



Identification of genomic runs of homozygosity and investigation of related genes in the dromedary camels using whole-genome sequencing data

Reza Khalkhali-Evrigh¹  | Nemat Hedayat² 

1. Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran. Email: rezakhalkhali110@gmail.com

2. Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran. Email: hedayatuma@gmail.com

Article Info	ABSTRACT
Article type: Research Article	Genetic diversity studies and evaluation of inbreeding levels in camel populations are essential, as well as, the first step for designing breeding programs in this species. In this regard, powerful tools such as next-generation sequencing technology have made it possible to decode the genome information in this species. Based on this research motive, the aim of the current study was to identify the genomic runs of homozygosity (ROH) and investigate related genes in dromedary camels using whole genome sequencing data. For this purpose, a total of 12 sequenced genomic data related to Iranian and non-Iranian dromedary camels were used. After bioinformatics analysis including quality assessment, data pre-preprocessing, alignment in the reference genome, and identification of variants, qualitative filter of variants, finally, ROH regions were identified. Based on the obtained results, 549 (137.3 regions per sample) and 1356 (169.5 regions per sample) ROH were identified in the genomes of Iranian dromedary camels and dromedaries from the Arabian Peninsula, respectively. The results of the annotation revealed that some important fertility-related genes such as FSHR and LHCGR are located in the ROH regions of Iranian dromedary camels. Also, investigation of gene ontology results revealed that some important genes including CXCL9, CXCL10 and CXCL11 (immune-related), STBD1 (related to energy metabolism), SCARB2 (related to lipid metabolism and fertility) and SHROOM3 (related to kidney function) are shared between Iranian and non-Iranian camels. Finally, as a summary, it seems that the controlling factor and the reason for the creation of ROHs in the genome of dromedary camels is natural selection to adapt to the desert environment.
Article history: Received: 25 March 2022 Received: 2 August 2022 Accepted: 5 August 2022 Published online: 22 June 2023	
Keywords: <i>Genomic data, Camel, Next-generation sequencing.</i>	

Cite this article: Khalkhali-Evrigh, R. & Hedayat, N. (2023). Identification of genomic runs of homozygosity and investigation of related genes in the dromedary camels using whole-genome sequencing data. *Iranian Journal of Animal Science*, 54 (2), 105-116. DOI: <http://doi.org/10.22059/IJAS.2022.340848.653883>



© The Author(s).

DOI: <http://doi.org/10.22059/IJAS.2022.340848.653883>

Publisher: The University of Tehran Press.

Extended Abstract

Introduction

Due to the specific ability of camels to survive in harsh desert conditions, they can be one of the main options for food supply in these areas. Despite this special feature and its importance for a country like Iran, as well as its potential role in becoming a farm species, no major effort has been made in the field of breeding this species. Studies of different aspects of genomic characteristics can be the key to solving the backwardness problem in camel studies. This study aimed to identify the genomic runs of homozygosity (ROH) and investigate related genes in dromedary camels using whole genome sequencing data.

Material and methods

In the present study, we used a total of 12 sequenced genomic data related to Iranian and non-Iranian dromedary camels for the identification of genome-wide ROHs. Prior to ROH identification, quality assessment, data pre-preprocessing, mapping to reference genome, identification of variants, and qualitative filter of variants were implemented, respectively. Also, the following criteria were used to define a genomic

region as ROH: genomic regions with at least 40 SNPs (homozyg-snp) and two heterozygous SNPs (homozyg-window-het), also, the minimum number of SNP with missing genotype (homozyg-window-missing) was set to four. In addition, the minimum distance between SNPs to be in two separate ROHs is equal to one million base pairs (homozyg-gap) and the minimum SNP density within an ROH is equal to one SNP per 50 kb (homozyg-density). In order to annotate the ROHs and identify the genes in these regions, the BEDtools program was used. After identification of the genes located in the ROH regions, g:Profiler was implemented for gene ontology (GO) and KEGG pathway analysis.

Results and discussion

Here, we identified 549 (137.3 regions per sample) and 1356 (169.5 regions per sample) ROHs in the genomes of Iranian dromedary camels and dromedaries from the Arabian Peninsula, respectively. Based on the GO analysis results, we detected some important fertility-related genes such as FSHR and LHCGR that are located in the ROH regions of Iranian dromedary camels. Also, several important genes including CXCL9, CXCL10, and CXCL11 (immune-related), STBD1 (related to energy metabolism), SCARB2 (related to lipid metabolism and fertility), and SHROOM3 (related to kidney function) are common between Iranian and non-Iranian camels.

Conclusion

Considering that breeding and artificial selection on camel ecotypes have not been done in different regions of the world, including Iran, therefore, it seems that the creation of ROHs in the genome of camels is a product of positive natural selection for adaptation to harsh desert conditions. It seems that the use of genomic data of a larger number of Iranian camels can provide researchers with a profile of these regions to be used for purposes such as designing appropriate breeding strategies.



شناسایی رشته‌های هموزیگوت ژنومی و بررسی ژن‌های مرتبط در شترهای تک‌کوهانه با استفاده از داده‌های توالی‌یابی کل ژنوم

رضا خلخالی ایوریک^۱ | نعمت هدایت^۲ ✉

۱. گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: rezakhalkhali110@gmail.com

۲. گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران. رایانامه: hedayatuma@gmail.com

چکیده	اطلاعات مقاله
مطالعات تنوع ژنتیکی و بررسی میزان همخونی در جمعیت‌های شتر امری ضروری و همچنین، گام اولیه برای طراحی برنامه‌های اصلاح نژادی در این گونه محسوب می‌شود. در این راستا، ابزارهای قدرتمندی مانند توالی‌یابی نسل بعد، امکان رمزگشایی اطلاعات کل ژنوم در این گونه را فراهم نموده است. با این انگیزه تحقیقاتی، هدف از پژوهش حاضر، شناسایی رشته‌های هموزیگوت ژنومی و بررسی ژن‌های مرتبط در شترهای تک‌کوهانه با استفاده از داده‌های توالی‌یابی کل ژنوم می‌باشد. بدین منظور، در مجموع از ۱۲ داده ژنوم توالی‌یابی شده مربوط به شترهای ایرانی و غیرایرانی استفاده شد. پس از طی مسیر تجزیه و تحلیل بیوانفورماتیکی شامل تعیین کیفیت، پیش‌پردازش داده، همترازی در ژنوم مرجع، شناسایی واریانت‌ها، فیلتر کیفی واریانت‌ها، در نهایت شناسایی نواحی ROH صورت پذیرفت. نتایج پژوهش حاضر منجر به شناسایی ۵۴۹ (۱۳۷/۳) ناحیه به ازای هر نمونه) و ۱۳۵۶ (۱۶۹/۵) ناحیه به ازای هر نمونه) رشته هموزیگوت به ترتیب در ژنوم شترهای ایرانی و شترهای شبه‌جزیره عربستان گردید. نتایج حاصل از حاشیه‌نویس (annotation)، حضور ژن‌های مهم مرتبط با صفات باروری مانند FSHR و LHCGR در داخل رشته‌های هموزیگوت شترهای تک‌کوهانه ایرانی را نشان داد. همچنین، تجزیه و تحلیل بیوانفورماتیکی هستی‌شناسی در مناطق رشته‌های هموزیگوت ژنومی شناسایی شده، حضور ژن‌های مهمی مانند (CXCL10, CXCL9, CXCL11) مرتبط با عملکرد سیستم ایمنی، (STBD1) (مرتبط با متابولیسم انرژی)، (SCARB2) مرتبط با متابولیسم چربی و باروری) و (SHROOM3) مرتبط با عملکرد سیستم کلیوی) بین هر دو جمعیت شترهای ایرانی و غیرایرانی را نشان داد. در نهایت، بعنوان جمع‌بندی به نظر می‌رسد عامل کنترل و پیدایش رشته‌های هموزیگوت در ژنوم شترها، انتخاب طبیعی برای سازگاری با محیط بیابان می‌باشد.	<p>نوع مقاله: مقاله پژوهشی</p> <p>تاریخ دریافت: ۱۴۰۱/۰۱/۰۵ تاریخ بازنگری: ۱۴۰۱/۰۵/۱۱ تاریخ پذیرش: ۱۴۰۱/۰۵/۱۴ تاریخ انتشار: ۱۴۰۲/۰۴/۰۱</p> <p>کلیدواژه‌ها: داده‌های ژنومی، شتر، توالی‌یابی نسل بعد.</p>

استناد: خلخالی ایوریک، رضا و هدایت، نعمت (۱۴۰۲). شناسایی رشته‌های هموزیگوت ژنومی و بررسی ژن‌های مرتبط در شترهای تک‌کوهانه با استفاده از داده‌های توالی‌یابی کل ژنوم. نشریه علوم دامی ایران، ۵۴ (۲)، ۱۰۵-۱۱۶. DOI: <http://doi.org/10.22059/IJAS.2022.340848.653883>

DOI: <http://doi.org/10.22059/IJAS.2022.340848.653883>



© نویسندگان.

DOI: <http://doi.org/10.22059/IJAS.2022.340848.653883>

ناشر: مؤسسه انتشارات دانشگاه تهران.

مقدمه

شتر تک‌کوهانه، یکی از مقاومترین دامها برای زندگی در شرایط سخت بیابان‌های گرم و خشک جهان به شمار می‌رود. از لحاظ پراکنش جغرافیایی، این گونه در قسمت‌هایی از قاره‌های آفریقا، خاورمیانه و آسیای میانه پراکنده بوده و کشور ایران نیز به دلیل دارا بودن اقلیم گرم و خشک، یکی از زیستگاه‌های مهم آن می‌باشد. علی‌رغم اهمیت ویژه این گونه دامی، در تامین احتیاجات غذایی برای ساکنان نواحی بیابانی ایران و همچنین، نقش بالقوه آن در تبدیل به عنوان یک گونه مزرعه‌ای، تاکنون اقدامات اساسی در زمینه اصلاح نژاد این گونه صورت نگرفته است. با این حال، پراکنش در نواحی مختلف با شرایط اقلیمی کم و بیش متفاوت، تاثیراتی بر ژنوم جمعیت‌های مختلف گذاشته است که با استفاده از فناوریهای نوین مانند توالی‌یابی نسل بعدی، می‌توان اقدام به شناسایی و مطالعه نواحی ژنومی مذکور که تحت تاثیر عوامل محیطی بوده‌اند، پرداخت. رشته‌های هموزیگوت (Runs of Homozygosity - ROH) یکی از نواحی ژنومی هستند که در سالهای اخیر، مورد توجه محققان قرار گرفته‌اند. در واقع، ROHها توالی‌های به هم پیوسته هموزیگوتی هستند که یک فرد از هر دو والدین خود به ارث برده است و منشاء آن احتمالاً، یک جد مشترک می‌باشد (Ceballos *et al.*, 2018). در اغلب موارد، انتخاب افرادی با صفات خاص و آمیزش آنها با یکدیگر منجر به افزایش فراوانی ROHها خاصی در جمعیت شده و آن را به سمت همخونی سوق می‌دهد. بخاطر همین مسئله، علاوه بر اینکه شناسایی ROHها می‌تواند، ما را به نشانه‌هایی از انتخاب در جمعیت برساند (Purfield *et al.*, 2017)، برای اندازه‌گیری همخونی در جمعیت‌ها نیز به کار گرفته می‌شود (McQuillan *et al.*, 2008). علیرغم وجود مطالعاتی در مورد ژنوم شترهای تک‌کوهانه ایرانی (Khalkhali-Evrigh *et al.*, 2018; Khalkhali-Evrigh *et al.*, 2022) مطالعه‌ای بر روی نواحی ROH در شترهای تک‌کوهانه ایرانی صورت نگرفته است. هدف پژوهش، استفاده از داده‌های توالی‌یابی شده کل ژنوم به منظور شناسایی و بررسی ROHهای موجود در ژنوم شترهای تک‌کوهانه ایرانی و مقایسه آنها با شترهای تک‌کوهانه از شبه‌جزیره عربستان بود. همچنین، شناسایی ژن‌های واقع شده در ROHها نیز صورت پذیرفت که این امر می‌تواند اطلاعاتی از نشانه‌های انتخاب در دو گروه مذکور در اختیار ما قرار دهد.

مواد و روش‌ها

نمونه‌ها، پالایش کیفی داده‌ها و همترازی در ژنوم مرجع

در پژوهش حاضر از داده‌های توالی‌یابی شده کل ژنوم ۱۲ شتر تک‌کوهانه شامل چهار نفر شتر ایرانی و هشت نفر شتر از شبه‌جزیره عربستان استفاده شد. داده‌های ژنومی شترهای تک‌کوهانه ایرانی مربوط به مطالعه Khalkhali-Evrigh و همکاران (۲۰۱۸) و Ming و همکاران (۲۰۲۰) بود. داده‌های ژنومی شترهای شبه‌جزیره عربستان نیز مربوط به مطالعه Holl و همکاران (۲۰۱۷) و همچنین، Fitak و همکاران (۲۰۲۰) بود.

پیش از انجام هرگونه آنالیز، داده‌های ژنومی نمونه‌های مورد مطالعه با استفاده از برنامه FastQC (نسخه ۱۱/۹) مورد سنجش کیفی قرار گرفتند. در گام بعدی، از نسخه ۰/۳۹ برنامه Trimmomatic (Bolger *et al.*, 2014) و با استفاده از رویه Sliding Windows (5:20) به پیرایش کیفی داده‌ها پرداخته شد. در برخی از نمونه‌ها از گزینه HEADCROP برای حذف سه تا پنج نوکلئوتید ابتدایی توالی‌ها به دلیل کیفیت بسیار پایین آنها استفاده شد. در نمونه‌های دارای آلودگی آداپتوری نیز از گزینه ILLUMINA CLIP برای حذف آداپتورها بهره گرفته شد. در نهایت، توالی‌هایی که معیارهای فیلتراسیون را گذرانده بودند و طول آنها بالای ۴۰ باز بود، برای آنالیزهای بعدی انتخاب شدند. داده‌های پیرایش شده با استفاده از الگوریتم mem در برنامه BWA (نسخه ۰/۱۷؛ Li & Durbin, 2009)، در ژنوم مرجع مربوط به شتر تک‌کوهانه (GCF_000803125.2) نگاهشته شدند. در مرحله بعد و با استفاده از برنامه Picard (<http://broadinstitute.github.io/picard>) فایل‌های BAM حاصله از گام پیشین، منظم شده و سپس خوانش‌های تکرار شده (Duplicated Reads) حذف شدند. نهایتاً با استفاده از دو

رویه "RealignerTargetCreator" و "IndelRealigner" در برنامه GATK (نسخه ۳/۷؛ McKenna *et al.*, 2010) همترازی مجدد در اطراف حذف/درج‌ها (ایندل) اجرا شد.

شناسایی و فیلتر کیفی واریانت‌ها

در پژوهش حاضر، از الگوریتم "HaplotypeCaller" برنامه GATK و با استفاده از گزینه "ERC GVCF" به منظور تولید فایل‌های gVCF برای هر یک از نمونه‌های مورد مطالعه استفاده شد. سپس، به کمک ابزار "GenotypeGVCFs" برنامه GATK، شناسایی واریانت‌ها برای چهار نمونه شتر ایرانی و هشت نمونه شتر شبه‌جزیره عربستان به طور همزمان اجرا شده و برای هر کدام از گروه‌ها یک فایل VCF حاوی واریانت‌های خام، تولید شد. برای ادامه آنالیزها SNP‌های شناسایی شده از ایندل‌ها جدا شده و با استفاده از برنامه‌های VCFtools، BCftools (Danecek *et al.*, 2011) و GATK فیلتر شدند. پارامترهای در نظر گرفته شده برای فیلتراسیون SNP‌ها شامل $QD < 2$ ، $FS > 60$ ، $MQ < 40$ ، $MQRankSum < -12.5$ ، $ReadPosRankSum < -8$ و $SOR > 3$ ، $cluster-size$ 3 و $cluster-window-size$ 10 بودند. در گام بعدی فیلتراسیون، SNP‌های دو آلی با فراوانی آلل مینور بزرگتر یا مساوی با ۰/۰۵ که حداقل در ۸۰ درصد افراد دارای ژنوتیپ مشخص بودند، حفظ شده و مابقی حذف شدند. در گام پایانی، SNP‌های موجود در کروموزوم جنسی (کروموزوم X) و کانتیگ‌هایی که در توالی‌های کروموزومی جای نگرفته بودند حذف شده و فایل نهایی SNP‌ها بدست آمد.

شناسایی نواحی ROH

برای شناسایی نواحی ROH در ژنوم شترهای تک‌کوهانه ایرانی و شبه‌جزیره عربستان از ابزار homozyg برنامه PLINK 1.9 (Purcell *et al.*, 2007) استفاده شد. از آنجایی که عدم تعادل لینکاژی می‌تواند منجر به شکل‌گیری ROH‌های کوتاه در سطح ژنوم شود، به همین منظور و برای اجتناب از شناسایی این نواحی توسط برنامه PLINK 1.9، حداقل طول ROH‌ها برای این مطالعه 0.5 Mb (homozyg-kb) در نظر گرفته شد. همچنین، در مطالعه حاضر، برای اینکه یک ناحیه ژنومی به عنوان ROH معرفی شود باید دارای معیارهای رو به رو می‌بود: نواحی ژنومی با حداقل ۴۰ SNP (homozyg-het)، حداکثر دو SNP هتروزیگوت (homozyg-window-het) و حداکثر چهار SNP با ژنوتیپ گمشده (homozyg-window-missing). علاوه بر این، حداقل فاصله بین SNP‌ها برای اینکه در دو ROH جداگانه قرار بگیرند برابر با یک میلیون جفت‌باز (homozyg-gap) و حداقل تراکم SNP در داخل یک ROH برابر با یک SNP به ازای هر 50 kb (homozyg-density) در نظر گرفته شد. پس از شناسایی ROH‌ها در سطح ژنوم، آنها به سه گروه با اندازه کوتاه (1 Mb - 0.5)، متوسط ($1 - 5 \text{ Mb}$) و طویل ($> 5 \text{ Mb}$) تقسیم‌بندی شدند. همچنین، برای شناسایی ROH‌هایی که حداقل بین دو فرد دارای همپوشانی بودند، از ابزار homozyg-group در برنامه PLINK 1.9 بهره گرفته شد. برای یافتن ژن‌های موجود در ROH‌های شناسایی شده برای شترهای تک‌کوهانه ایرانی و غیر ایرانی، از ROH‌های همپوشان بین افراد هر گروه استفاده شد.

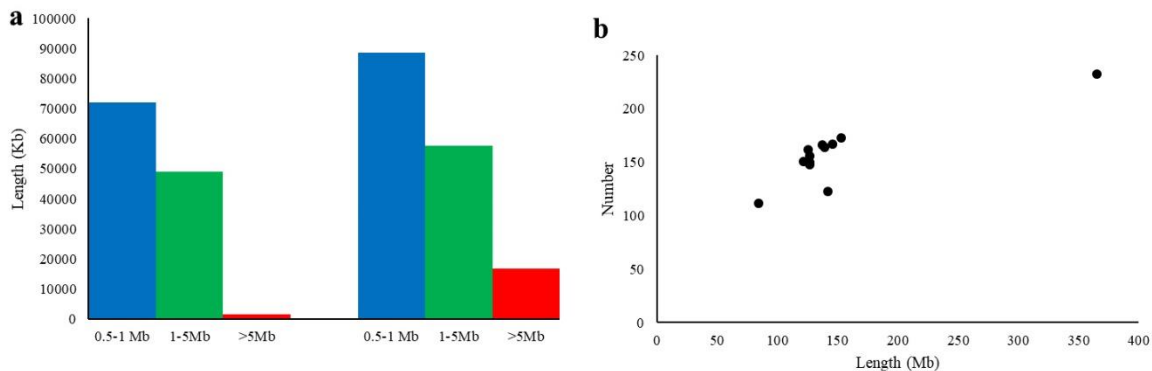
حاشیه‌نویسی مناطق ROH و آنالیزهای پایین دستی

در مطالعه کنونی و به منظور حاشیه‌نویسی مناطق ROH و شناسایی ژن‌های موجود در این نواحی، از برنامه BEDtools (نسخه ۲/۲۷/۱؛ Quinlan & Hall, 2010) و فایل GTF مربوط به ژنوم رفرنس استفاده شد. پس از شناسایی ژن‌های واقع شده در نواحی ROH، برنامه تحت وب g:Profiler برای آنالیز ماهیت‌شناسی ژنی و مسیرهای KEGG به کار گرفته شد. آنالیز ماهیت‌شناسی ژنی به ما این امکان را می‌دهد که ژن‌ها را در گروه‌های عملکردی طبقه‌بندی کرده و درک درستی از

روابط و عملکرد آنها به دست آوریم. به طور کلی ماهیت‌شناسی ژنی عبارات را در سه دسته فرآیندهای زیستی (BP)، جایگاه سلولی (CC) و عملکرد مولکولی (MF) کلاس‌بندی می‌کند. برای شناسایی عبارات معنادار ماهیت‌شناسی ژنی و مسیرهای KEGG از آستانه P -value تصحیح شده (با استفاده از روش Benjamini-Hochberg) کمتر و مساوی ۰/۰۵ استفاده شد.

نتایج و بحث

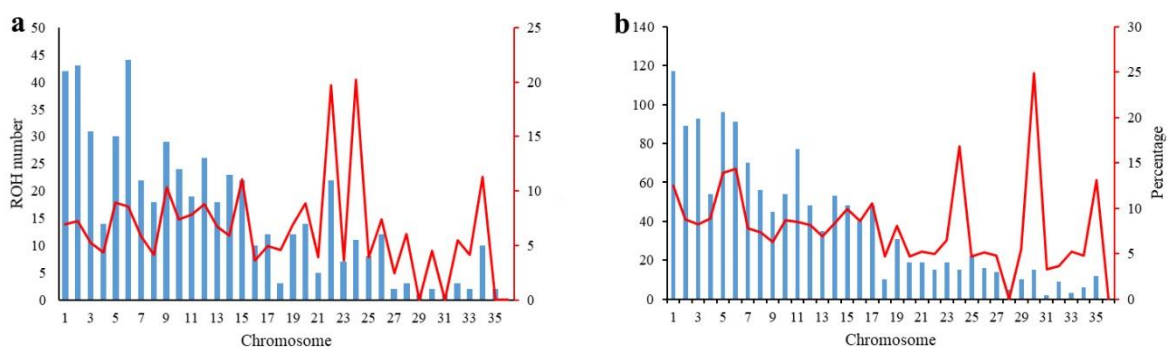
براساس نتایج بدست آمده در این مطالعه، تعداد ۵۴۹ (۱۳۷/۳ ناحیه به ازای هر فرد) و ۱۳۵۶ (۱۶۹/۵ ناحیه به ازای هر فرد) ناحیه ROH به ترتیب برای شترهای ایرانی و شترهای شبه‌جزیره عربستان شناسایی شد. میانگین طول کل ROHهای موجود در ژنوم شترهای ایرانی برابر با ۸۹۲/۶ Kb (محدوده ۵۰۰/۲ تا ۵۹۲۹/۴ Kb) و در شترهای شبه‌جزیره عربستان برابر با ۹۶۱/۲ Kb (محدوده ۵۰۰/۱ تا ۹۶۱۹/۵ Kb) بود. براساس دسته‌بندی صورت گرفته، در هر دو گروه از شترها، ROHهای کوتاه، بیشترین درصد طول ROHهای شناسایی شده را به خود اختصاص دادند به طوریکه ۵۸/۸ و ۵۴/۴ درصد از ROHهای شترهای ایرانی و شترهای غیر ایرانی به ترتیب مربوط به این دسته بودند (شکل ۱a). براساس همین نتایج، سهم ROHهای با طول متوسط و طویل در رتبه‌های بعدی قرار گرفتند. همچنین نتایج نشان دادند که به طور کلی، بین تعداد ROHهای شناسایی شده برای هر فرد و مجموع طول آنها همبستگی مثبتی وجود دارد (شکل ۱b). به جز یک نمونه مربوط به شترهای شبه‌جزیره عربستان، تعداد ROHهای شناسایی شده برای مابقی نمونه‌ها در محدوده ۱۰۰ تا ۱۷۵ و با مجموع طول تقریباً ۸۵ تا ۱۵۳ Mb بودند. به طور کلی، ROHهای طولانی نشان دهنده وقوع همخونی در نسل‌های اخیر هستند در حالی که ROHهای با طول کوتاه این مفهوم را می‌رسانند که وقوع این ROHها در زمان‌های قدیمی‌تری اتفاق افتاده و کوتاه شدن طول آنها می‌تواند در اثر فرآیند شکسته شدن قطعات کروموزومی و نوترکیبی باشد (Kirin et al., 2010).



شکل ۱. اندازه کل ROHهای شناسایی شده در سه دسته طولی در شترهای ایرانی (چپ a) و غیر ایرانی (راست a) و همبستگی بین طول کل ROHها و

تعداد آنها در هر فرد (دایره‌های سیاه) (b)

پراکنش ROHها بر روی کروموزوم‌های مختلف الگوی یکسانی را نشان نداد. اما، تا حدودی، تعداد ROHها تابعی از طول کروموزومها بود، هرچند درصد پوشش ROHها در هر کروموزوم همبستگی آشکاری با طول کروموزومها نشان نداد (شکل ۲). در هر دو گروه مورد مطالعه، درصد طولی ROHها نسبت به طول کروموزوم، با کوتاه‌تر شدن کروموزومها تا حدودی افزایش یافت.



شکل ۲. تعداد ROHها (ستون آبی) و میانگین درصد پوشش طولی آنها (خط قرمز) در هر کروموزوم در شترهای تک‌کوهانه ایرانی (a) و شبه‌جزیره عربستان (b)

نتایج شناسایی ROHهای مشترک (حداقل در دو نمونه) بین نمونه‌های مورد مطالعه در هر گروه، منجر به شناسایی هفت ROH (پراکنده در شش کروموزوم) برای شترهای ایرانی و ۴۰ ROH (پراکنده در ۱۶ کروموزوم) برای شترهای شبه‌جزیره عربستان شد. حاشیه‌نویسی این نواحی برای یافتن ژن‌های موجود در آنها به ترتیب باعث شناسایی ۴۱ و ۴۱۱ ژن کد کننده پروتئین برای شترهای تک‌کوهانه ایرانی و غیر ایرانی شد. آنالیز ماهیت‌شناسی ژنی بر روی ژن‌های مذکور نشان داد که ۱۶ عبارت (یک عبارت BP، ۱۱ عبارت MF و چهار عبارت KEGG) برای شترهای ایرانی (جدول ۱) و هشت عبارت (شش عبارت BP و دو عبارت MF) نیز برای شترهای غیرایرانی (جدول ۲) از نظر آماری معنادار هستند ($\text{adjusted } P\text{-value} \leq 0.05$).

جدول ۱. نتایج آنالیز ماهیت‌شناسی ژنی بر روی ژن‌های واقع شده در ROHهای شناسایی شده برای شترهای تک‌کوهانه ایرانی

کلاس	شماره هستی‌شناسی ژن	عبارت ماهیت‌شناسی ژن	نام ژن	P-Value تصحیح شده
BP	GO:0001541	ovarian follicle development	KIT,LHCGR,FSHR	۰/۰۳۲
MF	GO:0008009	chemokine activity	CXCL9,CXCL10,CXCL11	۰/۰۱۰
MF	GO:0042379	chemokine receptor binding	CXCL9,CXCL10,CXCL11	۰/۰۱۰
MF	GO:0016500	protein-hormone receptor activity	LHCGR,FSHR	۰/۰۱۸
MF	GO:0004963	follicle-stimulating hormone receptor activity	FSHR	۰/۰۳۵
MF	GO:0004964	lutinizing hormone receptor activity	LHCGR	۰/۰۳۵
MF	GO:0005020	stem cell factor receptor activity	KIT	۰/۰۳۵
MF	GO:0005125	cytokine activity	CXCL9,CXCL10,CXCL11	۰/۰۳۵
MF	GO:0035472	choriognadotropin hormone receptor activity	LHCGR	۰/۰۳۵
MF	GO:0038106	choriognadotropin hormone binding	LHCGR	۰/۰۳۵
MF	GO:0047412	N-(long-chain-acyl)ethanolamine deacylase activity	NAAA	۰/۰۳۵

کلاس	شماره هستی‌شناسی ژن	عبارت ماهیت‌شناسی ژن	نام ژن	P-Value تصحیح شده
MF	GO:0001664	G protein-coupled receptor binding	CXCL9,CXCL10,CXCL11	۰/۰۴۲
KEGG	KEGG:04061	Viral protein interaction with cytokine and cytokine receptor	CXCL9,CXCL10,CXCL11	۰/۰۰۶
KEGG	KEGG:04620	Toll-like receptor signaling pathway	CXCL9,CXCL10,CXCL11	۰/۰۰۶
KEGG	KEGG:04062	Chemokine signaling pathway	CXCL9,CXCL10,CXCL11	۰/۰۱۹
KEGG	KEGG:04913	Ovarian steroidogenesis	LHCGR,FSHR	۰/۰۲۳

جدول ۲. نتایج آنالیز ماهیت‌شناسی ژنی بر روی ژن‌های واقع شده در ROH‌های شناسایی شده برای شترهای تک‌کوهانه از شبه‌جزیره عربستان

کلاس	شماره هستی‌شناسی ژن	عبارت ماهیت‌شناسی ژن	نام ژن	P-Value تصحیح شده
BP	GO:0000054	ribosomal subunit export from nucleus	SDAD1,RIOK2,XPO1	۰/۰۲۴
BP	GO:0007178	transmembrane receptor protein serine/threonine kinase signaling pathway	NCEH1, RGMB,ITGB6,SNW1,SKI,PRDM16,ZFYVE9, ADAM17,SPTBN1,TGIF2,MAPK14, ONECUT2,SMAD4,SMAD7,SMAD2	۰/۰۲۴
BP	GO:0007179	transforming growth factor beta receptor signaling pathway	ITGB6, SNW1,SKI,PRDM16,ZFYVE9,ADAM17, ONECUT2,SMAD4,SMAD7,SMAD2	۰/۰۲۴
BP	GO:0017015	regulation of transforming growth factor beta receptor signaling pathway	SNW1, SKI,PRDM16,ADAM17,ONECUT2, SMAD4,SMAD7,SMAD2	۰/۰۲۴
BP	GO:0033750	ribosome localization	SDAD1,RIOK2,XPO1	۰/۰۲۴
BP	GO:1903844	regulation of cellular response to transforming growth factor beta stimulus	SNW1, SKI,PRDM16,ADAM17,ONECUT2, SMAD4,SMAD7,SMAD2	۰/۰۲۴
MF	GO:0005031	tumor necrosis factor-activated receptor activity	TNFRSF18,TNFRSF4,TNFRSF11A	۰/۰۳۸
MF	GO:0005035	death receptor activity	TNFRSF18,TNFRSF4,TNFRSF11A	۰/۰۳۸

به طور کلی عبارات معنادار شده برای شترهای ایرانی را می‌توان به دو دسته عبارات مرتبط با ایمنی و باروری تقسیم کرد. از آنجایی که فشار مثبت انتخاب (انتخاب طبیعی یا مصنوعی) می‌تواند یکی از دلایل ایجاد ROHها باشد (Gorssen *et al.*, 2021)، به نظر می‌رسد احتمالاً تعدادی از ژن‌های مرتبط با ایمنی و باروری در شترهای تک‌کوهانه ایرانی تحت تاثیر فرآیند انتخاب قرار گرفته باشند. دو ژن FSHR و LHCGR به عنوان ژن‌های مرتبط با صفات باروری، بر روی کروموزوم ۱۵ شترهای تک‌کوهانه و در کنار یکدیگر قرار دارند که در مطالعه حاضر، در داخل یکی از ROH‌های شناسایی شده قرار گرفتند. ژن FSHR، یک گیرنده پروتئینی با همین نام را کد می‌کند که بر سطح سلول‌های گرانولوزای (Granulosa cells) فولیکول قرار گرفته و در اثر اتصال با هورمون FSH، منجر به شروع فرآیندهایی می‌شود که نهایتاً باعث توسعه و بلوغ فولیکول در تخمدان می‌شوند (Simoni *et al.*, 1997). در مطالعات مختلفی بر روی تنوع‌های موجود در ژن FSHR در گونه‌های مختلفی مانند گاو (Cory *et al.*, 2013)، گاو میش (Al-Hamedawi *et al.*, 2017)، گوسفند (Chu *et al.*, 2012) و بز (Guo *et al.*, 2013) تاثیر مستقیم این ژن در صفات مرتبط با باروری نشان داده شده است. پس از هورمون FSH که در بلوغ فولیکول‌ها نقش دارد، هورمون LH نیز در تخمک‌گذاری و کمک به تداوم تولید هورمون پروژسترون در دوران آبستنی پستانداران دخیل است. هورمون LH برای تاثیر گذاری، به یک گیرنده پروتئینی نیاز دارد که محصول ژن LHCGR می‌باشد، به طوریکه مطالعات نشان داده‌اند که تخمک‌گذاری و لوتئینیزاسیون (ایجاد جسم زرد) فولیکول‌های غالب با میزان سطح بیان

LHCGR مرتبط است (Minegishi *et al.*, 2008). همچنین مطالعه‌ای بر روی گاو (Widmer *et al.*, 2021) نشان داد که QTL در برگیرنده این دو ژن، به دلیل نقشی که این ژن‌ها در سوپراوولاسیون دارند، با صفت دوقلو زایی مرتبط می‌باشد. در مورد نتایج ماهیت‌شناسی ژنی شترهای غیرایرانی نیز چند عبارت مرتبط با عملکرد دستگاه ایمنی مشاهده شد که شامل ژن‌هایی از جمله TNFRSF4، TNFRSF18 و TNFRSF11A بودند. این سه ژن از اعضای خانواده ژنی گیرنده‌های TNF به شمار می‌روند و در عملکرد دستگاه ایمنی و سلول‌های T نقش دارند به طور مثال شواهد نشان می‌دهند که وجود واریانت‌های خاصی در ژن TNFRSF11A منجر به بیماری استخوانی پاژه (Paget) در انسان می‌شود (Nakatshka *et al.*, 2003; Chung *et al.*, 2010). به نظر می‌رسد انتخاب ژن‌های مرتبط با ایمنی متفاوت در جمعیت‌های مختلف، در اثر عواملی بیماری‌زای گوناگونی باشد که در نقاط مختلف جهان وجود دارند و منجر به انتخاب مثبت ژن‌های خاصی شده‌اند (Kjærner-Semb *et al.*, 2016).

تعداد ۲۸ ژن شناسایی شده در هر دو گروه، مشترک بودند. اجرای آنالیز ماهیت‌شناسی ژنی بر روی این ژن‌ها آشکار ساخت که نه عبارت از ۱۰ عبارت معنادار شده (شش عبارت MF و چهار عبارت KEGG) مرتبط با ایمنی و سلامتی می‌باشند (جدول ۳). تمام عبارات مرتبط با ایمنی، حاوی سه ژن CXCL9، CXCL10 و CXCL11 بودند. ژن‌های مذکور، پروتئین‌های کوچکی به نام شیموکین‌ها (chemokine) را تولید می‌کنند که در فعال‌سازی لوکوسیت‌ها نقش دارند (Colvin *et al.*, 2004). این سه شیموکین دارای یک گیرنده واحد به نام CXCR3 هستند و تحت تاثیر اینترفرون گاما بیان می‌شوند و در پاسخ‌های التهابی، ایمنی ذاتی، ایمنی اکتسابی، پاسخ به عفونت و ایمنی دستگاه عصبی نیز دارای نقش هستند (Müller *et al.*, 2010). یکی دیگر از ژن‌های مهم مشترک بین هر دو گروه، ژن STBD1 بود. این ژن در انتقال گلیکوژن به لیزوزیم در سلول‌های کبدی نقش دارد که این فرآیند با نام گلیکوفازی شناخته می‌شود (Sun *et al.*, 2016). فرآیند گلیکوفازی در حفظ هموستازی گلوکز بافت‌هایی مانند قلب، ماهیچه و کبد در نوزدان تازه متولد شده نقش دارد (Zhao *et al.*, 2018). یکی دیگر از ژن‌های مشترک، SCARB2 بود که در انتقال و ذخیره چربی، جذب کلسترول و صفات باروری به ایفای نقش می‌پردازد (Bertevello *et al.*, 2018). از آنجایی که کلسترول یکی از چربی‌های مهم موجود در مایع فولیکولی به شمار می‌رود، لذا، این ژن با کمک به جذب کلسترول توسط تخمدان، به تولید استروئیدها و در نهایت بهبود باروری کمک می‌کند (Zhou *et al.*, 2021). یکی از ژن‌های مرتبط با عملکرد کلیه نیز در لیست ژن‌های مشترک قرار داشت. براساس شواهد، ژن SHROOM3 در توسعه ساختاری سلول‌های پودوسیت در کپسول بومن کلیه نقش دارد به طوری که اختلال در عملکرد این ژن، منجر به آسیب سیستم کلیوی و نشت مولکول‌های بزرگی مانند آلبومین خون به ادرار می‌شود (Matsuura *et al.*, 2020).

جدول ۳. نتایج آنالیز ماهیت‌شناسی ژنی بر روی ژن‌های مشترک شناسایی شده برای شترهای ایرانی و غیرایرانی

شماره هستی‌شناسی ژن	عبارت ماهیت‌شناسی ژن	نام ژن	P-Value تصحیح شده	کلاس
GO:0042379	chemokine receptor binding	CXCL9,CXCL10,CXCL11	۰/۰۰۲	MF
GO:0008009	chemokine activity	CXCL9,CXCL10,CXCL11	۰/۰۰۳	MF
GO:0005125	cytokine activity	CXCL9,CXCL10,CXCL11	۰/۰۱۷	MF
GO:0001664	G protein-coupled receptor binding	CXCL9,CXCL10,CXCL11	۰/۰۳۱	MF
GO:0047412	N-(long-chain-acyl)ethanolamine deacylase activity	NAAA	۰/۰۳۸	MF
GO:0005126	cytokine receptor binding	CXCL9,CXCL10,CXCL11	۰/۰۴۰	MF
KEGG:04061	Viral protein interaction with cytokine and cytokine receptor	CXCL9,CXCL10,CXCL11	۰/۰۰۰۲	KEGG
KEGG:04620	Toll-like receptor signaling pathway	CXCL9,CXCL10,CXCL11	۰/۰۰۰۲	KEGG
KEGG:04062	Chemokine signaling pathway	CXCL9,CXCL10,CXCL11	۰/۰۰۱	KEGG
KEGG:04060	Cytokine-cytokine receptor interaction	CXCL9,CXCL10,CXCL11	۰/۰۰۳	KEGG

نتیجه‌گیری

مطالعه انجام شده برای شناسایی ROH‌های موجود در ژنوم شترهای تک‌کوهانه نشان داد که این نواحی می‌توانند هدف مطالعاتی مناسبی برای شناسایی ژن‌های مهم عملکردی در شترها باشند. با توجه به اینکه اصلاح‌نژاد و انتخاب خاصی بر روی اکوتیپ‌های شتر در نواحی مختلف جهان از جمله ایران صورت نگرفته است لذا به نظر می‌رسد ایجاد ROH‌ها در ژنوم شترها، محصول انتخاب طبیعی مثبت برای آداپته کردن این حیوانات به شرایط سخت بیابانی باشد. به نظر می‌رسد استفاده از داده‌های ژنومی تعداد بیشتری از شترهای ایرانی می‌تواند پروفایلی از این نواحی را در اختیار محققان قرار دهد تا برای اهدافی مانند طراحی استراتژی‌های پرورشی و اصلاحی مورد استفاده قرار گیرد.

سپاسگزاری

بدینوسیله از صندوق حمایت از پژوهشگران و فناوران کشور (INSF) به جهت حمایت از این پروژه (به شماره: ۹۹۰۲۳۲۹۷) کمال تشکر را داریم.

REFERENSES

- Al-Hamedawi, T., AL-Mutar, H. A. K. & Al-Yasiri, E. (2017). Investigation of the polymorphism in FSHR gene associated with fertility in pregnant and non-pregnant Iraq buffaloes. *Current Research in Microbiology and Biotechnology*, 5(6), 1323-1327.
- Bertevello, P. S., Teixeira-Gomes, A. P., Seyer, A., Vitorino Carvalho, A., Labas, V., Blache, M. C., Banliat, C., Cordeiro, L. A. V., Duranthon, V., Papillier, P. & Maillard, V. (2018). Lipid identification and transcriptional analysis of controlling enzymes in bovine ovarian follicle. *International Journal of Molecular Sciences*, 19(10), 3261.
- Bolger, A. M., Lohse, M. & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.
- Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M. & Wilson, J. F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nature Reviews Genetics*, 19(4), 220-234.
- Chu, M. X., Guo, X. H., Feng, C. J., Li, Y., Huang, D. W., Feng, T., Cao, G. L., Fang, L., Di, R., Tang, Q. Q. & Ma, Y. H. (2012). Polymorphism of 5' regulatory region of ovine FSHR gene and its association with litter size in Small Tail Han sheep. *Molecular Biology Reports*, 39(4), 3721-3725.
- Chung, P. Y. J., Beyens, G., Riches, P. L., Van Wesenbeeck, L., De Freitas, F., Jennes, K., Daroszewska, A., Franssen, E., Boonen, S., Geusens, P. & Vanhoenacker, F. (2010). Genetic variation in the TNFRSF11A gene encoding RANK is associated with susceptibility to Paget's disease of bone. *Journal of Bone and Mineral Research*, 25(12), 2592-2605.
- Colvin, R. A., Campanella, G. S., Sun, J. & Luster, A. D. (2004). Intracellular domains of CXCR3 that mediate CXCL9, CXCL10, and CXCL11 function. *Journal of Biological Chemistry*, 279(29), 30219-30227.
- Cory, A. T., Price, C. A., Lefebvre, R. & Palin, M. F. (2013). Identification of single nucleotide polymorphisms in the bovine follicle-stimulating hormone receptor and effects of genotypes on superovulatory response traits. *Animal Genetics*, 44(2), 197-201.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T. & McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156-2158.
- Fitak, R.R., Mohandesan, E., Corander, J., Yadamsuren, A., Chuluunbat, B., Abdelhadi, O., Raziq, A., Nagy, P., Walzer, C., Faye, B. & Burger, P.A. (2020). Genomic signatures of domestication in Old World camels. *Communications Biology*, 3(1), 1-10.
- Gorssen, W., Meyermans, R., Janssens, S. & Buys, N. (2021). A publicly available repository of ROH islands reveals signatures of selection in different livestock and pet species. *Genetics Selection Evolution*, 53(1), 1-10.

- Guo, X., Li, Y., Chu, M., Feng, C., Di, R., Liu, Q., Feng, T., Cao, G., Huang, D., Fang, L. & Tang, Q. (2013). Polymorphism of 5' regulatory region of caprine FSHR gene and its association with litter size in Jining Grey goat. *Turkish Journal of Veterinary and Animal Sciences*, 37(5), 497-503.
- Holl, H., Isaza, R., Mohamoud, Y., Ahmed, A., Almathen, F., Youcef, C., Gaouar, S., Antczak, D.F. & Brooks, S. (2017). A frameshift mutation in KIT is associated with white spotting in the Arabian camel. *Genes*, 8(3), 102.
- Kirin, M., McQuillan, R., Franklin, C. S., Campbell, H., McKeigue, P. M. & Wilson, J. F. (2010). Genomic runs of homozygosity record population history and consanguinity. *PloS one*, 5(11), p.e13996.
- Kjærner-Semb, E., Ayllon, F., Furmanek, T., Wennevik, V., Dahle, G., Niemelä, E., Ozerov, M., Vähä, J. P., Glover, K. A., Rubin, C. J. & Wargelius, A. (2016). Atlantic salmon populations reveal adaptive divergence of immune related genes—a duplicated genome under selection. *BMC Genomics*, 17(1), 1-12.
- Khalkhali-Evrigh, R., Hafezian, S. H., Hedayat-Evrigh, N., Farhadi, A. & Bakhtiarizadeh, M. R. (2018). Genetic variants analysis of three dromedary camels using whole genome sequencing data. *PloS one*, 13(9), p.e0204028.
- Khalkhali-Evrigh, R., Hedayat-Evrigh, N., Hafezian, S. H., Farhadi, A. & Bakhtiarizadeh, M. R., (2019). Genome-wide identification of microsatellites and transposable elements in the dromedary camel genome using Whole-Genome sequencing data. *Frontiers in Genetics*, p.692.
- Khalkhali-Evrigh, R., Hedayat, N., Ming, L., & Jirimutu. (2022). Identification of selection signatures in Iranian dromedary and Bactrian camels using whole genome sequencing data. *Scientific Reports*, 12(1), 1-10.
- Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), 1754-1760.
- Matsuura, R., Hiraishi, A., Holzman, L. B., Hanayama, H., Harano, K., Nakamura, E., Hamasaki, Y., Nangaku, M. & Noiri, E., (2020). SHROOM3, the gene associated with chronic kidney disease, affects the podocyte structure. *Scientific Reports*, 10(1), 1-11.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297-1303.
- McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A. & MacLeod, A. K. (2008). Runs of homozygosity in European populations. *The American Journal of Human Genetics*, 83(3), 359-372.
- Minegishi, T., Nakamura, K., Yamashita, S., Ikeda, S. & Kogure, K. (2008). Regulation of human luteinizing hormone receptor in the ovary. *Reproductive Medicine and Biology*, 7(1), 11-16.
- Ming, L., Yuan, L., Yi, L., Ding, G., Hasi, S., Chen, G., Jambl, T., Hedayat-Evrigh, N., Batmunkh, M., Badmaevna, G.K. & Gan-Erdene, T. (2020). Whole-genome sequencing of 128 camels across Asia reveals origin and migration of domestic Bactrian camels. *Communications Biology*, 3(1), 1-9.
- Müller, M., Carter, S., Hofer, M.J. & Campbell, I. L. (2010). the chemokine receptor CXCR3 and its ligands CXCL9, CXCL10 and CXCL11 in neuroimmunity—a tale of conflict and conundrum. *Neuropathology and Applied Neurobiology*, 36(5), 368-387.
- Nakatsuka, K., Nishizawa, Y. & Ralston, S. H. (2003). Phenotypic characterization of early onset Paget's disease of bone caused by a 27-bp duplication in the TNFRSF11A gene. *Journal of Bone and Mineral Research*, 18(8), 1381-1385.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J. & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559-575.
- Purfield, D. C., McParland, S., Wall, E. & Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS One*, 12(5), p.e0176780.

- Quinlan, A.R. & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841-842.
- Simoni, M., Gromoll, J. & Nieschlag, E. (1997). The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology. *Endocrine Reviews*, 18(6), 739-773.
- Sun, T., Yi, H., Yang, C., Kishnani, P. S. & Sun, B. (2016). Starch binding domain-containing protein 1 plays a dominant role in glycogen transport to lysosomes in liver. *Journal of Biological Chemistry*, 291(32), 16479-16484.
- Widmer, S., Seefried, F. R., von Rohr, P., Häfliger, I. M., Spengeler, M. & Drögemüller, C. (2021). A major QTL at the LHCGR/FSHR locus for multiple birth in Holstein cattle. *Genetics Selection Evolution*, 53(1), 1-15.
- Zhao, H., Tang, M., Liu, M. & Chen, L. (2018). Glycophagy: An emerging target in pathology. *Clinica Chimica Acta*, 484, 298-303.
- Zhou, J., Du, J., Yue, S., Xue, B., Wang, L., Peng, Q. & Xue, B. (2021). N-carbamylglutamate promotes follicular development by modulating cholesterol metabolism in Yak ovaries. *Agriculture*, 11(9), 825.