

## "Research Paper"

### Genome Wide Association Study Based on Pathway Analysis Relate to Total Number of Lambs Born and First Lambing Age in Prolificacy Sheep

Hossein Mohammadi<sup>1</sup> and Mohammad Shamsollahi<sup>2</sup>

1- Assistant Professor, Department of Animal Sciences, Faculty of Agriculture and Natural Resources, Arak University, Arak, Iran, (Corresponding author: mohammadi13371364@gmail.com)

2- Assistant Professor, Department of Animal Sciences, Faculty of Agriculture, University of Ilam, Ilam, Iran

Received: 28 April, 2023 Accepted: 31 July, 2023

#### Extended Abstract

**Introduction and Objectives:** Genetic architecture of sheep reproduction is increasingly gaining scientific interest due to the major impact on sheep production systems. The present study aimed to conduct for identifying the loci associated with total number of progeny born and first lambing age in highly prolific Chios sheep using the 50K arrays.

**Material and Methods:** In this research, phenotypic and genotypic data related to total number of lambs born and first lambing age in lifetime and first lambing age were obtained from 538 highly prolific sheep. The gene set analysis consists basically in three different steps: the assignment of SNPs to genes, the assignment of genes to functional categories and finally the association analysis between each functional category and the phenotype of interest. Genome wide association study was performed with reproductive traits using GEMMA software. Using the *biomaRt2* R package the SNP were assigned to genes if they were within the genomic sequence of the gene or within a flanking region of 15 kb up- and downstream of the gene. Finally, a gene enrichment analysis was performed with the KOBAS platform from online bioinformatics databases for the assignment of the genes to functional categories, the GO, KEGG, DAVID and PANTHER databases were used.

**Results:** In this research, 11 SNP markers on chromosomes 1, 2, 5, 7, 8, 10, 13, 15, 17, 22 and 23 were identified. Also, we identified different sets of candidate genes related to reproductive traits: *QULP1*, *AURKA*, *TEX12*, *DLGIT*, *ACVR1*, *LCP2* and *PPFIA2* in Chios sheep. Some of these genes identified in this research were consistent with previous studies. In the enrichment analysis of gene sets, 19 biological pathways related to reproduction traits were identified. Some of these pathways have important role in fertilization, ovulation rate, estrogen biosynthesis, conception rate, skeletal muscle development, early development of the fetus and puberty.

**Conclusion:** The results of this research can further reveal the role of genetic mechanisms affecting reproductive traits. Considering the conformation of previous genomic region and the new genomic regions identified in this study, the findings of this research can be used more confidently in breeding programs for important economic traits in sheep.

**Keywords:** Candidate gene, Enrichment analysis, Reproduction, Sheep, SNP

**"مقاله پژوهشی"****مطالعه پویش کل ژنومی بر پایه آنالیز مسیبر مرتبط با صفات تعداد کل بره متولد شده و سن اولین زایش در گوسفندان چند قلوزا**حسین محمدی<sup>۱</sup> و محمد شمس‌اللهی<sup>۲</sup>

۱- استادیار گروه علوم دامی، دانشکده کشاورزی و محیط زیست، دانشگاه اراک، اراک، ایران، (نویسنده مسوول: mohammadi13371364@gmail.com)

۲- استادیار گروه علوم دامی، دانشکده کشاورزی، دانشگاه ایلام، ایلام، ایران

تاریخ دریافت: ۱۴۰۲/۲/۸ تاریخ پذیرش: ۱۴۰۲/۵/۹

صفحه: ۷۰ تا ۷۹

**چکیده مسبوط**

**مقدمه و هدف:** بررسی معماری ژنتیکی صفات تولیدمثلی به دلیل این که تأثیر عمده‌ای بر سیستم‌های تولیدی در صنعت گوسفنداری دارد، مورد توجه محققین قرار گرفته است. هدف پژوهش حاضر، شناسایی مناطق ژنومی مؤثر بر تعداد کل نتاج متولد شده و سن اولین بره‌زایی میش‌های نژاد پربازده کاپاس با استفاده از آرایه‌های ژنومی 50K بود.

**مواد و روش‌ها:** در این تحقیق، از اطلاعات ژنوتیپی و فنوتیپی ۵۲۸ رأس از میش‌های با باروری بالا شامل تعداد کل بره متولد شده در طول عمر و سن اولین زایش استفاده شد. آنالیز در سه مرحله شامل تعیین مکان SNPهای معنی‌دار با ژن، ارتباط ژن‌ها به طبقات عملکردی و مسیرهای بیوشیمیایی و در ادامه آنالیز همبستگی بین هر یک از مسیرهای عملکردی با صفات فنوتیپی مورد نظر انجام گرفت. ارزیابی پویش ژنومی برای صفات تولید مثلی در نرم‌افزار GEMMA انجام شد و سپس با استفاده از بسته نرم‌افزاری *biomart2* برنامه R ژن‌های معنی‌داری که در داخل و یا ۱۵ کیلوباز بالا و پایین‌دست نشانگرهای معنی‌دار قرار داشتند، شناسایی شدند. در نهایت، آنالیز غنی‌سازی مجموعه‌های ژنی با برنامه برخط KOBAS با هدف شناسایی عملکرد بیولوژیکی ژن‌های نزدیک به مناطق انتخابی از طریق پایگاه‌های برخط GO، KEGG، Reactome، BioCyc و PANTHER انجام شد.

**یافته‌ها:** در این پژوهش تعداد ۱۱ نشانگر تک نوکلئوتیدی روی کروموزوم‌های ۱، ۲، ۵، ۷، ۸، ۱۰، ۱۳، ۱۵، ۱۷، ۲۲ و ۲۳ شناسایی شدند. در این مناطق ژن‌های *PPFIA2* و *LCP2*، *ACVRI*، *DLG1*، *TEX12*، *AURKA*، *QULP1* که با صفات تولیدمثلی در ارتباط هستند، قرار داشتند. برخی از این ژن‌های شناسایی شده در این تحقیق با مطالعات قبلی هم‌خوانی داشتند. در تحلیل غنی‌سازی مجموعه‌های ژنی، تعداد ۱۹ مسیر بیولوژیکی مرتبط با صفات تولید مثلی شناسایی شدند. برخی از مسیرهای زیستی شناسایی شده در باروری، نرخ تخمک‌اندازی، بیوسنتز استروژن، نرخ آبستنی، رشد و توسعه عضلات اسکلتی، رشد ابتدایی جنین و سن بلوغ نقش مهمی دارند.

**نتیجه‌گیری:** نتایج حاصل از این تحقیق می‌تواند نقش مکانیسم‌های ژنتیکی مؤثر بر صفات تولید مثلی بیشتر آشکار سازد. با توجه به تأیید مناطق قبلی پویش ژنومی و نیز شناسایی مناطق ژنومی جدید در این مطالعه، می‌توان با اطمینان بیشتر از یافته‌های این پژوهش در برنامه‌های اصلاح نژادی برای صفات اقتصادی مهم در گوسفند استفاده نمود.

**واژه‌های کلیدی:** آنالیز غنی‌سازی، تک نوکلئوتیدی، چندشکلی تولیدمثلی، ژن کاندیدا، گوسفند

**مقدمه**

صفات تولیدمثلی، مهم‌ترین صفات مؤثر بر سودآوری اقتصادی سیستم‌های پرورش گوسفند هستند و سالانه به‌طور قابل ملاحظه‌ای تحت تأثیر توانایی میش در تولید بره‌های بیشتر قرار می‌گیرند (Abdoli et al., 2019). از سوی دیگر، تولیدمثلی یک فرآیند پیچیده زیست‌شناختی بوده که علاوه بر شرایط محیطی تحت تأثیر فاکتورهای ژنتیکی نیز قرار می‌گیرد. یکی از مهم‌ترین عواملی که کارایی تولیدمثلی دام را نشان می‌دهد، سن اولین زایش (فاصله تولید تا زایش) است، که نشان دهنده میزان رشد جنس ماده است. سن اولین زایش شامل دوره‌ای است که یک دام برای رسیدن به بلوغ و تولید مثل برای اولین بار نیاز دارد (Xu et al., 2018). سن اولین بره‌زایی یکی دیگر از عوامل مؤثر بر عملکرد تولید مثلی می‌باشد و معمولاً به‌عنوان یک صفت شاخص برای بلوغ جنسی گوسفند ماده در برنامه‌های اصلاح نژادی مورد توجه قرار می‌گیرد. ورود دیر هنگام به دوره تولیدمثلی، ارزش اقتصادی حیوانات را کاهش می‌دهد، زیرا تعداد فرزندان بالقوه تولید شده در طول عمر آنها کاهش می‌یابد. این پارامتر، به‌عنوان اولین پارامتر در برنامه‌های بهداشتی گله در نظر گرفته شده که ممکن است تحت تأثیر اندازه بدن و شروع

فعالیت‌های هورمونی سیستم تولیدمثل واقع شود (Salazar-Carranza et al., 2014).

نژاد مورد استفاده در پژوهش حاضر، یکی از نژادهای گوسفند ممتاز کشور یونان از نظر صفات تولیدی و تولیدمثلی می‌باشد. به‌طوریکه این نژاد دارای تولید شیر بالا و عملکرد تولیدمثلی مطلوبی در سیستم‌های پرورشی صنعتی و نیمه‌صنعتی است (Tsartsianidou et al., 2021).

از طرفی تعداد نتاج متولد شده در هر زایش یکی از مهمترین صفت اقتصادی در صنعت پرورش گوسفند است. در واقع یکی از اجزای بیولوژیکی در تولید، صفات تولیدمثلی می‌باشد، لذا بازده پرورش گوسفند به‌مقدار زیادی تابع توان تولیدمثلی میش‌ها می‌باشد. به‌عبارت دیگر افزایش تعداد بره‌های متولد شده به‌ازای هر میش در یک سال سبب افزایش بازده تولید در پرورش گوسفند خواهد شد (Moosanezhad Khabisi et al., 2022).

در طی سال‌های اخیر، تکنیک‌های بیولوژی مولکولی با سرعت شگرفی توسعه پیدا کرده و موجب گردیده است تا پروژه‌های مختلف توالی‌یابی کل ژنوم، از جمله گوسفند تکمیل شود و متعاقب آن پانل‌های مختلفی از آرایه‌های نانویی با تراکم متفاوت طراحی شده و انجام مطالعات پویش کل ژنومی را امکان‌پذیر نماید. مطالعه پویش ژنومی می‌تواند

اسپریم شناسایی شده بود که از این بین، مسیرهای purine cellular calcium ion nucleotide metabolic process و homeostasis regulation of fertilization نقش مهمی با میزان جنیندگی و تعداد اسپرم متحرک در هر انزال داشتند (Mohammadi et al., 2022).

همچنین در مطالعه‌ی پویش ژنومی بر مبنای مسیر با استفاده از تفسیر مجموعه‌های ژنی روی صفات مرتبط با تعداد بره متولد شده در هر زایش متعلق به نژادهای مختلف گوسفندان جهان انجام شده است. نتایج به‌دست آمده از آنالیز مسیر، منجر به شناسایی ۳۰ طبقه مختلف عملکردی هستی‌شناسی ژن و مسیرهای زیستی KEGG معنی‌دار مرتبط شامل Oxytocin, TGF- $\beta$  signaling pathway, Estrogen signaling pathway, signaling pathway, Insulin signaling و Prolactin signaling pathway pathway و ژن‌های کاندیدای *DHCR24*, *BMP5*, *PLCB1*, *SMAD1* و *ESR2*, *ESR1*, *BMP1B*, *KCNMA1* و *EGFR*, *BMP7*, *PTGS2*, *INSR*, *SMAD2* شده بود (Khaltabadi Farahani et al., 2020).

در پژوهشی دیگر، مطالعه پویش ژنومی بر مبنای مسیر با استفاده از تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی روی خصوصیات لخته‌شدگی شیر، تولید پنیر و استحکام دل‌مه انجام شد. براساس این نتایج خروجی‌ها براساس مسیر، منجر به شناسایی ۲۱ طبقات مختلف عملکردی هستی‌شناسی ژن و ۱۷ مسیر بیوشیمیایی KEGG معنی‌دار مرتبط با این صفات شد که شامل ژن‌های کاندیدای *CSN1S1*, *LAP3*, *Dadousis et al.*, *HSF1* و *RPL8*, *DGAT1* بودند (2017).

داده‌های مورد استفاده در این پژوهش، قبلاً با هدف آنالیز پویش کل ژنومی براساس مدل‌های خطی تک نشانگری با در نظر گرفتن ضریب تصحیح بنفرونی برای تعیین آستانه معنی‌داری و جلوگیری از خطای نوع اول مرتبط با صفات تولیدمثلی استفاده شده بود. نتایج این تحقیق نشان داد ژن‌های کاندیدای *IL5RA*, *TRNT1*, *CRBN*, *CFH* و *THEGL* ارتباط معنی‌داری با صفات مورد بررسی دارند (Tsartsianidou et al., 2023). لذا هدف از انجام پژوهش حاضر، تجزیه و تحلیل مجموعه‌های ژنی و مسیرهای زیستی مرتبط با صفات تولیدمثلی در گوسفندان پربازده کایاس براساس پویش کل ژنوم بر مبنای مسیر می‌باشد.

### مواد و روش‌ها

در پژوهش حاضر از اطلاعات ژنوتیپی و فنوتیپی تعداد ۵۳۸ رأس از گوسفندان غیر خویشاوند مربوط به نژاد چند قلوهای کایاس<sup>۲</sup> استفاده شد. نمونه‌ها با استفاده از آرایه‌های شرکت ایلومینا Ovine SNP50 Bead Chip برای ۵۴۲۴۱ جایگاه نشانگری تعیین ژنوتیپ شده بودند. داده‌های مورد استفاده این پژوهش در پایگاه برخط Mendeley (<https://data.mendeley.com/datasets/nvjn8mvtwg/>) در دسترس است. در این پژوهش از اطلاعات فنوتیپی مرتبط با صفات تولیدمثلی شامل تعداد کل نتاج متولد شده به‌ازای هر

به شناسایی دقیق ژن‌های مرتبط با صفات اقتصادی کمک کرده و یافته‌های آن برای انتخاب به کمک نشانگر مفید باشد (Johnston et al., 2011).

در مطالعات پویش کل ژنومی به‌طور معمول از ضریب تصحیح بنفرونی برای تعیین آستانه معنی‌داری و جلوگیری از خطای نوع اول استفاده می‌گردد. یکی از ایرادات تحقیقات مطالعات پویش ژنومی در نظر گرفتن آستانه معنی‌داری برای جلوگیری از بروز خطای نوع اول است. در حالی‌که پرهیز از خطای نوع اول سبب افزایش خطای نوع دوم و در نظر نگرفتن SNP‌های دارای اثر معنی‌دار پایین‌تر از آستانه می‌شود (Marjanovic & Calus, 2020).

یک جایگزین مناسب برای حل این مشکل، انجام مطالعات پویش کل ژنومی بر مبنای مسیر<sup>۱</sup> با استفاده از تجزیه و تحلیل مجموعه‌های ژنی است. در واقع در این روش به‌جای انجام تجزیه برای یک SNP یا یک ژن، ارتباط بین صفت مورد مطالعه و واریانت‌های ژنتیکی در یک دسته یا گروه ژنی که به‌طور عملکردی با هم مرتبط هستند بررسی می‌شود. به عبارت دیگر ارتباط معنی‌دار بین یک مجموعه ژنی با فنوتیپ، مورد آزمون قرار می‌گیرد (Wang et al., 2011).

اخیراً پژوهشی با عنوان مطالعه ارتباط‌سنجی در سطح ژنوم برپایه تجزیه و تحلیل غنی‌سازی مجموعه ژنی برای شناسایی جایگاه‌های ژنی مؤثر بر نرخ دو قلوهای در ۹۶ گوسفند بلوچی که براساس نرخ دو قلوهای بالا و پایین انتخاب شده بودند، موفق به شناسایی ژن‌های کاندیدا شامل *CTH*, *LRRC40*, *LDHB*, *PTGER3*, *SRSF11*, *ANKRD13C* و *NTRK2* روی کروموزوم‌های شماره ۱، ۳، ۱۰، ۱۵ و ۲۵ شدند. همچنین مسیرهای زیستی defense response, cell adhesion و cell junction گزارش شده بودند، نتایج نشان داد روش پویش ژنومی بر مبنای مسیر کارایی بالاتری برای یافتن مناطق ژنومی و درک بهتری از معماری ژنتیکی صفت تعداد نتاج متولد شده نسبت به آنالیز پویش ژنومی بر پایه تک نشانگری دارد (Esmaeili-Fard et al., 2021). مطالعه‌ی مروری با هدف شناسایی ژن‌های کاندیدای و مسیرهای زیستی مرتبط با تعداد بره متولد شده در گوسفند براساس آنالیز مسیر انجام شده است، از میان ۲۱ ژن کاندیدای شناسایی شده ژن‌های *CCL2* و *FLT1* و از میان ۲۰ مسیر زیستی مرتبط با تعداد نتاج، مسیر زیستی negative regulation of vascular endothelial cell proliferation بیشترین ارتباط را با تعداد نتاج متولد شده در هر زایش را داشتند (Ghiasi & Abdollahi-Arpanahi, 2021).

در مطالعه پویش کل ژنومی در گاوهای نر هلشتاین چینی با هدف شناسایی ژن‌های کاندیدا و مسیرهای زیستی مرتبط با خصوصیات اسپرم انجام شده است. نتایج این پژوهش نشان داد ژن‌های کاندیدای *SLC24A*, *SEPT12*, *SPEF2*, *RPM1*, *MYH3*, *DSCAML1*, *ERBB4*, *ETNK1*, *BSP3* و *RPM2* و *DMRT1* و *PPAR $\gamma$*  با صفات کمی و کیفی اسپرم مرتبط بودند. همچنین در آنالیز غنی‌سازی مجموعه ژنی، تعداد ۱۵ مسیر هستی‌شناسی ژنی با صفات

میش در طول عمر و سن میش در اولین بهره‌زایی استفاده  
گردید. جدول ۱ آمار توصیفی صفات مورد بررسی مرتبط با  
جدول ۱- آمار توصیفی صفات تولیدمثلی بررسی شده در گوسفند کاپاس  
صفات تولیدمثلی در این پژوهش را نشان می‌دهد.

Table 1. Descriptive statistics for reproductive traits in Chios sheep

Trait صفت	Number of records تعداد رکورد	Mean میانگین	SD (%) انحراف معیار	Min حداقل	Max حداکثر
تعداد کل نتاج متولد شده به‌ازای هر میش Total number of lambs born per ewe	538	4.31	3.20	1.00	25.00
سن میش در اولین بهره‌زایی Age at 1st lambing (days)	538	595.88	232.05	345.00	1454.00

نشانه‌گر SNP موردنظر در داخل آن ژن و یا ۱۵kb بالادست یا پایین‌دست آن ژن قرار داشت، ارتباط داده شدند (Clancey et al., 2019). ۲- ارتباط ژن‌ها به طبقات عملکردی و مسیرهای بیوشیمیایی؛ جهت تعیین طبقات عملکردی ژنی و مسیرهای متابولیکی و تنظیمی ژن‌های معنی‌دار از پایگاه‌های اطلاعاتی شامل GO (<http://www.geneontology.org>)، KEGG Panther (<http://www.genome.jp/kegg>)، BioCyc (<http://www.pantherdb.org>) و Reactome (<http://www.reactome.org>) جهت تعیین طبقات عملکردی و مسیرهای بیوشیمیایی استفاده شدند. در این مرحله فرض بر این است که ژن‌هایی که در یک طبقه عملکردی خاص از سه مجموعه ژنی هستی شناسی قرار دارند می‌توانند به‌عنوان یک گروه از ژن‌هایی که دارای ویژگی‌های خاص هستند، قلمداد نمود. ۳- پویس کل ژنومی بر پایه آنالیز مسیر؛ ارتباط‌های معنی‌دار مسیرهای عملکردی مرتبط با صفات تولیدمثلی با استفاده از توزیع فوق هندسی<sup>۱</sup> و آماره Fisher's exact test مورد آزمون قرار گرفت. تجزیه غنی‌سازی مجموعه ژنی با استفاده از بسته نرم‌افزاری KOBAS (Bu et al., 2021) انجام گرفت. در نهایت برای تفسیر بهتر عملکرد ژن‌های به دست آمده از پایگاه‌های اطلاعاتی آنلاین GeneCards (<http://www.genecards.org>) و UniProtKB (<http://www.uniprot.org>) استفاده شدند.

### نتایج و بحث

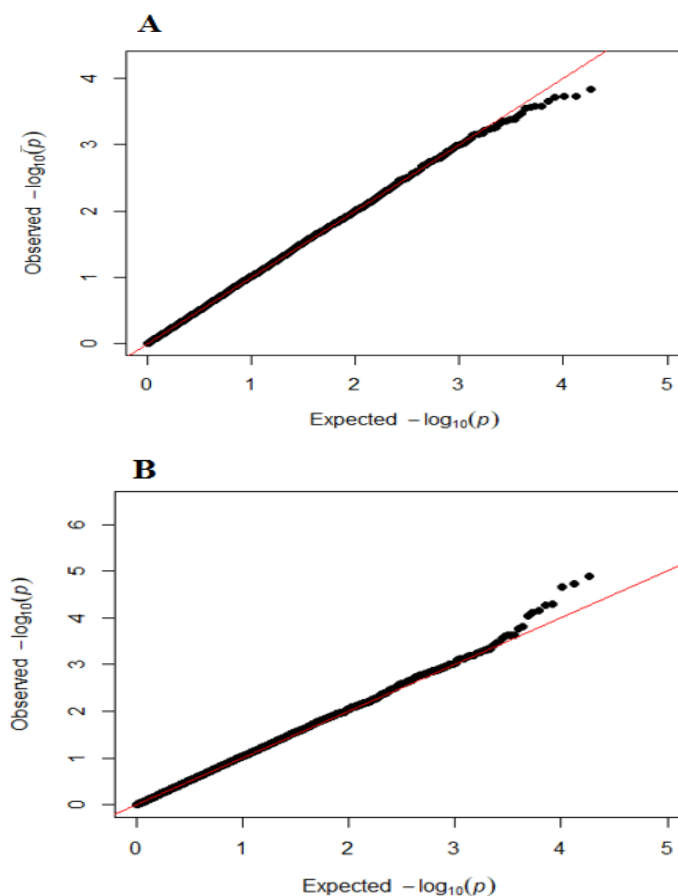
پلات‌های Q-Q و منهن مرتب با صفات تعداد کل نتاج متولد شده و سن اولین بهره‌زایی در شکل ۱ و ۲ ارائه شده است. چندین روش برای تخمین کنترل تورم ژنومیک (λ) وجود دارد که در این تحقیق روش تخمین گر میانه اجرا شد. فاکتور تورم کنترل جمعیتی بزرگتر از ۱ (λ > 1) وجود لایه‌بندی جمعیتی و یا خطای تعیین ژنوتیپ را نشان می‌دهد. فاکتور لامبدا از طریق تجزیه و تحلیل پیوستگی در PLINK برای صفات مورد مطالعه محاسبه شد که برای صفات تعداد کل نتاج متولد شده و سن اولین بهره‌زایی به ترتیب ۱/۰۳۴ و ۱/۰۵۱ بود که همان‌طور که مشاهده می‌شود تقریباً برابر با ۱ بودند و نشان دهنده عدم وجود لایه‌بندی جمعیتی برای اجرای تجزیه و تحلیل‌های پویس کل ژنومی بود. در مطالعاتی که لامبدا کوچکتر از ۱/۱ یا مساوی ۱ است نیازی به تصحیح اثرات ساختار جمعیتی نیست (Hinrichs et al., 2009).

جهت اطمینان از کیفیت داده‌های حاصل از تعیین ژنوتیپ در تجزیه نهائی، ویرایش داده‌ها و مراحل مختلف کنترل کیفیت با استفاده از نرم‌افزار PLINK (نسخه ۱/۹) روی داده‌های اولیه تعیین ژنوتیپ شده انجام شد (Purcell et al., 2007). در نخستین مرحله نمونه‌هایی با بیش از ده درصد ژنوتیپ از دست رفته از آنالیزهای بعدی حذف شدند. سپس از دو عامل فراوانی آلی نادر (MAF) و نرخ تعیین ژنوتیپ (Call rate) برای هر نشانه‌گر SNP محاسبه شده و SNP‌هایی که در مجموع دام‌ها دارای MAF و Call rate به ترتیب کمتر از دو درصد و ۹۰ درصد بودند از مراحل بعدی آنالیزها کنار گذاشته شدند. سپس نشانه‌گرهای باقیمانده که به‌شدت از تعادل هاردی-واینبرگ انحراف داشتند به‌عنوان معیار دیگری از خطای تعیین ژنوتیپ از آنالیز نهایی کنار گذاشته شدند. همچنین SNP‌هایی که روی کروموزوم جنسی یا موقعیت ناشناخته قرار داشتند از آنالیزهای بعدی کنار گذاشته شدند. در نهایت بعد از کنترل کیفیت تعداد ۵۲۸ رأس دام و تعداد ۴۵۹۳۶ SNP برای آنالیزهای پویس کل ژنومی بر پایه تجزیه و تحلیل غنی‌سازی مجموعه ژنی باقی ماندند. همچنین در این تحقیق برای بررسی وجود یا عدم وجود لایه‌بندی جمعیتی با استفاده از روش کنترل ژنومیک از طریق فاکتور تورم کنترل جمعیتی (λ) در نرم‌افزار PLINK و ترسیم پلات‌های Q-Q در نرم‌افزار SNPEVWin\_3.2 بررسی شد. جهت ارتباط فنوتیپ‌ها با ژنوتیپ‌ها از نرم‌افزار GEMMA (Zhou & Stephens, 2012) نسخه ۰/۹۸ استفاده شد. مدل مورد استفاده بر پایه مدل خطی مختلط به شکل زیر بود:

$$y = Wa + x\beta + Zu + e$$

که  $y$ : بردار مشاهدات فنوتیپی،  $a$  بردار اثرات ثابت (شامل سن مادر (۲-۵ سالگی)، سال تولد (۲۰۰۵-۲۰۱۹) و شکم زایش (۱-۴)،  $x$ : بردار ژنوتیپ‌های نشانه‌گر تحت آزمون،  $u$  بردار اثرات پلی‌ژنیک،  $e$  بردار اثر باقی‌مانده‌های تصادفی و  $\beta$ ،  $W$  و  $Z$  ماتریس‌های طرح ارتباط دهنده.

آنالیز پویس ژنومی بر پایه تجزیه و تحلیل مجموعه‌های ژنی در سه مرحله انجام گردید: ۱) تعیین مکان SNP‌های معنی‌دار با ژن ۲) ارتباط ژن‌ها به طبقات عملکردی و مسیرهای بیوشیمیایی ۳) پویس کل ژنومی بر پایه آنالیز مسیر ۱- تعیین مکان SNP‌ها با ژن‌ها: SNP‌هایی که مقدار  $P$ -value آنها کمتر از ۰/۰۵ بود با استفاده از بسته نرم‌افزاری  $biomaRt2$  (Durinck et al., 2009) در محیط R و با استفاده از ژنوم مرجع گوسفند ( $Oar_v3.1$ ) به ژن‌هایی که

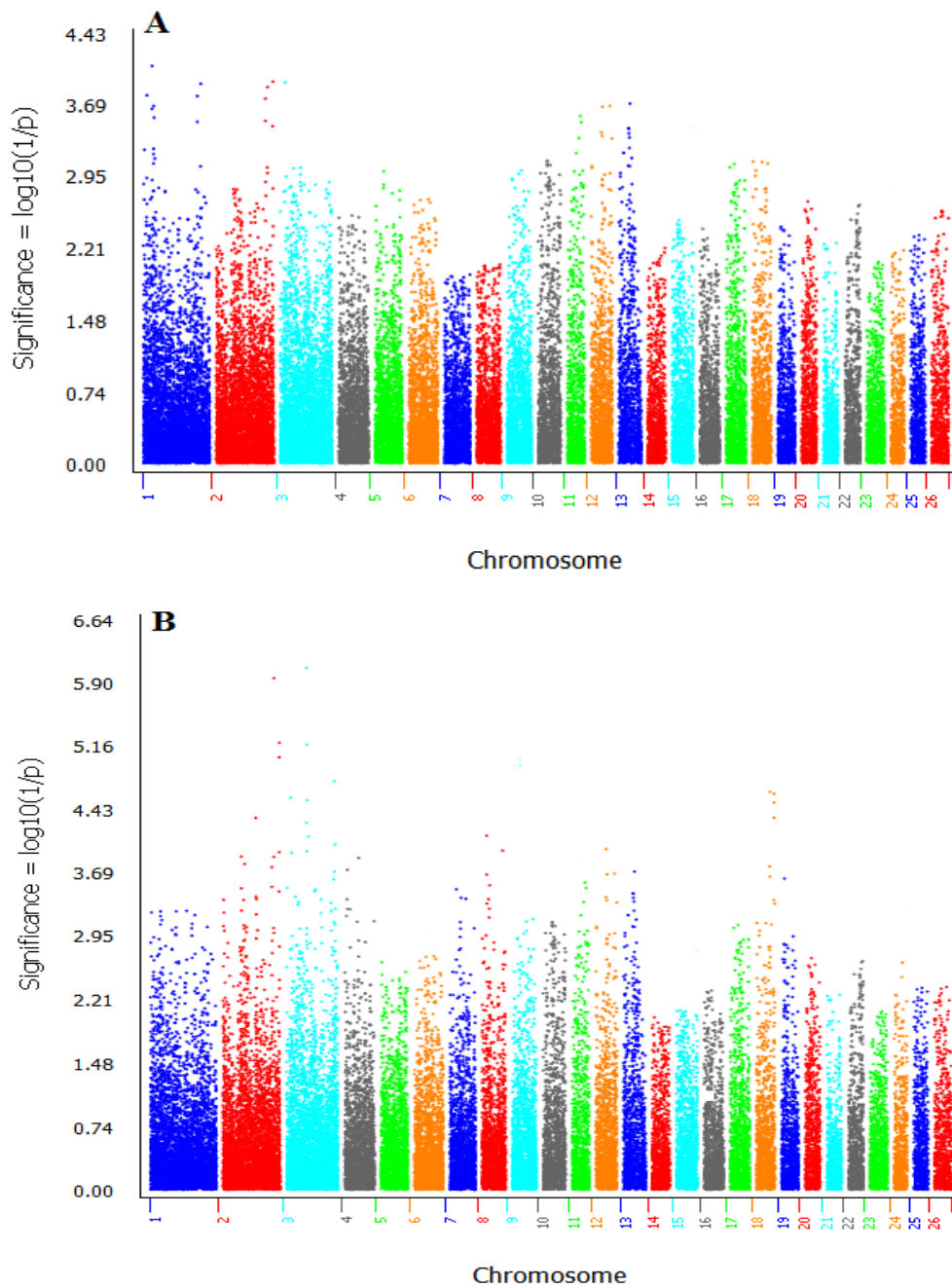


شکل ۱- پلات‌های Q-Q برای صفات تعداد کل نتاج متولد شده (A) و سن اولین بره‌زایی (B) در گوسفند کایاس  
Figure 1. Plots Q-Q for (A) total number of lambs born and (B) age at first lambing traits in Chios sheep.

براساس تحلیل هستی‌شناسی ژن (GO)، فرآیندهای زیستی مختلفی برای ژن‌های مؤثر بر تعداد کل نتاج متولد شده مشاهده شد که مطابق با نتایج به‌دست آمده از تحقیقات پیشین بود. جزئیات کامل ترم‌های هستی‌شناسی به‌همراه اسامی ژن‌های کاندیدا در جدول ۲ ارائه شده است. نتایج حاصل از این تحلیل نشان دهنده این است که ژن‌های *Ovarian follicle development* و ژن *DLG1* با فرآیند *Calcium channel activity* مرتبط با تعداد کل نتاج متولد شده با بیشترین مقدار ضریب غنی‌سازی هستند، در گوسفندان نژاد کایاس مشاهده شد. برخی از ژن‌های این فرآیند زیستی در مطالعات مختلف بررسی ژن‌های کاندیدای مرتبط با تولیدمثل به‌خصوص چند قلو‌زایی در ارتباط می‌باشند.

#### شناسایی ژن‌های کاندیدا در مناطق ژنومی

تعداد ۲۶۱۰۹ عدد ژن از ۲۷۰۵۴ ژن حاشیه‌نویسی شده در گوسفند به‌وسیله نشانگرهای SNP پوشش داده شد که در این میان تعداد ژن‌های معنی‌دار در نژاد مورد بررسی ۳۴۰۶ ژن بودند، یعنی حداقل یک نشانگر با P-value کمتر از ۰/۰۵ در داخل و یا در بالا یا پایین‌دست این ژن تا فاصله ۱۵ kb قرار گرفت. این ژن‌ها به‌عنوان ژن‌های معنی‌دار مرتبط با صفات تعداد کل نتاج متولد شده و سن اولین بره‌زایی برای تجزیه GSEA-SNP انتخاب شدند. تعداد مجموعه‌های ژنی حاصل از پایگاه‌های داده‌ای مختلف شامل ۲۶۴ طبقات هستی‌شناسی به دست آمد و ۱۷ مسیر زیستی KEGG مشاهده شد. مسیرهای که بیشتر از سه ژن و کمتر از ۳۰۰ ژن داشتند گزارش شدند.



شکل ۲- پلات منهنن برای صفات تعداد کل نتاج متولد شده و سن اولین بره‌زایی در گوسفند کایاس. محور X مکان آل‌های ژنوتیپی روی کروموزوم‌ها و محور Y منفی لگاریتم بر مبنای ۱۰ ارزش P-value

Figure 2. Manhattan plot for (A) total number of lambs born and (B) age at first lambing traits in Chios sheep. X axis, genotypes alleles positions on chromosomes, Y axis, -Log10 P-value

ژن *AURKA* گزارش شده بود (Wang et al., 2021). ژن کاندیدای *DLGI* مرتبط با محرک تقسیم میتوز سلول‌های گرانولوزا بوده و تولید پروتسترون تخمدانی می‌شود و یک ژن درگیر در لانه‌گزینی بوده و در اپیتلیوم رحم چسبندگی آغازین را بر عهده دارد. در مطالعه پویش کل ژنومی جهت شناسایی ژن‌های کاندیدای مرتبط با تعداد نتاج متولد شده در هر زایش در گوسفندان نژاد Pelibuey، ژن کاندیدای *DLGI* گزارش

ژن کاندیدای *AURKA* در مطالعه پویش کل ژنومی برپایه آنالیز غنی سازی مجموعه‌های ژنی در گوسفندان نژاد بلوچی مرتبط با دو قلو زایی در مسیر زیستی phosphotransferase activity و transferase activity گزارش شده است (Esmaili-Fard et al., 2021). همچنین در مطالعه‌ی پویش کل ژنومی مرتبط با تعداد بزغاله‌های متولد شده در هر زایش نیز ژن کاندیدای

همچنین این ژن تنظیم کننده کلیدی سیگنال‌دهی فاکتور رشد تغییر دهنده بتا (TGF-B) با میانجی‌گری *LRP1* به‌شمار می‌رود و نقش مؤثری در تنظیم سیگنال‌های TGF-B در سلول‌های تخمدان دارد.

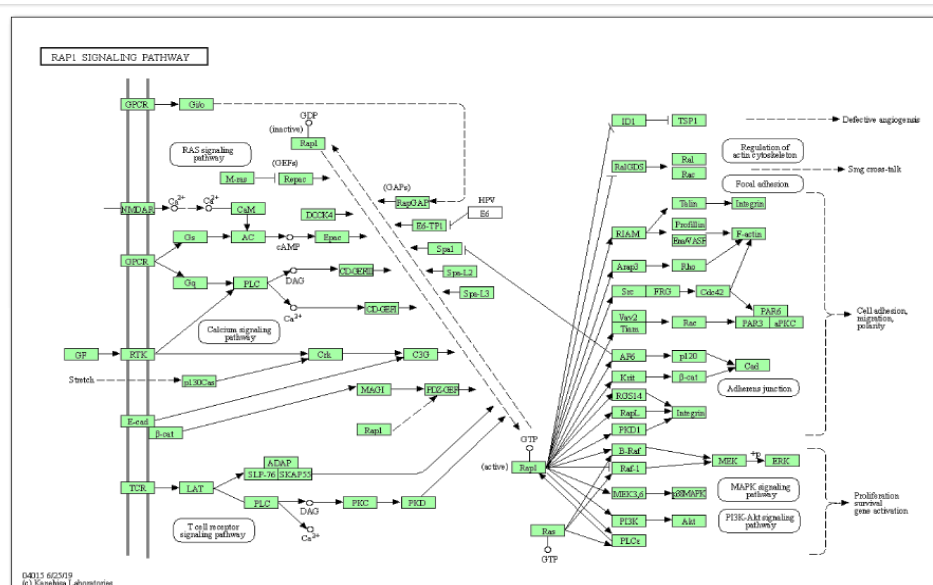
شده است (Hernández-Montiel et al., 2020). همچنین نتایج تحقیق حاضر، دو نشانگر روی کروموزوم ۲ شناسایی شد که با چندقلوژی مرتبط بودند. تجزیه بیشتر در پایگاه داده نشان داد که این نشانگرها با ژن *QULP1* در ارتباط هستند.

جدول ۲- تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی معنی‌دار ( $p < 0.05$ ) مرتبط با صفات تولیدمثلی

P-adjust P-ارزش تصحیح شده	Candidate genes ژن‌های کاندیدا	No. genes in the GO term تعداد ژن موجود در مسیر	Term GO نام مسیر هستی‌شناسی	GO ID طبقات هستی‌شناسی
Biological process				
فرآیند زیستی				
0.014	<i>SCN1A, PTGIS</i>	19	Reproductive structure development	GO:0048608
0.011	<i>HSD17B12</i>	9	Embryonic skeletal system development	GO: 0048706
0.015	<i>HOXC9, HOXC11</i>	8	Embryonic skeletal system morphogenesis	GO: 0048704
0.026	<i>DLG1, INHBA, AURKA</i>	16	Ovarian follicle development	GO: 0001541
0.001	<i>FGF2</i>	8	Regulation of protein processing	GO:0070613
0.029	<i>DGKE, PPFIA2</i>	7	Positive regulation of synapse assembly	GO: 0032926
0.032	<i>MAG11, MAG12</i>	12	Negative regulation of activin receptor signaling pathway	GO:0055074
0.045	<i>TEX11, TEX12</i>	9	Histone methylation	GO:0031062
0.014	<i>CHRN2, GRIN3A,</i>	11	Actin polymerization or depolymerization	GO:0097305
0.005	<i>ING2, QULP1</i>	6	Calcium channel activity	GO:0035556
0.008	<i>ACVR1, CACNB2, CACNA2D1</i>	13	Steroid hormone mediated signaling pathway	GO:0043401
	<i>CHTOP, SNW1</i>	11	Regulation of acute inflammatory response	GO: 0002673
0.015	<i>ACVR1, ABCA1, ABCG1</i>	19	Estrogen response element binding	GO:0034056
0.031	<i>PIK3CD, PLCG1</i>	21	Defense response	GO:0051607
0.0015	<i>ARHGEF4, DGKI,</i>		Intracellular signal transduction	GO:0035556
Cellular component				
اجزای سلولی				
0.006	<i>CDH13, FRMD4A</i>	6	Cell junction	GO:0030054
0.003	<i>BMP4, PDGFB</i>	8	Cation channel complex	GO:0034703
KEGG Pathways				
مسیرهای KEGG				
0.008	<i>FGF2, FGF6, LCP2</i>	13	Rap1 signaling pathway	oas04015
0.012	<i>PPP3R2, ITPR1</i>	7	Calcium signaling pathway	oas04020

۱- ژن‌های کاندیدا مرتبط با صفت به صورت پر رنگ نشان داده شده‌اند.

1. Candidate genes associated with trait are shown in bold.



شکل ۳- مسیر سیگنالی Rap1 و ژن‌های کاندیدی مرتبط با صفت تولیدمثلی (پایگاه داده KEGG)

Figure 3. Rap1 signaling pathway and candidate genes related to reproductive traits (KEGG database).

فرآیندهای تنظیمی را تعیین می‌کنند که بین ژنوم نر و ماده متفاوت است (Bonde et al., 2016). ژن کاندیدی *TEX12* از بین ژن‌های موجود در این مسیر بیشترین ارتباط با باروری را داشت (Boroujeni et al., 2018).

مسیر مهم معنی‌دار دیگر مرتبط با تعداد کل نتاج متولد شده مسیر بیولوژیکی Histone methylation بود که یک فرآیند بیوشیمیایی ضروری و حیاتی است که در تنظیم گامتوزن، رشد جنین و جفت نقش دارند (Menezo et al., 2016). فرآیندهای متیلاسیون هیستون‌ها ویژگی‌های

مهم‌ترین مسیرهای مؤثر بر فرآیندهای تخمدان معرفی شده است. مسیر ذکر شده در حمایت از سوخت و ساز تخمک در حال رشد، یون‌ها برای تنظیم pH تخمک و cGMP مورد نیاز برای نگهداری و حفظ تخمک در حالت میوزی نقش دارد (Genecards). مطالعه‌ی پویس کل ژنومی در گاوهای بومی کلمبیا (Romosinuano and Costeño con Cuernos) با عملکرد تولیدمثلی بالا ارتباط معنی‌داری بین ژن *FRMD4A* را با سن اولین گوساله‌زایی و فاصله گوساله‌زایی گزارش شده است (Fernández et al., 2019).

تحلیل مسیرهای زیستی KEGG نشان‌دهنده این است که ژن‌های *FGF2* و *LCP2* به‌طور معنی‌داری با مسیر زیستی Rap1 signaling pathway در ارتباط می‌باشند (شکل ۳). این مسیر زیستی نقش مهمی در مکانیسم تولیدمثلی به‌خصوص سن بلوغ و رشد بدن دارد. در شکل ۳ مسیر سیگنالی Rap1 و ژن‌های مربوطه شناسایی شده از پایگاه برخط KEGG را نشان می‌دهد. ژن *FGF2* نقش کلیدی در رشد و توسعه سلول‌های عضلات بدن در دوره جنینی دارد (Mohammadabadi et al., 2021). ژن *LCP2* به‌عنوان یک فاکتور التهابی، نقش واسطه‌ای بین بافت چربی و بافت هدف دارد. از این رو می‌تواند موجب افزایش رهاسازی تخمک در اویداکت گردد. ارتباط معنی‌داری بین چندشکلی در ژن کاندیدای *LCP2* با رشد جنین در دوره آبستی در بز گزارش شده است (Luo et al., 2021).

### نتیجه‌گیری کلی

بررسی مناطق ژنومی شناسایی شده با استفاده از پایگاه داده BioMart، GeneCards و UniProtKB نشان داد که این مناطق با صفات تعداد بره متولد شده و سن اولین زایش مرتبط هستند. با توجه به عملکرد ژن‌های کاندیدای شناسایی شده در مسیرهای زیستی این پژوهش، به‌نظر می‌رسد می‌توان کارایی روش تجزیه و تحلیل غنی‌سازی مجموعه ژنی برای پویس ژنومی صفات تولیدی اقتصادی را نیز مورد تأیید قرار داد. در نهایت مسیرهای شناسایی شده در این تحقیق می‌تواند ما را به سمت ژن‌های مؤثر بر صفت مهم تولیدمثلی هدایت کند و بررسی بیشتر نواحی مهم ژنومی شناسایی شده با استفاده از آزمون‌های آزمایشگاهی مختلف می‌تواند در تأیید نتایج به‌دست آمده در این پژوهش مؤثر باشند. از طرفی در این تحقیق به‌دلیل عدم دسترسی به رکوردهای فنوتیپی و اطلاعات ژنوتیپی مرتبط با صفت سن اولین زایش در گوسفندان بومی کشور، استفاده از نتایج این تحقیق در گوسفندان بومی کشور نیاز به مطالعات بیشتر دارد تا در این جمعیت‌ها نیز تأیید شوند.

ژن *ACVRI* به‌علت مؤثر بودن آن در مسیرهای پاسخ BMP در طول رشد فولیکولی نقش کلیدی دارد. *ACVRI* از اعضا و گیرنده‌های ابر خانواده TGF به‌عنوان تنظیم‌کننده‌های مهم توسعه سیستم تولیدمثل و هموستاز گزارش شده است (Zhang, 2017). ارتباط ژن *ACVRI* با تعداد بره متولد شده در نژاد گوسفند بومی چین گزارش شده است (McBride et al., 2012). این ژن‌ها در دو مسیر مختلف فرآیند زیستی شامل Steroid hormone mediated signaling و Estrogen response element binding pathway مشاهده شد.

تجزیه و تحلیل هستی‌شناسی مرتبط با سن اولین بره‌زایی در جدول ۲ ارائه شده است. نتایج حاصل از این تحلیل نشان‌دهنده این است که ژن‌های *HSD17B12* با فرآیند embryonic skeletal system development و *PPFIA2* با فرآیند positive regulation of synapse assembly در ترم فرآیندهای زیستی بوده که مرتبط با صفت سن اولین بره‌زایی هستند. برخی از ژن‌های این فرآیند زیستی در مطالعات مختلف بررسی ژن‌های کاندیدای مرتبط با صفات تولید مثل به خصوص باروری و سن اولین زایش در ارتباط می‌باشند.

ژن *HSD17B12* نقش کلیدی در بیوسنتز استروژن دارد، به‌طوری‌که ژن *HSD17B12* کد کننده آنزیم تبدیل استرون به استرادیول است و همچنین در سنتز اسید آراشیدونیک نقش کلیدی داشته که برای زنده‌مانی جنین لازم است (Genecards). در مطالعه پویس کل ژنوم، کاندیدای *HSD17B12* ارتباط معنی‌داری با صفت تولیدمثلی نرخ آبستی در گاوهای هلشتاین گزارش شده است (Cochran et al., 2013). همچنین در تحقیق دیگر گاوهای هلشتاین با هدف شناسایی ژن‌های کاندیدای مرتبط با نرخ آبستی، تلیسه‌ها به‌عنوان صفت شاخص برای پیش‌بینی شایستگی ژنتیکی عملکرد تولیدمثل ژن *HSD17B12* را گزارش کردند (Ortega et al., 2016). ژن *PPFIA2* نیز مهم دیگری مرتبط با تولیدمثل است که تأثیر این ژن با نرخ باروری در مطالعات اخیر مشاهده است. در مطالعه‌ی پویس کل ژنومی گاوهای نلور در ارتباط با سن اولین گوساله‌زایی ژن *PPFIA2* روی کروموزوم ۵ گزارش شده است (Mota et al., 2020). ژن *PPFIA2* نقش مهمی در بسیاری از فرآیندهای تولیدمثلی و توسعه غدد پستانی را بر عهده دارد (Genecards). عملکردهای مولکولی و اجزای سلولی مشاهده شده شامل ژن‌های *SCNIA* با فرآیند Reproductive structure development و ژن *FRMD4A* با فرآیند سلولی Cell-cell junction بود. مسیر Cell-cell junction یکی از

### منابع

- Abdoli, R., Zamani, P., Mirhosseini, S. Z., Ghavi Hossein-Zadeh, N., & Almasi, M. (2019). Genetic parameters and trends for litter size in Markhoz goats. *Revista Colombiana de Ciencias Pecuarias*, 32(1), 58-63.
- Bezdicsek, J., Stadnik, L., & Louda, F. (2015). The influence of inbreeding depression on reproduction traits-age at first calving in dairy cows. *Bulgarian Journal of Agricultural Science*, 21(3), 680-686.
- Boroujeni, P.B., Sabbaghian, M., Totonchi, M., Sodeifi, N., Sarkardeh, H., Samadian, A., Sadighi-Gilani M. A., & Gourabi, H. (2018). Expression analysis of genes encoding TEX11, TEX12, TEX14 and

- TEX15 in testis tissues of men with non-obstructive azoospermia. *JBRA Assist Reproduction*, 22(3), 185-192
- Bonde, J. P., Flachs, E. M., Rimborg, S., Glazer, C. H., Giwercman, A., Ramlau-Hansen, C. H., & Bräuner, E. V. (2016). The epidemiologic evidence linking prenatal and postnatal exposure to endocrine disrupting chemicals with male reproductive disorders: A systematic review and meta-analysis. *Human Reproduction Update*, 23, 104-125.
- Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., Wu, Y., Zhao, L., & Guo, J. (2021). KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Research*, 49, 317-325.
- Clancey, E., Kiser, J. N., Moraes, J. G. N., Dalton, J. C., Spencer, T. E., & Neibergs, H. L. (2019). Genome-wide association analysis and gene set enrichment analysis with SNP data identify genes associated with 305-day milk yield in Holstein dairy cows. *Animal Genetics*, (3), 254-258.
- Cochran, S. D., Cole, J. B., Null, D. J., & Hansen, P. J. (2013). Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. *BMC Genetic*, 7(14): 49.
- Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/bioconductor package biomaRt. *Nature Protocols*, 4, 1184-1191.
- Dadousis, C., Pegolo, S., Rosa, G. J. M., Gianola, D., Bittante, G., & Cecchinato, A. (2017). Pathway-based genome-wide association analysis of milk coagulation properties, curd firmness, cheese yield, and curd nutrient recovery in dairy cattle. *Journal of Dairy Science*, 100, 1223-1231.
- Esmaili-Fard, S. M., Gholizadeh, M., Hafezian, S. H., & Abdollahi-Arpanahi, R. (2021). Genome-wide association study and pathway analysis identify NTRK2 as a novel candidate gene for litter size in sheep. *PLoS One*, 16(1), e0244408.
- Fernández, J. C., Pérez, J. E., Herrera, N., Martínez, R., Bejarano, D., & Rocha J. F. (2019). Genomic association study for age at first calving and calving interval in Romosinuano and Costeño con Cuernos cattle. *Genetics and Molecular Research*, 18(2), 1-13.
- Gholizadeh, M., & Esmaili-Fard, S. M. (2022). Multi-population joint genome-wide association study to detect genomic regions associated with litter size in sheep. *Animal Production Research*, 11(3), 15-26. In Persian.
- Ghiasi, H., & Abdollahi-Arpanahi, R. (2021). The candidate genes and pathways affecting litter size in sheep. *Small Ruminant Research*, 205, 106546.
- Hernández-Montiel, W., Martínez-Núñez, M. A., Ramón-Ugalde, J. P., Román-Ponce, S. I., Calderón-Chagoya, R., & Zamora-Bustillos, R. (2020). Genome-Wide Association Study Reveals Candidate Genes for Litter Size Traits in Pelibuey Sheep. *Animals (Basel)*, 10(3), 434.
- Hinrichs, A. L., Larkin, E. K., & Suarez, B. K. (2009). Population Stratification and Patterns of Linkage Disequilibrium. *Genetic Epidemiology*, 33, 88-92.
- Khaltabadi Farahani, A. H., mohammadi, H., & Moradi, H. (2020). Gene set enrichment analysis using genome-wide association study to identify genes and pathways associated with litter size in various sheep breeds. *Animal Production*, 22(3), 325-335. In Persian.
- Johnston, S. E., McEwan, J. C., Pickering, N. K., Kijas, J. W., Beraldi, D., Pilkington, J. G., Pemberton, J. M., & Slate, J. (2011). Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, 20, 2555-2566.
- Luo, N., Cheng, W., Zhou, Y., Gu, B., Zhao, Z., & Zhao, Y. (2021). Screening candidate genes regulating placental development from trophoblast transcriptome at early pregnancy in Dazu black goats (*Capra hircus*). *Animals*, 11, 2132.
- Marjanovic, J., & Calus. M. P. L. (2020). Factors affecting accuracy of estimated effective number of chromosome segments for numerically small breeds. *Journal of Animal Breeding and Genetics*, 138, 151-160.
- McBride, D., Carré, W., Sontakke, S. D., Hogg, C. O., & Law. A. (2012). Identification of miRNAs associated with the follicular-luteal transition in the ruminant ovary. *Reproduction*, 144, 221-233.
- Menezo, Y. J., Silvestris, E., Dale, B., & Elder, K. (2016). Oxidative stress and alterations in DNA methylation: Two sides of the same coin in reproduction. *Reproduction*, 33, 668-683.
- Mohammadi, H., Najafi, A., & Javanmard, A. (2022). Genome-wide Association Study Related to Semen Traits Based on Gene-set Enrichment Analysis in Holstein Bulls. *Research on Animal Production*, 13 (35), 168-175 (In Persian).
- Mohammadabadi, M., Bordbar, F., Jensen, J., Du, M., & Guo, W. (2021). Key genes regulating skeletal muscle development and growth in farm animals. *Animals*, 11, 835.
- Moosanezhad Khabisi, M., Esmailzadeh, A., & Asadi Fozi, M. (2022). Evaluation of genomic inbreeding rate in Iranian native sheep using dense SNP markers (600K). *Research on Animal Production*, 13 (35), 158-167 (In Persian).
- Mota, L. F. M., Lopes, F. B., Fernandes, G. A., Rosa, G. J. M., Magalhães, A. F. B., Carvalheiro, R., & Albuquerque, L.G. (2020). Genome-wide scan highlights the role of candidate genes on phenotypic plasticity for age at first calving in Nellore heifers. *Scientific Reports*, 10(1), 6481.

- Ortega, M. S., Denicol, A. C., Cole, J. B., Null, D. J., & Hansen, P. J. (2016). Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. *Animal Genetics*, 47(3), 288-97.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559-575.
- Salazar-Carranza, M., Castillo-Badilla, G., Murillo-Herrera, J., Hueckmann-Voss, F., & Romero-Zúñiga, J. J. (2014). Effect of age at first calving on first lactation milk yield in holstein cows from Costa Rican specialized dairy herds. *Open Journal of Veterinary Medicine*, 4(09), 197.
- Tsartsianidou, V., Pavlidis, A., Tosiou, E., Arsenos, G., Banos, G., & Triantafyllidis, A. (2023). Novel genomic markers and genes related to reproduction in prolific Chios dairy sheep: a genome-wide association study. *Animal*, 17, 100723.
- Wang, L., Jia, P., & Wolfinger, R. D. (2011). Gene set analysis of genome-wide association studies: Methodological issues and perspectives. *Genomics*, 98, 1-8.
- Wang, K., Liu, X., Qi, T., Hui, Y., Yan, H., Lang, X., & Pan, C. (2021). Whole-genome sequencing to identify candidate genes for litter size and to uncover the variant function in goats (*Capra hircus*). *Genomics*, 13(1), 142-150.
- Xu, S. S., Gao, L., Xie, X. L., Ren, Y. L., Shen, Z. Q., Wang, F., Shen, M., Eyþórsdóttir, E., Hallsson, J. H., Kiseleva, T., Kantanen, J., & Li, M. H. (2018). Genome-wide association analyses highlight the potential for different genetic mechanisms for litter size among sheep breeds. *Frontiers Genetics*, 9, 118.
- Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Method gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biology*, 11, 14-23.
- Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44, 821.
- Zhang, Y. E. (2017). Non-Smad signaling pathways of the TGF- family. *Cold Spring Harbor Perspectives in Biology*, 9, 56-71.