Research on Animal Production, Vol. 16, Issue 1, 2025 p: 121-131

Research Paper

Identification of Loci under Selection by Combining Different Methods in Thoroughbred and Turkmen Horses

Milad Hosseini[1] and Hossein Moradi-Shahrbabak[2]

1. PhD in Genetics and Animal Breeding, Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran, (Corresponding Author: milad.hosseiny88@gmail.com)
2. Associate Professor, Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran, (Corresponding Author: hmoradis@ut.ac.ir)

Received: 16 April, 2024 Revised: 03 July, 2024 Accepted: 04 August, 2024

Extended Abstract

Background: In recent years, wide access to molecular polymorphism data has increased attention to identifying patterns within the genome that indicate positive selections occurring continuously in populations. The basic view about positive selection is the phenomenon of Genetic Hitchhiking. When a beneficial allele is the target of positive selection during different times, it creates signs at the genome level and these signs can be identified and tracked by different methods. The meta-analysis of different methods of selection signature at the genome level can lead to a much more accurate identification of these signature, because each of the different methods of identifying selection signature follow different criteria for identifying marks in the genome of organisms. In fact, the purpose of performing different methods of identification of selection signature is to improve the discovery and identification by using the combination of information and increase the accuracy and precision of the results.

Methods: The number of samples in this study included 111 Thoroughbred and Turkoman horses (44 and 67 respectively). Blood samples were obtained from the veins of the vedic and subcaudal veins of horses. In this research, single nucleotide polymorphism data (SNP chips, 60k) were used. Genotyping of the samples was done using 60 K arrays from Illumina company at the University of Kentucky. To ensure the quality of the genotyped genomic data, quality control analyzes were performed with a series of filters including MAF, SNP-CR, Animal-CR, HWE test. One of the important points in identifying signature of selection is to consider subpopulations in each population, for this reason, the principal components in this section was obtained based on the kinship matrix. Principal components analysis was done in R software. In this research, Fst, iHS (EHH), Rsb methods were used to identify the selection signature, and finally meta-analysis was performed. The regions under genome selection and the genomic positions obtained from whole genome linkage studies were further investigated to find the genes present in these regions. The genes reported in these regions in horses and related regions in cattle were identified using the ENSEMBL database and based on the assembly of the latest horse genome version available from the NCBI and BioMart databases. The identification code related to the identified genes was done in DAVID software to check the biological pathways related to the selected regions.

Results: Primary quality control was performed on 111 animals from 2 populations (67 Turkoman and 44 Thoroughbred). 3 Turkoman animals were excluded from further analysis due to the missing genotype rate of more than 9%, and finally 108 samples remained. A total of 59,349 SNP markers, 52,036 markers remained after quality control. A total of 118 SNPs were excluded due to the minimum allelic frequency of less than 1%, 704 SNPs out of Hardy-Weinberg equilibrium, 6493 SNPs with a missing genotype rate greater than 5%. The results of PCA analysis show that the samples examined in this study were placed in two completely separate groups. Finally, 5 regions on the genome were selected for further investigation. 5 regions with theta value higher than 0.28 between Turkmen and Thoroughbred horse breeds are located on chromosomes 4, 5, 10, 13, and 15, respectively. The highest number of selected markers were located on chromosome 5, and the lowest number of identified markers were located on chromosomes 13 and 15. According to the results obtained from the iHS analysis, the number of 8 and 23 snps, , in Thoroughbred and Turkmen breeds had the highest iHS value respectively, and the number of 3, 2, 2, 1 snps were respectively on chromosomes 28, 21, 1 and The highest iHS value is for the BIEC2_548092 snp located on chromosome 21 of this breed. The results of meta-analysis showed that these regions are located on chromosomes 10, 12, 14, 15, 16, 17, 32. Rsb analysis was performed between two populations of Thoroughbred and Turkoman horses, and the results

Copyright ©2025 Hosseini & Moradi Shahrbabak. Published by Sari Agricultural Sciences and Natural Resources University.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 Unported License which allows users to read, copy, distribute and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.

[Downloaded from rap.sanru.ac.ir on 2026-10-23]

[DOI: 10.61882/rap.16.1.121]

showed that selection markers are located on chromosomes 12, 16, 17, 31, 32. After examining the genes in the obtained regions, the most important genes were involved in glycosylation, cellular response to stress, metabolic process of cellular macromolecules, repair of broken peptides.

Conclusion: The present research, taking into account the various methods of identifying selection marks in Turkmen and Thoroughbred horses and combining the results of each of the methods, implements a new approach to infer the positive points of selection marks in the horse genome. From the results of this research, it can be concluded that the combination of methods to identify the signs of selection increases the strength and accuracy of the results. Also, the genes identified in this research are involved in important pathways such as glycosylation, cellular response to stress, metabolic process of peptides and organization of protein subunits, cellular macromolecules catabolic process and extracellular transport.

Keywords: Gene ontology, Rsb analysis, Meta-analysis, Thoroughbred horse, Turkmen horse

How to Cite This Article: Hosseini, M., & Moradi-Shahrbabak, H. (2025). Identification of Loci under Selection by Combining Different Methods in Thoroughbred and Turkmen Horses. *Res Anim Prod,* 16(1), 121-131. DOI: 10.61186/rap.16.1.121

[Downloaded from rap.sanru.ac.ir on 2026-10-23]

[DOI: 10.61882/rap.16.1.121]

مقاله پژوهشی

شناسایی جایگاه‌های تحت انتخاب به‌وسیله‌ی ترکیب روش‌های مختلف در اسب‌های تروبرد و ترکمن

سید میلاد حسینی[1] و حسین مرادی شهربابک[2]

۱- دانش آموخته دوره دکتری، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران، (نویسنده مسوول: milad.hosseiny88@gmail.com)
۲- استادیار، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران، (نویسنده مسوول: hmoradis@ut.ac.ir)

تاریخ دریافت: ۱۴۰۳/۰۱/۲۸ تاریخ ویرایش: ۱۴۰۳/۰۴/۱۳ تاریخ پذیرش: ۱۴۰۳/۰۵/۱۴
صفحه: ۱۲۱ تا ۱۳۱

چکیده مبسوط

مقدمه و هدف: در سال‌های اخیر، دسترسی وسیع به داده‌های چندشکلی مولکولی، توجه به شناسایی الگوهای درون ژنوم که نشان‌دهنده انتخاب‌های مثبتی به‌وقوع پیوسته در جمعیت‌ها است را افزایش داده‌است. دیدگاه کلیدی در این خصوص پدیده انتقال همراه می‌باشد. هنگامی‌که یک آلل سودمند در طی زمان‌های مختلف هدف انتخاب مثبت قرار می‌گیرد، باعث ایجاد نشانه‌هایی در سطح ژنوم می‌گردد که این نشانه‌ها با روش‌های مختلف قابل شناسایی و پیگیری می‌باشند. فراتحلیل روش‌های مختلف نشانه‌های انتخاب در سطح ژنوم می‌تواند به پیگیری و شناسایی بسیار دقیق‌تر این نشانه‌ها منجر شود چراکه هر کدام از روش‌های مختلف شناسایی نشانه‌های انتخاب، معیارهای متفاوتی را برای پیگیری نشانه‌ها در ژنوم موجودات دنبال می‌کنند. در واقع هدف از انجام روش‌های مختلف شناسایی نشانه‌های انتخاب، بهبود مرحله کشف و شناسایی با استفاده از ترکیب اطلاعات و افزایش صحت و دقت نتایج است.

مواد و روش‌ها: تعداد نمونه‌ها در این مطالعه شامل ۱۱۱ راس از اسب‌های نژاد تروبرد و ترکمن (به‌ترتیب با تعداد ۴۴ و ۶۷ راس) بود. نمونه‌های خون از رگ‌های وداجی و زیردمی اسب تهیه شد. در این پژوهش از داده‌های چندشکلی تک نوکئوتیدی (SNP chips, 60k) استفاده شد. تعیین ژنوتایپ نمونه‌ها با استفاده از آرایه‌های K ۶۰ مربوط به شرکت Illumina در دانشگاه کنتاکی انجام شد. برای اطمینان از کیفیت داده‌های ژنومی تعیین ژنوتایپ شده آنالیزهای کنترل کیفیت با یک‌سری فیلترها شامل MAF، SNP-CR، Animal-CR، آزمون HWE انجام شد. به‌دلیل اینکه یکی از مباحث مهم در بحث نشانه‌های انتخاب در نظر گرفتن زیر جمعیت‌های هر نژاد می‌باشد، مولفه اصلی در این بخش بر اساس ماتریس خویشاوندی به‌دست‌آمد. آنالیز مؤلفه‌های اصلی در محیط R صورت گرفت. در این پژوهش به‌منظور شناسایی نشانه‌های انتخاب از روش‌های F_{st}، iHS (EHH)، Rsb و در نهایت آنالیز فراتحلیل انجام شد. مناطق تحت انتخاب ژنوم و جایگاه‌های ژنومی به‌دست‌آمده از مطالعات پیوستگی کل ژنوم جهت یافتن ژن‌های موجود در این نواحی مورد بررسی بیشتر قرار گرفت. ژن‌های گزارش شده در این مناطق در اسب و مناطق متناظر در گاو با استفاده از پایگاه اطلاعاتی ENSEMBL و براساس سرهم‌سازی آخرین نسخه ژنومی در دسترس اسب از پایگاه اطلاعاتی NCBI و BioMart شناسایی شد. کد شناسایی مربوط به ژن‌های شناسایی شده جهت بررسی مسیرهای بیولوژیکی مرتبط با مناطق تحت انتخاب در نرم‌افزار DAVID انجام گرفت.

یافته‌ها: کنترل کیفیت اولیه بر روی ۱۱۱ حیوان از ۲ جمعیت (تعداد ۷۶ حیوان نژاد ترکمن و ۴۴ حیوان نژاد تروبرد) مورد مطالعه انجام شد. تعداد ۳ حیوان از نژاد ترکمن به‌دلیل نرخ ژنوتایپ از دست رفته‌ی بیشتر از ۹ درصد از ادامه‌ی آنالیزها حذف شدند و تعداد ۱۰۸نمونه باقی ماندند. از تعداد کل ۵۹۳۴۹ نشانگر SNP، تعداد ۵۲۰۳۶ نشانگر بعد از کنترل کیفیت باقی ماندند. در مجموع تعداد ۱۱۸ SNP به‌دلیلحداقل فراوانی آللی کمتر از ۱ درصد، تعداد ۷۰۴ SNP خارج از تعادل هاردی واینبرگ، تعداد ۶۴۹۳ SNP با نرخ ژنوتیپ گم شده بیشتر از ۵ درصد حذف شدند. نتایج آنالیز PCA نشان دهنده این موضوع است که نمونه‌های مورد بررسی در این مطالعه در ۲ گروه کاملا جدا از هم قرار گرفتند. در نهایت ۵ ناحیه بر روی ژنوم جهت بررسی‌های بعدی انتخاب شدند که توانستند از حد آستانه بالاتر باشند. ۵ منطقه که آماره تتا ارزش عددی بالاتر از ۰.۲۸ را بین دو نژاد اسب ترکمن و تروبرد داشتند به‌ترتیب روی کروموزوم‌های ۴، ۵، ۱۰، ۱۳، ۱۵ قرار دارند. بیشترین تعداد نشانگرهایی که ارزش تتای آن‌ها از حد آستانه بالاتر بود بر روی کروموزوم ۵ قرار داشت و کمترین تعداد نشانگر شناسایی شده بر روی کروموزوم ۱۳ و ۱۵ قرار داشتند. با توجه به نتایج به‌دست‌آمده از آنالیز iHS، تعداد ۸ و ۲۳ اسنیپ به ترتیب در نژادهای تروبرد و ترکمن دارای بالاترین ارزش iHS بودند که تعداد ۳، ۲، ۲، ۱ اسنیپ به‌ترتیب روی کروموزوم‌های ۲۸، ۲۱، ۱ و ۷ نژاد تروبرد قرار داشتند و بالاترین مقدار ارزش iHS برای اسنیپ BIEC2_548092 واقع بر کروموزوم ۲۱ این نژاد است. نتایج آنالیز فراتحلیل نشان داد این مناطق بر روی کروموزوم‌های ۱۰، ۱۲، ۱۴، ۱۵، ۱۶، ۱۷، ۳۲ قرار گرفته‌اند. آنالیز Rsb بین دو جمعیت اسب‌های کرد و ترکمن انجام شد و نتایج نشان داد که نشانه‌های انتخاب بر روی کروموزوم‌های ۱۲، ۱۶، ۱۷، ۳۱، ۳۲، قرار دارند. پس از بررسی ژن‌های در مناطق به‌دست‌آمده، مهم‌ترین ژن‌ها در گلیکوزیلاسیون، پاسخ سلولی به استرس، فرآیند متابولیک ماکرومولکول سلولی، تعمیر شکستگی دو رشته پپتیدها دخالت داشتند.

نتیجه‌گیری: تحقیق حاضر با در نظر گرفتن روش‌های متنوع به‌منظور شناسایی نشانه‌های انتخاب در اسب‌های ترکمن و تروبرد و ترکیب نتایج حاصل از هر یک از روش‌ها، رویکرد جدیدی را برای استنتاج نقاط مثبت نشانه‌های انتخاب در ژنوم اسب اجرا می‌کند. از نتایج این تحقیق می‌توان نتیجه‌گیری کرد که ترکیب روش‌های شناسایی نشانه‌های انتخاب قدرت و صحت نتایج را بالا می‌برد. همچنین ژن‌های شناسایی شده در این تحقیق در مسیرهای مهم از جمله گلیکوزیلاسیون، پاسخ سلولی به استرس، فرایند متابولیک پپتیدها، سازماندهی زیر واحدهای پروتئین‌ها، فرآیندکاتابولیک ماکرومولکول‌های سلولی و انتقال از فضای خارج سلولی نقش دارند.

واژه‌های کلیدی: اسب ترکمن، اسب تروبرد، آنالیز فراتحلیل، آنالیز Rsb، ژن آنتولوژی

مقدمه

اسب یکی از باهوش‌ترین و تربیت‌پذیرترین حیوانات اهلی است که با انسان اُنس گرفته و مورد انقیاد و تمکین او قرار می‌گیرد (Mostafavi *et al.*, 2019). اسب دارای پراکنشی جهانی است ولی باتوجه به شرایط جغرافیایی خاص هر منطقه و کشور، نژادهای مختص به شرایط همان منطقه شکل گرفته و تکثیر شده‌اند. نژادهای بومی بعد از هزاران سال انتخاب طبیعی و گذر از موانع بسیار و غلبه بر ناملایمات و شرایط نامساعد محیطی همچنان به حیات خویش ادامه داده و به تکثیر و ازدیاد نسل پرداخته‌اند (Moazemi *et al.*, 2020). از این‌رو این موجودات در هر کشوری به‌عنوان سرمایه ملی و ذخایر کلیدی بوده و حفظ و تکثیر آن‌ها از ارزش و اهمیت بالایی برخوردار است (Saadatabadi *et al.*, 2023). پویش گستره ژنوم امروزه غالباً از طریق بررسی تعداد بسیار زیاد SNP یا

[Downloaded from rap.sanru.ac.ir on 2026-10-23]
[DOI: 10.61882/rap.16.1.121]

همان چندشکلی‌های تک نوکلئوتیدی، که در طول ژنوم پراکنده‌اند، صورت می‌پذیرد و بستر مناسبی برای افزایش درک بشر از معماری ژنتیکی صفات پیچیده، انتخاب ژنومیک، تحقیقات پیوستگی در سطح ژنوم و همچنین بررسی سیگنال‌های انتخاب در حیوانات مزرعه‌ای فراهم کرده‌است. عبارت ژنومیک جمعیت اولین‌بار برای آنالیز تنوع جمعیتی درون و بین جمعیت‌ها براساس داده‌های کل ژنوم استفاده شد و خصوصیات خاص جمعیت‌ها از قبیل اندازهٔ مؤثر جمعیت و تمایز جمعیت‌ها با استفاده از روش‌های آنالیز ژنومیک جمعیت استنباط گردید (Gondro *et al*., 2013). یکی از رایج‌ترین و کارآمدترین تکنیک‌ها در شناسایی مناطق ژنومی موثر بر صفات کمی در حوزه‌های مختلف ژنتیک از جمله ژنتیک دامی، شناسایی نشاه‌های انتخاب است. انتخاب، فرکانس واریانت‌های سودمند و مکان‌های چندشکلی همسایه‌شان را تغییر می‌دهد، ژنوم را فرا می‌گیرد و الگوهایی را که علی‌رغم نوترکیبی کروموزوم در یک جمعیت رایج می‌شوند، باقی می‌گذارد (Sabeti *et al*., 2002). این الگوها به‌طور گسترده به‌عنوان نشانه‌های (یا ردپای) انتخاب نامیده می‌شوند و روش‌های زیادی برای شناسایی آن‌ها در داده‌های ژنومی توسعه یافته‌اند (Oleksyk *et al*., 2010). این نشانه‌ها شامل کاهش مکانی تنوع ژنتیکی، انحراف از توزیع فراوانی اللی[۶] و افزایش الگوی عدم تعادل لینکاژی و توسعه ساختار هاپلوتایپی[۷] می‌شود (Grossman *et al*., 2010). با وجود این‌که روش‌های موجود برای شناسایی نشانه‌های انتخاب بر روی داده‌های ژنومی در فرض‌ها و اجرا متفاوت است و همبستگی‌های ضعیفی بین آن‌ها مورد انتظار است ولی مناطقی که با روش‌های مختلف تحت عنوان نشانه‌های انتخاب شناسایی می‌شوند، شواهد قوی دال‌بر وجود انتخاب در این نواحی هستند (Utsunomiya *et al*., 2013). از مزایای اصلی این روش‌ها این است که اجرای آن‌ها تنها با استفاده از اطلاعات حاصل از داده‌های ژنومی و در غیاب رکوردهای فنوتیپی نیز وجود دارد. وقتی یک جهش مطلوب در یک جمعیت رخ می‌دهد فراوانی این آلل در جمعیت در طی نسل‌های بعدی افزایش پیدا می‌کند. به‌دلیل این‌که DNA ترکیب خطی از مولکول‌ها است و نیز احتمال نوترکیبی نسبت معکوس با فاصله ژنومی بین دو جایگاه دارد بنابراین فراوانی نوکلئوتیدهای مجاور جهش مطلوب نیز افزایش پیدا می‌کند. این اتفاق باعث باقی ماندن نشانه‌هایی در سطح ژنوم می‌شود که به‌وسیله‌ی توزیع معنی‌دار آماری نوکلئوتیدها در مجاورت جهش مطلوب مشخص می‌شود (Gurgul *et al*., 2019). در حال حاضر ابزارها و روش‌های مختلفی برای شناسایی این نشان‌های انتخاب در سطح ژنوم جمعیت‌های مختلف، به‌طور موفقیت‌آمیز توسعه یافته و استفاده می‌گردد (Qanbari *et al*., 2010). با این‌که روش‌های موجود برای شناسایی نشانه‌های انتخاب روی داده‌های ژنومیکی در فرض‌ها و اجرا متفاوت هستند و انتظار می‌رود رابطهٔ ضعیفی بین نتایج آن‌ها وجود داشته باشد، مناطقی که با روش‌های مختلف به‌عنوان نشانه‌های انتخاب شناسایی می‌شوند، براساس شواهد قطعی نشان‌دهندهٔ وجود انتخاب در این نواحی هستند (Utsunomiya *et al*., 2013). فراتحلیل در ابتدا به‌عنوان روشی برای ادغام نتایج حاصل از مجموعه‌ای از آزمایشات بالینی مشابه توسعه یافت. متعاقباً برای ترکیب داده‌های حاصل از مطالعات مشاهده‌ای، سپس برای ترکیب مطالعات ژنتیکی ژن‌های کاندید استفاده شد، و اکنون فراتحلیل به‌طور معمول برای تجمیع نتایج حاصل از مطالعات ارتباط ژنومی (GWAS) استفاده می‌شود. علی‌رغم این گسترش حوزه کاربرد، روش‌های مورد استفاده تغییر بسیار کمی داشته‌اند، بنابراین کم و بیش همان رویکرد فراتحلیل اولیه برای داده‌های ژنومی مورد استفاده قرار می‌گیرد که برای اولین‌بار برای آزمایش‌های بالینی توسعه داده‌شد (Lean *et al*., 2009). یکی از ویژگی‌های متاآنالیزهای ژنومی این است که می‌توان آن‌ها را به‌دلایل مختلفی انجام داد و هر هدف به‌شکل کمی متفاوت از تجزیه و تحلیل نیاز دارد. هنگامی که انگیزه برای متاآنالیز واضح نیست، انتخاب یک تحلیل مناسب می‌تواند دشوار باشد و این دشواری زمانی بدتر می‌شود که یک متاآنالیز واحد دارای چندین مؤلفه مجزا با اهداف ضمنی متفاوت باشد. بنابراین، اولین وظیفه ما این است که اشکال مختلف متاآنالیز در سطح ژنوم و آنچه که آن‌ها در تلاش برای دستیابی هستند را در نظر بگیریم (Vesterinen *et al*., 2014). فراتحلیل روش‌های مختلف نشانه‌های انتخاب در سطح ژنوم می‌تواند به پیگیری و شناسایی بسیار دقیق‌تر این نشانه‌ها منجر شود چراکه هر کدام از روش‌های مختلف شناسایی نشانه‌های انتخاب، معیارهای متفاوتی را برای پیگیری نشانه‌ها در ژنوم موجودات دنبال می‌کنند. یعنی یک آنالیز فراتحلیل که داده‌های GWAS را ترکیب می‌کند، می‌تواند با در نظر گرفتن ویژگی‌های بیولوژیکی ژن‌ها و SNPها در سراسر ژنوم از پایگاه‌های اطلاعاتی بیوانفورماتیک یا سایر منابع تجربی (مانند مطالعات بیان)، باعث افزایش چشم‌گیر صحت و دقت نتایج شود. در واقع هدف بهبود مرحله کشف و شناسایی با استفاده از ترکیب اطلاعات خواهد بود (Cantor *et al*., 2010). فراتحلیل در واقع یک روش آماری است که نتایج حاصل از مطالعات مختلف ولی با یک هدف مشخص ترکیب می‌کند و نتایج دقیق و حتی ناشناخته را ارائه می‌دهد (Ramasamy *et al*., 2008). این روش‌ها همچنین امکان ارزیابی کمی سازگاری یا ناسازگاری (ناهمگونی) نتایج را در میان مجموعه‌های داده متعدد فراهم می‌کنند. روش‌های متاآنالیز برای چندین دهه در زمینه‌های مختلف علمی مورد استفاده قرار گرفته‌اند و در حال‌حاضر چندین کتاب درسی و کتاب راهنما برای آن‌ها وجود دارد (Higgins *et al*., 2007; Sutton *et al*., 2000). هدف این پژوهش انجام آنالیز فراتحلیل مجموعه داده‌های نشانه‌های انتخاب حاصل از آماره‌های مختلف به‌منظور افزایش قدرت تشخیص سیگنال‌های ارتباطی، افزایش صحت و دقت شناسایی ردپاهای انتخاب در اسب‌های ترکمن و تروبرد با بررسی انواع بیشتری از این نشانه‌ها در سراسر ژنوم نسبت به هر مجموعه داده به‌تنهایی است.

مواد و روش‌ها

در این پژوهش از داده‌های چندشکلی تک نوکئوتیدی (SNP chips, 60k) استفاده شد. این داده‌ها توسط شرکت ایلومینا تولید شد. تعداد نمونه‌ها در این مطالعه شامل ۱۱۱ راس

[Downloaded from rap.sanru.ac.ir on 2026-10-23]
[DOI: 10.61882/rap.16.1.121]

از اسب‌های نژاد تروبرد و ترکمن (به‌ترتیب با تعداد ۴۴ و ۶۷ راس) بود. همچنین تعداد ۲۱ نمونه مادیان و ۲۳ نمونه نریان در اسب‌های تروبرد و تعداد ۵۶ نمونه مادیان و ۵۵ نمونه نریان در اسب‌های ترکمن وجود داشت. نمونه‌های خون از رگ‌های وداجی و زیردمی اسب تهیه شده و نمونه‌های گرفته شده تا زمان استخراج در دمای ۲۰- سانتی‌گراد نگهداری شدند. استخراج DNA از خون با استفاده از روش بهینه نمکی (Grimberg *et al*., 1989) انجام‌گرفت. تعیین کمیت و کیفیت DNA به‌وسیله‌ی دستگاه نانودراپ و ژل آگارز ۱/۲ درصد صورت گرفت. سپس جهت تعیین ژنوتایپ نمونه‌ها به دانشگاه کنتاکی ارسال شدند و با استفاده از آرایه‌های ۶۰ K مربوط به شرکت Illumina تعیین ژنوتایپ شدند. برای اطمینان از کیفیت داده‌های ژنومی تعیین ژنوتایپ شده آنالیزهای کنترل کیفیت با یک‌سری فیلترها شامل MAF، SNP-CR، Animal-CR، آزمون HWE انجام شد. SNPهایی که در مجموع حیوانات دارای CR کمتر از ۹۵ درصد بودند، حذف شدند (Barendse *et al*., 2009). همچنین حیواناتی که در مجموع بیش از ۹ درصد ژنوتیپ از دست‌رفته داشتند، به‌منظور جلوگیری از کاهش صحت و دقت نتایج حذف شدند. سپس SNPهایی که سطح معنی‌داری تعادل هاردی-واینبرگ برای آن‌ها کمتر از (۱×۱۰^{-۶}) بود حذف شدند (Teo *et al*., 2007). سطح معنی‌داری به‌عنوان معیاری از خطای تعیین ژنوتایپ در نظر گرفته شد. کنترل کیفیت داده‌ها با استفاده از نرم‌افزار PLINK صورت گرفت. ژنوتیپ‌های به‌دست‌آمده از این بخش جهت آنالیز PCA مورد استفاده قرار گرفتند. یکی از شیوه‌های آماری برای آزمون ارتباط بین جوامع و اختصاص افراد به آن‌ها با استفاده از ماتریس فاصله، استفاده از آنالیز مولفه‌های اصلی (PCA) است (Jombart & Ahmed, 2011). به‌دلیل اینکه یکی از مباحث مهم در بحث نشانه‌های انتخاب در نظر گرفتن زیر جمعیت‌های هر نژاد می‌باشد، مولفه اصلی در این بخش بر اساس ماتریس خویشاوندی به‌دست‌آمد. آنالیز مؤلفه‌های اصلی در محیط R صورت گرفت. در این پژوهش به‌منظور شناسایی نشانه‌های انتخاب از روش‌های زیر استفاده شد.

روش‌های آماری مبتنی بر تمایز جمعیتی

یکی از متداول‌ترین روش‌ها برای شناسایی نشانه‌های انتخاب، آماره‌ی F_{st} است. جهت بررسی الگوی ژنومی انتخاب مثبت در جمعیت‌های اسب‌های نژاد ترکمن و تروبرد، ارزش‌های F_{st} برای هر مارکرSNP به‌روش برآوردگر ناریب تتا در محیط R محاسبه گردید. روش ناریب تتا غالبا در تحقیقاتی با اندازه نمونه کم و مقایسه جمعیت‌هایی با اندازه متفاوت استفاده می‌شود. برای محاسبه‌ی این آماره از فرمول زیر استفاده شد (Weir & Cockerham, 1984).

$$F_{st} = \frac{H_T - H_S}{H_T}$$

H_T هتروزیگوسیتی مورد انتظار در کل جمعیت و H_Sهتروزیگوسیتی داخل زیر جمعیت است.

روش‌های مبتنی بر عدم تعادل پیوستگی

علاوه بر روش‌های تمایز جمعیتی، آماره EHH که مبتنی بر عدم تعادل پیوستگی می‌باشد، جهت شناسایی مناطق ژنومی تحت انتخاب مورد استفاده قرارگرفت. در واقع EHH یک هاپلوتیپ گسترده شده خاص بدون نوترکیبی است که به‌صورت زیر محاسبه می‌شود:

$$EHH_{s,t} = \frac{1}{n_{a_s}(n_{a_s}-1)} \sum_{k=0}^{k_{a_s,t}} n_k(n_k - 1)$$

$N_{a\ s}$ فراوانی هاپلوتیپ‌هایی است که حامل هاپلوتیپ مرکزی a_s هستند، $K_{as,t}$ تعداد هاپلوتیپ‌های منحصربه فردی است که حامل هاپلوتیپ مرکزی از اسنیپ t تا اسنیپS در طول کروموزوم هستند و nk تعداد هاپلوتیپ‌های منحصربه فرد است. آزمون iHS یکی از متداول‌ترین روش‌های مبتنی بر روش‌های عدم تعادل لینکاژی است. این آماره یک معیار استاندارد از کاهش EHH آلل‌های مشتق‌شده در یک منطقه (برای مثال یک SNP) نسبت به آلل‌های اجدادی را فراهم می‌کند (Voight *et al*., 2006). مناطقی با روند آهسته‌ی کاهش EHH در آلل‌های مشتق‌شده (که دارای طول هاپلوتیپی زیاد نسبت به هاپلوتیپ‌های مورد انتظار آلل‌های اجدادی هستند) به‌عنوان نشانه‌ای از انتخاب تلقی می‌گردد. برای محاسبه‌ی این آماره از پکیج rehh در برنامه R استفاده شد. فایل‌های ورودی این پکیج دو فایل hap و map می‌باشند که فایل hap حاوی اطلاعات ژنوتیپ‌ها و فایلmap نیز حاوی اطلاعات نقشه اسنیپ‌ها و همچنین آلل‌های اجدادی و مشتق‌شده برای هر اسنیپ می‌باشد.

$$iHs = \frac{\ln\left(\frac{iHHA}{iHHD}\right) - Ep\left[\ln\left(\frac{iHHA}{iHHD}\right)\right]}{SDp\left[\ln\left(\frac{iHHA}{iHHD}\right)\right]}$$

A آلل اجدادی و D آلل اشتقاقی است. انحراف استاندارد مورد انتظار ln(iHHA/iHHD) از توزیع تجربیSNP برآورد شده که p فراوانی آلل مشتق‌شده با فراوانی snp هسته منطبق شده‌است.

Rsb روشی است که طول هاپلوتیپ‌های مرتبط با همان آلل را بین جمعیت‌ها مقایسه می‌کند. روش Rsb پروفایل‌های EHH از یک آلل را بین جمعیت‌ها به‌صورت جفتی مقایسه می‌کند. این روش به‌عنوان لگاریتم نسبت بین iESpop1 و iESpop2 تعریف شد، که در آن iES نشان‌دهنده EHHS یکپارچه (EHH مختص یک منطقه) برای هر دو آلل هر SNP در هر جمعیت است. Rsb بین جمعیت‌ها (AB) محاسبه شد. این روش به هیچ اطلاعاتی از آلل‌های اجدادی و مشتق‌شده نیاز ندارد. مقادیر مثبت Rsb نشان می‌دهد که iESpop1 بزرگ‌تر از iESpop2 است، یعنی جمعیت ۱ هاپلوتیپ طولانی‌تری نسبت به جمعیت ۲ دارد، بنابراین انتخاب مثبت در جمعیت جایگزین (جمعیت ۱) را پیشنهاد می‌کند. برعکس، نمره منفی نشان‌دهنده انتخاب مثبت در یک جمعیت مرجع (جمعیت ۲) است (Tang et al., 2007).

$$Rsb = \ln\left(\frac{iESpop1}{iESpop2}\right)$$

[Downloaded from rap.sanru.ac.ir on 2026-10-23]
[DOI: 10.61882/rap.16.1.121]

روش فراتحلیل

آزمون فراتحلیل آزمون دیگری است که بر روی داده‌های حیوانات اهلی مورد استفاده قرار گرفته است. در این آزمون از ترکیب P-Valueهای به‌دست‌آمده از آزمون‌های شناسایی نشانه‌های انتخاب استفاده شد. از آن‌جایی که همه روش‌های شناسایی نشانه‌های انتخاب دارای مقادیر P بودند که از توزیع‌های نرمال با پارامترهای یکسان (میانگین و واریانس ۱) بازیابی‌شده بودند، نسخه وزن‌دار روش استوفر برای ترکیب مقادیر P استفاده شد. برای هر نشانگر و هر آزمون i، P-value مربوطه با Zi =2w21(12pi) به یک Z-score تبدیل شد.

آزمایش‌های درون‌جمعیتی فقط یک‌بار برای هر نژاد انجام شد، از این‌رو وزن مربوطبه آن‌ها vi روی ۱ تنظیم شد. سپس، آمار ترکیبی K تست، برای هر SNP در هر نژاد، به‌صورت زیر تعریف شد:

$$meta-SS=\frac{\sum_{i=1}^{k}\omega_i Z_i}{\sqrt{\sum_{i=1}^{k}\omega_i{}^2}}$$

نمرات فراتحلیل به توزیع نرمال استاندارد ارجاع داده‌شد تا مقادیر معنی‌داری ترکیبی به‌دست‌آید. سطح معنی‌داری برای مقادیر P فراتحلیل در سطح ژنوم براساس آستانه بونفرونی (a = 0.05/nSNP) بود.

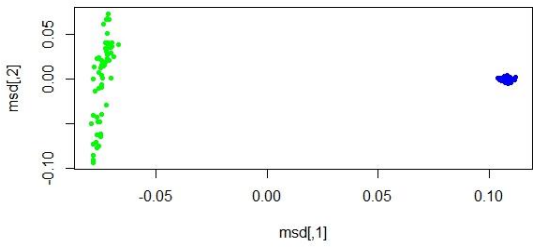
شناسایی ژن‌ها و مسیرهای بیولوژیکی مرتبط با مناطق ژنومی تحت انتخاب

مناطق تحت انتخاب ژنوم و جایگاه‌های ژنومی به‌دست‌آمده از مطالعات پیوستگی کل ژنوم جهت یافتن ژن‌های موجود در این نواحی مورد بررسی بیشتر قرار گرفت. ژن‌های گزارش‌شده در این مناطق در اسب و مناطق متناظر در گاو با استفاده از پایگاه اطلاعاتی ENSEMBLو بر اساس سرهمسازی آخرین نسخه ژنومی در دسترس اسب از پایگاه اطلاعاتی NCBI و BioMart شناسایی شد. کد شناسایی مربوط به ژن‌های شناسایی‌شده جهت بررسی مسیرهای بیولوژیکی مرتبط با مناطق تحت انتخاب در نرم‌افزار DAVID انجام گرفت.

جدول ۱- کنترل کیفی داده‌های ژنومی مربوط به اسب‌های نژاد تروبرد و ترکمن
Table 1. Quality control of genomic data related to Thoroughbred and Turkoman breed horses

تروبرد Thoroughbred	ترکمن Turkoman	جمعیت population
44	67	تعداد حیوانات هر جمعیت Number of animal in populations
59349	59349	تعداد کل نشانگرها قبل از کنترل کیفیت Number of markers before quality control
118	118	تعداد SNPها با فراوانی آلل نادر کمتر از ۰/۰۱ The number of SNPs with rare allele frequency less than 0.01
704	704	تعداد SNPهای خارج از تعادل هاردی-وینبرگ The number of SNPs discharge the Hardy Weinberg equilibrium
6493	6493	تعداد SNPها با نرخ ژنوتیپ گم شده بیشتر از ۰/۰۵ The number of SNPs with a missing genotype rate greater than 0.05
0	3	تعداد افراد با نرخ ژنوتیپ از دست رفته ۰/۰۹ The number of animals with a missing genotype rate of 0.09
0.954	0.954	نرخ ژنوتایپینگ در افراد باقیمانده Genotyping rate in the remaining sample
52036	52036	تعداد کل SNPها بعد از کنترل کیفیت Number of SNPs after quality control

آنالیز مولفه‌های اصلی ساختار جمعیتی ۲ نژاد اسب تروبرد و ترکمن در محیط بررسی شد. نتایج این بررسی نشان دهنده این موضوع است که نمونه‌های مورد بررسی در این مطالعه در ۲ گروه کاملاً جدا از هم قرار گرفتند. این دو گروه شامل اسب‌های ترکمن (رنگ سبز) و اسب‌های تروبرد (رنگ آبی) هستند. همچنین هرچه فاصله‌ی بین دو گروه بیشتر باشد نشان دهنده‌ی خلوص نژادی و تفکیک بیشتر دو گروه است (شکل ۱).



شکل ۱- آنالیز مؤلفه‌های اصلی مربوط به نژادهای اسب تروبرد (آبی) و ترکمن (سبز) با استفاده از اطلاعات چندشکلی‌های تک نوکلئوتیدی
Figure 1. PCA analysis of Thoroughbred horse breeds (blue) and Turkoman (green) using SNP

نتایج و بحث

جهت بررسی الگوی ژنومی انتخاب مثبت در جمعیت‌های اسب‌های نژاد ترکمن و تروبرد، ارزش‌های F_{st} برای هر مارکرSNP به روش برآوردگر نااریب تتا در محیط R محاسبه گردید (Weir & Cockerham, 1984). باتوجه به این‌که این روش نااریب است احتمال به‌دست‌آوردن ارزش‌های منفی نیز

[Downloaded from rap.sanru.ac.ir on 2026-10-23]
[DOI: 10.61882/rap.16.1.121]

وجود دارد. ارزش‌های نااریب تتا برای هر SNP و به‌طور جداگانه‌ای برای کروموزوم‌های اتوزومی و کروموزوم جنسی X، محاسبه شد. جهت شناسایی بهتر سیگنال‌های انتخاب در سطح ژنوم به‌جای در نظر گرفتن ارزش عددی هر SNP از میانگین‌گیری ارزش‌های عددی SNPهای مجاور به‌روش CW[1] (Qanbari *et al*., 2011) استفاده شد و تنها ۱٪ مناطقی از ژنوم که در آن تمام نشانگرهای مجاور دارای ارزش بالایی بودند، به‌عنوان نشانه‌های انتخاب شناسایی و تعیین شدند. حد آستانه در این بررسی سخت‌گیرانه بود و فقط مناطقی از ژنوم که بیشترین تمایز را بین دو نژاد نشان دادند، انتخاب گردیدند. مزیت این روش نسبت به‌روش پایه‌ای F_{st} که توسط رایت ارائه شد این است که در روش نااریب تتا اندازه‌ی نمونه‌ها در فرمول لحاظ گردیده و در واقع خطای نمونه‌گیری در نظر گرفته شده‌است. این حد آستانه در مطالعات مختلف از ۰/۰۱ درصد تا ۲/۵ درصد متغیر بود (Moradi *et al*., 2012; Qanbari *et al*., 2010). در نهایت ۵ ناحیه بر روی ژنوم جهت بررسی‌های بعدی انتخاب شدند که توانستند از حد آستانه بالاتر باشند. ۵ منطقه که آماره تتا ارزش عددی بالاتر از ۰/۲۸ را بین دو نژاد اسب ترکمن و تروبرد داشتند به‌ترتیب روی کروموزوم‌های ۴، ۵، ۱۰، ۱۳، ۱۵ قرار دارند. بیشترین تعداد نشانگرهایی که ارزش تتای آن‌ها از حد آستانه بالاتر بود بر روی کروموزوم ۵ قرار داشت و کمترین تعداد نشانگر شناسایی شده بر روی کروموزوم ۱۳ و ۱۵ قرار داشتند. پس از به‌دست‌آوردن آماره iHS ، اسنیپ‌هایی که دارای بالاترین مقدار iHS هستند به‌عنوان اسنیپ‌هایی که در هر نژاد تحت انتخاب قرار گرفتند معرفی شدند (جدول ۲ و جدول ۳). اگر هاپلوتایپ مشتق‌شده بزرگ‌تر از هاپلوتایپ اجدادی باشد، مقدار iHS منفی است و اگر هاپلوتایپ اجدادی بزرگ‌تر از هاپلوتایپ مشتق‌شده باشد مقدار iHS مثبت است.

جدول ۲- نشانگرهای دارای بالاترین ارزش iHS در نژاد تروبرد

Table 2. Markers with the highest iHS value in Thoroughbred breed

نام اسنیپ SNP Name	کروموزوم Chromosome	موقعیت ژنومی Genomic Position	iHS
BIEC2_548092	21	5030403	-4.35
BIEC2_745459	28	45097396	-4.34
BIEC2_745385	28	44680548	-3.93
BIEC2_47180	1	112650539	-3.91
BIEC2_548091	21	5026970	-3.88
BIEC2_745469	28	45186697	-3.87
BIEC2_47251	1	112751114	-3.82
BIEC2_1007560	7	77913275	-3.66

جدول۳- نشانگرهای دارای بالاترین ارزش iHS در نژاد ترکمن

Table 3. Markers with the highest iHS value in Turkoman breed

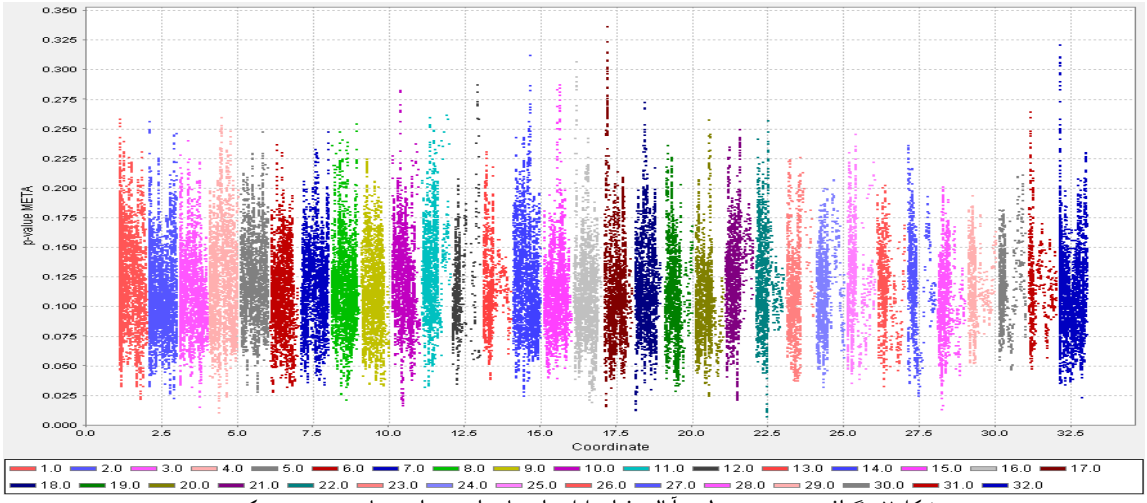
نام اسنیپ SNP Name	کروموزوم Chromosome	موقعیت ژنومی Genomic Position	iHS
BIEC2_867460	4	58264781	4.38
BIEC2_1027164	8	8498964	-4.34
BIEC2_665593	25	24590167	4.28
BIEC2_411742	18	42935127	4.20
BIEC2_894650	5	15211001	4.10
BIEC2_882951	4	104984674	4.02
BIEC2_1027098	8	8138630	-3.97
BIEC2_885665	4	108136351	3.96
BIEC2_867482	4	58352096	3.95
BIEC2_38058	1	91269455	3.95
BIEC2_891593	5	10763970	3.94
BIEC2_867049	4	56134845	3.94
BIEC2_12789	1	28826556	3.87
BIEC2_238237	14	1109903	3.86
BIEC2_898305	5	23868607	3.82
BIEC2_891633	5	10954654	3.79
BIEC2_898277	5	23674756	3.74
BIEC2_892648	5	12437611	3.71
BIEC2_1027310	8	8722189	-3.71
BIEC2_1027146	8	8470201	-3.70
BIEC2_663913	25	21373000	3.67
BIEC2_754760	29	14979173	3.67
BIEC2_702123	27	7756449	3.67

1- Creeping Window

[Downloaded from rap.sanru.ac.ir on 2026-10-23]

[DOI: 10.61882/rap.16.1.121]

باتوجه به نتایج به‌دست‌آمده از آنالیز iHS، تعداد ۸ و ۲۳ اسنیپ به‌ترتیب در نژادهای تروبرد و ترکمن دارای بالاترین ارزش iHS بودند که تعداد ۳، ۲، ۲، ۱ اسنیپ به‌ترتیب روی کروموزوم‌های ۲۸، ۲۱، ۱ و ۷ نژاد تروبرد قرار داشتند و بالاترین مقدار ارزش iHS برای اسنیپ BIEC2_548092 واقع بر کروموزوم ۲۱ این نژاد است. همچنین در نژاد ترکمن تعداد ۵، ۴، ۲، ۱، ۶، ۲، ۱، ۱، ۱، اسنیپ به‌ترتیب بر روی کروموزوم‌های ۴، ۸، ۲۵، ۱۸، ۵، ۱، ۱۴، ۲۷، ۲۹ واقع شده‌اند و بالاترین مقدار ارزش iHS را اسنیپ BIEC2_548092 واقع بر کروموزوم ۴ دارد. به‌منظور شناسایی چندشکلی‌های تک نوکلئوتیدی که تحت انتخاب قرار گرفته‌اند، آماره‌ی Rsb محاسبه شد. این آماره نشانه‌های انتخاب را بین دو جمیت بررسی می‌کند. به‌همین منظور آنالیز Rsb بین دو جمعیت اسب‌های تروبرد و ترکمن انجام شد و نتایج نشان داد که نشانه‌های انتخاب بر روی کروموزوم‌های ۱۲، ۱۶، ۱۷، ۳۱، ۳۲، قرار دارند. آنالیز فراتحلیل مربوط به همه‌ی روش‌های شناسایی نشانه‌های انتخاب در اسب‌های نژاد ترکمن و تروبرد انجام شد و به‌طور کلی مناطقی که دارای بیشترین تمایز بین این دو نژاد بودند، به‌طور سخت گیرانه‌ای شناسایی و معرفی شد. نتایج آنالیز فراتحلیل نشان داد این مناطق بر روی کروموزوم‌های ۱۰، ۱۲، ۱۴، ۱۵، ۱۶، ۱۷، ۳۲ قرار گرفته‌اند. گراف منهتن برای آنالیز فراتحلیل اسب‌های نژاد ترکمن و تروبرد در شکل ۲ آورده شده‌است. همچنین نتایج آنالیزهای مربوط به شناسایی ژن‌ها در این مناطق در جدول ۴ ارائه شده‌است.



شکل۲- گراف منهتن مربوط به آنالیز فراتحلیل داده‌های اسنیپ اسب‌های تروبرد و ترکمن

Figure 2. Manhattan graph for meta-analysis of SNP-data of Thoroughbred and Turkoman horses

جدول ۴- اطلاعات مربوط به هستی‌شناسی ژن‌های شناسایی‌شده در مناطق تحت انتخاب مثبت

Table 4. Ontology information of genes identified in regions under positive selection

P-value	ژن‌های کاندید در Term Candidate Genes In Term	تعداد ژن‌ها در هر TermCount	نام مسیر Pathway Name	کد هستی‌شناسی Ontology Code
0/089	ALG11- COG6	2	گلیکوزیلاسیون Glycosylation	GO:0070085
0/068	SEM1-SETMAR-ITPR1	3	پاسخ سلولی به استرس Cellular response to stress	GO:0033554
0/048	SEM1-SETMAR-BHLHE40-CRBN-SUMF1- TRNT1	6	فرآیند متابولیک ماکرومولکول سلولی Cellular macromolecule metabolic process	GO:0044260
0/071	SEM1-SETMAR	2	تعمیر شکستگی دو رشته‌ای double-strand break repair	GO:0006302
0/084	LXN- PTX3- RARRES1	3	انتقال از فضای خارج سلولی Extracellular transport	GO:0005615
0/047	ACCS- ALKBH3- AMBRA1-ATG13- CHST1- F2- DGKZ-EXT2-HSD17B12-MAPK8IP1-SLC35C1	11	فرآیند متابولیک تک ارگانیسم single organism Metabolic process	GO:0044710
0/007	CDC42EP3-NDUFAF7-CYP1B1- EIF2AK2-SPAST	5	سازماندهی زیر واحدهای پروتئین‌ها protein subunits Organization	GO:0071822
0/0202	MCTS1- UPF3B- LAMP2-RPL39	4	فرایند متابولیک پپتیدها peptide Metabolic process	GO:0006518
0/065	FBXL17- NUDT12- PJA2	3	فرآیندکاتابولیک ماکرومولکول‌های سلولی cellular macromolecules Catabolic process	GO:0044265

ژن *ALG11*[۲] روی کروموزوم ۱۶ قرار داشت و در فرآیند گلیکوزیلاسیون نقش دارد. مانوزیل ترانسفراز در آخرین مراحل سنتز الیگوساکارید هسته *(Man5GlcNAc(2)-PP-dolichol)* روی صفحه سیتوپلاسمی شبکه آندوپلاسمی نقش دارد. افزودن چهارمین و پنجمین باقیمانده مانوز به زنجیره الیگوساکارید متصل به دولیکول را کاتالیز می‌کند. دو دسته اصلی گلیکوزیلاسیون وجود دارد: *N-glycosylation* و *O-glycosylation*. نقص‌های ژنتیکی در مجموعه N–گلیکان،

[2] Alpha-1,2-Mannosyltransferase

[Downloaded from rap.sanru.ac.ir on 2026-10-23]

[DOI: 10.61882/rap.16.1.121]

یعنی سنتز و اتصال گلیکان‌ها به گلیکوپروتئین‌ها و گلیکولیپیدها، متعلق‌به خانواده بزرگ اختلالات مادرزادی گلیکوزیلاسیون (CDG) است (Péanne *et al*., 2018). تقریباً ۱۳۰ اختلال با نقص در گلیکوزیلاسیون که باعث اختلال عملکرد پروتئین یا لیپید می‌شود، مرتبط بوده‌است (Freeze *et al*., 2015). ناهنجاری‌های مرتبط با CDG تقریباً بر هر سیستم اندامی تأثیر می‌گذارد. نقص N–گلیکوزیلاسیون معمولاً به دو نوع طبقه‌بندی می‌شود. CDG نوع I در سنتز یا انتقال یک الیگوساکارید مرتبط با لیپید (دولیکول) (LLO) به پروتئین در شبکه آندوپلاسمی دچار کمبود هستند، در حالی‌که نوع II ناشی از پردازش N–گلیکان‌های متصل به پروتئین است که عمدتاً وابسته به گلژی است (Dhamija & Chambers, 2016). ژن *COG6*[3] روی کروموزوم ۱۶ قرار داشت و یک ژن کدکننده پروتئین است. این ژن زیرواحد کمپلکس الیگومری حفاظت‌شده گلژی را کد می‌کند که برای حفظ ساختار طبیعی و فعالیت دستگاه گلژی مورد نیاز است. پروتئین کدگذاری شده با اجزای ۵، ۷ و ۸ کمپلکس الیگومری حفاظت‌شده گلژی در یک مجموعه فرعی به نام لوب B سازماندهی شده‌است. پیوند جایگزین منجر به انواع رونوشت‌های متعدد می‌شود (Li *et al*., 2019). مجموعه ژن‌های SEM1، SETMAR، ITPR1 در فرآیندهای پاسخ به استرس نقش دارند و روی کروموزوم ۱۷ قرار گرفتند. این مجموعه با حذف پروتئین‌های تاخورده یا آسیب‌دیده، که می‌تواند عملکرد سلولی را مختل کند، و با حذف پروتئین‌هایی که دیگر به عملکرد آن‌ها نیازی نیست، نقش کلیدی در حفظ هموستاز پروتئین ایفا می‌کند (Jani *et al*., 2012). همچنین مجموعه ژن‌های SEM1، SETMAR، BHLHE40، CRBN، SUMF1، TRNT1 روی کروموزوم ۱۷ قرار دارند و در سوخت و ساز ماکرومولکول‌ها نقش دارند. ژن BHLHE40 یکی از مهم‌ترین ژن‌های این گروه است. این ژن یک پروتئین پایه مارپیچ–حلقه–مارپیچ را کد می‌کند که در بافت‌های مختلف بیان می‌شود. پروتئین کدگذاری شده می‌تواند با ARNTL تعامل داشته باشد یا برای مکان‌های اتصال E-box در پروموتر PER1 رقابت کند و فعال‌سازی CLOCK/ARNTL PER1 را سرکوب کند. اعتقاد بر این است که این ژن در کنترل ریتم شبانه روزی و تمایز سلولی نقش دارد (Kawamoto *et al*., 2004). ژن LXN و RARRES1 روی کروموزوم ۱۰ قرار داشتند. این دو ژن پارالوگ یکدیگرند و تنها بازدارنده پروتئین شناخته‌شده متالوکاربوکسی پپتیدازهای وابسته به روی را کد می‌کنند.

پروتئین کدگذاری شده، لاتکسین، اندازه جمعیت سلول‌های بنیادی خونساز را کاهش می‌دهد. این پروتئین در سلول‌های سرطانی به‌دلیل هیپرمتیلاسیون پروموتر کاهش می‌یابد. همچنین ممکن است در التهاب نقش داشته باشد. ژن PTX3در تنظیم مقاومت ذاتی در برابر عوامل بیماری‌زا، واکنش‌های التهابی، باروری نقش دارد (Lui *et al*., 2023). مجموعه ۱۱ ژن ACCS، ALKBH3، AMBRA1، ATG13، CHST1، F2، DGKZ، EXT2، HSD17B12، MAPK8IP1، SLC35C1 روی کروموزوم ۱۲ قرار داشتند و در فرآیندهای بیوسنتز (ACCS، ALKBH3)، در تنظیم رونویسی (AMBRA1)، در تنظیم فعالیت کیناز و تکثیر سلولی (ATG13)، ایمنی سلولی (CHST1، F2)، نقش بسیار مهم در تولید مجموعه‌های چربی (DGKZ)، نقش مهم در هموستاز و سیگنال‌دهی بافت (EXT2)، کاتالیز تبدیل استرون (E1) به استرادیول (E2) و تشکیل استروژن (HSD17B12)، سیگنال‌دهی سلولی در پایانه‌های عصبی (MAPK8IP1)، تامین انرژی (SLC35C1) نقش دارند (Dhamija & Chambers, 2016; Dominissini *et al*., 2016; Duncan *et al*., 2002). ژن‌های CDC42EP3، NDUFAF7، CYP1B1، EIF2AK2، SPAST، همگی روی کروموزوم ۱۵ قرار گرفته‌اند. یکی از مهم‌ترین ژن‌های این مجموعه CYP1B1 است که در متابولیسم اکسیداتیو، به‌ویژه تبدیل هیدروکربن‌های آروماتیک چند حلقه‌ای و آریل آمین‌های هتروسیکلیک پروسرسینوژن‌ها به محصولات آسیب‌رسان به DNA نقش دارد. احتمالاً با متابولیزه کردن محصولات اکسیژن دار تولید شده در طول استرس اکسیداتیو نقش مهمی در رشد عروق شبکیه دارد. تحت شرایط هیپراکسی، رگ‌زایی شبکیه و مورفوژنز مویرگی را افزایش می‌دهد. همچنین از طریق تعدیل بیان (بر اساس شباهت) به هموستاز اکسیداتیو و سازماندهی فراساختاری و عملکرد بافت شبکه ترابکولار کمک می‌کند (Shimada *et al*., 1999). ژن‌های MCTS1، UPF3B، LAMP2، RPL39 روی کروموزوم ۳۲ قرار داشتند و در فرآیند متابولیک پپتیدها نقش داشتند. همچنین پروتئین کدگذاری شده توسط ژن LAMP2 عضوی از خانواده گلیکوپروتئین‌های غشایی است. ممکن است در متاستاز سلول‌های تومور نقش داشته باشد. همچنین ممکن است در محافظت، نگهداری و چسبندگی لیزوزوم عمل کند. پیوند جایگزین این ژن منجر به انواع رونوشت‌های متعددی می‌شود که پروتئین‌های متمایز را کد می‌کند (Cuervo & Dice, 1996). ژن FBXL17 برای توانایی از بین بردن دایمرهای آسیب‌دیده برای تمایز و بقای تاج عصبی و سلول‌های عصبی مورد نیاز است. این ژن روی کروموزوم ۱۴ قرار داشت (Mena *et al*., 2018).

نتیجه‌گیری کلی

تحقیق حاضر با در نظر گرفتن روش‌های متنوع به‌منظور شناسایی نشانه‌های انتخاب در اسب‌های ترکمن و تروبرد و ترکیب نتایج حاصل از هر یک از روش‌ها، رویکرد جدیدی را برای استنتاج نقاط مثبت نشانه‌های انتخاب در ژنوم اسب اجرا می‌کند. نتایج این تحقیق می‌تواند منبع اطلاعاتی با ارزش و مفیدی به‌منظور شناسایی ژن‌های متمایزکننده و شناسایی نواحی ژنومی کاندید برای بسیاری از صفات مهم اقتصادی موجود در اسب‌های تروبرد و ترکمن را فراهم آورد.

3- Component Of Oligomeric Golgi Complex 6

[Downloaded from rap.sanru.ac.ir on 2026-10-23]

[DOI: 10.61882/rap.16.1.121]

References

Barendse, W., Harrison, B. E., Bunch, R. J., Thomas, M. B., & Turner, L. B. (2009). Genome wide signatures of positive selection: the comparison of independent samples and the identification of regions associated to traits. *BMC Genomics*, *10*, 1-15.

Cantor ,R. M., Lange, K., & Sinsheimer, J. S. (2010). Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *The American Journal of Human Genetics*, *86*(1), 6-22.

Cuervo, A. M., & Dice, J. F. (1996). A receptor for the selective uptake and degradation of proteins by lysosomes. *Science*, *273*(5274), 501-503.

Dhamija, R., & Chambers, C. (2016). Clinical and Molecular Characterization of ALG1-CDG. *Pediatric Neurology Briefs*, 14-14.

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., Clark, W. C. (2016). The dynamic N 1-methyladenosine methylome in eukaryotic messenger RNA. *Nature*, *530*(7591), 441-446.

Duncan, T., Trewick, S. C., Koivisto, P., Bates, P. A., Lindahl, T & ,.Sedgwick, B. (2002). Reversal of DNA alkylation damage by two human dioxygenases. *Proceedings of the National Academy of Sciences*, *99*(26), 16660-16665.

Freeze, H. H., Eklund, E. A., Ng, B. G., & Patterson, M. C. (2015). Neurological aspects of human glycosylation disorders. *Annual Review of Neuroscience*, *38*, 105-125.

Gondro, C., Van der Werf, J., & Hayes, B. (2013). *Genome-Wide Association Studies and Genomic Prediction*, (Vol. 1019). Springer.

Grimberg, J., Nawoschik, S., Belluscio, L., McKee, R., Turck, A., & Eisenberg, A. (1989). A simple and efficient non-organic procedure for the isolation of genomic DNA from blood. *Nucleic Acids Research*, *17*(20), 8390.

Grossman, S. R., Shylakhter, I., Karlsson, E. K., Byrne, E. H., Morales, S., Frieden, G ,. Zuk, O. (2010). A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*, *327*(5967), 883-886.

Gurgul, A., Jasielczuk, I., Semik-Gurgul, E., Pawlina-Tyszko, K., Stefaniuk-Szmukier, M., Szmatoła, T., Bugno-Poniewierska, M. (2019). A genome-wide scan for diversifying selection signatures in selected horse breeds. *PloS one*, *14*(1), e0210751.

Higgins, J. P., Little, J., Ioannidis, J. P., Bray, M. S., Manolio, T. A., Smeeth, L., ... & Khoury, M. J. (2007). Turning the pump handle: evolving methods for integrating the evidence on gene-disease association. *American Journal of Epidemiology*, 166(8), 863-866.

Jani, D., Lutz, S., Hurt, E., Laskey, R. A., Stewart, M., & Wickramasinghe, V. O. (2012). Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic Acids Research*, *40*(10), 4562-4573 .

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070-3071.

Kawamoto, T., Noshiro, M., Sato, F., Maemura, K., Takeda, N., Nagai, R., Miyazaki, K. (2004). A novel autofeedback loop of Dec1 transcription involved in circadian rhythm regulation. *Biochemical and Biophysical Research Communications*, *313*(1), 117-124.

Lean, I., Rabiee, A., Duffield, T. F., & Dohoo, I. (2009). Invited review: Use of meta-analysis in animal health and reproduction: Methods and applications. *Journal of Dairy Science*, *92*(8), 3545-3565.

Li, G., Xu, Y., Hu, X., Li, N., Yao, R., Yu, T., Wang, J. (2019). Compound heterozygous variants of the COG6 gene in a Chinese patient with deficiency of subunit 6 of the conserved oligomeric Golgi complex (COG6-CDG). *European Journal of Medical Genetics*, *62*(1), 44-46.

Lui, K., Huang, Y., Sheikh, M. S., Cheung, K.-K., Tam, W. Y., Sun, K.-T., Loh, A. W.-K. (2023). The oncogenic potential of Rab-like protein 1A (RBEL1A) GTPase: The first review of RBEL1A research with future research directions and challenges. *Journal of Cancer*, *14*(17), 3214.

Mena, E. L., Kjolby, R. A., Saxton, R. A., Werner, A., Lew, B. G., Boyle, J. M., Rape, M. (2018). Dimerization quality control ensures neuronal development and survival. *Science*, *362*(6411), eaap8236.

Moazemi, I., Mohammadabadi, M., Mostafavi, A., Esmailizadeh, A., Babenko, O., Bushtruk, M., Klopenko, N. (2020). Polymorphism of DMRT3 gene and its association with body measurements in horse breeds. *Russian Journal of Genetics*, *56*, 1232-1240.

Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K. G., & McEwan, J. C. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics*, *13*, 1-15.

Mostafavi, A., Fozi, M. A., Koshkooieh, A. E., Mohammadabadi, M., Babenko, O. I., & Klopenko, N. I. (2019). Effect of LCORL gene polymorphism on body size traits in horse populations. *Acta Scientiarum. Animal Sciences*, *42*, e47483.

Oleksyk, T. K., Smith ,M. W., & O'Brien, S. J. (2010). Genome-wide scans for footprints of natural selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 185-205.

[Downloaded from rap.sanru.ac.ir on 2026-10-23]

[DOI: 10.61882/rap.16.1.121]

Péanne, R., De Lonlay, P., Foulquier, F., Kornak, U., Lefeber, D. J., Morava, E., Van Schaftingen, E. (2018). Congenital disorders of glycosylation (CDG): Quo vadis? *European Journal of Medical Genetics*, *61*(11), 643-663.

Qanbari, S., Gianola, D., Hayes, B., Schenkel, F., Miller, S., Moore, S., Simianer, H. (2011). Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC Genomics*, *12*, 1-12.

Qanbari, S., Pimentel, E., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A. R., & Simianer, H. (2010). A genome-wide scan for signatures of recent selection in Holstein cattle. *Animal Genetics*, *41*(4), 377-389.

Ramasamy, A., Mondry, A., Holmes, C. C., & Altman, D. G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Medicine*, *5*(9) ,e184.

Saadatabadi, L. M., Mohammadabadi, M., Nanaei, H. A., Ghanatsaman, Z. A., Stavetska, R. V., Kalashnyk, O., Kucher, D. M. (2023). Unraveling candidate genes related to heat tolerance and immune response traits in some native sheep using whole genome sequencing data. *Small Ruminant Research*, *225*, 107018.

Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., McDonald, G. J. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, *419*(6909), 832-837.

Shimada, T., Watanabe, J., Kawajiri, K., Sutter, T. R., Guengerich, F. P., Gillam, E. M., & Inoue, K. (1999). Catalytic properties of polymorphic human cytochrome P450 1B1 variants. *Carcinogenesis*, *20*(8), 1607, 14-16.

Sutton, A. J., Abrams, K. R., Jones, D. R., Sheldon, T. A., & Song, F. (2000). *Methods for Meta-analysis in Medical Research* (Vol. 348). Wiley Chichester.

Tang, K., Thornton, K. R., & Stoneking, M. (2007). A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology*, *5*(7), e171.

Teo, Y. Y., Fry, A. E., Clark, T. G., Tai, E., & Seielstad, M. (2007). On the usage of HWE for identifying genotyping errors. *Annals of Human Genetics*, *71*(5), 701-703.

Utsunomiya, Y. T., Pérez O'Brien, A. M., Sonstegard, T. S., Van Tassell, C. P., do Carmo, A. S., Meszaros, G., Garcia, J. F. (2013). Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods .*PloS One*, *8*(5), e64280.

Vesterinen, H., Sena, E., Egan, K., Hirst, T., Churolov, L., Currie, G., Macleod, M. (2014). Meta-analysis of data from animal studies: a practical guide. *Journal of Neuroscience Methods*, *221*, 92-102.

Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*, *4*(3), e72.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358-1370.

[Downloaded from rap.sanru.ac.ir on 2026-10-23]

[DOI: 10.61882/rap.16.1.121]