

Research Paper

A Genome-wide Association Study based on Pathway Analysis Related to the Angle Wing in Pekin Ducks using Whole-Genome Sequencing Data

Hossein Mohammadi 

Assistant Professor, Department of Animal Sciences, Faculty of Agriculture and Natural Resources, Arak University, Arak, Iran, (Corresponding author: h-mohammadi64@araku.ac.ir)

Received: 02 August, 2023

Accepted: 28 October, 2023

Extended Abstract

Background: Angel wing is a developmental wing deformity that can influence breeding and reproduction in the commercial duck industry. Therefore, genetic diversity studies and detection of the genomic region related to angel wings in duck populations are essential. In this regard, powerful tools, such as next-generation sequencing technology, have made it possible to decode genome information in this species. The genome-wide association study (GWAS) has been a powerful tool in detecting loci associated with complex traits and diseases; however, it also has some limitations. Complex traits are controlled by many genes, and hence, significant SNPs in general represent only a small fraction of genetic variation. Moreover, studies often report only the most significant SNPs and their neighboring genes, hence some smaller genetic variants and disease risks are unlikely to be detected. Alternatively, pathway-based analysis has been proposed as a complementary approach to investigate complex traits from a genetic and biological perspective. In contrast to a GWAS, pathway-based analysis considers factors that contribute simultaneously to the complex trait and looks beyond the most significant SNPs and genes. To complement GWAS studies, it is becoming common to use gene-set enrichment and pathway analyses. Such an approach helps alleviate problems related to GWAS (e.g., GWAS ignores the fact that genes work together in networks in various biological pathways), and to deepen the understanding of the biological pathways affecting quantitative traits. Integration of F, GWAS, and pathways analyses might address some aforementioned issues and has been already used in human studies, whereas its potential application in livestock breeding and genetics remains still unexplored. In addition, studies are available that performed GWAS or GWAS plus pathway analysis.

Methods: A total of 63 adult purebred Pekin ducks from the same population were selected for this study, of which 33 were ducks that could be identified as having angel wings (case) and 30 were ducks with normal wings (control). Genomic DNA was extracted from blood samples by DNA extraction using a kit (QIAampR DNA Blood Mini Kit; QIAGEN), following the manufacturer's protocol. Whole-genome re-sequencing data were generated on the Illumina HiSeq 4000 platform with 150 bp paired-end reads. Single-nucleotide polymorphism (SNP) calling was performed using the GATK (v4.1), and all parameters were kept at default settings, except for stand_callconf 30. VCFtools (v0.1.16) and plink (v 1.90) were used for the quality control of the data. The 14 064 984 SNPs passed quality control that excluded SNPs using the following criteria: --min- alleles 2, --max- alleles 2, --minDP 3 --minQ 30 with VCFtools, minor allele frequency >0.01, and SNP call rate \geq 0.95 with plink. An independent SNP set was used via the plink command --indep- pairwise 50 5 0.2 for principal component analysis. After quality control, 686 449 SNPs were used for the GWAS. The gene set analysis consists basically of three different steps: (i) the assignment of SNPs to genes, (ii) the assignment of genes to functional categories, and (iii) an association analysis between each functional category and the phenotype of interest. 1. The SNPs were assigned to bovine genes based on the CAU_duck1.0 duck genome sequence assembly using the Bioconductor R package biomaRt2. A given SNP was assigned to a particular gene if it was located within the gene or at most 15 kb either upstream or downstream of the gene. An arbitrary threshold of P-value \leq 0.005 was used to define significant SNPs (based on the results of the GWAS); in this context, significant genes were defined as those genes that contained at least one significant SNP. 2. The databases Gene Ontology (GO) and Medical Subject Headings (MeSH) were used to define functional categories of genes. The idea is that genes assigned to the same functional category can be considered the members of a group of genes that share some particular properties, typically their involvement in the same biological or molecular process. 3. The significant association of a given term with angle wing was analyzed using Fisher's exact



test. Finally, a gene enrichment analysis was performed with the GOstats Bioconductor from R software for the assignment of the genes to functional categories.

Results: The random effect was estimated from the groups clustered based on the kinship among all accessions, and the first two PCs, including PC1 with normal wings and PC2 with angel wings derived from whole-genome SNPs, were used as fixed effects in the mixed model to correct for stratification. In this research, SNP markers were identified on chromosomes 1, 2, 3, 6, 8, 11, 18, 20, 27, and 31. Different sets of candidate genes related to the angle wing trait, namely *ATP11A*, *UBE2E2*, *ITPR2*, *GUCA1C*, *ATP2C1*, *PLCG1*, and *BMPRIA*, were also identified in ducks. Some of the found genes are consistent with some of the previous studies related to wing traits. According to pathway analysis, 21 pathways from GO and biological pathways were associated with the angle wing trait. Some of the detected genes are consistent with some previous studies and are involved in biological pathways related to skeletal muscle growth and development, calcium ion response, bone growth and development, bone minerals, and calcium signaling pathway activation.

Conclusion: The results of our research can be used to understand the genetic mechanism controlling the angle wing trait. This study supports previous results from the GWAS of reproductive traits, revealing additional regions. Using these findings could potentially be useful for genetic selection in ducks.

Keywords: Duck, Enrichment analysis, Genomic data, Next generation sequencing, Wing deformity

How to Cite This Article: Mohammadi, H. (2024). A Genome-wide Association Study based on Pathway Analysis Related to the Angle Wing in Pekin Ducks using Whole-Genome Sequencing Data. *Res Anim Prod*, 15(2), 32-43. DOI: [10.61186/rap.15.2.32](https://doi.org/10.61186/rap.15.2.32)



مقاله پژوهشی

پویش ژنومی بر پایه آنالیز مسیر مرتبط با صفت بال فرشته‌ای در اردک‌های نژاد پیکین با استفاده از داده‌های توالی‌یابی کل ژنوم

حسین محمدی^{ID}

استادیار گروه علوم دامی، دانشکده کشاورزی و محیط زیست، دانشگاه اراک، اراک، ایران، (نویسنده مسوول: h-mohammadi64@araku.ac.ir)

تاریخ پذیرش: ۱۴۰۲/۸/۶

تاریخ دریافت: ۱۴۰۲/۵/۱۱

صفحه: ۳۲ تا ۴۳

چکیده مسوط

مقدمه و هدف: بال فرشته یک بدشکلی توسعه یافته بال است که می‌تواند بر پرورش و تولیدمثل در گله‌های تجاری اردک تأثیر گذار باشد. از این رو مطالعه تنوع ژنتیکی و شناسایی مناطق ژنومی مؤثر بر صفت بال فرشته‌ای در جمعیت‌های اردک، امری ضروری در این گونه محسوب می‌شود. در این راستا، ابزارهای قدرتمندی مانند نسل جدید توالی‌یابی امکان رمزگشایی اطلاعات کل ژنوم در این گونه را فراهم آورده است. به‌طور معمول در مطالعات پویش کل ژنومی در نظر گرفتن تصحیحات سخنگیرانه برای جلوگیری از نتایج مثبت کاذب ضروری می‌باشد، ولی اعمال این نوع تصحیحات موجب از دست رفتن نشانگرهای SNP با اثر کوچکتر مؤثر بر صفات کمی می‌گردد که نتیجه آن شناسایی SNP‌های می‌ماند. یکی از روش‌های جایگزین استفاده از آنالیزهای غنی‌سازی مجموعه‌های ژنی می‌باشد. در این روش ارتباط بین صفت مورد مطالعه و نشانگرهای ژنتیکی را در یک دسته یا گروه ژنی که به‌طور عملکردی با هم مرتبط هستند بررسی می‌کند. در حقیقت در این روش به دنبال ژن‌هایی هستیم که به‌تنهایی اثر آنها بر صفت مورد نظر معنی‌دار نشده، ولی اثر تجمعی آنها روی صفت دارای اثر معنی‌دار است. علاوه بر این، یکی از دلایل اصلی آنالیز غنی‌سازی مجموعه‌های ژنی تعداد کم بودن SNP‌های معنی‌دار می‌باشد که موجب عدم شناسایی مناطق ژنومی مرتبط با صفات مهم اقتصادی می‌گردد. در نتیجه روش پویش ژنومی بر مبنای مسیر کارایی بهتری برای یافتن مناطق ژنومی، درک بهتر مکانیسم و معماری ژنتیکی را دارا می‌باشد. بنابراین، هدف از پژوهش حاضر شناسایی مناطق ژنومی و ژن‌های کاندیدای مرتبط با صفت بال فرشته‌ای در اردک با استفاده از پویش کل ژنوم بر پایه آنالیز مسیر می‌باشد.

مواد و روش‌ها: بدین منظور از اطلاعات داده ژنوم توالی‌یابی شده ۶۳ قطعه اردک شامل ۳۰ قطعه دارای بال نرمال (شاهد) و ۳۳ قطعه دارای بال فرشته‌ای (موردی) استفاده شد. توالی‌یابی کل ژنوم توسط شرکت ایلومینا 4000 Hiseq انجام شده بود. ابتدا نشانگرهای SNP شناسایی شده از ایندل جدا شدند و با استفاده از برنامه GATK فیلتر شدند. سپس با استفاده از برنامه PLINK، SNP‌های دو آللی با حداقل فراوانی آللی بزرگتر یا مساوی ۰/۰۱ که حداقل در ۹۵ درصد افراد دارای ژنوتیپ مشخص بودند، حفظ شده و مابقی حذف شدند. که در نهایت ۱۴۰۶۴۹۸۴ نشانگر SNP بعد از مراحل مختلف کنترل کیفیت باقی ماندند. در گام بعدی برای شناسایی SNP‌های مستقل از نرم‌افزار PLINK استفاده شد. برای این منظور با حذف SNP‌هایی که در حالت عدم تعادل پیوستگی بالایی با یکدیگر قرار داشتند، در پنجره‌هایی شامل SNP ۵۰ و با حرکت SNP ۵ رو به جلو در هر مرحله، SNP‌های دارای r^2 (معیار عدم تعادل پیوستگی) بیش از ۰/۲ (دستور `--indep-pairwise 50 5 0.2`) با یکدیگر از مجموعه داده حذف شدند. در نهایت بعد از کنترل کیفیت تعداد ۶۸۶۴۴۹ SNP برای آنالیزهای پویش کل ژنومی بر پایه تجزیه و تحلیل غنی‌سازی مجموعه ژنی باقی ماندند. اساساً آنالیز پویش ژنومی بر پایه تجزیه و تحلیل مجموعه‌های ژنی در سه مرحله انجام می‌گردد: (۱) تعیین مکان SNP‌های معنی‌دار با ژن (۲) ارتباط ژن‌ها به طبقات عملکردی و مسیرهای زیستی (۳) پویش کل ژنومی بر پایه آنالیز مسیر. -۱) تعیین مکان SNP‌ها با ژن‌ها: SNP‌هایی که مقدار P-value آنها کمتر از ۰/۰۰۵ بود با استفاده از بسته نرم‌افزاری *biomaR2* در محیط R و با استفاده از فرانس ژنومی اردک (*CAU_duck1.0*) به ژن‌هایی که نشانگر SNP موردنظر در داخل آن ژن و یا ۱۵ kb بالادست یا پایین‌دست آن ژن قرار داشت، ارتباط داده شدند. -۲) ارتباط ژن‌ها به طبقات عملکردی و مسیرهای بیوشیمیایی: جهت تعیین طبقات عملکردی ژنی و مسیرهای متابولیکی و تنظیمی ژن‌های معنی‌دار از ۵ پایگاه‌های اطلاعاتی شامل هستی‌شناسی ژن (<http://www.geneontology.org>)، (GO)، مسیرهای بیوشیمیایی (<http://www.kegg.jp/kegg>)، Panther (<http://www.pantherdb.org>)، Metacyc (<http://www.metacyc.org>) و Reactome (<http://www.reactome.org>) جهت تعیین طبقات عملکردی و مسیرهای بیوشیمیایی استفاده گردید. -۳) پویش کل ژنومی بر پایه آنالیز مسیر: ارتباط‌های معنی‌دار مسیرهای عملکردی با صفت بال فرشته‌ای با استفاده از توزیع فوق هندسی و آماره Fisher's exact test مورد آزمون قرار گرفت.

یافته‌ها: نتایج آنالیز PCA نشان داد که با PC1 گروه جمعیت سالم و PC2 گروه جمعیت بیمار را به‌خوبی از یکدیگر تفکیک و جدا کردند. در این پژوهش نشانگرهای تک نوکلئوتیدی واقع روی کروموزوم‌های ۱، ۲، ۳، ۴، ۵، ۶، ۸، ۱۱، ۱۸، ۲۰، ۲۷ و ۳۱ شناسایی شدند که با ژن‌های *ATP11A*، *UBE2E2*، *TPR2*، *BMPRIA* و *PLCG1* مرتبط بودند. در تفسیر مجموعه ژنی، تعداد ۲۱ مسیر هستی‌شناسی ژنی و بیوشیمیایی با صفت بال فرشته‌ای شناسایی شد ($P < 0.05$). از این بین، مسیرهای *wing disc response to calcium ion*، *skeletal muscle myosin thick filament assembly*، *Calcium signaling pathway* و *morphogenesis* مهمی را در ارتباط با رشد و توسعه عضلات اسکلتی، پاسخ به یون کلسیم، رشد و توسعه استخوان، حجم مواد معدنی استخوان و فعال‌سازی مسیر سیگنال‌دهی کلسیم بر عهده داشتند

نتیجه‌گیری: با توجه به نقش کلیدی و مؤثر ژن‌های کاندیدای شناسایی شده بر صفت بال فرشته‌ای در این تحقیق، در صورت تأیید به‌وسيله مطالعات تکمیلی می‌توان از آنها در انتخاب ژنتیکی گله‌های تجاری اردک مورد استفاده قرار گیرند.

واژه‌های کلیدی: آنالیز غنی‌سازی، اردک، بدشکلی بال، توالی‌یابی نسل بعد، داده‌های ژنومی

مقدمه

بودن پرورش این پرنده به‌دلیل تولید گوشت (به‌طور ویژه سینه) و تخم از مزایای این پرنده می‌باشد (Huang et al., 2013). در سال‌های اخیر اردک به‌عنوان یک پرنده اقتصادی در کشورهای مختلف پرورش داده می‌شود. سودآوری در صنعت تولید و پرورش اردک به هزینه‌های ورودی و خروجی بستگی دارد. بنابراین، بهبود بازده خوراک، مقدار غذای موردنیاز برای

اردک‌ها گونه‌ای از پرندگان متعلق به خانواده‌ی آناتید، زیر خانواده‌ی آنسرین و دسته‌ی اردک‌ها هستند. این پرنده دارای خصوصیات رشد بالا، پایین بودن هزینه‌های بهداشتی به‌دلیل مقاومت بسیار زیاد به بیماری‌ها، تولید محصول ارگانیک به‌دلیل عدم استفاده از دارو و واکسن در دوره پرورش و چند منظوره

این روش به دنبال ژن‌هایی هستیم که به‌تنهایی اثر آنها بر صفت موردنظر معنی‌دار نشده، ولی اثر جمعی آنها روی صفت دارای اثر معنی‌دار است (Dadousis *et al.*, 2017). علاوه‌براین، یکی از دلایل اصلی آنالیز غنی‌سازی مجموعه‌های ژنی تعداد کم بودن SNP‌های معنی‌دار می‌باشد که موجب عدم شناسایی مناطق ژنومی مرتبط با صفات مهم اقتصادی می‌گردد. در نتیجه روش پویش ژنومی بر مبنای مسیر کارایی بهتری برای یافتن مناطق ژنومی، درک بهتر مکانیسم و معماری ژنتیکی را دارا می‌باشد (Srikanth *et al.*, 2020).

برای اولین بار Peñagaricano و همکاران (2013) نشان دادند که تجزیه و تحلیل پویش ژنومی بر مبنای مسیر دقت شناسایی مناطق ژنومی مؤثر بر صفت نرخ باروری گاوهای نر را بالا برده است، زیرا با استفاده از این روش تمامی نشانگرهای معنی‌دار در سطح ۰/۰۵ آنالیز می‌شوند و در نتیجه میزان خطای نوع اول و بیش برآوردها کاهش پیدا می‌کند. در تحقیق دیگر مطالعه پویش ژنومی بر مبنای مسیر با استفاده از آنالیز غنی‌سازی مجموعه‌های ژنی روی صفات رشد در چهار نژاد مرغ شامل چوآ، سیلک، لنگشن و ببرد انجام شد. براساس این نتایج آنالیز براساس مسیر، منجر به شناسایی ۲۱ طبقات مختلف عملکردی هستی شناسی ژن و ۱۷ مسیر بیوشیمیایی KEGG معنی‌دار مرتبط با این صفات شد (Khalatabadi Farahani *et al.*, 2020).

داده‌های مورد استفاده در این پژوهش اولین بار توسط سان و همکاران (Sun *et al.*, 2023) برای صفت بال فرشته‌ای، در اردک‌های نژاد پیکین مورد آنالیز قرار گرفته‌اند که در آن تحقیق از مدل‌های خطی مختلط تک نشانگری و از تصحیح بنفرونی برای تعیین آستانه معنی‌داری و جلوگیری از خطای نوع اول استفاده شده است. در مجموع ۲۴ نشانگر تک نوکلئوتیدی معنی‌دار مرتبط با صفت روی کروموزوم‌های ۲، ۲۳، ۲۶، ۲۷ و ۲۸ شناسایی شده بود که ژن‌های کاندیدای *OPCML*، *FEMIA* و *SLC18A1 ATP6V1B2* در این مناطق قرار داشتند. اخیراً مطالعه پویش کل ژنومی بر پایه آنالیز مسیر جهت شناسایی جایگاه‌های ژنی مؤثر بر صفات مهم اقتصادی در بلدرچین ژاپنی انجام شده است و مسیرهای *cytoskeletal MAPK signaling*، *protein binding*، *anatomical structure development pathway* و *Myoblast differentiation* که شامل ژن‌های کاندیدای *MYF5*، *BMPR1B*، *PTPRN2*، *DRD2*، *NPY*، *ACOX1*، *BMP4*، *LDB2*، *FGF2*، *MYO1E*، *IGF2BP1*، *PLCB1*، *PLCB4*، *PCK1* و *PNPLA2* بودند، را گزارش کرده‌اند (Mohammadi *et al.*, 2023).

با توجه به اینکه مطالعات پویش کل ژنومی روی اردک به‌منظور شناسایی نواحی ژنومی مؤثر بر صفات مهم اقتصادی به‌طور بسیار معدودی انجام گرفته، لذا هدف پژوهش حاضر تجزیه و تحلیل مجموعه‌های ژنی و مسیرهای زیستی مرتبط با صفت بال فرشته‌ای در اردک براساس پویش کل ژنومی می‌باشد.

مواد و روش‌ها

رشد، هزینه تولید و مقدار ضایعات نیتروژنی را کاهش خواهد داد (Zhou *et al.*, 2018).

استراتژی‌های اصلی موجود در برنامه‌های پرورش طیور تجاری با هدف افزایش سرعت رشد و وزن بدن همراه است. رشد حیوان که شامل فرآیندهای فیزیولوژیکی است، یکی از عوامل بسیار مهم و تاثیرگذار در علم پرورش طیور است. یکی از بیماری‌های اردک، که بسیار هم شایع است بیماری بال فرشته‌ای می‌باشد. این بیماری به این دلیل اتفاق می‌افتد که وزن پرهای موجود در بال زیاد بوده و فشار زیادی به عضلات و مچ پای اردک وارد می‌کند. این بیماری در پرندگانی که رشد سریع‌تری دارند، شایع‌تر است و می‌تواند متأثر از عوامل ژنتیکی، تغذیه‌ای (پروتئین بالا و کمبود کلسیم و ویتامین D₃) و مدیریتی رخ دهد (Lin *et al.*, 2015; Lin *et al.*, 2018).

با پیشرفت‌های اخیر در تهیه تراشه‌های SNP طیور و استفاده از چندشکلی در تک نوکلئوتیدی‌ها به‌عنوان نشانگر، امکان پوشش کامل ژنوم و نقشه‌یابی با وضوح بالاتر که با روش‌های پیشین از قبیل نشانگرهای ریزماهواره و یا سایر نشانگرهای بیولوژیکی مقدور نبود، امکان‌پذیر گردید که بهبود صحت مکان‌یابی QTL‌ها و نواحی ژنومی مؤثر بر صفات تولیدی را در پی داشته است. استفاده از داده‌های کل ژنوم یا به‌طور کلی داده‌های DNA و به‌کارگیری آن می‌تواند زمینه ساز ورود به عصر پسازنومی در اصلاح نژاد پرندگان اهلی گردد (Ding *et al.*, 2019).

در مطالعات پویش کل ژنوم (GWAS) از وجود تعداد زیادی چندشکلی در سطح ژنوم به‌همراه فنوتیپ و اطلاعات شجره جهت آنالیز همبستگی بین ژنوتیپ و فنوتیپ استفاده می‌شود و ژن‌ها و عناصر تنظیم‌کننده مرتبط با صفات مهم اقتصادی را شناسایی می‌گردند (Tang *et al.*, 2022).

به‌علاوه، مطالعات و بررسی‌های به‌عمل آمده در اواخر دهه ۸۰ میلادی روشن نمود که مکانیسم‌های مولکولی در زمره مهم‌ترین فرایندهای ژنتیکی (مشتمل بر همانندسازی DNA، رونویسی، ترجمه و حتی نحوه تنظیم ژن‌ها) هستند (Mohammadabadi, 2020; Safaei *et al.*, 2022). یکی از اقدامات اساسی در به‌نژادی مولکولی مطالعه ژن‌ها و پروتئین‌های مرتبط با صفات مهم و مطالعه آن‌ها در سطح سلولی یا کروموزومی است (Mohammadabadi and Soflaei 2020; Shamsavari *et al.*, 2022).

به‌طور معمول در مطالعات پویش کل ژنومی در نظر گرفتن تصحیحات سختگیرانه (عمدتاً تصحیح بنفرونی) برای جلوگیری از نتایج مثبت کاذب ضروری می‌باشد، ولی اعمال این نوع تصحیحات موجب از دست رفتن نشانگرهای SNP با اثر کوچکتر مؤثر بر صفات کمی می‌گردد که نتیجه آن شناسایی SNP‌هایی می‌گردد که تنها بخش کوچکی از تنوع ژنتیکی صفت را نشان می‌دهد و از این‌رو در اکثر مواقع بخش عمده واریانس ژنتیکی پنهان باقی می‌ماند. یکی از روش‌های جایگزین استفاده از آنالیزهای غنی‌سازی مجموعه‌های ژنی می‌باشد. در این روش ارتباط بین صفت مورد مطالعه و نشانگرهای ژنتیکی را در یک دسته یا گروه ژنی که به‌طور عملکردی با هم مرتبط هستند بررسی می‌کند. در حقیقت در

مدل مورد استفاده بر پایه مدل خطی مختلط تک صفت به شکل زیر بود:

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

که y : بردار مشاهدات فنوتیپی، W ماتریس ضرایب مربوط به اثر ثابت جنس، P ماتریس ضرایب مربوط به اثرات پنج مؤلفه اصلی با بیشترین واریانس توجیه شده، x : بردار ژنوتیپ‌های نشانگر تحت آزمون، u بردار اثرات تصادفی پلی‌ژنیک، Z ماتریس روابط ژنومی، e بردار اثر تصادفی باقی‌مانده، بردارهای α ، c و β به ترتیب در برگزیده اثرات ثابت، اثر مؤلفه‌های اصلی و اثر نشانگر SNP هستند.

آنالیز پویش کل ژنومی بر اساس مجموعه‌های ژنی (GSEA-SNP)

اساساً آنالیز پویش ژنومی بر پایه تجزیه و تحلیل مجموعه‌های ژنی در سه مرحله انجام می‌گردد: (۱) تعیین مکان SNP‌های معنی‌دار با ژن (۲) ارتباط ژن‌ها به طبقات عملکردی و مسیرهای زیستی (۳) پویش کل ژنومی بر پایه آنالیز مسیر

۱- تعیین مکان SNP‌ها با ژن‌ها: SNP‌هایی که مقدار P-value آنها کمتر از ۰/۰۰۵ بود (Srikanth *et al.*, 2020) با استفاده از بسته نرم‌افزاری biomaR2 در محیط R و با استفاده از رفرانس ژنومی اردک (CAU_duck1.0) به ژن‌هایی که نشانگر SNP موردنظر در داخل آن ژن و یا 15 kb بالادست یا پایین دست آن ژن قرار داشت، ارتباط داده شدند.

۲- ارتباط ژن‌ها به طبقات عملکردی و مسیرهای بیوشیمیایی: جهت تعیین طبقات عملکردی ژنی و مسیرهای متابولیکی و تنظیمی ژن‌های معنی‌دار از ۵ پایگاه اطلاعاتی شامل هستی‌شناسی ژن (<http://www.geneontology.org/GO>)

مسیرهای بیوشیمیایی (<http://www.genome.jp/kegg/KEGG>) Panther (<http://www.pantherdb.org>) Metacyc (<http://www.metacyc.org>) Reactome (<http://www.reactome.org>) جهت تعیین طبقات عملکردی و مسیرهای بیوشیمیایی استفاده گردید.

۳- پویش کل ژنومی بر پایه آنالیز مسیر: ارتباط‌های معنی‌دار مسیرهای عملکردی با صفت بال فرشته‌ای با استفاده از توزیع فوق هندسی و آماره Fisher's exact test مورد آزمون قرار گرفت. P-value مسیرهای عملکردی که تعداد k ژن معنی‌دار در آن قرار دارد با فرمول زیر محاسبه شد:

$$P - value = 1 - \sum_{i=1}^{k-1} \frac{\binom{S}{i} \binom{N-S}{m-i}}{\binom{N}{m}}$$

در این فرمول، k برابر با تعداد ژن‌های معنی‌دار در طبقه عملکردی، S برابر با تعداد کل ژن‌های معنی‌دار مرتبط با صفت مورد بررسی، N برابر با کل تعداد ژن‌هایی که در این مطالعه آنالیز شدند و m برابر با تعداد کل ژن‌های موجود در مسیر عملکردی. تجزیه غنی‌سازی مجموعه ژنی با استفاده از بسته نرم‌افزاری *GOstats Bioconductor* در محیط R انجام

در پژوهش حاضر از داده‌های توالی‌یابی شده کل ژنوم ۶۳ قطعه اردک نژاد پیکین استفاده شد (Sun *et al.*, 2023)، که در آن تحقیق از اردک‌های خالص در سن ۴۲ روزگی شامل ۳۴ قطعه اردک نر و ۲۹ اردک ماده استفاده شده بود. میانگین وزن زنده اردک‌های نر و ماده به ترتیب برابر با ۳۸۷۷/۹±۲۱۴/۸ و ۳۶۶۲/۱۷۵±۶/۸ گرم بودند. انتخاب دو گروه براساس ظاهر فنوتیپی انجام شده بود، به طوری که اردک‌های بدون هرگونه ظهور بال فرشته‌ای (شاهد) به تعداد ۳۰ قطعه و با ظاهر کامل با نقص بال فرشته‌ای (موردی) به تعداد ۳۳ قطعه گروه‌بندی شده بودند.

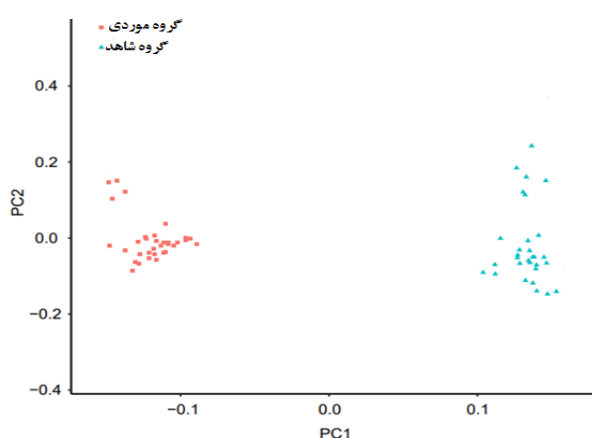
توالی‌یابی کل ژنوم به صورت paired-end یا دو سویه توسط شرکت ایلومینا 4000 Hiseq انجام شده بود. داده‌های تعیین ژنوتیپ شده مورد استفاده در این پژوهش در پایگاه برخط NCBI (<https://www.ncbi.nlm.nih.gov/bioproject>) با کد دسترسی (PRJNA886058) و پایگاه برخط figshare (<https://doi.org/10.6084/m9.figshare.21631868.v1>) آدرس در دسترس می‌باشد.

ابتدا نشانگرهای SNP شناسایی شده از ایندل جدا شدند و با استفاده از برنامه GATK (McKenna *et al.*, 2010) فیلتر شدند. سپس با استفاده از برنامه PLINK (Purcell *et al.*, 2007)، SNP‌های دو آللی با حداقل فراوانی آللی بزرگتر یا مساوی ۰/۰۱ که حداقل در ۹۵ درصد افراد دارای ژنوتیپ مشخص بودند، حفظ شده و مابقی حذف شدند. که در نهایت ۱۴۰۶۴۹۸۴ نشانگر SNP بعد از مراحل مختلف کنترل کیفیت باقی ماندند. در گام بعدی برای شناسایی SNP‌های مستقل از نرم‌افزار PLINK استفاده شد. برای این منظور با حذف SNP‌هایی که در حالت عدم تعادل پیوستگی بالایی با یکدیگر قرار داشتند، در پنجره‌هایی شامل SNP ۵۰ و با حرکت SNP ۵ رو به جلو در هر مرحله، SNP‌های دارای r^2 (معیار عدم تعادل پیوستگی) بیش از ۰/۲ (دستور 50 5 0.2 --indep-pairwise) با یکدیگر از مجموعه داده‌ها حذف شدند. در نهایت بعد از کنترل کیفیت تعداد ۶۸۶۴۴۹ SNP برای آنالیزهای پویش کل ژنومی بر پایه تجزیه و تحلیل غنی‌سازی مجموعه ژنی باقی ماندند.

همچنین در این پژوهش برای بررسی وجود یا عدم وجود لایه‌بندی جمعیتی با استفاده از روش تجزیه و تحلیل مؤلفه‌های اصلی (PCA) با استفاده از نرم‌افزار PLINK.9 محاسبه و گراف PCA توسط برنامه R نسخه (نسخه ۳/۶/۱) ترسیم شد. هدف از تجزیه به مؤلفه‌های اصلی آن است که واریانس موجود در داده‌های چند متغیره را به مؤلفه‌هایی تجزیه کند که اولین مؤلفه تا آنجا که ممکن است علت بیشترین واریانس موجود در داده‌ها باشد. دومین مؤلفه علت بیشترین واریانس ممکن بعد از مؤلفه اول و الی آخر باشد.

جهت انجام تجزیه و تحلیل پویش کل ژنومی با استفاده از نرم‌افزار GEMMA (Zhou & Stephens., 2012) نسخه ۰/۹۸ استفاده شد. این نرم‌افزار توانایی در نظر گرفتن هر کدام مشاهدات فنوتیپی داده‌های باینتری را به صورت متغیر جداگانه دارا می‌باشد (Zhuang *et al.*, 2020).

مثبت کاذب (خطای نوع اول) شود (Liu *et al.*, 2022). نتایج آنالیز PCA در دو گروه جمعیتی مورد بررسی در شکل ۱ ارائه شده است. نتایج آنالیز PCA نشان داد که با PC1 گروه جمعیت سالم و PC2 گروه جمعیت بیمار را به خوبی از یکدیگر تفکیک و جدا کردند و در هیچ نقطه‌ای با هم همپوشانی ندارند. پلات‌های Q-Q و منهتن مرتبط با صفت بال فرشته‌ای در شکل‌های ۲ و ۳ ارائه شده است. چندین روش برای تخمین کنترل تورم ژنومیکی (λ) وجود دارد که در این تحقیق روش تخمین گر میانه اجرا شد. فاکتور تورم کنترل جمعیتی بزرگتر از ۱ ($\lambda > 1$) وجود لایه‌بندی جمعیتی را نشان می‌دهد. فاکتور لامبدا برای صفت بال فرشته‌ای برابر با ۱/۰۵۷ بود که همان‌طور که مشاهده می‌شود تقریباً برابر با ۱ بود.



شکل ۱- خوشه‌بندی پرندگان به دو گروه جداگانه جمعیتی با بال فرشته‌ای (موردی) و جمعیت نرمال (شاهد) براساس آنالیز PCA با استفاده از اطلاعات ژنوتیپی تمام پرندگان

Figure 1. Principal components analysis (PCA) clustered individuals separately that belong to a different group of angle wing (case) and normal duck populations using genotyping information animals.

تحلیل نشان دهنده این است که ژن *ATP11A* با فرآیند و skeletal muscle myosin thick filament assembly response to calciumion در فرآیند زیستی *UBE2E2* مرتبط با صفت بال فرشته‌ای با بیشترین مقدار ضریب غنی‌سازی هستند، در اردک‌های نژاد پیکین مشاهده شد. در مطالعه پویش ژنومی مرتبط با اندازه بدن و قطعات لاشه، ژن کاندیدای *ATP11A* در ارتباط با وزن بال گزارش شده است (Deng *et al.*, 2019). در مطالعه پویش ژنومی مرتبط با رفتار مصرف خوراک، ژن کاندیدای *UBE2E2* با صفت مصرف خوراک روزانه گزارش شده است (Li *et al.*, 2021). همچنین در مطالعه پویش ژنومی گزارش شده است که ژن *UBE2E2* با سندروم‌های متابولیکی در انسان ارتباط دارد (El Yaagoubi *et al.*, 2017).

گردید. برای تفسیر بهتر عملکرد ژن‌های به دست آمده از پایگاه‌های اطلاعاتی آنالیز UniProtKB و GeneCards (<http://www.genecards.org>) استفاده شد. در نهایت برای ترسیم گراف‌های منهتن مربوطه از نرم‌افزار SNPEVG2 استفاده گردید (Wang *et al.*, 2012).

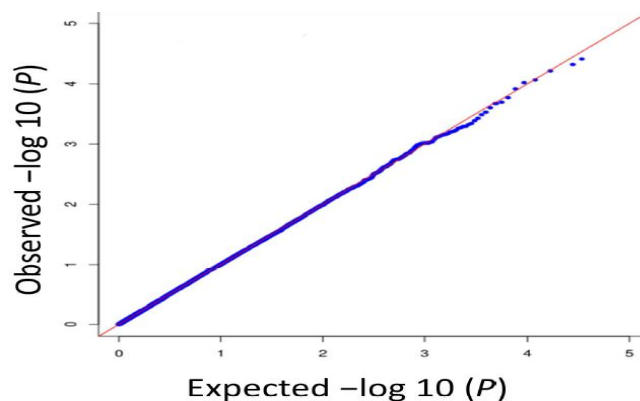
نتایج و بحث

لایه‌بندی جمعیتی که ناشی از تفاوت در فراوانی آلی زیر جمعیت‌ها به دلیل تفاوت ژنتیکی جد مشترک و همچنین میزان خویشاوندی می‌تواند نتایج یک مطالعه ارتباط ژنومی را با مشکل مواجه نماید چرا که در مطالعات پویش ژنومی فرض می‌شود که جامعه مورد مطالعه همگن می‌باشد. علاوه بر این، در نظر گرفتن اثرات خویشاوندی می‌تواند منجر به کاهش نتایج

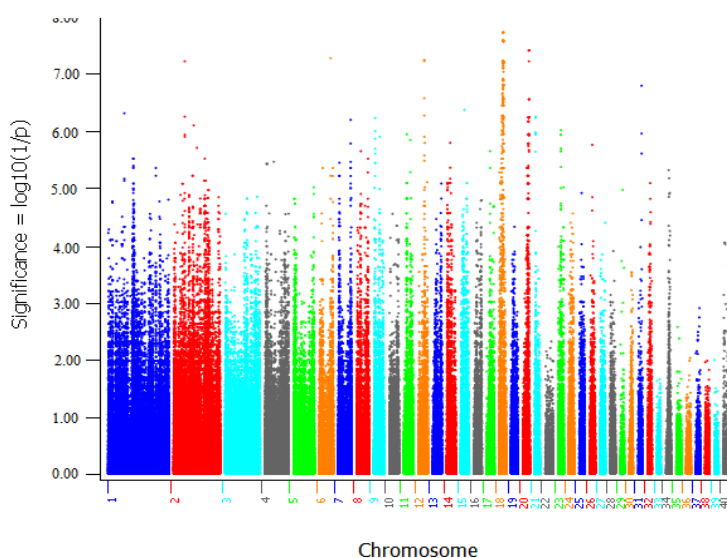
شناسایی ژن‌های کاندیدا در مناطق ژنومی

تعداد ۲۴۶۷۱ عدد ژن از ۲۵۹۵۲ ژن حاشیه‌نویسی شده در اردک به وسیله نشانگرهای SNP پوشش داده شد که در این میان تعداد ژن‌های معنی‌دار در نژاد مورد بررسی ۳۱۰۸ ژن بودند، به عبارتی حداقل یک نشانگر با P-value کمتر از ۰/۰۰۵ در داخل و یا در بالا یا پایین دست این ژن تا فاصله ۱۵ kb قرار گرفت. این ژن‌ها به عنوان ژن‌های معنی‌دار مرتبط با صفت بال فرشته‌ای برای تجزیه GSEA-SNP انتخاب شدند. تعداد مجموعه‌های ژنی حاصل از پایگاه‌های داده‌ای مختلف شامل ۲۳۲ طبقات هستی‌شناسی به دست آمد و ۱۱ مسیر زیستی KEGG مشاهده شد. مسیرهای که بیشتر از سه ژن و کمتر از ۳۰۰ ژن داشتند گزارش شده‌اند.

جزئیات کامل ترم‌های هستی‌شناسی به همراه اسامی ژن‌های کاندیدا در جدول ۱ ارائه شده است. نتایج حاصل از این



شکل ۲- پلات Q-Q برای صفت بال فرشته‌ای در اردک نژاد پکین
Figure 2. Plots Q-Q for angel wing trait in pekin duck



شکل ۳- پلات منهن برای صفت بال فرشته‌ای در اردک. محور X مکان آلل‌های ژنوتیپی روی کروموزوم‌ها و محور Y منفی لگاریتم بر مبنای ۱۰ ارزش P-value
Figure 3. Manhattan plot for angle wing trait in duck. X axis, genotypes alleles positions on chromosomes, Y axis, -Log10 P-value

(McMullan *et al.*, 2003). ارتباط معنی‌داری بین ژن *GUCA1C* با حجم معدنی استخوان گزارش شده است (Li *et al.*, 2020).

مسیر *regulation of glucose import in response to insulin stimulus* با تعداد ۴۱ ژن معنی‌دار از دیگر مسیرهایی بود که در ارتباط با صفت مورد بررسی شناسایی شد. این مسیر در مسیر فرآیند زیستی طبقه‌بندی می‌شود. از بین ژن‌های معنی‌دار در این مسیر، ژن کاندیدای *FGF9* از خانواده ژنی فاکتور رشد فیروبلاستی یا (FGF) شامل ۲۲ ژن می‌باشد (GeneCards). این خانواده ژنی در بسیاری از فرآیندهای سلولی نظیر تمایز و رشد سلول‌های جنینی، بازیابی بافت، تقسیمات سلولی و متابولیسم انرژی دخالت دارند (UniProtKB). ژن *FGF9* در فرآیندهای بیولوژیکی توسعه سیستم عضلات اسکلتی جنین و تفرق استئوبلاست نقش دارد (Halgrain *et al.*, 2023).

مسیر مهم معنی‌دار دیگر مرتبط با صفت بال فرشته‌ای مسیر بیولوژیکی *response to muscle stretch* بود که یک فرآیند بیوشیمیایی ضروری و حیاتی است که در تنظیم گامتوز، رشد جنین و جفت نقش دارند (UniProtKB). ژن کاندیدای *ITPR2* نقش کلیدی در رشد و توسعه سیستم عصبی و مورفولوژی سلولی دارد (Chen *et al.*, 2023). همچنین گزارش شده است ژن *ITPR2* نقش کلیدی در مسیر سیگنال‌دهی کلسیم و بازدهی خوراک در دوره تخم‌گذاری مرغان تخمگذار دارد (Yuan *et al.*, 2015).

از دیگر مسیرهای بیولوژیکی معنی‌دار مرتبط با صفت بال فرشته‌ای می‌توان به مسیر *wing disc morphogenesis* اشاره کرد که جزء مسیرهای فرآیندهای زیستی می‌باشد و از بین ۲۰ ژن معنی‌دار وجود در این مسیر زیستی، ژن کاندیدای *GUCA1C* دارای نقش بیولوژیکی مستقیمی با رشد و توسعه بال داشت. ژن کاندیدای *GUCA1C* نقش اساسی در تنظیم تفرق سلول‌های کراتینوسیت‌ها از مسیر سیگنال‌دهی Rho دارد

مرتبط با صفات مهم اقتصادی، ژن کاندیدای *ATP2C1* ارتباط معنی‌داری با رشد عضلات سینه مرغ در مراحل مختلف دوره رشد گزارش شده است (Tang et al., 2020).

مسیر دیگر معنی‌دار مرتبط با صفت بال فرشته‌ای را می‌توان *carboxylic acid biosynthetic process* نام برد که از بین ۳۱ ژن، ژن کاندیدای *ATP2C1* در مطالعه پویس ژنومی با هدف شناسایی مناطق ژنومی و ژن‌های کاندیدای

جدول ۱- تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی معنی‌داری ($P < 0.05$) مرتبط با صفت بال فرشته‌ای در اردک
Table 2. Gene set enrichment analysis significantly ($P < 0.05$) associated with angle wing trait in duck

GO ID طبقات هستی‌شناسی	Term GO نام مسیر هستی‌شناسی	No. genes in the GO term تعداد ژن موجود در مسیر	Candidate genes1 ژن‌های کاندیدا	P-adjust P-ارزش تصحیح شده
Biological process فرآیند زیستی				
GO:0006936	muscle contraction	19	<i>LAMA2, LAMC2, ITGB3, ITGA2B, LAMB4, GP1BA, ITGB8, LAMA1, LAMB2, LAMB3</i>	0.00399
GO:0035994	response to muscle stretch	24	<i>LOC101803633, AKR1D1, SRD5A3, HSD11B2, HSD17B2, SRD5A1, LOC101795569, LOC101803865, LOC101790812, ITPR2</i>	0.00587
GO:0007472	wing disc morphogenesis	20	<i>HEXA, GUCA1C, GNPDA1, TSTA3, PGM3, FPGT, UGDH, NANS, GALK1, GMPPP, GNPDA2, NANP</i>	0.00587
GO:0030241	skeletal muscle myosin thick filament assembly	37	<i>EHHADH, ACOX3, SCD, ACSBG1, ATP11A, OXSM, ACSBG2, ACSL1, ACOX1, TECR, ACSL5, ACAT2, HACD1, FADS1, FADS2, SCD5, ELOVL2, ACADL</i>	0.0175
GO:0051592	response to calcium ion	24	<i>LOC101790524, ALDOB, ALDH7A1, ALDH18A1, ASL, MAT2B, ASS1, UBE2E2, CTH, PFKM, PFKL, AADAT, GAPDH, CPS1</i>	0.0218
GO:0010658	striated muscle cell apoptotic process	27	<i>PPARGC1A, PRKAB2, G6PC2, PRKAR2A, PDE3B, SH2B2, GYS2, SOCS4, TSC1, IKBKB, GSK3B, SORBS1, INSR, SOS1, PYGL, SHC4, MAP2K1</i>	0.02327
GO:0010829	negative regulation of glucose transport	32	<i>MPI, ALDOB, SORD, TSTA3, FPGT, PFKL, FUK, GMPPP, TIGAR</i>	0.02394
GO:2000505	regulation of energy homeostasis	21	<i>SQLE, LOC101801210, FDFT1, MSMO1, LOC101796874</i>	0.02913
GO:2001273	regulation of glucose import in response to insulin stimulus	41	<i>MAPK11, CACNA1S, PDGFB, MAP3K14, DUSP3, ILIR2, FGF9, RRAS2, RAPGEF2, DUSP8, PLA2G4B, DUSP10, DDI3, ARRB1, EGFR, SOS1, FGF19, EGF</i>	0.03035
GO:0044381	glucose import in response to insulin stimulus	38	<i>PLA2G3, ENPP2, LA2G1B, PLB1, ENPP6, PLA2G4F, PLD4, PLA2G4B</i>	0.03209
GO:0042149	cellular response to glucose starvation	15	<i>SLC27A4, PLTP, PLIN2, ACSBG1, ACOX1, ACSL5, LOC101796631, FADS2, SCD5, ACADL, RXRG</i>	0.00177
GO:0043406	positive regulation of MAP kinase activity	24	<i>POP1, TPR, TACC3, GEMIN8, XPO5, NUP214, TGS1, EIF3B, EIF2B3, THOC5, CYFIP2, NDC1, DDX20, GEMIN5</i>	0.03848
GO:0006464	protein modification process	18	<i>SPDYA, DE3B, BUB1, PGR, CPEB2, CDK1, ANAPC5, CDC25A, PLK1</i>	0.04159
GO:0046394	carboxylic acid biosynthetic process	31	<i>TRADD, CAMKK1, PPARGC1A, PRKAB2, G6PC2, AGRP, ATP2C1, LEPR, ACSBG2, STK11, ACSL1, JAK2, ACSL5, PRKAG2, NFKBIE, POMC, CAMKK2, RXRG</i>	0.043394
Cellular component اجزای سلولی				
GO:0030054	Cell junction	25	<i>SRC, MYH9, JAM2, TIP1, SHROOM2, MYH10, CLDN2</i>	0.04566
GO:0034703	Cation channel complex	17	<i>EPB41, MAGI2, MPDZ, MAGI1, EPB41L2, PRKCD, RRAS2</i>	0.04037
KEGG Pathways مسیرهای KEGG				
apla04020	Calcium signaling pathway	68	<i>ITPR2, ATP2A3, ADRB3, SLC25A6, ADCY8, ADCY2, ITPR1, PLCD1, SLC8A3, ERBB4, PLCG1, PLCE1, PDGFRB, PLCB4, ADRB1</i>	0.02331

ادامه جدول ۱- تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی معنی‌داری ($P < 0.05$) مرتبط با صفت بال فرشته‌ای در اردک
Continue Table 2. Gene set enrichment analysis significantly ($P < 0.05$) associated with angle wing trait in duck

GO ID طبقات هستی‌شناسی	Term GO نام مسیر هستی‌شناسی	No. genes in the GO term تعداد ژن موجود در مسیر	Candidate genes ¹ ژن‌های کاندیدا	P-adjust P-ارزش تصحیح شده
Biological process فرآیند زیستی				
apla04310	Wnt signaling pathway	61	<i>SERPINF1, WNT8B, CAMK2G, PLCB2, CTBP2</i>	0.048414
apla04350	TGF-beta signaling pathway	75	<i>TGIF2, BMPRIA, BAMBI</i>	0.049095

¹ژن‌های کاندیدا مرتبط با صفت به صورت پر رنگ نشان داده شده‌اند.

¹Candidate genes associated with trait are shown in bold.

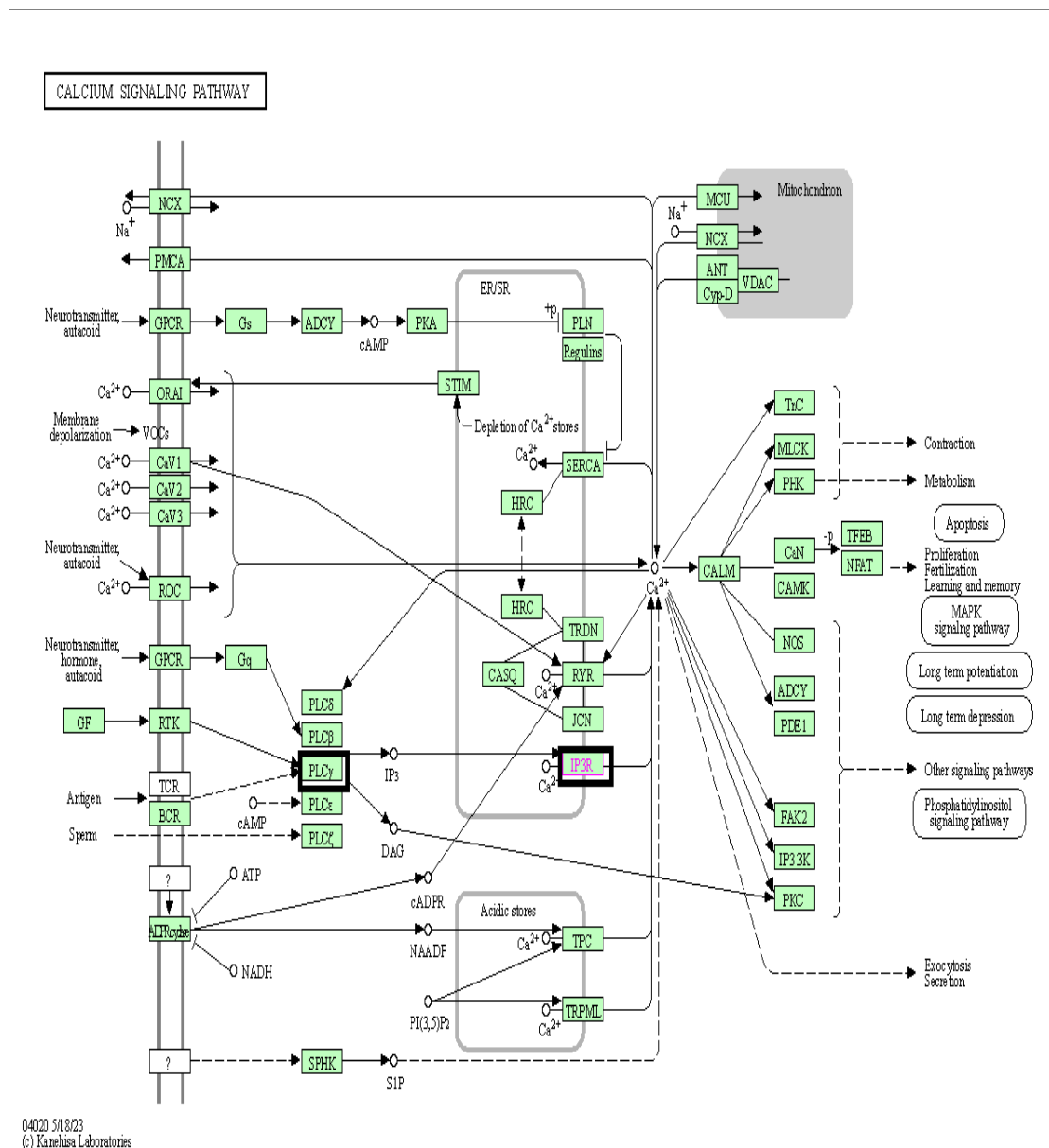
در این پژوهش اغلب ژن‌ها و جایگاه‌های ژنومی یافت شده در اردک جدید می‌باشند و هیچگونه تحقیقی بر روی این ژن‌ها در اردک انجام نشده است. با توجه به عملکرد بیولوژیکی ژن‌های کاندیدای شناسایی شده در این تحقیق که منطبق با گزارشات پوشش ژنومی پیشین مخصوصاً در گونه مرغ می‌باشد، می‌توان بروز فنوتیپی هریک از صفات مورد مطالعه را توجیه کرد. البته باید توجه کرد که یکی از محدودیت‌های اجرای مطالعات پوشش ژنومی، کافی نبودن تعداد رکوردهای فنوتیپی و جمعیت غیر خویشاوند برای آنالیز پیوستگی می‌باشد که منجر به افزایش خطای نوع اول یا به عبارتی شناسایی نشانگرهای معنی‌دار مثبت کاذب می‌شود (Seabury *et al.*, 2017). با توجه به تعداد کم نمونه مورد استفاده در پژوهش حاضر نیز باید در استفاده از ژن‌های کاندیدای شناسایی شده مرتبط با صفات مهم اقتصادی در برنامه‌های اصلاحی با احتیاط عمل کرد. البته می‌توان با ادغام داده‌های تحقیقات مشابه جدید و استفاده از آنالیزهای آماری جامع‌تر برای تأیید نتایج پژوهش حاضر استفاده کرد.

نتیجه‌گیری کلی

مناطق ژنومی شناسایی شده مرتبط با صفت بال فرشته‌ای در مطالعه حاضر روی کروموزوم‌های ۱، ۲، ۳، ۴، ۶، ۸، ۱۱، ۱۸، ۲۰، ۲۷ و ۳۱ قرار داشتند. با توجه به عملکرد ژن‌های کاندیدای شناسایی شده می‌توان کارایی روش تجزیه و تحلیل غنی‌سازی مجموعه ژنی برای پوشش ژنومی صفت بال فرشته‌ای را نیز مورد تأیید قرار داد. استفاده از آزمون‌های آزمایشگاهی مختلف می‌تواند در تأیید نتایج به‌دست آمده در این پژوهش مؤثر باشند. ژن‌های عمده اثر شناسایی شده در پژوهش حاضر و سایر پژوهش‌های مرتبط می‌تواند در این زمینه کاربردی باشد.

تحلیل مسیرهای زیستی KEGG نشان‌دهنده این است که ژن‌های *ITPRI* و *PLCG1* به‌طور معنی‌داری با مسیر Calcium signaling pathway در ارتباط می‌باشند. این مسیر زیستی نقش مهمی در رشد و توسعه استخوان دارد. در شکل ۳ مسیر سیگنالی کلسیم و ژن‌های مربوطه شناسایی شده از پایگاه برخط KEGG ارائه شده است.

ژن‌های کاندیدای *ITPRI* و *PLCG1* جزوی از خانواده پروتئینی فسفولیپازها بوده و نقش کاتالیزوری در تشکیل اینوزیتول تری فسفات (IP3) از فسفات دی‌اینوزیتول بی‌فسفات (PIP2) دارند. حضور یون کلسیم جهت انجام این واکنش ضروری می‌باشد و نقش محوری در بسیاری از سیگنال‌های خارج سلولی دارند (Liang *et al.*, 2021). گزارش شده است ژن کاندیدای *ITPRI* نقش کلیدی در تفرق و تمایز میوبلاست دارد (Faenza *et al.*, 2004). در مطالعه ژنومی با هدف شناسایی ژن‌های کاندیدای مرتبط با پوکی استخوان، ژن *PLCG1* به عنوان ژن کاندیدا گزارش شده است (Guo *et al.*, 2019). با مطالعه پوشش کل ژنومی با هدف شناسایی ژن‌های کاندیدای مرتبط با قابلیت جذب فسفر، ژن *PLCG1* گزارش شده است (Vollmar *et al.*, 2021). همچنین تحلیل مسیر زیستی KEGG نشان‌دهنده این است که ژن *BMPRIA* به‌طور معنی‌داری با مسیر زیستی TGF-beta signaling pathway در ارتباط می‌باشند. این مسیر زیستی نقش مهمی در مکانیسم تولیدمثلی به‌خصوص سن بلوغ و رشد بدن دارد. در مطالعه پوشش ژنومی با هدف شناسایی جایگاه‌های ژنی مرتبط با توسعه سیستم اسکلتی مرتبط با اندازه بدن، ژن کاندیدای *BMPRIA* گزارش شده است (Kominakis *et al.*, 2017).



شکل ۳- مسیر سیگنالی کلسیم و ژن‌های کاندیدای مرتبط با صفت بال فرشته‌ای (پایگاه داده KEGG)
 Figure 3. Calcium signaling pathway and candidate genes related to angle wing trait (KEGG database)

References

Chen, H., Luo, K., Wang, C., Xuan, R., Zheng, S., Tang, H., Li, Y., Xiong, Y., Wu, Y., Wang, L., Ouyang, J., & Yan, X. (2023). Genomic characteristics and selection signals of Zhongshan ducks. *Animal*, 17(5), 100797.

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.

Deng, M.T., Zhu, F., Yang, Y.Z., Yang, F.X., Hao, J.P., Chen, S.R., & Hou, Z.C. (2019). Genome-wide association study reveals novel loci associated with body size and carcass yields in Pekin ducks. *BMC Genomics*, 20(1), 1.

El Yaagoubi, F.L., Charoute, H., Morjane, I., Sefri, H., Rouba, H., & Ainahi, A. (2017). Association analysis of genetic variants with metabolic syndrome components in the Moroccan population. *Current Research in Translational Medicine*, 65(3), 121-5.

Faenza, I., Bavelloni, A., & Fiume, R. (2004). Expression of phospholipase C beta family isoenzymes in C2C12 myoblasts during terminal differentiation. *Journal Cell Physiology*, 200, 291-6.

- Guo, L., Han, J., Guo, H., Lv, D., & Wang, Y. (2019). Pathway and network analysis of genes related to osteoporosis. *Molecular Medicine Reports*, (2), 985-994.
- Halgrain, M., Bernardet, N., Hennequet-Antier, C., Hincke, M., & Réhault-Godbert, S. (2023). RNA-seq analysis of the active chick embryo chorioallantoic membrane reveals genes that encode proteins assigned to ion transport and innate immunity. *Genomics*, 115(2), 110564.
- Huang, Y., Li, Y., Burt, D.W., Chen, H., Zhang, Y., & Qian, W. (2013). The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nature Genetics*, 45, 776–783.
- Kominakis, A., Hager-Theodorides, A.L., Zoidis, E., Saridaki, A., Antonakos, G., & Tsiamis, G. (2017). Combined GWAS and 'guilt by association'-based prioritization analysis identifies functional candidate genes for body size in sheep. *Genetic Selection Evolution*, 49(1), 41.
- Khaltabadi Farahani, A.H., Mohammadi, H., Moradi, M.H., Ghasemi, H.A., & Hajkhodadadi, I. (2020). Gene set enrichment analysis to identify genes and biological pathways associated with body weight in chicken. *Animal Production Research*, 9(3), 47-57. In Persian.
- Liang, S., Guo, H., Ma, K., Li, X., Wu, D., Wang, Y., Wang, W., Zhang, S., Cui, Y., Liu, Y., Sun, L., Zhang, B., Xin, M., Zhang, N., Zhou, H., Liu, Y., Wang, J., & Liu, L. (2021). A PLCB1-PI3K-AKT Signaling Axis Activates EMT to Promote Cholangiocarcinoma Progression. *Cancer Research*, 81(23), 5889-5903
- Li, G.S., Zhu, F., Zhang, F., Yang, F.X., Hao, J.P., & Hou, Z.C. (2021). Genome-wide association study reveals novel loci associated with feeding behavior in Pekin ducks. *BMC Genomics*, 22(1), 334.
- Li, G.S., Liu, W.W., Zhang, F., Zhu, F., Yang, F.X., Hao, J.P., & Hou, Z.C. (2020). Genome-wide association study of bone quality and feed efficiency-related traits in Pekin ducks. *Genomics*, 112(6), 5021-5028.
- Lin, M.J., Chang, S.C., Lin, T.Y., Cheng, Y.S., Lee, Y.P., & Fan, Y.K. (2015). Factors Affecting the Incidence of Angel Wing in White Roman Geese: Stocking Density and Genetic Selection. *Asian-Australasian Journal of Animal Sciences*, 29, 901 - 907.
- Lin, F.B., Zhu, F., Hao, J.P., Yang, F.X., & Hou, Z.C. (2018). In vivo prediction of the carcass fatness using live body measurements in Pekin ducks. *Poultry Science*, 97, 2365–2371.
- Liu, Z., Li, H., Zhong, Z., & Jiang, S. (2022). A Whole Genome Sequencing-Based Genome-Wide Association Study Reveals the Potential Associations of Teat Number in Qingping Pigs. *Animals (Basel)*, 12(9), 1057.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., & Kernytsky, A. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20, 1297–1303.
- Mohammadi, H., Khaltabadi Farahani, A.H., & Moradi, M. H. (2023). Genome-wide association study based on gene-set enrichment analysis of economically important traits in Japanese quail. *Animal Production Research*, 12(1), 65-78. In Persian.
- Mohammadabadi, M. (2021). Tissue-specific mRNA expression profile of ESR2 gene in goat. *Agriculture Biotechnology Journal*, 12 (4), 167-181. In Persian.
- Mohammadabadi, M., & Soflaei, M. (2020). Tissue-specific mRNA expression profile of BMP15 gene in goat. *Agriculture Biotechnology Journal*, 12, 191-208. In Persian.
- Peñagaricano, F., Weigel, K.A., Rosa, G.J., & Khatib, H. (2013). Inferring quantitative trait pathways associated with bull fertility from a genome-wide association study. *Frontiers in Genetics*, 3, 307-314.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., & Bender, D. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *The American Journal of Human Genetics*, 81, 559-575.
- Safaei, S.M.H., Dadpasand, M., & Mohammadabadi, M. (2022). An Origanum majorana Leaf Diet Influences Myogenin Gene Expression, Performance, and Carcass Characteristics in Lambs. *Animals*, 13 (1), e14.
- Shahsavari, M., Mohammadabadi, M., & Khezri, A. (2022). Effect of Fennel (*Foeniculum Vulgare*) Seed Powder Consumption on Insulin-like Growth Factor 1 Gene Expression in the Liver Tissue of Growing Lambs. *Gene Expression*, 21 (2), 21-26.
- Seabury, C.M., Oldeschulte, D.L., Saatchi, M., Beever, J.E., Decker, J.E., Halley, Y.A., Bhattarai, E.K., Molaie, M., Freetly, H.C., Hansen, S.L., Spangler, M.L., Weaber, R.L., Garrick, D.J., & Taylor, J.F. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics*, 18(1), 386-396.
- Srikanth, K., Lee, S.H., Chung, K.Y., Park, J.E., Jang, G.W., Park, M.R., Kim, N.Y., Kim, T.H., Chai, H.H., Park, W.C., & Lim, D. (2020). A Gene-Set Enrichment and Protein-Protein Interaction Network-Based GWAS with Regulatory SNPs Identifies Candidate Genes and Pathways Associated with Carcass Traits in Hanwoo Cattle. *Genes (Basel)*, 11(3), 316.
- Tang, H., Zhang, H., Liu, D., Wang, Z., Yu, D., Fan, W., Guo, Z., Huang, W., Hou, S., & Zhou, Z. (2022). Genome-wide association study reveals the genetic determinism of serum biochemical indicators in ducks. *BMC Genomics*, 23(1), 856.

- Vollmar, S., Haas, V., Schmid, M., Preuß, S., Joshi, R., Rodehutsord, M., & Bennewitz, J. (2021). Mapping genes for phosphorus utilization and correlated traits using a 4k SNP linkage map in Japanese quail (*Coturnix japonica*). *Animal Genetics*, 52(1), 90-98.
- Wang, S., Dvorkin, D., & Da, Y. (2012). SNPEVG: A graphical tool for GWAS graphing with mouse clicks. *BMC Bioinformatics*, 13, 319.
- Yuan, J., Wang, K., & Yi, G. (2015). Genome-wide association studies for feed intake and efficiency in two laying periods of chickens. *Genetic Selection Evolution*, 47, 82.
- Zhou, Z., Li, M., Cheng, H., Fan, W., Yuan, Z., & Gao, Q. (2018). An intercross population study reveals genes associated with body size and plumage color in ducks. *Nature Communications*, 9, 2648.
- Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44, 821.
- Zhuang, Z., Ding, R., Peng, L., Wu, J., Ye, Y., Zhou, S., Wang, X., Quan, J., Zheng, E., Cai, G., Huang, W., Yang, J., & Wu, Z. (2020). Genome-wide association analyses identify known and novel loci for teat number in Duroc pigs using single-locus and multi-locus models. *BMC Genomics*, 21(1), 344.