

Research Paper

Exploring the Caspian Horse Population Structure in Iranian and Worldwide Populations using Pedigree Data

Seyedeh Fatemeh Mousavi¹, Hamid-Reza Seyedabadi², Ali Javanrouh³, and
Mahmoud Amiri Roudbar⁴ 

- 1- Department of Animal Science, Faculty of Agriculture Engineering, University of Kurdistan, Sanandaj, Iran
2- Department of Biotechnology, Animal Science Research Institute of Iran, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran
3- Department of Animal Breeding and Genetic, Animal Science Research Institute of Iran, Agricultural Research, Education & Extension Organization (AREEO), Karaj, Iran
4- Department of Animal Science, Safiabad-Dezful Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEO), Dezful, Iran,
(Corresponding author: Mahmood.amiri225@gmail.com)

Received: 28 March, 2025

Revised: 05 July, 2025

Accepted: 15 August, 2025

Extended Abstract

Background: Investigating genetic diversity and the population structure using genealogical information and molecular markers can reveal patterns of genetic differences resulting from past breeding methods, geographic separation, and human-mediated selection. Understanding the population structure is crucial for implementing corrective, conservation, breeding, and management programs. One of the important tools for investigating the structure of populations is the analysis of pedigree information, which plays a pivotal role in identifying genetic diversity and tracking changes across generations. It also provides estimates of essential population parameters, including increasing homozygosity, decreasing heterozygosity, inbreeding, and effective population size. While numerous studies have explored the population structure of various horse breeds worldwide using pedigree information, research specific to Iranian horses has been limited to Iranian Asil (Iranian Arab) and Turkmen horses. Furthermore, the Caspian horse, recognized as one of Iran's significant native breeds, holds the distinction of being one of the world's oldest horse breeds. Surprisingly, despite its historical significance, no study has yet explored the population structure of this breed using pedigree information. In this research, this gap is addressed by examining the Caspian horse population structure based on two information sources: pedigree records from the Caspian Conservation Society (CCS) and the International Caspian Society (ICS).

Methods: This study examined data from 1034 and 2264 horses registered in CCS and ICS pedigrees. After correcting errors in the pedigrees, important demographic parameters were calculated for each of the information sources separately. Initially, general information were extracted from the pedigrees using CFC software. Furthermore, inbreeding coefficients were estimated using the AGHmatrix package in the R programming language for both populations, and their changes were investigated over different years. The effective population size (N_e) in both pedigrees was calculated based on the method of individual increase in inbreeding using purgeR. The N_e was estimated using two distinct scenarios. All individuals in the pedigree were included in the first case, while the second case considered only the horses that were currently alive. Subsequently, the number of founders (N_f) and ancestors (N_a), the number of founder genomes equivalents (N_g), and the deviation from Hardy-Weinberg were calculated for both pedigrees. The generation distance was calculated based on the average age of parents for four selected paths: 1) stallion to male foal, 2) stallion to female foal, 3) mare to male foal, and 4) mare to female foal. Finally, the pedigree completeness index (PCI) was calculated using the optiSel package in the R language for both pedigrees.

Results: The value of PCI calculated based on CCS and ICS information was above 70% until the 2nd and 4th generations, respectively, but its value showed a great decrease in the later generations. The rate of inbreeding observed for the whole population was estimated at 0.794% (for CCS) and 4.373% (for ICS), and a positive trend of changes was observed in inbreeding in both sources of information. The estimated observed inbreeding in the CCS pedigree was lower than that in the ICS pedigree. This discrepancy can be attributed to the incomplete genealogical information available in the CCS data source. Overall, the N_e calculated for both pedigrees was less than 100. When considering the entire pedigree, the N_e for the ICS pedigree (68.33 ± 13.03)



was nearly twice that estimated for the CCS pedigree (32.86 ± 0.88). The estimated N_e values for the CCS and ICS populations, considering only live horses, were 66.43 ± 14.39 and 32.40 ± 0.93 , respectively. Notably, the N_e estimates based on genealogical information aligned with those from previous studies that utilized genomic data. The values of N_a and N_f calculated in the two pedigrees showed a huge difference, which indicates the relatively strong influence of a population bottleneck in the past. Additionally, given the substantial differences observed in N_g (110.12 for CCS versus 13.25 for ICS) and other metrics, such as N_f (333 for CCS and 155 for ICS) and N_a (516 for CCS and 996 for ICS), it is evident that genetic drift has significantly impacted reducing population diversity. The estimated deviation from the Hardy-Weinberg equilibrium was observed to be positive in two pedigrees, which indicates the presence of a small amount of positive mating between relatives in both pedigrees.

Conclusion: Although the origin of this breed is in Iran, the registered genealogy was of lower quality than its international pedigree, indicating a serious need for more accurate genealogy registration in Iran. According to N_e 's estimates, the Caspian horse population is in an endangered condition, and a specific program should be designed to conserve the genetic diversity of this breed. Despite the smaller number of horses registered in the CCS pedigree, the study of the reference population structure reveals that the genetic diversity is superior within the Iranian population. Consequently, this genetic resource holds promise for enhancing the overall condition of the Caspian horse worldwide. Moreover, the presence of inbreeding can worsen the existing situation; thus, a controlled breeding program to reduce inbreeding for both populations can be considered a solution.

Keywords: Caspian horse, Pedigree analysis, Population structure, Inbreeding

How to Cite This Article: Mousavi, S. F., Seyedabadi, H. R., Javanrouh, A., & Amiri Roudbar, R. (2025). Exploring the Caspian Horse Population Structure in Iranian and Worldwide Populations using Pedigree Data. *Res Anim Prod*, 16(4), 9-17. DOI: 10.61882/rap.2025.1468

مقاله پژوهشی

بررسی ساختار جمعیت اسب کاسپین در جمعیت‌های ایران و جهان با استفاده از اطلاعات شجره

سیده فاطمه موسوی^۱، حمیدرضا سید آبادی^۲، علی جوانروح^۳ و محمود امیری رودبار^۴ ^۱ - گروه علوم دامی، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران^۲ - بخش بیوتکنولوژی، موسسه تحقیقات علوم دامی کشور، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران^۳ - بخش ژنتیک و اصلاح نژاد دام و طیور، موسسه تحقیقات علوم دامی کشور، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران^۴ - بخش علوم دامی، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی صفی‌آباد-دزفول، سازمان تحقیقات، آموزش و ترویج کشاورزی، دزفول، ایران، (نویسنده مسوول: Mahmood.amiri225@gmail.com)

تاریخ پذیرش: ۱۴۰۴/۰۵/۲۴

تاریخ ویرایش: ۱۴۰۴/۰۴/۱۴
صفحه ۹ تا ۱۷

تاریخ دریافت: ۱۴۰۴/۰۱/۰۸

چکیده مبسوط

مقدمه و هدف: بررسی تنوع ژنتیکی و ساختار جمعیتی توسط اطلاعات شجره‌ای و نشانگرهای مولکولی می‌تواند الگوهای تفاوت ژنتیکی را که توسط روش‌های پرورشی در گذشته، جدایی جغرافیایی و انتخاب با واسطه انسان ایجاد شده‌اند را نشان دهد. اطلاعات از ساختار جمعیتی می‌تواند برای اجرای برنامه‌های اصلاحی، حفاظتی، پرورشی و مدیریتی مورد استفاده قرار گیرد. یکی از ابزارهای مهم جهت بررسی ساختار جمعیت‌ها تجزیه و تحلیل اطلاعات شجره است، که برای شناسایی تنوع ژنتیکی و تغییرات ژنتیکی منتقل شده از نسلی به نسل دیگر مورد استفاده قرار می‌گیرد. همچنین، تحلیل شجره قادر است تا برآوردی را از پارامترهای مهم جمعیت مانند افزایش هموزیگوسیتی و کاهش هتروزیگوسیتی، همخوانی و اندازه مؤثر جمعیت ارائه دهد. با وجود مطالعات زیادی که در خصوص ساختار جمعیت اسب در دنیا با کمک اطلاعات شجره صورت گرفته‌اند، اما گزارشات در این خصوص برای اسب‌های ایرانی تنها محدود به اسب اصیل ایرانی (عرب ایرانی) و ترکمن هستند. علاوه بر این، اسب کاسپین یکی از مهم‌ترین نژادهای بومی ایران است و به‌عنوان یکی از قدیمی‌ترین نژادهای اسب در جهان شناخته می‌شود، اما مطالعه‌ای در خصوص ساختار جمعیت این نژاد با کمک اطلاعات شجره صورت نگرفته است. بنا بر این، در این پژوهش ساختار جمعیت اسب کاسپین با کمک دو منبع اطلاعاتی شامل شجره انجمن حفاظت از اسب کاسپین (CCS) و انجمن جهانی کاسپین (ICS) بررسی شد.

مواد و روش‌ها: در این پژوهش، از اطلاعات ۱۰۳۴ و ۲۲۶۴ راس اسب ثبت‌شده در شجره‌های CCS و ICS استفاده شد. پس از تصحیح خطاهای موجود در شجره‌ها، پارامترهای مهم جمعیتی برای هر یک از منابع اطلاعاتی به‌صورت جداگانه محاسبه گردیدند. ابتدا با نرم‌افزار CFC اطلاعات کلی از ساختار شجره‌های مورد مطالعه استخراج شد. همچنین، ضریب همخوانی به کمک بسته نرم‌افزاری AGHmatrix در زبان برنامه‌نویسی R در هر دو جمعیت برآورد شد و روند تغییرات آن طی سال‌های مختلف بررسی گردید. اندازه مؤثر جمعیت (N_e) در هر دو شجره براساس روش افزایش فردی در همخوانی با کمک purgeR محاسبه گردید. برآورد N_e بر اساس دو حالت مختلف شامل حالت اول استفاده از تمام افراد موجود در شجره، و حالت دوم استفاده از تنها اسب‌هایی که زنده بودند، انجام شد. در ادامه به‌وسیله همین بسته نرم‌افزاری تعداد افراد پایه‌گذار (N_f) و اجداد (N_a)، تعداد معادل‌های ژنوم پایه‌گذار (N_g) و همچنین میزان انحراف از هاردی-واینبرگ برای هر دو شجره محاسبه شدند. فاصله نسل براساس میانگین سن والدین برای چهار مسیر انتخابی شامل: (۱) سیلیمی به کره، (۲) سیلیمی به کره ماده، (۳) مادریان به کره، و (۴) مادریان به کره ماده محاسبه شد. در نهایت، آماره شاخص کامل بودن شجره (PCI) به وسیله بسته نرم‌افزاری optiSel در زبان R برای هر دو شجره محاسبه شد.

یافته‌ها: مقدار PCI محاسبه‌شده براساس اطلاعات CCS و ICS به‌ترتیب تا نسل ۲ و ۴ بالای ۷۰٪ بود، اما در نسل‌های بعد مقدار آن کاهش بسیار زیادی نشان داد. مقادیر همخوانی مشاهده شده برای کل جمعیت برابر ۰/۷۹۴ (برای CCS) و ۴/۳۷۳ درصد (برای ICS) برآورد گردیدند و روند تغییرات همخوانی مشاهده شده در هر دو منبع اطلاعاتی مثبت بود. میزان همخوانی مشاهده شده در شجره CCS کمتر از ICS برآورد گردید که دلیل اصلی آن ناقص بودن بیشتر اطلاعات شجره‌ای در این منبع اطلاعاتی بود. به‌طور کلی، N_e محاسبه‌شده برای هر دو شجره کمتر از ۱۰۰ بود. همچنین، N_e در زمانی که از کل شجره استفاده گردید برای شجره ICS ($68/33 \pm 13/03$) تقریباً دو برابر N_e برآورد شده برای شجره CCS ($32/86 \pm 0/88$) بود. مقادیر N_e برآوردشده در حالتی که از اسب‌های زنده استفاده گردید، برای جمعیت‌های CCS و ICS به‌ترتیب $66/43 \pm 14/39$ و $32/40 \pm 0/93$ برآورد شدند. در کل، مقادیر N_e برآوردشده با کمک اطلاعات شجره‌ای با برآوردهای انجام شده در مطالعات قبلی که با کمک اطلاعات ژنومی انجام شده بود، همخوانی نشان دادند. مقادیر N_f و N_a محاسبه شده در دو شجره اختلاف بسیار زیادی را نشان دادند که نشان‌دهنده تأثیر نسبتاً زیاد تنگه‌های جمعیتی در گذشته است. همچنین، با توجه به اختلاف بالای مشاهده شده بین N_g (۱۱۰/۱۲) برای CCS و ۱۳/۲۵ برای ICS) و دیگر آماره‌ها شامل N_f (۳۳۳ برای CCS و ۱۵۵ برای ICS) و N_a (۵۱۶ برای CCS و ۹۹۶ برای ICS)، میزان تأثیر رانش ژنتیکی بر کاهش تنوع جمعیت نیز قابل‌توجه بود. مقدار انحراف برآوردشده از تعادل هاردی-واینبرگ در دو شجره مثبت بود، که نشان از وجود مقدار کمی آمیزش جور شده مثبت بین افراد خویشاوند در هر دو شجره است.

نتیجه‌گیری: با وجود این که خاستگاه این نژاد در ایران است ولی شجره ثبتی دارای کیفیت کمتری نسبت به شجره جهانی آن است و نیاز جدی برای ثبت دقیق‌تر شجره در ایران را نشان می‌دهد. با توجه به برآوردهای N_e می‌توان گفت که جمعیت اسب کاسپین در شرایط بحرانی است و باید در خصوص حفظ تنوع ژنتیکی این نژاد برنامه مشخصی طراحی گردد. نتایج بررسی ساختار جمعیت مرجع نشان می‌دهند که با وجود تعداد کمتر اسب‌های ثبت شده در شجره CCS، وضعیت تنوع ژنتیکی در جمعیت ایران بهتر است و می‌توان از این منبع ژنتیکی برای بهبود شرایط اسب کاسپین در کل دنیا بهره جست. همچنین، وجود آمیزش‌های بین خویشاوندی می‌تواند وضعیت موجود را بدتر کند که انجام یک برنامه تلاقی کنترل‌شده برای کاهش آمیزش‌های خویشاوندی برای هر دو جمعیت می‌تواند به‌عنوان یک راه حل در نظر گرفته شود.

واژه‌های کلیدی: اسب کاسپین، آنالیز شجره، ساختار جمعیت، همخوانی

مقدمه

نشان‌دهنده روابط پیچیده بین افراد درون و بین نژادها است. مطالعات زیادی تاکنون برای بررسی تنوع ژنتیکی و ساختار جمعیتی نژادهای اسب توسط اطلاعات شجره‌ای و نشانگرهای

ساختار جمعیتی در اسب‌ها نتیجه تلاقی پیچیده‌ای از تنوع ژنتیکی، تاریخ جمعیتی و روش‌های پرورش انتخابی است که

(عرب ایرانی) با کمک اطلاعات شجره، تنها میزان همخونی در این نژاد بررسی گردید (Gharahveysi & Irani, 2011). میانگین ضریب همخونی کل جمعیت و همچنین میانگین ضریب همخونی در اسب‌های همخون برای این نژاد به ترتیب ۲/۱ و ۷/۸ درصد برآورد گردید. در مطالعه دیگر در اسب ترکمن (Saghi & Mobaraki, 2018)، ساختار جمعیت این اسب با کمک اطلاعات بیش از ۲۰۰۰ اسب که طی سال‌های ۱۹۷۰ تا ۲۰۱۶ متولد شده بودند مورد بررسی قرار گرفت. نتایج این تحقیق نشان دادند که میانگین ضریب همخونی در جمعیت اسب‌های اصیل ترکمن ایران (۰/۰۶٪) نسبت به مقادیر منتشر شده در نژادهای دیگر بسیار پایین‌تر بود که به احتمال زیاد به دلیل عدم وجود اطلاعات مشخص از تعدادی والدین حیوانات و در نتیجه ناقص بودن شجره است. با وجود اهمیت بسیار بالای نژاد اسب کاسپین از منظر ذخایر ژنتیکی، تاکنون گزارشی در خصوص بررسی ساختار جمعیت اسب کاسپین با کمک اطلاعات شجره ارائه نشده است. بنا بر این، هدف این پژوهش بررسی ساختار جمعیت اسب کاسپین با کمک اطلاعات شجره‌ای بود.

مواد و روش‌ها

شبیه‌سازی جمعیت و ژنوم

در کل، از دو منبع شجره اسب کاسپین شامل انجمن حفاظت از اسب کاسپین (Caspian conservation society (CCS)) و همچنین کتاب کاسپین بین‌المللی (the international Caspian studbook) ارائه شده توسط انجمن جهانی کاسپین (international Caspian society (ICS)) استفاده گردید. تعداد ۱۰۳۴ راس اسب در شجره CCS ثبت شده بود، که ۶۰۵ راس ماده و ۴۲۷ راس نر بودند. همچنین، وجود خطاهایی مانند ثبت اشتباهی پدر به عنوان مادر و بالعکس (۳ مورد خطا)، عدم تطابق سن پدر (۱۵ مورد خطا) و مادر (۸ مورد خطا) با در نظر گرفتن حداقل و حداکثر سن ۳ و ۳۰ سال در زمان تولد کره‌ها و فواصل زایش مادر حداقل ۱۱ ماه (۲۹ مورد) بررسی شد. پس از تصحیح اطلاعات شجره CCS، از این تغییرات برای تصحیح اطلاعات شجره ICS نیز استفاده گردید. در کل، برای شجره ICS تعداد ۱۲۱۶ راس نر و ۱۰۴۸ راس دیگر ماده بودند. برای بررسی ساختار جمعیتی بر اساس اطلاعات شجره، اطلاعات ساختار شجره با نرم‌افزار CFC نسخه ۱.۰ استخراج شد (Sargolzaei et al., 2006).

برآورد ضریب همخونی و روند آن طی سال‌های مختلف

همخونی به عنوان احتمال این که یک فرد دارای دو آلل یکسان به دلیل جد مشترک باشد، تعریف می‌شود. روش محاسبه براساس تجزیه ماتریکس روابط ژنتیکی افزایشی (A) که توسط (Meuwissen & Luo, 1992) توضیح داده شده است انجام شد. پس از برآورد همخونی در تمام افراد، روند همخونی در سال‌های مختلف در جمعیت اسب کاسپین ایران و جهان بررسی شد. از بسته نرم‌افزاری AGHmatrix برای محاسبه میزان همخونی در بین افراد استفاده شد (Amadeu et al., 2023).

برآورد اندازه مؤثر جمعیت

مولکولی (مانند SNP و STR) انجام شده‌اند، که الگوهای تفاوت ژنتیکی که توسط روش‌های اصلاحی در گذشته، جدایی جغرافیایی و انتخاب شکل گرفته با واسطه انسان را نشان می‌دهند (da Silva Faria et al., 2018; MacCluer et al., 1983; Machmoum et al., 2020; Mousavi et al., 2023; Petersen et al., 2013). این تغییرات در ساختار ژنتیکی جمعیت‌های می‌توانند از طریق تغییر در محتوای ژنتیکی جمعیت موجب تغییرات در بروز صفات عملکردی شوند (Ghafouri-Kesbi, 2023). درک ساختار جمعیتی برای حفاظت، برنامه‌های اصلاحی و مدیریت ژنتیکی حیاتی است، که حفظ تنوع ژنتیکی درون‌نژادی و شناسایی افراد یا جمعیت‌هایی با اهمیت ژنتیکی ویژه را ممکن می‌سازد. اسب کاسپین به عنوان جد اولیه همه اسب‌های خون‌گرم دنیا نیز معرفی شده است که با اسب‌های عرب تشابهات زیادی دارد (Firouz, 1969). بررسی ساختار جمعیتی اسب کاسپین می‌تواند به درک بهتر این نژاد کمک کند و برای مطالعات آینده مورد استفاده قرار گیرد.

واکاوی اطلاعات شجره ابزار مهمی برای شناسایی تنوع ژنتیکی و تغییراتی است که از نسلی به نسل دیگر انتقال می‌یابند (Bokor et al., 2013). تنوع ژنتیکی در یک جمعیت را می‌توان برحسب تنوع آلل‌ها و ژنوتیپ‌های فعلی تعریف کرد (Boichard et al., 1997). علاوه بر این، نتایج واکاوی شجره می‌تواند مقادیر مورد انتظار در افزایش هموزیگوسیتی و کاهش هتروزیگوسیتی را برای کل ژنوم برآورد نمایند (Falconer & Mackay, 1996). این اطلاعات می‌توانند در ارائه راهبردهای مناسب برای مدیریت برنامه‌های جفت‌گیری و همچنین نشان دادن تنوع ژنتیکی مورد استفاده قرار گیرند (Valera et al., 2005). افزون بر این، تجزیه و تحلیل شجره ممکن است ارزیابی ساختار یک جمعیت و سطوح همخونی را ممکن سازد (Gutiérrez & Goyache, 2005)، که برای جمعیت‌های تحت فشار انتخاب بالا و همچنین جمعیت‌های بسته برای سال‌های طولانی مهم هستند. تاکنون، مطالعات زیادی در مورد ساختار جمعیت و تنوع ژنتیکی نژادهای اسب در دنیا انجام شده‌اند (Faria et al., 2021; Głazewska & Jezierski, 2004; Klein et al., 2022; Medeiros et al., 2014; Zechner et al., 2002). برای مثال، پژوهشی که بر روی اسب لیپیزان (Lipizzan) در هشت گله اروپایی صورت گرفت، نشان داد که تمام اسب‌های این نژاد مربوط به چند اسب در گذشته بودند و عمق شجره این نژاد نیز ۳۲ نسل برآورد گردید (Zechner et al., 2002). میانگین همخونی برآورد شده برای این نژاد از ۸/۶ تا ۱۴/۴٪ متفاوت بودند. در مطالعه دیگری که بر روی اسب عرب لهستانی (Polish Arabian) با اطلاعات ۵۰ سال شجره انجام شد نشان داد که این نژاد نتیجه ۲۱۳ اسب بنیان‌گذار بود و یک کاهش در تنوع ژنتیکی در طول زمان در این نژاد مشاهده گردید (Głazewska & Jezierski, 2004).

با وجود مطالعات زیاد انجام شده برای بررسی ساختار شجره اسب در دنیا، اطلاعات در خصوص ساختار جمعیت اسبان نژادهای ایرانی با کمک بررسی ساختار شجره بسیار محدود است. برای مثال، در مطالعه انجام شده برای اسب اصیل ایرانی

در بررسی جمعیت‌ها، دانستن میزان آمیزش‌های غیرتصادفی می‌تواند مفید باشد. این پارامتر به‌وسیله انحراف از تعادل هاردی-واینبرگ (Hardy-Weinberg equilibrium) قابل اندازه‌گیری خواهد بود، که از فرمول زیر محاسبه می‌شود:

$$\alpha = \frac{F-f}{1.0-f} \quad (۳)$$

در این معادله، F میانگین همخوانی محاسبه‌شده برای جمعیت، و f برابر با میانگین ضریب همبستگی (coancestry coefficient) هستند. مقادیر منفی نشان‌دهنده وجود مدیریت در آمیزش‌های افراد غیرخویشاوند در جمعیت مورد مطالعه هستند. مقادیر مثبت و صفر نیز به ترتیب آمیزش‌های جورشده مثبت بین افراد خویشاوند و آمیزش‌های تصادفی در جمعیت را نشان می‌دهند. محاسبه این پارامتر توسط بسته نرم‌افزاری purgeR در زبان برنامه‌نویسی R انجام شد (López-Cortegano, 2022).

برآورد فاصله نسل

فاصله نسل براساس میانگین سن والدین در زمانی که فرزندان انتخاب‌شده آن‌ها (کره‌هایی که برای تولید نسل بعد استفاده شده‌اند) به دنیا آمدند، محاسبه شد. فاصله نسل برای چهار مسیر انتخابی شامل: (۱) سیلیمی به کره نر، (۲) سیلیمی به کره ماده، (۳) مادپان به کره نر، و (۴) مادپان به کره ماده محاسبه شد. محاسبه این آماره‌ها با زبان برنامه‌نویسی R انجام شد.

شاخص کامل بودن شجره

آماره شاخص کامل بودن شجره (pedigree completeness index (PCI) معیار مهمی است که برای بررسی کیفیت شجره مورد استفاده قرار می‌گیرد، نشان‌دهنده درصد اجداد مشخص‌شده در هر نسل به‌صورت مجزا است و میزان احتمال شناسایی همخوانی را در شجره مشخص می‌نماید (Sørensen *et al.*, 2005). برای محاسبه این آماره از معادله زیر استفاده گردید:

$$I_d = \frac{4I_{dpat}I_{dmat}}{I_{dpat} + I_{dmat}} \quad (۴)$$

که در آن، I_{d_k} ($k = pat, mat$) برابر است با:

$$I_{d_k} = 1/a \sum_{i=1}^d a_i \quad (۵)$$

در اینجا، k نشان‌دهنده لاین پدری یا مادری هر فرد، و a_i درصد اجداد شناسایی شده در نسل i است. همچنین، d تعداد کل نسل‌هایی هستند که برای محاسبه کامل بودن شجره در نظر گرفته شده‌اند. میزان PIC از ۱ تا ۰ متغیر است. اگر تمام اجداد یک فرد برای یک نسل خاص (d) شناخته شده باشد در این صورت $I_d = 1$ است، یا اگر یکی از والدین نامشخص باشد، در این صورت هیچ یک از اجداد شناسایی نمی‌شوند و $I_d = 0$ خواهد بود (MacCluer *et al.*, 1983). برای محاسبه PCI از بسته نرم‌افزاری optiSel در زبان R استفاده گردید (Wellmann, 2019).

نتایج و بحث

ساختار کلی شجره و همچنین وضعیت همخوانی جمعیت در جدول ۱ نشان داده شده‌اند.

اندازه موثر جمعیت (N_e) را می‌توان براساس افزایش فردی در همخوانی (ΔF) محاسبه کرد (Gutiérrez *et al.*, 2008) که ΔF برابر با:

$$N_e = \frac{1}{2\Delta F} \quad (۱)$$

است. برای محاسبه ΔF در هر فرد می‌توان از فرمول زیر:

$$\Delta F_i = 1 - \sqrt[t_i-1]{1 - F_i} \quad (۲)$$

استفاده نمود، که در آن F_i ضریب همخوانی فرد i و t_i شماره نسلی است که فرد به آن تعلق دارد. برای برآورد اندازه موثر جمعیت از بسته نرم‌افزاری purgeR در زبان برنامه‌نویسی R استفاده گردید (López-Cortegano, 2022).

برآورد تعداد اجداد و پایه‌گذار

تعداد افراد پایه‌گذار (N_f) (number of founders) و اجداد (N_a) (number of ancestors) و همچنین تعداد معادل‌های ژنوم پایه‌گذار (number of founder genomes) (N_g) (equivalents) برای هر دو شجره محاسبه گردیدند. این پارامترها می‌توانند اطلاعاتی در خصوص تنگناهای جمعیتی (population bottlenecks) به دلیل مشارکت نامتعادل پایه‌گذار یا اجداد و همچنین رانش جمعیتی فراهم آورند. این برآوردها براساس احتمال محاسبات منشا ژنی هستند و اطلاعات جمعیت پایه را نشان می‌دهند (Boichard *et al.*, 1997; Caballero & Toro, 2000; Tahmoorespur & Sheikhloo, 2011). N_f از طریق شمارش تعداد پایه‌گذار جمعیت مرجع محاسبه شد (تعداد افراد بدون اطلاعات پدر و مادر)، اما تعداد موثر افراد پایه‌گذار N_{fe} برابر با تعداد پایه‌گذار مشارکت‌کننده به‌صورت برابر بود که براساس تنوع ژنتیکی مشاهده شده در جمعیت مرجع تصحیح شدند. N_a برابر با تعداد کل اجداد شجره است که برای تشکیل جمعیت مرجع (که هم می‌تواند پایه‌گذار باشد یا خیر) مشارکت داشته است. در مقابل، تعداد مؤثر اجداد N_{ae} برابر با تعداد حداقل اجداد (که هم می‌تواند پایه‌گذار باشد یا خیر) مورد نیاز برای تشکیل تنوع ژنتیکی مشاهده شده در جمعیت مرجع است. N_g مشابه با N_{fe} است، اما برای محاسبه آن از شبیه‌سازی تفرق آلی با روش Monte Carlo استفاده شد که به‌صورت مؤثری نه‌تنها کاهش تنوع ژنتیکی به دلیل تنگناهای جمعیتی را در نظر می‌گیرد بلکه کاهش تنوع به دلیل عوامل تصادفی مانند رانش را نیز منظور می‌کند (Boichard *et al.*, 1997; Caballero & Toro, 2000). همیشه مقدار N_{ae} از N_f کوچکتر است و نسبت این دو میزان کاهش تنوع به دلیل تنگناهای جمعیتی بین جمعیت پایه و جمعیت مرجع را نشان می‌دهد (Tahmoorespur & Sheikhloo, 2011). در مقابل، N_g همواره کمترین مقدار را دارد که نه‌تنها نشان‌دهنده کاهش تنوع به دلیل مشارکت نامتعادل پایه‌گذارها یا اجداد است بلکه کاهش تنوع به دلیل رانش را نیز نشان می‌دهد. برای برآورد این پارامترهای جمعیت مرجع از بسته نرم‌افزاری purgeR در زبان برنامه‌نویسی R استفاده گردید (López-Cortegano, 2022).

انحراف از هاردی-واینبرگ

جدول ۱- پارامترهای محاسبه‌شده برای ساختار شجره‌ای جمعیت‌های اسب کاسپین براساس اطلاعات شجره حاصل از حفاظت از اسب کاسپین (CCS) و انجمن جهانی کاسپین (ICS).

Table 1. Calculated parameters for the pedigree structure of Caspian horse populations based on pedigree information from Caspian Horse Conservation (CCS) and International Caspian Society (ICS).

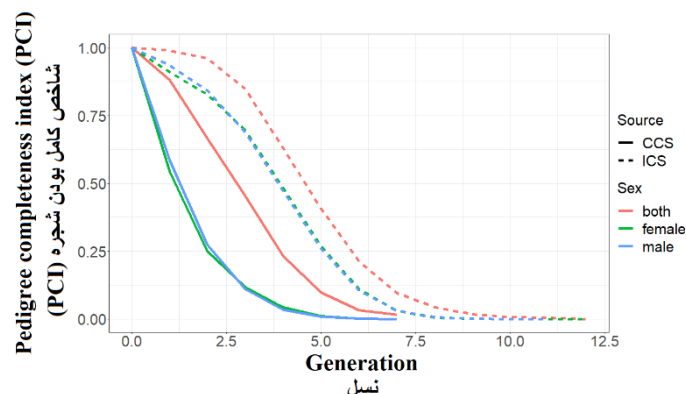
ICS	CCS	پارامتر
2264	1032	تعداد کل Total number
1500	89	تعداد کل افراد همخون Inbreds in total
4.373	0.794	میانگین درصد ضریب همخونی کل جمعیت Average inbreeding coefficients
6.601 (0.20-32.666)	9.202 (0.781-28.125)	میانگین درصد ضریب همخونی جمعیت همخون (دامنه) Average inbreeding coefficients in inbreds (range)
318 (2080)	122 (547)	تعداد کل پدر (تعداد کل فرزندان پدران) Total number of sires (total number of fathers' progeny)
640 (2093)	247 (611)	تعداد کل مادر (تعداد کل فرزندان مادران) Total number of mothers (total number of progeny of mothers)
1306	663	تعداد بدون فرزند Number without progeny
158	375	تعداد افراد پایه Number of founders
371	75	تعداد برادران و خواهران تنی Number of full-sib brothers and sisters
2.49 (2-6)	2.32 (2-5)	میانگین اندازه خانواده (دامنه) Average family size (range)
5.048 (0-12)	1.624 (0-7)	میانگین عمق شجره (دامنه) Average pedigree depth (range)
3.337 (0-6.991)	0.987 (0-4.03)	میانگین متوسط تعداد نسل (دامنه) Average number of generations (range)
2067	501	تعداد افراد با هر دو والد مشخص Number of individuals with known sire and dam
0.0779 (0.0774)	0.0100 (0.0110)	میانگین رابطه خویشاوندی (بدون در نظر گرفتن رابطه فرد با خود) Average relatedness (without considering self-relatedness)
9.19	9.25	فاصله نسل ^۱ کل Total generation interval
8.77	9.08	فاصله نسل مادر به دختر Generation interval dam-daughter
8.90	8.35	فاصله نسل مادر به پسر Generation interval dam-son
9.57	9.02	فاصله نسل پدر به پسر Generation interval sire-daughter
9.56	10.12	فاصله نسل پدر به دختر Generation interval sire-daughter

^۱ It is calculated based on the year.

^۱ بر اساس سال محاسبه شده‌است.

تا نسل هفتم بالای ۷۰٪ برآورد گردید، ولی پیش از این نسل کاهش زیادی نشان داده است (Duru, 2017) و یا در مقابل، برای عرب اسپانیایی میزان PCI تا نسل چهارم بالای ۷۴٪ گزارش شد (Cervantes *et al.*, 2008). برای این تحقیق، مقادیر PCI محاسبه شده بر اساس اطلاعات CCS و ICS به ترتیب تا نسل ۲ و ۴ بالای ۷۰٪ بودند، که پس از این نسل‌ها کاهش بسیار زیادی در PCI مشاهده گردید. همان طور که شاخص PCI نشان می‌دهد، میزان کیفیت شجره ICS نسبت به ICC بیشتر است و درصد بیشتری از اجداد در آن مشخص هستند. بنا بر این، با وجود این که خاستگاه این نژاد ایران است ولی شجره ثبتی دارای کیفیت کمتری نسبت به شجره جهانی آن است و نیاز جدی جهت ثبت دقیق‌تر شجره در ایران را نشان می‌دهد.

نتایج آنالیز میزان کیفیت شجره برای دو منبع اطلاعاتی در شکل ۱ نشان داده شده‌اند. به‌طور کلی، درجه کیفیت اطلاعات شجره بر برآورد پارامترهای مهمی همچون میانگین ضریب همخونی تأثیر دارد و عمق شجره نیز تأثیر مهمی بر تعداد مؤثر اجداد دارد. برای مثال، اگر شجره ثبت‌شده برای یک جمعیت ناقص باشد، مقادیر ضرایب همخونی تا حد زیادی کمتر از مقدار واقعی برآورد می‌شوند (Cassell *et al.*, 2003). نه تنها این دو پارامتر، بلکه تمام پارامترهایی که منشأ ژن احتمالی را توصیف می‌کنند نیز تحت تأثیر عمق شجره قرار می‌گیرند. برآورد تنوع ژنتیکی به مقدار بسیار زیادی به میزان در دسترس بودن اطلاعات شجره که توسط PCI نشان داده می‌شود، بستگی دارد. به‌طور کلی، میزان PCI برآوردشده و میزان کاهش آن طی نسل‌ها برای نژادهای مختلف متفاوت است. برای مثال، میزان کامل بودن شجره در نژاد عرب ترک

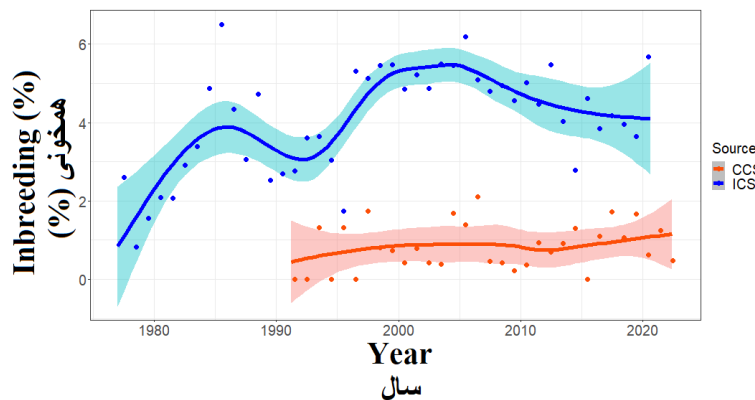


شکل ۱- محاسبه شاخص کامل بودن شجره (PCI) برای بررسی میزان کیفیت شجره جمعیت‌های اسب کاسپین بر اساس اطلاعات حاصل از حفاظت از اسب کاسپین (CCS) و انجمن جهانی کاسپین (ICS).

Figure 1. Calculation of the pedigree completeness index (PCI) to examine the pedigree quality of Caspian horse populations based on the information obtained from the Caspian Horse Conservation Society (CCS) and the Caspian World Society (ICS).

برای شجره ICS همخونی از زیر یک درصد (برای سال ۱۹۷۶) تا بالای پنج درصد (برای سال ۱۹۹۵) افزایش دارد اما دوباره به حدود چهار درصد (برای سال ۲۰۱۸) کاهش نشان می‌دهد. به‌طور کلی، میزان همخونی مشاهده شده در شجره CCS کمتر از ICS است که دلیل اصلی آن ناقص بودن بیشتر اطلاعات شجره‌ای در این منبع اطلاعاتی است به‌طوری‌که درصد بیشتری از پدرها و مادرها مشخص نبودند.

در ادامه، تغییرات میزان همخونی در سال‌های مختلف برای هر دو منبع اطلاعاتی مورد بررسی قرار گرفتند (شکل ۲). برای محاسبه روند تغییرات همخونی از اطلاعات سال‌هایی استفاده شد که حداقل ۱۰ زایش برای آن سال ثبت شده بود. همان‌طور که در شکل مشاهده می‌شود، روند تغییرات همخونی مشاهده شده در هر دو منبع اطلاعاتی مثبت است به‌طوری‌که در شجره CCS از زیر نیم درصد همخونی (برای سال ۱۹۹۰) به بالای یک درصد (برای سال ۲۰۲۱) افزایش دارد. همچنین،



شکل ۲- نمودار تغییرات میزان همخونی بر اساس سال تولد کره‌ها برای دو منبع اطلاعاتی مختلف شامل انجمن حفاظت از اسب کاسپین (CCS) و انجمن جهانی کاسپین (ICS) نشان داده شده است. نقطه‌های روی نمودار مربوط به میانگین درصد همخونی برآورد شده برای هر سال هستند.

Figure 2. A graph of changes in inbreeding rates by year of birth of foals is shown for two different data sources, the Caspian Horse Conservation Society (CCS) and the International Caspian Society (ICS). The points on the graph correspond to the average percentage of inbreeding estimated for each year.

وضعیت آن‌ها برآورد شد (جدول ۲). اندازه موثر جمعیت براساس دو حالت برآورد گردید. در حالت اول، تمام افراد موجود در شجره استفاده شد. اما چون تنها اسب‌هایی که در حال حاضر موجود هستند برای ایجاد نسل‌های بعد می‌توانند مورد استفاده قرارگیرند، در حالت دوم تنها اسب‌هایی که زنده بودند برای محاسبه استفاده شدند. در حالت دوم، برآورد اندازه موثر واقعی‌تر بود و تنها از اسب‌های نسل‌های انتهایی استفاده شد که به تازگی برای برآورد همخونی ناشی از اجدادشان استفاده شده بودند. به‌طور کلی، اندازه موثر محاسبه شده برای شجره ICS

اندازه موثر جمعیت نشان‌دهنده تعداد حیوانات مولد در یک جمعیت تحت شرایط ایده‌آل است (Wright, 1931). کاهش در اندازه موثر جمعیت به‌عنوان یک نتیجه از کاهش تنوع ژنتیکی در درون یک جمعیت، منجر به کاهش معنی‌دار در توانایی جمعیت به سازگاری با مشکلاتی از قبیل شرایط آب و هوایی نامناسب، بیماری و عفونت می‌شود (Palstra & Ruzzante, 2008) که این امر به نوبه‌ی خود سبب افزایش ریسک انقراض می‌شود (Newman & Pilson, 1997). اندازه موثر جمعیت برای هر دو شجره مورد مطالعه جهت بررسی

وجود، اندازه موثر برآوردشده با کمک شجره CCS نزدیک‌تر به برآورد ژنومی بود. با توجه به این که تنوع ژنتیکی با کاهش N_e به کمتر از ۱۰۰ با سرعت فزاینده‌ای کاهش می‌یابد، توصیه شده است که سعی شود تا N_e کمتر از ۲۵۰ نباشد (Henson, 1992). البته، باید در نظر داشت که تعیین میزان بحرانی اندازه موثر جمعیت که کمتر از آن جمعیت در طولانی مدت با مشکل همخونی مواجه می‌شود، مشکل است، اما به نظر می‌رسد که بر سر مقادیر ۱۰۰-۵۰۰ اجماعی وجود داشته باشد و با یک حاشیه اطمینان می‌توان ۱۰۰ را توصیه کرد (Meuwissen, 2009). بنا بر این، برآورد N_e انجام‌شده براساس هر دو شجره که کمتر از ۱۰۰ است نشان می‌دهد که وضعیت جمعیت اسب کاسپین در شرایط بحرانی است و باید در خصوص حفظ ذخایر ژنتیکی آن اقدام صورت گیرد.

کمتر از نصف مقدار برآوردشده برای شجره CCS بود. شاید بتوان گفت که مهمترین دلیل این بیشتر بودن با توجه به این که زادگاه اصلی این نژاد ایران است و برآورد اندازه موثر نیز نشان‌دهنده وضعیت بهتر جمعیت اسب کاسپین ایران است، به نظر می‌رسد که ذخایر ژنتیکی اسب کاسپین ایران می‌تواند برای بهبود شرایط اسب کاسپین جهان مورد استفاده قرار گیرد. در مطالعه قبلی جدید که از اطلاعات ژنومی برای برآورد اندازه موثر جمعیت اسب‌های بومی ایران استفاده شد، مقدار آن برای جمعیت اسب‌های برای نژادهای ترکمن، کاسپین، اصیل ایرانی و کرد به ترتیب ۵۹، ۹۸، ۱۰۲ و ۱۱۳ برآورد شد (Mousavi et al., 2023). اندازه موثر برآوردشده براساس اطلاعات شجره‌ای برای هر دو منبع اطلاعاتی در این تحقیق کمتر از مقدار برآوردشده آن مطالعه براساس اطلاعات ژنومی بود. با این

جدول ۲. برآورد اندازه موثر (N_e) و میزان خطای استاندارد بر اساس اطلاعات شجره‌ای انجمن حفاظت از اسب کاسپین (CCS) و انجمن جهانی کاسپین (ICS).

جمعیت Population	N_e برای کل شجره (خطا استاندارد) N_e for the entire pedigree (standard error)	N_e برای جمعیت موجود (خطا استاندارد) N_e For the existing population (standard error)
CCS	68.33 (13.03)	66.43 (14.39)
ICS	32.86 (0.88)	32.40 (0.93)

می‌دهند که با وجود تعداد کمتر اسب‌های ثبت‌شده در شجره CCS، اما وضعیت تنوع ژنتیکی در این جمعیت بهتر است و می‌توان از این منبع ژنتیکی برای بهبود شرایط اسب کاسپین در کل دنیا بهره جست. همچنین، مقدار انحراف برآوردشده از تعادل هاردی-واینبرگ در دو شجره مثبت بود، که این مثبت بودن احتمالاً نشان دهنده وجود مقدار کمی آمیزش جور شده مثبت بین افراد خویشاوند در هر دو شجره است. این آمیزش جور شده مثبت در شجره ICS تقریباً دو برابر جمعیت ثبت‌شده در شجره CCS است. بنا بر این، ارائه یک برنامه تلاقی کنترل‌شده برای کاهش آمیزش‌های بین خویشاوندی برای هر دو جمعیت احساس می‌شود.

نتایج برآورد پارامترهای جمعیت مرجع بر اساس دو منبع اطلاعاتی مورد استفاده در جدول ۳ نشان داده شده‌اند. اختلاف مقادیر محاسبه‌شده برای N_f و N_{ae} در دو شجره نشان‌دهنده تاثیر نسبتاً زیاد تنگه‌های جمعیتی در گذشته است که مقدار اثر آن در کاهش تنوع جمعیتی برای جمعیت ICS بسیار بیشتر است. مطالعه قبلی که بر روی این نژاد با کمک اطلاعات STR انجام شد نیز وجود این تنگه‌ی ژنتیکی را نشان داد (Amirinia et al., 2007). با توجه به مقادیر به دست‌آمده برای N_g و اختلاف آن با دیگر آماره‌ها، میزان تاثیر رانش ژنتیکی بر کاهش تنوع جمعیت نیز قابل توجه است. نتایج نشان می‌دهند که تاثیر رانش در جمعیت ICS به نسبت جمعیت ایران (CCS) بسیار بیشتر است. در کل، نتایج بررسی ساختار جمعیت مرجع نشان

جدول ۳. برآورد پارامترهای مرتبط با جمعیت مرجع و انحراف از تعادل هاردی-واینبرگ بر اساس اطلاعات شجره‌ای انجمن حفاظت از اسب کاسپین (CCS) و انجمن جهانی کاسپین (ICS).

پارامتر ^۱ Parameter	CCS	ICS
N_f (N_{fe})	333 (109.31)	155 (20.92)
N_a (N_{ae})	516 (137.31)	996 (20.22)
N_g (SE_{Ng})	110.12 (7.76)	13.25 (1.53)
انحراف هاردی-واینبرگ جمعیت Hardy-Weinberg deviation of the population	0.0024	0.0050

^۱ N_f : تعداد پایه‌گذار؛ N_{fe} : تعداد موثر پایه‌گذار؛ N_a : تعداد اجناد؛ N_{ae} : تعداد موثر اجناد؛ N_g : تعداد معادل‌های ژنوم پایه‌گذار؛ SE_{Ng} : خطا استاندارد تعداد معادل‌های ژنوم پایه‌گذار.

^۱ N_f : number of founders; N_{fe} : effective number of founders; N_a : number of ancestors; N_{ae} : effective number of ancestors; N_g : number of founder genomes equivalents; SE_{Ng} : standard error number of founder genomes equivalents.

توجه به N_e بسیار پایین که احتمالاً به دلیل وجود تنگه ژنتیکی در گذشته است، می‌توان گفت که جمعیت اسب کاسپین در شرایط بحرانی است و باید در خصوص حفظ تنوع ژنتیکی این نژاد برنامه مشخصی طراحی گردد. همچنین، وجود آمیزش‌های

نتیجه‌گیری کلی

نتایج پژوهش ما نشان می‌دهند که شجره ثبتی در ایران دارای کیفیت کمتری نسبت به شجره جهانی آن است و نیاز جدی جهت ثبت دقیق‌تر شجره در ایران را نشان می‌دهد. با

تشکر و قدردانی

نویسندگان از انجمن حفاظت از اسب کاسپین (Caspian Conservation Society) برای در اختیار قرار دادن اطلاعات جمعیت اسب کاسپین ایران تشکر و قدردانی می‌نمایند.

خویشاوندی می‌تواند وضعیت موجود را بدتر کند که انجام یک برنامه تلاقی کنترل شده برای کاهش آمیزش‌های بین خویشاوندی برای هر دو جمعیت می‌تواند به‌عنوان یک راه حل در نظر گرفته شود.

References

- Amadeu, R. R., Garcia, A. A. F., Munoz, P. R., & Ferrão, L. F. V. (2023). AGHmatrix: genetic relationship matrices in R. *Bioinformatics*, 39(7), btad445.
- Amirinia, C., Seyedabadi, H., Banabazi, M. H., & Kamali, M. A. (2007). Bottleneck study and genetic structure of Iranian Caspian horse population using microsatellites. *Pakistan Journal of Biological Sciences*, 10(9), 1540-1543.
- Boichard, D., Maignel, L., & Verrier, E. (1997). The value of using probabilities of gene origin to measure genetic variability in a population. *Genetics Selection Evolution*, 29(1), 5-23.
- Caballero, A., & Toro, M. A. (2000). Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genetics Research*, 75(3), 331-343.
- Ghafour-Kesbi, F. (2023). Assessing the performance of Ridge regression method-6 in genomic evaluation of discrete threshold traits with additive and dominance genetic architecture. *Research on Animal Production*, 14(42), 102-113. <https://doi.org/10.61186/rap.14.42.102> [In Persian]
- Gutiérrez, J. P., Cervantes, I., Molina, A., Valera, M., & Goyache, F. (2008). Individual increase in inbreeding allows estimating effective sizes from pedigrees. *Genetics Selection Evolution*, 40(4), 359-378.
- Henson, E. L. (1992). *In situ conservation of livestock and poultry*. Food and Agriculture Organization of the United Nations Rome.
- López-Cortegano, E. (2022). purgeR: inbreeding and purging in pedigreed populations. *Bioinformatics*, 38(2), 564-565.
- MacCluer, J. W., Boyce, A. J., Dyke, B., Weitkamp, L. R., Pfenning, D. W., & Parsons, C. J. (1983). Inbreeding and pedigree structure in Standardbred horses. *Journal of Heredity*, 74(6), 394-399.
- Meuwissen, T. (2009). Genetic management of small populations: A review. *Acta Agriculturae Scandinavica, Section a Animal Science*, 59(2), 71-79. <https://doi.org/10.1080/09064700903118148>
- Meuwissen, T. H. E., & Luo, Z. (1992). Computing inbreeding coefficients in large populations. *Genetics Selection Evolution*, 24(4), 305. <https://doi.org/10.1186/1297-9686-24-4-305>
- Mousavi, S. F., Razmkabir, M., Rostamzadeh, J., Seyedabadi, H.-R., Naboulsi, R., Petersen, J. L., & Lindgren, G. (2023). Genetic diversity and signatures of selection in four indigenous horse breeds of Iran. *Heredity*, 131(2), 96-108. <https://doi.org/10.1038/s41437-023-00624-7>
- Newman, D., & Pilson, D. (1997). Increased probability of extinction due to decreased genetic effective population size: *experimental populations of Clarkia pulchella* (Vol. 51).
- Palstra, F. P., & Ruzzante, D. E. (2008). Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, 17(15), 3428-3447. <https://doi.org/10.1111/j.1365-294x.2008.03842.x>
- Saghi, D., & Mobaraki, A. (2018). Estimation of inbreeding and survey of the pedigree structure of Iranian Turkmen horses population. *Research on Animal Production*, 9(22), 131-137. [In Persian]
- Sørensen, A. C., Sørensen, M. K., & Berg, P. (2005). Inbreeding in Danish dairy cattle breeds. *Journal of Dairy Science*, 88(5), 1865-1872.
- Tahmoorespur, M., & Sheikhloo, M. (2011). Pedigree analysis of the closed nucleus of Iranian Baluchi sheep. *Small Ruminant Research*, 99(1), 1-6.
- Wellmann, R. (2019). Optimum contribution selection for animal breeding and conservation: the R package optiSel. *BMC Bioinformatics*, 20(1), 25. <https://doi.org/10.1186/s12859-018-2450-5>
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(2), 97-159. <https://doi.org/10.1093/genetics/16.2.97>