی شود؟ در پاسخ به این سوال، نتایج این مطالعه نشان داد که فرآیند انتخاب در اسب‌های ایران با منشاء جغرافیایی متفاوت موجب ایجاد الگوی تفرق نژادی مرتبط با فاصله جغرافیایی شده است. نژادهای ترکمن و کاسپین از نظر ژنتیکی به هم نزدیک‌تر بوده، اما با دو نژاد کرد و اصیل تفاوت قابل مشاهده‌تری داشتند. در تایید این مطلب، نتایج سه آنالیز PCA، NJ و اختلاط جمعیتی، دو نژاد شمالی (کاسپین و ترکمن) قرابت بیشتری نسبت به دو نژاد جنوب غرب و غرب (اصیل و کرد) نشان دادند. این نتایج با نتایج بدست آمده از مطالعات قبلی که بر روی اسب‌های ایران صورت گرفته هم‌خوانی دارد (Sadeghi et al. 2019; Salek Ardestani et al. 2022). در مقابل، مطالعه‌ی روش‌های مبتنی بر صفات ظاهری و جداسازی اسب‌ها با این معیار، نتایج قابل قبولی را نشان ندادند. برای مثال در مطالعه انجام شده در ۳ نژاد کرد، کاسپین و ترکمن بر اساس ۲۳ صفت ظاهری، نژاد کرد و ترکمن قرابت بیشتری نشان دادند که با منشاء جغرافیایی آن‌ها هم‌خوانی نشان نمی‌دهد (Hosseini et al. 2016). مطابقت نداشتن نتایج بدست آمده بر پایه صفات ظاهری می‌تواند به دلیل ارتباط تعداد جایگاه‌های کم با صفات مورد مطالعه بوده باشد، همچنین وجود عدم ارتباط شکل ظاهری با منشا جغرافیایی می‌تواند دلیلی دیگر برای گرفتن این نتایج باشد. این مطالعه و سایر مطالعات مشابه نشان می‌دهند که استفاده از اطلاعات ژنومی منبع مناسب‌تری برای درک بهتر تنوع و تاثیر عوامل محیطی بر ساختار ژنتیکی جمعیت‌های یک گونه خواهد بود. مطالعات قبلی نشان دادند که دانستن نواحی تحت انتخاب در نژادهای مختلف یک گونه دامی می‌تواند برای شناسایی نواحی ژنوم که با صفات مهم اقتصادی و همچنین سازگاری با شرایط محیطی همبستگی دارند، مفید باشد (Gurgul et al. 2019; Ghoreishifar et al. 2020). هر دو دسته صفات مرتبط با سازگاری و اقتصاد (عملکرد) که در بسیاری از مطالعات مورد بررسی قرار می‌گیرند، صفات کمی هستند که تحت تاثیر تعداد زیادی ژن می‌باشند. انتخاب برای چنین صفات پیچیده‌ای ممکن است به صورت همزمان در چندین جایگاه صورت گرفته باشد که موجب ایجاد سیگنال‌های ضعیفی در سطح ژنوم گردد (Kemper et al. 2014). با این وجود، انتخاب طبیعی و یا مصنوعی برای برخی آلل‌ها که دارای اثرات نسبتاً بزرگی هستند ممکن است الگوهای قابل مشاهده‌ای را از خود بجای بگذارند. برای یافتن این الگوهای قابل مشاهده در جمعیت اسب‌های ایران، از ترکیب مقادیر p-values سه روش مهم تک جمعیتی استفاده گردید. استفاده از روش DCMS برای ترکیب سیگنال‌ها می‌تواند باعث شناسایی نواحی و ژن‌های تحت انتخاب با

دقت و قدرت آماری بالاتری شود. با بکارگیری روش DCMS در چهار جمعیت اسب ایران، تعداد زیادی ژن در هر جمعیت به عنوان نواحی تحت انتخاب طبیعی و یا مصنوعی پیشنهاد گردید. جایگاه‌های پیشنهادی در هر نژاد متفاوت بوده که این تفاوت می‌تواند به دلیل تفاوت در گذشته تکاملی آنها به دلیل شرایط محیطی و یا انتخاب‌های مصنوعی به دست بشر بوده باشد. تفکیک این دو عامل انتخابی در نتایج کار دشوار بوده اما با بررسی نقش ژن‌ها می‌توان تاحدودی آن‌ها را تفکیک نمود. برای مثال در نژاد کرد، ژن DPF3 به عنوان ژن تحت تاثیر انتخاب شناسایی گردید. این ژن نقشی بسیار مهمی در نمو صحیح قلب و ماهیچه‌های بدنی داشته (Lange et al. 2008) که با در نظر گرفتن شرایط محیطی منشاء اسب کرد (استان کردستان با ارتفاعات بلند و صعب‌العبور) می‌توان به وجود انتخاب طبیعی در این ژن طی فرآیند تکاملی پی برد. به نظر می‌رسد از تنوع موجود در این ژن می‌توان در اصلاح نژاد استفاده نمود و از این طریق به اسب‌هایی که توانایی بالایی در تحمل شرایط سخت محیطی دارند دست یافت. همچنین وجود فشار انتخاب بر روی ژن SNTG1 در نژاد اصیل و نقش آن در اندازه بدن (An et al. 2019) می‌تواند مثالی از تاثیرات انسان در تغییر ساختار ژنتیکی باشد. بنابراین این ژن می‌تواند به عنوان کاندیدا برای انتخاب اسب‌هایی با اندازه‌های بدن بزرگ تر مورد استفاده قرار گیرد. نکته قابل ذکر دیگر تعداد SNPها و ژن‌های مشترک بین جمعیت‌های اسب مورد مطالعه با توجه به منشا جغرافیایی آن‌ها می‌باشد. برای مثال بیشترین تعداد SNPهای مشترک مشاهده شده بین دو نژاد کرد و اصیل بوده ($n = 79$) و این تعداد برای کاسپین و ترکمن کمتر بوده است ($n = 44$). بررسی تعداد اشتراکات در سطح ژن‌ها، وجود حداکثر اشتراکات در نژادهای نزدیک از نظر جغرافیایی (کاسپین/ترکمن = ۷۴ ژن؛ کرد و اصیل = ۱۰۵ ژن) را نشان داد. با تمام این تفاسیر، تایید این نتایج نیازمند بررسی‌های بیشتر درون نژادی و وجود ارتباط این تغییرات ژنتیکی بر صفات مرتبط با سازگاری و یا عملکردی می‌باشد.

استفاده از نتایج مطالعات گذشته در خصوص QTLs مرتبط با صفات مهم و یا اطلاعات عملکردی ژن‌ها با استفاده از آنالیز غنی سازی GO می‌تواند در درک بهتر نتایج مطالعات نشانه‌های انتخاب و یافتن نیروهای انتخاب که باعث شکل‌دهی به ساختار ژنوم اسب‌های ایران شده‌اند کمک نماید. برای این منظور، ارتباط بین جایگاه‌های پیشنهادی تحت انتخاب و ارتباط آن‌ها با QTLs شناسایی شده و یا GOs مورد بررسی قرار گرفته شد. در این مطالعه تعداد محدودی GO معنی‌دار در هر نژاد یافت گردید. با این وجود یک GO مشترک (GO:0008392) بین کاسپین و ترکمن مشاهده گردید. در مطالعه‌ی جدید، از این GO به عنوان یکی از مسیرهای مهم مرتبط با پاسخ ایمنی یاد شده است (Qi et al. 2019). مطالعه غنی سازی ژن‌ها برای تعداد ۱۶ ژن مشترک کاندیدا انتخاب در هر چهار نژاد تحت مطالعه، ۱۱ GO معنی‌دار را برای آن‌ها مشخص نمود. از بین این GOهای معنی‌دار، دو مورد از آن‌ها مربوط به مسیرهای بیولوژیکی مقاومت به باکتری بودند (GO:0061844 و GO:0050829). مطالعه در سطح QTLs نتایج قابل تاملی را ارائه نمود. برای مثال تعداد ۱۱ نوع QTL مشترک در تمام جمعیت‌های مورد مطالعه مشاهده گردید. این QTLs در دسته‌های صفات مرتبط با سازگاری (مانند سطح ایمینوگلوبولین E)، شایستگی از طریق نقص‌های ژنتیکی (مانند

انسداد مکرر راه هوایی^{۴۴}، سارکوئید اسب^{۴۵}، تمپان کیسه‌ای روده‌ای^{۴۶} و یا نقص استخوانی استئوکندروز^{۴۷}، و یا رفتاری و ظاهری (مانند ارتفاع جدوگاه و یا نحوه راه رفتن)، قرار گرفتند. این نتایج نشان داد که ژن‌های کاندیدا احتمالاً به عنوان اهداف انتخابی در جمعیت اسب‌های ایرانی طی سازگاری با شرایط محیطی و یا انتخاب با واسطه انسان قرار داشته‌اند.

نتیجه‌گیری: در این مطالعه از چندین روش برای بررسی تنوع ژنتیکی چهار نژاد اصلی اسب ایران شامل کاسپین، ترکمن، کرد و اسیل استفاده گردید. در مجموع، یک طبقه بندی مجزای ژنتیکی بین نژادهای تحت مطالعه شناسایی گردید که به صورت مشخصی بیانگر منشاء جغرافیایی آن‌ها بود. این اطلاعات می‌تواند برای یافتن راه حل ژنتیکی بهبود روش‌های اصلاحی برای این نژادها مفید باشد. ژن‌های شناسایی شده که تحت تاثیر انتخاب قرار گرفته بودند می‌توانند برای مواردی مانند اصلاح اسب‌های با مقاومت بیشتر به شرایط سخت محیطی و یا تولید اسب‌هایی با قدهایی بلندتر مورد استفاده قرار گیرند. با ترکیب سه روش مختلف شناسایی نشانه‌های انتخاب در قالب مقادیر DCMS، تعداد زیادی جایگاه‌های کاندیدا که تحت تاثیر انتخاب طبیعی یا مصنوعی بودند معرفی گردید که با GOS و QTLs مختلف در ارتباط بودند. یافتن ژن‌های کاندیدا که به احتمال زیاد تحت فشار انتخابی هستند، احتمالاً می‌تواند در درک بهتر فرآیندهای تکامل ژنتیکی نژادهای اسب که باعث تغییرات در صفات کمی و فراوانی آلی شده، مفید باشد. نتایج این تحقیق می‌تواند در درک بهتر فرآیند انتخاب طبیعی و مصنوعی در اسب‌های ایران کمک نماید. همچنین با شناسایی ژن‌های کاندیدا تحت انتخاب بین نژادهای مختلف ایران که هر کدام دارای خصوصیات ظاهری و عملکردی منحصر به فردی هستند، می‌توان از این ژن‌های کاندیدا برای اهداف اصلاحی مختلف مانند افزایش اندازه بدن و یا مقاومت به شرایط محیطی سخت استفاده نمود.

سپاسگزاری: از معاونت محترم پژوهشی دانشگاه کردستان به خاطر حمایت مالی در اجرای پژوهش حاضر سپاسگزاری می‌شود. از داوران محترم به خاطر ارائه نظرهای ساختاری و علمی سپاسگزاری می‌شود. نگارندگان بر خود لازم می‌دانند از آقای دکتر محمود امیری رودبار به خاطر مطالعه متن مقاله حاضر و ارائه نظرهای ارزشمند سپاسگزاری نمایند.

منابع

عسکری ناهید، باقی زاده امین، محمدآبادی محمدرضا (۱۳۸۹). مطالعه تنوع ژنتیکی در چهار جمعیت بز کرکی راینی با استفاده از نشانگرهای ISSR. مجله ژنتیک نوین ۵، ۵۶-۴۹.

محمدی فر آمنه، فقیه ایمانی سید علی، محمدآبادی محمد رضا، سفلی محمد (۱۳۹۲) تأثیر ژن $TGF\beta 3$ بر ارزش های فنوتیپی و ارثی صفات وزن بدن در مرغ بومی استان فارس. مجله بیو تکنولوژی کشاورزی ۵(۴)، ۱۳۶-۱۲۵.

References

- 44 Recurrent airway obstruction
45 Equine sarcoids
46 Guttural pouch tympany
47 Osteochondrosis

- Ala-Amjadi M, Yeganeh H, Sadeghi M (2017) Study of Genetic variation in Iranian Kurdish horse using microsatellite marker. *Iran J Anim Sci* 48, 335-42.
- Alexander DH, Lange K (2011) Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinf* 12, 246.
- Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19, 1655-64.
- An B, Xia J, Chang, T, et al. (2019) Genome-wide association study reveals candidate genes associated with body measurement traits in Chinese Wagyu beef cattle. *Anim Genet* 50, 386-90.
- Andersson L, Georges M (2004) Domestic-animal genomics: deciphering the genetics of complex traits. *Nat Rev Genet* 5, 202-12.
- Anthony DW, Brown DR (2011) The Secondary Products Revolution, Horse-Riding, and Mounted Warfare. *J World Prehis* 24, 131.
- Askari N, A Baghizadeh, MR Mohammadabadi (2010) Study of genetic diversity in four populations of Raeini cashmere goat using ISSR markers. *Modern Genet J* 5 (2), 49-56 (In Persian).
- Askari N, Baghizadeh A, Mohammadabadi MR (2008) Analysis Of The Genetic Structure Of Iranian Indigenous Raeni Cashmere Goat Populations Using Microsatellite Markers. *Biotechnology* 2 (3), 1-4.
- Breunig MM, Kriegel H-P, Ng RT, Sander J (2000) LOF: identifying density-based local outliers. In: *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pp. 93-104.
- Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Nat Am J Hum Genet* 81, 1084-97.
- Ceballos FC, Joshi PK, Clark DW, et al. (2018) Runs of homozygosity: windows into population history and trait architecture. *Nat Rev Genet* 19, 220-34.
- Cirello V, Grassi ES, Pogliaghi G, et al. (2022) FAM83B is involved in thyroid cancer cell differentiation and migration. *Sci rep* 12, 8608.
- Colli L, Milanesi M, Talenti A, et al. (2018) Genome-wide SNP profiling of worldwide goat populations reveals strong partitioning of diversity and highlights post-domestication migration routes. *Genet Sel Evol* 50, 58.
- Cook DE, Andersen EC (2017) VCF-kit: assorted utilities for the variant call format. *Bioinformatics* 33, 1581-2.

- Cosgrove EJ, Sadeghi R, Schlamp F, et al. (2020) Genome Diversity and the Origin of the Arabian Horse. *Sci rep* 10, 9702.
- D'Avola A, Legrave N, Tajan M, et al. (2021) PHGDH is required for germinal center formation and is a therapeutic target in MYC- driven lymphoma. *The Journal of clinical investigation* 132 (9).
- Danecek P, Auton A, Abecasis G, et al. (2011) The variant call format and VCFtools. *Bioinf* 27, 2156-8.
- Eydivandi S, Roudbar MA, Ardestani SS, Momen M, Sahana G. (2021a) A selection signatures study among Middle Eastern and European sheep breeds. *J Anim Breed Genet* 138 (5), 574-588.
- Eydivandi S, Roudbar MA, Karimi MO, Sahana G (2021b) Genomic scans for selective sweeps through haplotype homozygosity and allelic fixation in 14 indigenous sheep breeds from Middle East and South Asia. *Sci Rep* 11, 2834.
- Firouz L (1998) The original ancestors of the Turkoman, Caspian horses. In: First international conference on Turkoman horses. Ashgabat, Turkmenistan.
- Forbis J (1976) The classic Arabian horse. Liveright Publishing Corporation, New York.
- Fotovati A (2000) Persian horse breeds from ancient time to present and their roles in development of world horse breeds. *Asian-Australasian J Anim Scis* 13, 401-.
- Gautier M, Vitalis R (2012) rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics* 28, 1176-7.
- Gharahveysi S, Irani M (2011) Inbreeding study on the Iranian Arab horse population. *World J Zoology* 6, 1-6.
- Ghasemi M, Baghizadeh A, Abadi MRM (2010) Determination of genetic polymorphism in Kerman Holstein and Jersey cattle population using ISSR markers. *Aust J Basic Appl Sci* 4 (12), 5758-5760.
- Ghoreishifar SM, Eriksson S, Johansson AM, et al. (2020) Signatures of selection reveal candidate genes involved in economic traits and cold acclimation in five Swedish cattle breeds. *Genet Sel Evol* 52, 52.
- Gurgul A, Jasielczuk I, Semik-Gurgul E, Pawlina-Tyszko K (2019) A genome-wide scan for diversifying selection signatures in selected horse breeds. *PLoS One* 14, e0210751.
- Hosseini M, Shahrabak HM, Zandi MB, Fallahi M (2016) A Morphometric Survey Among Three Iranian Horse Breeds with Multivariate Analysis. *Media Peternakan* 39, 155-60.
- Islam F, Gopalan V, Lam AK-y (2018) RETREG1 (FAM134B): A new player in human diseases: 15 years after the discovery in cancer. *Journal of Cellular Physiology* 233, 4479-89.

- Jiskrová I, Vrtková I, Prausová M (2016) Genetic diversity of populations of Akhal-Teke horses from the Czech Republic, Russia, Estonia and Switzerland. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* 64, 461-6.
- Kalbfleisch TS, Rice ES, DePriest MS, et al. (2018) Improved reference genome for the domestic horse increases assembly contiguity and composition. *Communications Biology* 1, 197.
- Karimi K, Strucken EM, Moghaddar N, et al. (2016) Local and global patterns of admixture and population structure in Iranian native cattle. *BMC Genet* 17, 108.
- Kasprzyk A (2011) BioMart: driving a paradigm change in biological data management. *Database* 2011.
- Kemper KE, Saxton SJ, Bolormaa S, Hayes BJ, Goddard ME. (2014) Selection for complex traits leaves little or no classic signatures of selection. *BMC Genomics* 15, 1-14.
- Lange M, Kaynak B, Forster UB, et al. (2008) Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev* 22, 2370-84.
- Lee S, Choi I (2021) Expression patterns and biological function of *Specc1* during mouse preimplantation development. *Gene Expression Patterns* 41, 119196.
- Lotterhos KE, Card DC, Schaal SM, et al. (2017) Composite measures of selection can improve the signal-to-noise ratio in genome scans. *Methods Ecol Evol* 8, 717-27.
- Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol Ecol* 24, 1031-46.
- Ma Y, Ding X, Qanbari S, et al. (2015) Properties of different selection signature statistics and a new strategy for combining them. *Heredity* 115, 426-36.
- Malomane DK, Reimer C, Weigend S, et al. (2018) Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genomics* 19, 22.
- Mészáros G, Eaglen S, Waldmann P (2014) A genome wide association study for longevity in cattle. *Open Journal of Genetics* 2014.
- Mohammadabadi M, Bordbar F, Jensen J, et al. (2021) Key genes regulating skeletal muscle development and growth in farm animals. *Animals* 11 (3), e835.
- Mohammadabadi MR (2017) Inter-Simple Sequence Repeat loci Associations with Predicted Breeding Values of Body Weight in Kermani Sheep. *Genet 3rd Millennium* 14 (4), 4383-4390.
- Mohammadifar A, Mohammadabadi MR (2011) Application of Microsatellite Markers for a Study of Kermani Sheep Genome. *Iran J Anim Sci* 42 (4), 337-344.

- Mohammadifar A, Faghieh Imani SA, Mohammadabadi MR, Soflaei M (2014) The effect of TGFb3 gene on phenotypic and breeding values of body weight traits in Fars native fowls. *Agric Biotechnol J* 5 (4), 125-136.
- Mohammadifar A, Mohammadabadi M (2018) Melanocortin-3 receptor (MC3R) gene association with growth and egg production traits in fars indigenous chicken. *Malays Appl Biol* 47 (3), 85-90.
- Mohammadabadi MR, Esfandyarpoor E, Mousapour A (2017) Using Inter Simple Sequence Repeat Multi-Loci Markers for Studying Genetic Diversity in Kermani Sheep. *J Res Develop* 5 (2), e154.
- Moridi M, Masoudi A, Vaez Torshizi R, Hill E (2013) Mitochondrial DNA D-loop sequence variation in maternal lineages of Iranian native horses. *Anim Genet* 44, 209-13.
- Nazari F, Seyedabadi H-R, Noshary A, Emamjomeh-Kashan N, Banabazi M-H. (2022) A Genome-Wide Scan for Signatures of Selection in Kurdish Horse Breed. *J Equine Vet Sci* 113, 103916.
- Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *PNAS* 76, 5269-73.
- Petersen JL, Mickelson JR, Cothran EG, et al. (2013) Genetic Diversity in the Modern Horse Illustrated from Genome-Wide SNP Data. *PLoS ONE* 8, e54997.
- Purcell S, Neale B, Todd-Brown K, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *The Nat Am J Hum Genet* 81, 559-75.
- Qi M, Tan B, Wang J, et al. (2019) Small intestinal transcriptome analysis revealed changes of genes involved in nutrition metabolism and immune responses in growth retardation piglets1. *J Anim Sci* 97, 3795-808.
- Rahimi-Mianji G, Nejati-Javaremi A, Farhadi A (2015) Genetic diversity, parentage verification, and genetic bottlenecks evaluation in Iranian Turkmen horse1. *Russ J Genet* 51, 916-24.
- Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW (2014) Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genet* 15, 1-19.
- Sadeghi R, Moradi-Shahrbabak M, Miraei Ashtiani SR, et al. (2019) Genetic Diversity of Persian Arabian Horses and Their Relationship to Other Native Iranian Horse Breeds. *J Hered* 110, 173-82.
- Salek Ardestani S, Zandi MB, Vahedi SM, Janssens S (2022) Population structure and genomic footprints of selection in five major Iranian horse breeds. *Anim Genet*, DOI: 10.1111/age.13243.

- Saravanan KA, Panigrahi M, Kumar H, et al. (2020) Selection signatures in livestock genome: A review of concepts, approaches and applications. *Livest Sci* 241, 104257.
- Schlamp F, van der Made J, Stambler R, et al. (2016) Evaluating the performance of selection scans to detect selective sweeps in domestic dogs. *Mol Ecol* 25, 342-56.
- Seyedabadi H, Amirinia S, BANA BM, Emrani H (2006) Parentage verification of Iranian Caspian horse using microsatellites markers, *Iran J Biotechnol* 4 (4), 260-264.
- Shukla SK, Liu W, Sikder K, et al. (2017) HMGCS2 is a key ketogenic enzyme potentially involved in type 1 diabetes with high cardiovascular risk. *Sci rep* 7, 4590.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585-95.
- Team RC (2013) R: A language and environment for statistical computing.
- Todorov V, Templ M, Filzmoser P (2011) Detection of multivariate outliers in business survey data with incomplete information. *Advances in Data Analysis and Classification* 5, 37-56.
- Vatsiou AI, Bazin E, Gaggiotti OE (2016) Detection of selective sweeps in structured populations: a comparison of recent methods. *Mol Ecol* 25, 89-103.
- Venables WN, Ripley BD (2013) *Modern applied statistics with S-PLUS*. Springer Science & Business Media.
- Verity R, Collins C, Card DC, et al. (2017) minotaur: A platform for the analysis and visualization of multivariate results from genome scans with R Shiny. *Mol Ecol Resour* 17, 33-43.
- Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. *PLoS biology* 4, e72.
- Wu T, Hu E, Xu S, et al. (2021) clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* 2, 100141.
- Yurchenko AA, Daetwyler HD, Yudin N, et al. (2018) Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Sci rep* 8, 12984.
- Zheng Y-Z, Liang L (2018) High expression of PXDN is associated with poor prognosis and promotes proliferation, invasion as well as migration in ovarian cancer. *Ann Diagn Pathol* 34, 161-5.