Sari Agriculture Sciences and Natural Resources University

Research on Animal Production, Vol. 17, Issue 2, 2026 p: 15 - 25

OPEN ACCESS

Research Paper

Functional Gene Enrichment Analysis Based on the GWAS of Growth Rate Traits in Yak

Shahrokh Jalali[1], Mohsen Gholizadeh[2], and Hasan Hafezian[3]

1- Ph.D. Student, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran
2- Professor, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran, (Corresponding author: M.gholizadeh@sanru.ac.ir)
3- Professor, Department of Animal Science, Faculty of Animal Science and Fisheries, Sari Agricultural Sciences and Natural Resources University, Sari, Iran

Received: 28 September, 2025 Revised: 12 January, 2026 Accepted: 24 January, 2026

Extended Abstract

Background: Growth traits are among the most important economic indicators in livestock production, directly linked to meat yield, resource efficiency, and profitability. In yak, which is well adapted to high-altitude regions, growth performance is often constrained by harsh environmental conditions. Understanding the genetic basis of growth traits can provide insights that contribute to molecular breeding and genetic improvement. Genome-wide association studies (GWAS) are valuable tools for identifying genomic regions influencing economically important traits in livestock and for advancing genomic breeding programs. However, GWAS has several limitations: markers are tested individually, and stringent statistical thresholds often result in many causal variants being undetected, particularly for polygenic traits. On high-density SNP chips, the effect of a gene may be distributed across multiple markers, preventing any from reaching genome-wide significance. Moreover, extensive linkage disequilibrium in livestock populations complicates the precise identification of causal mutations. Additionally, GWAS does not capture the real interactions among genes within biological pathways or functional networks, and due to the biallelic nature of SNPs, it cannot fully represent the effects of multi-allelic QTLs. To overcome these challenges, it has been suggested that analyses should extend beyond SNP-level to gene-level or gene-set approaches. In this framework, genes harboring significant SNPs that participate in common pathways or biological processes are tested for statistical enrichment. Such approaches enhance the power to detect causal variants and provide deeper insights into the genetic architecture of complex traits. Based on this rationale, the present study utilized publicly available genomic data to functionally analyze candidate genes associated with growth traits in yak and to biologically interpret GWAS results to provide more practical implications for breeding programs. The main objective was to identify genomic loci associated with post-weaning growth in Ashidan yaks through GWAS and to investigate functional pathways by means of gene enrichment analysis.

Methods: This study used previously published data from 354 female Ashidan yaks (Jia *et al*., 2020). Genotyping was performed using the Illumina BovineHD BeadChip, which included 777,962 SNP markers. Data quality control was conducted in PLINK. SNPs with more than 0.05 missing data, minor allele frequency (MAF) < 0.05, or significant deviation from the Hardy–Weinberg equilibrium ($P < 10^{-6}$) were excluded, as were individuals with > 0.05 missing data. Missing genotypes were imputed using the LD-kNNi algorithm in TASSEL. Population structure was evaluated by principal component analysis (PCA), and the first five PCs were included as covariates in the mixed linear model. GWAS was carried out in TASSEL v5.2 using a mixed linear model (MLM) incorporating both a kinship matrix (K) and population structure (Q). Gene Ontology (GO) analysis was performed to identify biological pathways associated with growth traits. All genes located within ±15 kb of SNPs with a $P < 0.05$ were extracted based on the Bos taurus UMD3.1 genome assembly. These candidate genes were then submitted to the g: Profiler tool for enrichment analysis in three GO domains: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC), as well as in KEGG pathways. Significance thresholds were set at a corrected $P < 0.05$.

Copyright ©2026 Jalali et al. Published by Sari Agricultural Sciences and Natural Resources University.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 Unported License which allows users to read, copy, distribute and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.

[Downloaded from rap.sanru.ac.ir on 2026-10-22]

[DOI: 10.61882/rap.2026.1565]

Results: Following quality control, one individual was excluded due to poor genotyping quality, leaving 353 yaks for further analysis. From the initial SNP dataset, 2,540 SNPs were removed due to high missingness, 65,122 SNPs were removed due to low MAF, and four SNPs deviating from the Hardy–Weinberg equilibrium were excluded. Ultimately, 32,573 high-quality SNPs remained. The average genotyping rate of the retained individuals was 99.6%. For the first trait (average daily gain from 6 to 12 months), no genomic regions reached genome-wide significance ($P < 1.5 \times 10^{-6}$). In contrast, for the second trait (average daily gain from 6 to 30 months), 22 significant SNPs ($P < 1.5 \times 10^{-6}$) were detected on chromosome 1, with the strongest signals appearing between 1.8 and 2.1 Mb. Within this region, the gene *C1H21orf62* emerged as the main candidate, harboring several significant SNPs. Additional candidate genes, such as *GCFC* and *SYNJ1*, were also identified downstream of this region. For the third trait (average daily gain from 12 to 30 months), a highly significant region was again detected on chromosome 1, overlapping with the same region as the second trait, with all significant SNPs closely associated with the C1H21orf62 gene. Functional enrichment analysis further revealed that the identified growth-associated genes were significantly enriched in multiple categories. At the molecular function level, significant enrichment was observed for ion binding (GO:0043167), small molecule binding (GO:0036094), transmembrane transporter activity (GO:0022857), cytoskeletal motor activity (GO:0003774), and hydrolase activity (GO:0016787). At the biological process level, candidate genes were enriched in cell adhesion (GO:0007155), monoatomic ion transport (GO:0006811), and nervous system development (GO:0007399). At the cellular component level, genes were predominantly enriched in cell junctions (GO:0030054), cell periphery (GO:0071944), and the plasma membrane (GO:0005886). KEGG pathway analysis revealed that the identified genes were mainly involved in motor protein-related pathways (KEGG:04814), emphasizing the regulatory complexity of growth-related genes.

Conclusion: This study successfully identified significant SNPs and genomic regions associated with post-weaning growth traits in Ashidan yaks. While no significant associations were detected for early growth (6–12 months), extended growth intervals (6–30 and 12–30 months) revealed several key loci and biologically relevant pathways. The integration of high-density genotyping, genotype imputation, mixed linear modeling, and gene enrichment analysis provided a comprehensive overview of the genetic architecture underlying growth in yaks. These findings lay a valuable foundation for future molecular studies and provide practical markers for use in yak breeding and selection programs.

Keywords: Gene Enrichment, Growth traits, GWAS, LD-kNNi, Mixed Linear Model, SNP, TASSEL, Yak

How to Cite This Article: Jalali, Sh., Gholizadeh, M., & Hafezian, H. (2026). Functional Gene Enrichment Analysis Based on the GWAS of Growth Rate Traits in Yak. *Res Anim Prod,* 17(2), 15-25. DOI: 10.61882/rap.2026.1565

[Downloaded from rap.sanru.ac.ir on 2026-10-22]

[DOI: 10.61882/rap.2026.1565]

مقاله پژوهشی

تحلیل غنی‌سازی عملکردی ژنی بر پایه‌ی مطالعه ارتباط ژنومی برای صفات نرخ رشد در یاک

شاهرخ جلالی[۱]، محسن قلی‌زاده[۲] و حسن حافظیان[۳]

۱- دانشجوی دکتری، گروه علوم دامی، دانشکده علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران
۲- استاد، گروه علوم دامی، دانشکده علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران، (نویسنده مسوول: m.gholizadeh@sanru.ac.ir)
۳- استاد، گروه علوم دامی، دانشکده علوم دامی و شیلات، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران

تاریخ دریافت: ۱۴۰۴/۰۷/۰۶ تاریخ ویرایش: ۱۴۰۴/۱۰/۲۲ تاریخ پذیرش: ۱۴۰۴/۱۱/۰۴
صفحه: ۱۵ تا ۲۵

چکیده مبسوط

مقدمه و هدف: صفات رشد از شاخص‌های اقتصادی مهم در تولید دام محسوب می‌شوند و به‌طور مستقیم با تولید گوشت، بهره‌وری منابع و سودآوری مرتبط هستند. در یاک که به نواحی مرتفع سازگار شده است، عملکرد رشد معمولاً تحت تأثیر شرایط سخت محیطی است. درک پایه‌های ژنتیکی صفات رشد می‌تواند بینشی را فراهم کند که به اصلاح‌نژاد مولکولی و افزایش بهره ژنتیکی کمک نماید. مطالعات ارتباط ژنومی (GWAS) ابزار مهمی برای شناسایی نواحی ژنومی مؤثر بر صفات اقتصادی دام و بهبود برنامه‌های اصلاح‌نژاد ژنومی هستند. با وجود این، این روش محدودیت‌هایی دارد چرا که هر نشانگر به‌تنهایی آزمون می‌شود و آستانه‌های آماری سخت‌گیرانه باعث می‌شوند که بسیاری از واریانت‌های علّی، به‌ویژه در صفات چندژنی، شناسایی نشوند. در تراشه‌های با تراکم بالای SNP نیز ممکن است اثر یک ژن میان چندین نشانگر توزیع شود و هیچ‌کدام به آستانه معنی‌داری نرسند. همچنین، پیوستگی ژنتیکی وسیع در دام‌ها، تعیین دقیق جهش‌های سببی را دشوارتر می‌کند. افزون بر این، GWAS تعامل واقعی ژن‌ها در مسیرهای زیستی و شبکه‌های عملکردی را نادیده می‌گیرد و به‌دلیل ماهیت دوآللیSNPها، قادر به نمایش کامل اثر QTLهای چندآللی نیست. برای رفع این محدودیت‌ها، پیشنهاد شده است که تحلیل‌ها از سطح SNP فراتر روند و به سطح ژن یا مجموعه ژنی گسترش یابند. در این روش، ژن‌هایی که حامل SNPهای معنادار هستند و در یک مسیر زیستی یا فرایند مشترک قرار دارند، از نظر غنای آماری بررسی می‌شوند. این رویکرد می‌تواند توان شناسایی واریانت‌های علّی را افزایش دهد و درک عمیق‌تری را از پایه ژنتیکی صفات کمی فراهم سازد. بر همین اساس، مطالعه حاضر با استفاده از داده‌های ژنومی منتشر شده، به تحلیل عملکردی ژن‌های کاندید مؤثر بر رشد در یاک و تفسیر بیولوژیکی نتایج GWAS پرداخت تا راهکارهای کاربردی‌تری برای اصلاح‌نژاد این گونه ارائه شوند. هدف این پژوهش، شناسایی جایگاه‌های ژنومی مرتبط با رشد پس از از شیرگیری در یاک‌های نژاد آشیدان با استفاده از مطالعه ارتباط کل ژنوم (GWAS) و بررسی مسیرهای عملکردی از طریق تحلیل غنای ژنی بود.

مواد و روش‌ها: در این مطالعه، از داده‌های منتشر شده مربوط به ۳۵۴ رأس یاک ماده آشیدان استفاده شد. ژنوتیپ‌گیری با استفاده از تراشه Illumina BovineHD BeadChip شامل ۷۷۷,۹۶۲ نشانگر SNP انجام شد. کنترل کیفیت داده‌ها در نرم‌افزار PLINK صورت گرفت. به این صورت که نشانگرهایی با بیش از ۰/۰۵ از دست رفته، فراوانی آلل فرعی (MAF) کمتر از ۰/۰۵، یا انحراف معنی‌دار از تعادل هاردی–واینبرگ ($P < 1e-6$) و افراد با بیش از ۰/۰۵ از دست رفته حذف شدند. برای برآورد ژنوتیپ‌های مفقود از الگوریتم LD-kNNi در نرم‌افزار TASSEL استفاده شد. ساختار جمعیت با تحلیل مؤلفه‌های اصلی در نرم‌افزار TASSEL (PCA) بررسی شد و پنج مؤلفه نخست به‌عنوان متغیر کمکی در مدل خطی مختلط لحاظ شدند. تحلیل GWAS با مدل خطی مختلط (MLM) در TASSEL v5.2 انجام شد که ماتریس خویشاوندی (K) و ساختار جمعیتی (Q) را در نظر می‌گرفت. به‌منظور شناسایی مسیرهای بیولوژیکی مرتبط با نرخ رشد در یاک، تجزیه و تحلیل هستی‌شناسی ژن (GO) نیز انجام شد. در این بخش، از تمام ژن‌هایی که در فاصله ۱۵± kb از SNP های دارای $P < 0.05$ قرار داشتند استفاده شد. استخراج موقعیت ژنی این SNPها نیز با مراجعه به نسخه ژنومی Bos taurus UMD3.1 انجام گرفت. فهرست ژن‌های شناسایی شده به ابزار آنلاین g:Profiler منتقل شد و تحلیل غنی‌سازی در سه دسته هستی‌شناسی ژن شامل فرایندهای زیستی(BP)، عملکرد مولکولی (MF)، و جزء سلولی (CC) و نیز مسیرهای KEGG صورت گرفت. سطوح معنی‌داری مسیرها بر اساس $P < 0.05$ تصحیح شده گزارش شدند.

یافته‌ها: با کنترل کیفیت داده‌های ژنومی با استفاده از نرم‌افزار PLINK، یک نمونه به‌دلیل کیفیت پایین ژنوتایپ حذف گردید و در نهایت ۳۵۳ فرد باقی ماندند. از مجموع نشانگرها، ۲۵۴۰ SNP به‌دلیل درصد بالای داده‌های گمشده، ۶۵۱۲۲ SNP به‌دلیل MAF پایین، و ۴ نشانگر به‌دلیل انحراف از تعادل هاردی–واینبرگ حذف شدند. در نهایت، ۳۲۵۷۳ نشانگر با کیفیت بالا برای تحلیل‌های بعدی باقی ماندند. میانگین نرخ ژنوتایپ در افراد باقی‌مانده برابر با ۹۹/۶ درصد بود. برای صفت اول (نرخ رشد از ۶ تا ۱۲ ماهگی) هیچ ناحیه ژنومی در سطح معنی‌داری ژنومی ($P < 1.5\times10^{-6}$) شناسایی نشد. در مقابل، برای صفت دوم (نرخ رشد از ۶ تا ۳۰ ماهگی)، تعداد ۲۲ SNP معنی‌دار ($P < 1.5\times10^{-6}$) روی کروموزوم ۱ شناسایی گردیدند. قوی‌ترین سیگنال‌ها در بازه ۱/۸ تا ۲/۱ مگاباز ظاهر شدند. در این ناحیه، ژن *C1H21orf62* به‌عنوان کاندید اصلی معرفی شد که دربرگیرنده چندین SNP معنی‌دار بود. علاوه بر این، در بخش پایین دست همین منطقه، ژن‌های *GCFC* و *SYNJ1* نیز شناسایی شدند. برای صفت سوم (نرخ رشد از ۱۲ تا ۳۰ ماهگی) نیز یک ناحیه بسیار معنی‌دار بر روی کروموزوم ۱ آشکار شد. تمامSNPهای معنی‌دار در این صفت در بازه مشابهی با صفت دوم قرار داشتند و همگی با ژن *C1H21orf62* در ارتباط بودند. تحلیل عملکرد مولکولی نشان داد که ژن‌های شناسایی شده در ارتباط با رشد یاک به‌طور معنی‌داری در دسته‌های متعددی غنی شدند. مهم‌ترین این دسته‌ها در سطح عملکرد مولکولی شامل اتصال یون‌ها (GO:0043167)، اتصال مولکول‌های کوچک (GO:0036094)، فعالیت ناقل غشایی (GO:0022857)، فعالیت موتور سیتواسکلتی (GO:0003774)، و فعالیت هیدرولازی (GO:0016787) بودند. در سطح فرایندهای زیستی، ژن‌های کاندید رشد در گروه‌های اصلی، از جمله چسبندگی سلولی (GO:0007155)، انتقال یون‌های ساده (GO:0006811) و رشد و تکوین سیستم عصبی (GO:0007399) غنی بودند. در بخش جزء سلولی، ژن‌ها به‌طور برجسته‌ای در پیوندگاه‌های سلولی (GO:0030054) پیرامون سلول (GO:0071944) و غشای پلاسمایی (GO:0005886) غنی شدند. بررسی مسیرهای زیستی در KEGG نشان داد که ژن‌های شناسایی شده به‌طور عمده در مسیرهای پروتئین‌های حرکتی (KEGG:04814) مشارکت داشتند.

نتیجه‌گیری: این مطالعهSNPها و نواحی ژنومی معناداری را در ارتباط با صفات رشد پس از ازشیرگیری در یاک آشی‌دان شناسایی کرد. در حالی‌که رشد اولیه (۶ تا ۱۲ ماه) فاقد ارتباط ژنومی معنادار بود، بازه‌های طولانی‌تر (۶ تا۳۰ و ۱۲ تا۳۰ ماه) چندین جایگاه و مسیر زیستی کلیدی را آشکار ساختند. ترکیب ژنوتیپ‌گیری با تراکم بالا، برآورد داده‌های مفقود، مدل‌های خطی مختلط و تحلیل غنای ژنی، تصویری جامع را از معماری ژنتیکی رشد در یاک فراهم کردند. این یافته‌ها پایه ارزشمندی برای مطالعات مولکولی آینده هستند و نشانگرهای عملی را برای برنامه‌های اصلاح نژادی و انتخاب در یاک ارائه می‌دهند.

واژه‌های کلیدی: آنالیز غنی‌سازی، صفات رشد، مدل خطی ترکیبی، یاک، GWAS، SNP، TASSEL

[Downloaded from rap.sanru.ac.ir on 2026-10-22]
[DOI: 10.61882/rap.2026.1565]

مقدمه

یاک (*Bos grunniens*) یکی از اعضای خانواده گاوسانان است که به‌طور عمده در فلات تبت زندگی می‌کند و به‌عنوان اصلی‌ترین گونه دام اهلی در این منطقه شناخته می‌شود. این حیوان نقش حیاتی را در معیشت مردم بومی و عشایر ایفا می‌کند و منبع مهمی از گوشت، شیر، پشم و نیروی حمل و نقل به‌شمار می‌رود. یاک‌ها به‌دلیل سازگاری بالا با شرایط سخت محیطی مانند ارتفاع زیاد، دمای پایین و کمبود اکسیژن، از ویژگی‌های فیزیولوژیکی و ظاهری خاصی برخوردارند. در یاک‌های گوشتی، صفات رشدی نظیر وزن بدن، ارتفاع جدوگاه، طول بدن و محیط سینه از مهم‌ترین شاخص‌ها برای افزایش بهره‌وری تولید و سود اقتصادی محسوب می‌شوند. یاک آشیدان (Ashidan yak) که از جمعیت یاک‌های بدون شاخ فلات چینگهای تبت منشأ گرفته است، یک واریته جدید اصلاح شده است که با ویژگی‌هایی هم‌چون بی‌شاخ بودن، خلق و خوی آرام و عملکرد تولیدی مطلوب شناخته می‌شود و در سال‌های اخیر توجه ویژه‌ای را در سیستم‌های پرورش نیمه متمرکز و متمرکز به‌خود جلب کرده است (Jia *et al*., 2019; Ge *et al*., 2020).

بررسی‌های ژنتیکی اخیر در یاک نشان داده‌اند که صفات رشدی مانند وزن بدن، قد، طول بدن و دور سینه تحت تأثیر چندین نشانگر ژنتیکی (SNP) قرار دارند. محققین با تمرکز بر یکی از صفات مرفولوژیک مهم در یاک‌های اهلی، یعنی صفت بی‌شاخی، تلاش کردند تا پایه ژنتیکی این ویژگی را شناسایی کنند. آن‌ها ۵۰ یاک بدون شاخ را از نژاد Datong برای گروه مورد و ۵۱ یاک شاخ‌دار را به‌عنوان گروه کنترل انتخاب کردند و با استفاده از روش‌های توالی‌یابی DNA و آزمون ذوب با دقت بالا، به بررسی پلی‌مورفیسم‌های ژنتیکی پرداختند. آنالیزهای آماری نشان دادند که یک ناحیه ۱۴۷ کیلوبازی در کروموزوم ۱ که شامل سه ژن *C1H21orf62*، *GCFC1* و *SYNJ1* است، به احتمال زیاد جایگاه جهش مؤثر در ایجاد صفت بی‌شاخی در یاک است (Liu *et al*., 2014). همچنین، با استفاده از GWAS چند لوکوسی در ۳۵۴ یاک ماده، تعداد قابل توجهی SNP مرتبط با صفات رشدی شناسایی شدند که بیشتر آن‌ها در فرآیندهای مرتبط با رشد عضلات، ذخیره چربی و توسعه سیستم عصبی دخیل هستند. این یافته‌ها بینش‌های جدیدی را درباره معماری مولکولی رشد یاک فراهم کرده‌اند (Jia *et al*., 2019). از سوی دیگر، با بررسی قابلیت پیش‌بینی صفات رشدی اولیه یاک با روش‌های انتخاب ژنومی و مقایسه پنج روش آماری مختلف، نشان داده شده است که صحت پیش‌بینی برای این صفات بین ۰/۱۴۷ تا ۰/۳۹۱ متغیر است. نتایج حاکی از امکان‌پذیری استفاده از انتخاب ژنومی در بهبود صفات رشد یاک بودند (Ge *et al*., 2020). علاوه بر این، مطالعه جیانگ و همکاران (Jiang *et al.,* 2022) با استفاده از تراشه با تراکم بالای SNP در ۴۸۰ یاک از سه نژاد مختلف، SNPهایی را شناسایی کرد که به طور معنی‌داری با وزن بدن مرتبط بودند. همچنین، این پژوهش به تأثیر احتمالی برخیSNPها بر روی ژن‌های کد کننده پروتئین از جمله *GRK4* پرداخت که ممکن است از طریق تأثیر بر هورمون‌های مرتبط با وزن بدن مانند PTHوADM اثرگذار باشد.

مطالعات ارتباط ژنومی (GWAS) برای شناسایی و بهره‌گیری از نواحی ژنومی و تقویت برنامه‌های اصلاح‌نژاد ژنومی به‌طور گسترده به‌کار گرفته شده‌اند (Gholizadeh *et al*., 2015; Pasandideh *et al*., 2018). با این‌که نشانگرهای ژنتیکی شناسایی شده در ارتباط با یک صفت می‌توانند اطلاعات مفیدی را ارائه دهند، اما معمولاً تنها بخش نسبتاً کوچکی از وراثت‌پذیری آن صفت را تبیین می‌کنند. این موضوع ناشی از معماری ژنتیکی پیچیده صفات و همچنین قدرت آماری محدود روش‌های به‌کار رفته در مطالعات GWAS است (Mackay *et al*., 2009). بنا بر این، این قبیل مطالعات، به‌ویژه روش‌هایی که هر نشانگر تک نوکلئوتیدی (SNP) را به‌صورت جداگانه آزمون می‌کنند، به‌دلیل آستانه‌های آماری سخت‌گیرانه‌ای که برای کنترل آزمون‌های چندگانه در این مطالعات استفاده می‌شوند تا نرخ خطای نوع اول کنترل شود، منجر به سخت‌گیری زیاد در آستانه معناداری و کاهش قدرت تشخیص واریانت‌های علّی می‌گردند و ممکن است تعداد قابل توجهی از نشانگرهای مهم، به‌ویژه در صفات چند ژنی، شناسایی نشوند (Peng *et al*., 2010). تصحیح برای آزمایش چندگانه[1]، امکان معنی‌دار بودن تعداد کمی از SNPها را فراهم می‌کند، در حالی‌که نشانگرهای واقعاً مرتبط که به آستانه معنی‌داری نمی‌رسند، شناسایی نمی‌شوند. رویکرد کلاسیک تک واریانت برای تعیین ارتباط بین یک واریانت و یک فنوتیپ محدودیت دارد و با توجه به این که اکثر صفات اقتصادی ناشی از تأثیر واریانت‌های ژنتیکی متعدد در ژن‌ها و مسیرهای متعدد هستند، قدرت تشخیص آماری را کاهش می‌دهد (Chimusa *et al*., 2016). افزون بر این، در پنل‌های ژنوتایپی با تراکم بالای SNP، ممکن است یک ژن توسط چندین SNP مجاور بیان شود و همین موضوع باعث شود تا اثر آن ژن بین چند نشانگر تقسیم شود، به‌طوری‌که هیچ‌یک از آن‌ها به‌تنهایی از آستانه معنی‌داری عبور نکند (Ha *et al*., 2015). همچنین، در بسیاری از گونه‌های دام، پیوستگی ژنتیکی، نواحی گسترده‌ای از ژنوم را در بر می‌گیرد. در نتیجه، ممکن است تعداد زیادی SNP در پیوستگی با ناحیه ژنومی سببی قرار گیرند که این موضوع، شناسایی دقیق جهش‌های سببی را دشوارتر می‌سازد (Hayes, 2013). علاوه بر این، اگرچه GWAS قادر است نشانگرهایی را که با صفات مورد نظر به‌طور معنی‌داری مرتبط هستند شناسایی کند، اما این روش از واقعیت زیستیِ تعامل ژن‌ها در قالب مسیرهای زیستی و شبکه‌های عملکردی صرف‌نظر می‌کند.

[1] Multiple testing correction

[Downloaded from rap.sanru.ac.ir on 2026-10-22]
[DOI: 10.61882/rap.2026.1565]

همچنین، به‌دلیل ماهیت دو آللیSNPها، تأثیر برخی QTLهای چندآللی ممکن است به‌طور کامل در این تحلیل‌ها منعکس نشود. در نتیجه، تکیه بر GWAS ممکن است درک محدودی را از ماهیت پیچیده صفات کمی ارائه دهد (Dadousis *et al*., 2017). برای غلبه بر این محدودیت‌ها و دستیابی به درک عمیق‌تر از زمینه ژنتیکی صفات پیچیده، پیشنهاد شده است که تحلیل از سطح SNP به سطوح بالاتری مانند ژن و مجموعه‌های ژنی گسترش یابد (Dadousis *et al*., 2017). در آنالیز مجموعه ژنی، گروهی از ژن‌های مرتبط (از قبیل ژن‌های متعلق به یک مسیر زیستی خاص یا هستی‌شناسی ژن) که حاملSNPهای معنی‌دار شناسایی شده در GWAS هستند، از نظر فراوانی در یک مسیر مشخص مورد آزمون قرار می‌گیرند تا مشخص شود که آیا بیش از حد انتظار در آن مسیر زیستی نمایان شده‌اند یا خیر (Wang *et al*., 2011). در سال‌های اخیر، علاقه‌مندی فزاینده‌ای به تحلیل مسیرهای زیستی در گونه‌های مختلف دام شکل گرفته است تا به‌عنوان مکملی برای تحلیل‌های GWAS در بررسی صفات کمی مورد استفاده قرار گیرد (Dadousis *et al*., 2017; Esmaeili-Fard *et al*., 2021). تحلیل گروهیِ SNPها که از نظر عملکردی با هم مرتبط هستند یا به لحاظ فیزیکی در نزدیکی یکدیگر قرار دارند، می‌تواند توان تشخیص واریانت‌های سببی را افزایش دهد و درک بهتری را از زمینه زیستی بیماری‌ها یا صفات پیچیده فراهم آورد (Chen *et al*., 2011). هدف مطالعه حاضر، تحلیل عملکردی ژن‌های کاندید مؤثر بر نرخ رشد در یاک از طریق آنالیز غنی‌سازی مجموعه ژنی مبتنی بر داده‌های GWAS بود. برای این منظور، داده‌های ژنومیِ منتشر شده (Jia *et al*., 2019) استخراج گردیدند و در چارچوب سه شاخص افزایش وزن روزانه شامل بازه‌های زمانی ۶ تا ۱۲ ماهگی، ۶ تا ۳۰ ماهگی و ۱۲ تا ۳۰ ماهگی مورد بررسی قرار گرفتند تا با تفسیر بیولوژیکی دقیق‌تر نتایج GWAS، زمینه برای استفاده هدفمندتر از اطلاعات ژنتیکی در برنامه‌های اصلاح‌نژاد یاک فراهم گردد.

مواد و روش‌ها

داده‌های این مطالعه شامل اطلاعات ژنوتیپی و فنوتیپی مربوط به ۳۵۴ رأس یاک ماده نژاد آشیدان هستند که از پایگاه داده‌های عمومی (Jia *et al*., 2019) استخراج گردیدند. تعیین ژنوتیپ با استفاده از چیپ Illumina BovineHDBeadChip (Illumina Inc.) شامل ۷۷۷۹۶۲ مارکر SNP انجام شد و داده‌ها در فرمت PED/MAP در دسترس بودند.

کنترل کیفیت و آماده‌سازی داده‌ها

برای کنترل کیفیت داده‌های ژنوتیپی از نرم‌افزار PLINKv1.9 (Purcell *et al*., 2007) استفاده شد. نمونه‌ها و نشانگرها با شرایط زیر حذف شدند: نرخ داده‌ی گم شده بیش از ۵ درصد در افراد (mind 0.05--)، نرخ داده‌ی گم شده بیش از ۵ درصد درSNPها (geno 0.05--) فراوانی آلل کمیاب کمتر از ۵ درصد (maf 0.05--) و انحراف از تعادل هاردی–واینبرگ ($P \leq$ 1e-6). پس از اعمال فیلترها، ۳۵۳ حیوان و ۳۲۵۷۳ SNP برای تحلیل‌های نهایی حفظ شدند. داده‌های باقیمانده برای تحلیل وارد نرم‌افزار TASSEL v5.2 شدند و با استفاده از الگوریتم LD-kNNi ایمپیوتیشن داده‌های گم شده صورت گرفت. مزیت اصلی LD-kNNi، استفاده از ساختار LD بین SNPها برای افزایش دقت و کاهش خطای ایمپیوتیشن است.

تجزیه و تحلیل GWAS مدل خطی ترکیبی

برای کنترل ساختار جمعیت، آنالیز مولفه‌های اصلی در TASSEL (Bradbury *et al*., 2007) انجام شد و پنج مؤلفه اصلی اول (PCs) به‌عنوان کوارییت وارد مدل شدند. تحلیل ارتباط ژنومی با استفاده از مدل خطی ترکیبی در TASSEL اجرا شد که به‌صورت زیر تعریف می‌شود:

$$Y = Xb + Zu + e$$

در این مدل، y بردار فنوتیپ‌ها، X ماتریس طرح اثرات ثابت شامل نشانگر SNP مورد آزمون و مؤلفه‌های اصلی (PCs) جهت کنترل ساختار جمعیت، b بردار ضرایب اثرات ثابت، Z ماتریس طرح اثرات تصادفی، و u، بردار اثرات ژنتیکی تصادفی پلی‌ژنی هستند ($u \sim N(0, K\sigma_u^2)$، که در آن K ماتریس خویشاوندی ژنومی برآورد شده از داده‌های SNP است. بردار e نشان دهنده خطای تصادفی با توزیع نرمال است؛ ($e \sim N(0, I\sigma_e^2)$ که در آن I ماتریس واحد است. در این مدل، ماتریس خویشاوندی (K) با روش Centered_IBS در TASSEL محاسبه شد و ماتریس Q شامل پنج PC اول حاصل از PCA بود که برای کنترل لایه‌بندی جمعیت وارد مدل شدند. سطح معنی‌داری بر اساس تصحیح بونفرونی، $\alpha = {}^{0.05}/_{32579} = 1.5 \times 10^{-6}$ در نظر گرفته شد. SNPهای با P کمتر از این مقدار به عنوان نشانگرهای معنی‌دار در نظر گرفته شدند و سپس ژن‌های واقع در فاصله ۱۵± کیلوباز از این نشانگرها با استفاده از مرورگر ژنومی NCBI بر اساس نسخه ژنومی UMD3.1 شناسایی شدند.

جهت نمایش نتایج تحلیل GWAS، نمودارهای Manhattan و QQ plot برای هر یک از صفات رشد با استفاده از بسته‌های CMplot (Yin *et al*., 2021) و qqman (Turner *et al*., 2014) در محیط R ترسیم شدند. همچنین، نمودار چگالی توزیع مارکرها بر روی کروموزوم‌ها با استفاده از CMplot (Yin *et al*., 2021) تهیه گردید.

تحلیل غنی‌سازی مسیرهای زیستی

به‌منظور شناسایی مسیرهای بیولوژیکی مرتبط با صفات رشد، تحلیل غنی‌سازی عملکردی نیز انجام شد. در این بخش، از تمامی ژن‌هایی که در فاصله ۱۵ kb± ازSNPهای دارای $P < 0.05$ قرار داشتند (صرف‌نظر از سطح بونفرونی)، استفاده شد. استخراج موقعیت ژنی این‌SNPها نیز با مراجعه به نسخه ژنومی Bos taurus UMD3.1 انجام گرفت. فهرست ژن‌های شناسایی شده به ابزار آنلاین g:Profiler (Raudvere *et al*., 2019) منتقل شد و تحلیل غنی‌سازی در سه دسته‌ی Gene Ontology (GO)، Cellular Component(CC) و نیز مسیرهای KEGG صورت گرفت. سطوح معنی‌داری مسیرها بر اساس adjusted $P < 0.05$ گزارش شدند.

[Downloaded from rap.sanru.ac.ir on 2026-10-22]
[DOI: 10.61882/rap.2026.1565]

جهت ارزیابی این که آیا مسیرهای زیستی یا دسته‌های عملکردی خاصی در میان ژن‌های شناسایی شده به‌طور معنی‌داری بیشتر از انتظار ظاهر شدند، از آزمون آماری مبتنی بر توزیع هیپرجئومتریک (Hypergeometric test) استفاده شد. این آزمون بررسی می‌کند که آیا تکرار مشاهده شده یک مجموعه ژنی) مثلاً مسیر زیستی یا دسته‌ی GO خاص در لیست ژن‌های استخراج شده، به‌صورت تصادفی رخ داده است یا نشان دهنده‌ی غنی‌سازی واقعی است.

$$P(X=k)=\frac{\binom{K}{k}\cdot\binom{N-K}{n-k}}{\binom{N}{n}}$$

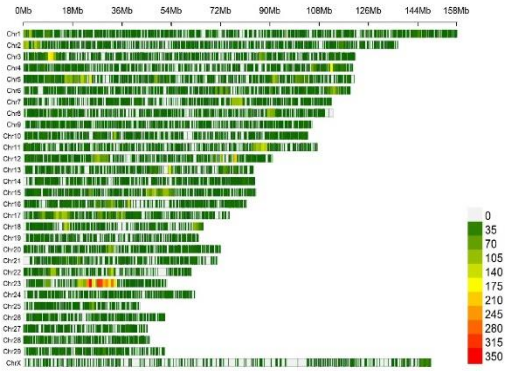
در این آزمون، N تعداد کل ژن‌ها در پس‌زمینه (مثلاً کل ژن‌های موجود در ژنوم مرجع گاوی)، K تعداد ژن‌های مرتبط با مسیر زیستی یا دسته عملکردی مورد نظر، N تعداد کل ژن‌های شناسایی شده در این مطالعه (برای مثال ژن‌های مجاورSNP های با $P < 0.05$ و K تعداد ژن‌های موجود در تقاطع این دو مجموعه است. احتمال مشاهده‌ی k یا بیشتر از این ژن‌ها تحت فرض تصادفی‌بودن، به‌صورت تجمعی محاسبه می‌شود و مقدار P مربوطه به‌دست می‌آید. این P بیانگر میزان غنی‌سازی آن مسیر یا دسته عملکردی در میان ژن‌های ورودی است:

$$p\text{-value}=\sum_{i=k}^{\min(n,K)}P(X=i)$$

در این مطالعه، محاسبات مذکور به‌صورت خودکار با استفاده از ابزار g:Profiler انجام شدند. برای تصحیح آزمایش چندگانه، از روش تنظیم سطح معنی‌داری بر اساس Benjamini Hochberg FDR یا الگوریتم اختصاصی g:SCS استفاده گردید. مسیرها و دسته‌هایی با P اصلاح شده کمتر از ۰/۰۵ به‌عنوان معنی‌دار گزارش شدند.

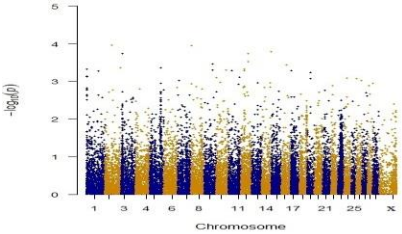
نتایج و بحث

در این مطالعه، با هدف شناسایی نشانگرهای ژنتیکی مؤثر بر صفات رشد پس از ازشیرگیری در یاک‌های ماده نژاد آشیدان، از داده‌های ژنوتیپی با تراکم بالا حاصل از آرایه Illumina BovineHD BeadChip و داده‌های فنوتیپی مربوط به نرخ رشد در سه بازه زمانی استفاده شد.



شکل ۱- چگالی نشانگرهای ژنتیکی روی هر یک از کروموزوم‌های یاک (بر اساس ژنوم مرجع UMD 3.1) پس از کنترل کیفیت
Figure 1. SNP density across each yak chromosome (UMD 3.1) after quality control

کنترل کیفیت داده‌های ژنومی با استفاده از نرم‌افزار PLIN انجام گرفت. پس از اعمال این فیلترها، یک نمونه به دلیل کیفیت پایین تعیین ژنوتیپ حذف گردید و در نهایت ۳۵۳ فرد باقی ماندند. از مجموع نشانگرها، ۲۵۴۰ SNP به دلیل درصد بالای داده‌های گم شده، ۶۵۱۲۲ SNP به دلیل MAF پایین، و چهار نشانگر SNP به دلیل انحراف از تعادل هاردی-واینبرگ حذف شدند. در نهایت، ۳۲۵۷۳ نشانگر با کیفیت بالا برای تحلیل‌های بعدی باقی ماندند. میانگین نرخ ژنوتایپ در افراد باقی‌مانده برابر با ۹۹/۶ درصد بود. شکل ۱ چگالی توزیع نشانگرها در طول کروموزوم‌های گاو را نمایش می‌دهد که حاکی از توزیع نسبتاً یکنواخت SNPها در سرتاسر ژنوم است.



شکل ۲- نمودار منهتن برای نرخ رشد روزانه در بازه ۶ تا ۱۲ ماهگی یاک‌ها بر اساس مدل خطی مختلط در TASSEL
Figure 2. The Manhattan plot of average daily gain from 6 to 12 months in yaks using MLM in TASSEL

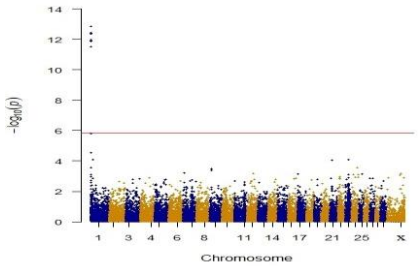
[Downloaded from rap.sanru.ac.ir on 2026-10-22]
[DOI: 10.61882/rap.2026.1565]

نتایج آنالیز ارتباط ژنومی (GWAS)

در این مطالعه، سه صفت مربوط به نرخ رشد یاک در بازه‌های مختلف سنی مورد بررسی قرار گرفتند. برای صفت اول (نرخ رشد از ۶ تا ۱۲ ماهگی) هیچ ناحیه ژنومی در سطح معنی‌داری ژنومی ($P < 1.5\times10^{-6}$) شناسایی نشد (شکل ۲). در مقابل، برای صفت دوم (نرخ رشد از ۶ تا ۳۰ ماهگی) تعداد ۲۲ SNP معنی‌دار ($P < 1.5\times10^{-6}$) روی کروموزوم ۱ آشکار گردیدند (شکل ۳). قوی‌ترین سیگنال‌ها در بازه ۱/۸ تا ۲/۱ مگاباز ظاهر شدند که شامل SNPهایی با مقدار p بسیار پایین‌تر از آستانه تصحیح بنفرونی به ویژه BovineHD0100000619 با $P = 1.42\times10^{-13}$بودند. در این ناحیه، ژن *C1H21orf62* به عنوان کاندید اصلی معرفی شد که در چندین SNP تکرار گردید. علاوه بر این، در بخش پایین‌دست همین منطقه، ژن‌های *GCFC* و *SYNJ1* نیز شناسایی شدند.

برای صفت سوم (نرخ رشد از ۱۲ تا ۳۰ ماهگی) نیز یک ناحیه بسیار معنی‌دار بر روی کروموزوم ۱ آشکار شد (شکل ۴). تمام SNPهای معنی‌دار در این صفت در بازه مشابهی با صفت دوم قرار داشتند و همگی با ژن *C1H21orf62* در ارتباط بودند. این هم‌پوشانی نشان می‌دهد که این ژن و ناحیه ژنومی مربوطه احتمالاً در رشد طولانی‌مدت یاک (از ۶ تا ۳۰ ماهگی) نقش کلیدی را ایفا می‌کند.

نکته‌ای قابل توجه در نتایج این مطالعه، شناسایی سه ژن *C1H21orf62*، *GCFC1* و *SYNJ1* به عنوان ژن‌های کاندید در صفات مرتبط با نرخ رشد بود؛ ژن‌هایی که پیش‌تر در چارچوب بررسی ژنتیکی صفت بی‌شاخی (polledness) در یاک گزارش شده بودند (Liu *et al*., 2014). اگر چه مطالعه مذکور این ژن‌ها را در یک ناحیه کاندید برای جهش مؤثر بر ایجاد فنوتیپ بدون شاخ شناسایی کرده بود، نتایج حاصل از آنالیز حاضر نشان می‌دهند که این ناحیه می‌تواند نقش گسترده‌تری در کنترل صفات رشد نیز ایفا کند. این هم‌پوشانی می‌تواند ناشی از اثر پلیتروپی ژن در بروز چندین صفت، یا نتیجه عدم پیوستگی ژنتیکی (LD) در آن ناحیه از ژنوم باشد. به ویژه ژن‌هایی مانند *SYNJ1* که در عملکردهای نورونی، سیگنال‌دهی سلولی و رشد بافت‌ها دخیل هستند، می‌توانند هم در رشد بدنی و هم در تکوین ساختارهایی نظیر شاخ نقش ایفا کنند. این یافته اهمیت بررسی عملکردی این ژن‌ها را در بافت‌ها و مراحل مختلف رشد یاک دوچندان می‌سازد و می‌تواند زمینه‌ساز مطالعات آینده در راستای درک بهتر از سازوکارهای ژنتیکی صفات اقتصادی این گونه باشد.

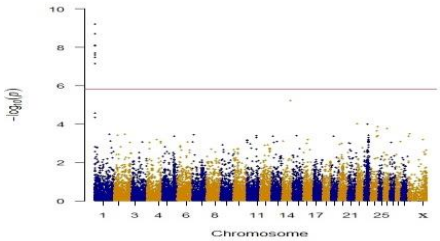


شکل ۳- نمودار منهتن برای نرخ رشد روزانه در بازه ۶ تا ۳۰ ماهگی یاک‌ها که چندین ناحیه ژنومی با سیگنال‌های معنی‌دار آشکار شدند.
Figure 3. The Manhattan plot of average daily gain from 6 to 30 months in yaks, showing several significant genomic regions.

عملکرد ژن‌های *C1H21orf62*، *GCFC1* و *SYNJ1* در گاو و یاک هنوز به‌طور کامل مشخص نشده است، اما این ژن‌ها در اغلب گونه‌ها، از جمله انسان، موش صحرایی و موش آزمایشگاهی، حفظ شده‌اند (Liu *et al*., 2013). همولوگ انسانی ژن C1H21orf62 در شبکه‌های سیگنال‌دهی شرکت دارد که رشد و تمایز سلول‌های بنیادی جنینی را کنترل می‌کنند (Brandenberger *et al*., 2004). ژن *GCFC1* که با نام‌های پروتئین باندینگ *PAX3* و *PAX7*، و *Paxbp1* نیز شناخته می‌شود (Genecards)، پروتئینی را کد می‌کند که به توالی‌های DNA غنی از باز GC متصل می‌شود؛ این موضوع نشان می‌دهد که ممکن است در تنظیم رونویسی ژن‌ها نقش داشته باشد (Hu *et al*., 2006). دو موج مشابه اما متمایز از میوژنز در مهره‌داران وجود دارند: میوژنز جنینی و میوژنز بالغ (Diao *et al*., 2012). در موش‌ها، موج اول با سلول‌های پیش‌ساز ماهیچه (MPCs) در میوتوم در حدود روز هشتم جنینی با بیان Myf5، یکی از چهار فاکتور تنظیمی میوژنز (MRFs) شروع می‌شود (Puri & Sartorelli, 200005)، در حالی‌که موج دوم توسط سلول‌های بنیادی ماهیچه، که سلول‌های ماهواره‌ای ماهیچه (MSCs) نیز نامیده می‌شوند، انجام می‌شود که از ناحیه مرکزی درمومیوتوم با بیان Pax7 و Pax3 منشأ می‌گیرند و بیان Pax7 و Pax3 در آن‌ها نشان دهنده هویت بنیادی و تعهد میوژنیک این سلول‌ها است (Ben-Yair & Kalcheim, 2005). در عضلات بالغ، سلول‌های بنیادی مزانشیمی (MSCs) خاموش هستند و بین لایه پایه و سارکولمای میوفیبرها قرار دارند و عمدتاً مسئول رشد عضلات پس از تولد هستند (Biressi & Rando, 2010; Buckingham & Relaix, 2007). ژن SYNJ1 که با نام سیناپتوجانین ۱ (Synaptojanin 1) نیز شناخته می‌شود می‌تواند از طریق آنزیم فسفاتاز phosphatidylinositol 4,5-bisphosphate بازآرایی رشته‌های اکتین را تنظیم کند (Sakisaka *et al*., 1997).

[Downloaded from rap.sanru.ac.ir on 2026-10-22]
[DOI: 10.61882/rap.2026.1565]

ژن SYNJ1 پروتئینی را کد می‌کند که سطح فسفاتیدیل اینوزیتول-۴،۵- بیس فسفات غشایی را تنظیم می‌کند و بر انتقال سیناپسی و عبور و مرور غشایی در سیستم عصبی تأثیر می‌گذارد (Mani *et al*., 2007).



شکل ۴- نمودار منهتن برای نرخ رشد روزانه در بازه ۱۲ تا ۳۰ ماهگی یاک که ناحیه‌ای برجسته روی کروموزوم ۱ شناسایی گردید.
Figure 4. The Manhattan plot of average daily gain from 12 to 30 months in yaks, highlighting a major association signal on chromosome 1.

نتایج آنالیز غنی‌سازی ژن‌ها

آنالیز غنی‌سازی ژن‌ها بر اساس ژن‌های استخراج شده از نواحی دربرگیرنده ۱۵ کیلوباز مجاور SNPهای با سطح معنی‌داری کمتر از ۰/۰۵ انجام شد. نتایج نشان دادند که در سطح عملکرد مولکولی (Molecular Function)، ترم‌های متعددی به‌طور معنی‌دار غنی‌سازی شدند که از مهم‌ترین آن‌ها می‌توان به اتصال یون، اتصال به مولکول‌های کوچک، فعالیت ناقل‌های غشایی، فعالیت موتورهای اسکلت سلولی، اتصال مولکول‌های چسبندگی سلولی و فعالیت‌های وابسته ATP اشاره کرد (جدول ۱). در گوسفند، تجزیه و تحلیل مجموعه ژن‌ها نشان داد که انتقال یون کلسیم به‌عنوان یک عامل مهم در رشد و نمو، صفات را در بدو تولد تحت تأثیر قرار می‌دهد و این نشان دهنده نقش آن در رشد اولیه است (Esmaeili-Fard *et al*., 2021). علاوه بر این، در گاوهای گوشتی، مطالعات انجام شده روی مصرف خوراک باقیمانده (RFI) معیاری که ارتباط نزدیکی با راندمان رشد دارد انتقال یون و انتقال کاتیون را از جمله غنی‌ترین فرآیندهای بیولوژیکی مرتبط با راندمان خوراک شناسایی کردند (Neibergs *et al*., 2016). این یافته‌ها تأکید می‌کنند که چگونه مکانیسم‌های انتقال یون ممکن است بر جذب مواد مغذی، هموستاز سلولی و متابولیسم انرژی که همگی برای فرآیندهای رشد در دام‌ها ضروری هستند تأثیرگذار باشند. فعالیت مولکولیِ اتصال به مولکول‌های کوچک به عملکردهایی که در آن‌ها پروتئین‌ها به مولکول‌های کوچک مانند متابولیت‌ها، یون‌ها یا لیگاندها متصل می‌شوند. این عملکرد در واقع می‌تواند با ویژگی‌های رشد مرتبط باشد، زیرا مولکول‌های کوچک (مانند هورمون‌ها، واسطه‌های متابولیکی، عوامل سیگنالینگ) اغلب فرآیندهای سلولی حیاتی برای رشد و نمو را تنظیم می‌کنند. مولکول‌های کوچک می‌توانند فاکتورهای رونویسی و بیان ژن‌های پایین‌دست را تعدیل کنند و بر مسیرهای تنظیم رشد تأثیر بگذارند (McNerney *et al*., 2018). برهمکنش‌های پروتئین-مولکول کوچک برای اکثر فرآیندهای بیولوژیکی، از جمله تنظیم آنزیم، سیگنالینگ و هموستاز، که همگی در رشد و نمو نقش دارند، موثر هستند (Wang *et al*., 2024). ناقل‌های غشایی با بهبود استفاده از مواد مغذی، انتقال قند و تحمل به تنش، رشد را افزایش می‌دهند که همگی در افزایش زیست‌توده و عملکرد نقش دارند (Gill *et al*., 2021). در حالی‌که تحقیقات در دام محدود هستند، ناقل‌های غشایی (به‌ویژه ناقل‌های SLC) جذب مواد مغذی و متابولیت‌ها را تنظیم می‌کنند و بنا بر این می‌توانند به‌طور غیر مستقیم بر صفات رشد تأثیر بگذارند (Zhang *et al*., 20219). فعالیت اگزو-آلفا-سیالیداز (GO:0004308) دسته‌ای از آنزیم‌ها را توصیف می‌کند که باقی‌مانده‌های اسید سیالیک انتهایی را از گلیکوکونژوگه‌هایی مانند گلیکوپروتئین‌ها، گلیکولیپیدها و الیگوساکاریدها جدا می‌کنند (Genecards). در پستانداران، آنزیم‌های مرتبط (مانند NEU1، NEU2، و NEU4) نشان داده‌اند که نقشی فراتر از کاتابولیسم صرف را ایفا می‌کنند. به‌عنوان مثال، مطالعات روی جوندگان نشان می‌دهند که NEU1 آنزیمی که فعالیت اگزو-آلفا-سیالیداز را فراهم می‌کند در تنظیم مثبت توسعه برون‌ریزی نورون‌ها و تکثیر میوبلاست‌ها نقش دارد (NCBI). این نشان می‌دهد که فعالیت اگزو-آلفا-سیالیداز ممکن است در فرآیندهای مهم برای رشد، از جمله تمایز سلولی و رشد عضلات، نقش داشته باشد.

در سطح فرآیند زیستی (Biological Process)، مهم‌ترین مسیرهای معنی‌دار شامل چسبندگی سلولی، انتقال یون‌های ساده، توسعه سیستم عصبی، فرآیندهای وابسته به فیلامان‌های اکتین، تنظیم ارتباطات سلولی و تنظیم کیفیت زیستی بودند. همچنین، مسیرهایی مانند تمایز نورونی و گسترش دندریت نیز معنی‌دار شدند که نشان دهنده دخالت مسیرهای رشد و سازماندهی عصبی در ویژگی‌های رشد یاک است. در سطح بخش سلولی (Cellular Component)، ترم‌های غنی شده عمدتاً مربوط به ساختارهای غشایی و اتصالات سلولی بودند. از مهم‌ترین آن‌ها می‌توان به اتصالات سلولی، پیرامون سلولی، زائده‌های سلولی، سیتوپلاسم و کمپلکس میوزین اشاره کرد.

[Downloaded from rap.sanru.ac.ir on 2026-10-22]
[DOI: 10.61882/rap.2026.1565]

همچنین ترم‌های مرتبط با کمپلکس‌های کانال‌های یونی و کمپلکس پروتئین‌های غشایی نیز به‌طور معنی‌دار مشاهده شدند. علاوه بر این، نتایج پایگاه‌های مسیرشناسی نشان دادند که مسیرهایی نظیر پروتئین‌های حرکتی و برهم‌کنش‌های پروتئین–پروتئین در سیناپس‌ها از جمله مسیرهای مهم دخیل در رشد بودند.

جدول ۱- نتایج آنالیز غنی‌سازی هستی‌شناسی ژن بر اساس SNP های شناسایی‌شده در GWAS (۱۵ کیلوباز بالادست و پایین‌دست).

Table 1. Results of gene ontology enrichment analysis based on GWAS-identified SNPs (±15 kb).

source	term_name	source	term_name
GO:MF	ion binding	GO:BP	transport
GO:MF	small molecule binding	GO:BP	homophilic cell adhesion via plasma membrane adhesion molecules
GO:MF	transmembrane transporter activity	GO:BP	nervous system development
GO:MF	transporter activity	GO:BP	transmembrane transport
GO:MF	exo-alpha-sialidase activity	GO:BP	monoatomic cation transport
GO:MF	alpha-sialidase activity	GO:BP	actin filament-based process
GO:MF	cytoskeletal motor activity	GO:BP	cell projection organization
GO:MF	calcium ion binding	GO:BP	cell-cell adhesion
GO:MF	metal ion transmembrane transporter activity	GO:BP	developmental process
GO:MF	metal ion binding	GO:BP	plasma membrane bounded cell projection organization
GO:MF	cation binding	GO:BP	neuron projection development
GO:MF	monoatomic ion transmembrane transporter activity	GO:BP	multicellular organism development
GO:MF	cell adhesion molecule binding	GO:BP	system development
GO:MF	inorganic molecular entity transmembrane transporter activity	GO:BP	regulation of metal ion transport
GO:MF	cadherin binding	GO:CC	cell junction
GO:MF	ATP-dependent activity	GO:CC	cell periphery
GO:MF	ATP hydrolysis activity	GO:CC	plasma membrane
GO:MF	inorganic cation transmembrane transporter activity	GO:CC	synapse
GO:MF	hydrolase activity	GO:CC	cell projection
GO:MF	calcium ion transmembrane transporter activity	GO:CC	plasma membrane bounded cell projection
GO:MF	monoatomic ion channel activity	GO:CC	presynapse
GO:MF	actin filament binding	GO:CC	cytoplasm
GO:BP	cell adhesion	GO:CC	transmembrane transporter complex
GO:BP	monoatomic ion transport	GO:CC	transporter complex
GO:BP	localization	GO:CC	membrane
GO:BP	establishment of localization	GO:CC	neuron projection
GO:BP	metal ion transport	KEGG	Motor proteins

نتیجه‌گیری کلی

نتایج این مطالعه نشان می‌دهند که تحلیل ژنومی با استفاده از تراشه‌های با چگالی بالا ابزار کارآمدی برای شناسایی نواحی ژنومی و ژن‌های مرتبط با صفات رشد در یاک است. شناسایی‌SNPهای معنی‌دار و ژن‌های نزدیک به آن‌ها به‌ویژه ژن‌هایی مانند *GCFC1*، *SYNJ1* و *C1H21orf62* شواهدی را از دخالت مسیرهای زیستی مهمی از جمله اتصال یون‌ها، فعالیت ناقل‌های غشایی، فرآیندهای عصبی و چسبندگی سلولی در کنترل رشد فراهم کرد. این یافته‌ها نه‌تنها درک ما را از پایه‌های ژنتیکی رشد در یاک افزایش می‌دهند، بلکه می‌توانند به‌عنوان منابعی ارزشمند برای طراحی نشانگرهای انتخابی در برنامه‌های به‌نژادی مورد استفاده قرار گیرند. با این‌حال، برای تأیید و کارکردی‌سازی این نتایج نیاز به مطالعات بیشتری شامل داده‌های فنوتیپی گسترده‌تر، جمعیت‌های مستقل و آزمایش‌های عملکردی وجود دارد. در نهایت، این پژوهش گام مهمی در جهت پیوند ژنومیک با بهبود کارایی تولیدی در یاک محسوب می‌شود.

References

Ben-Yair, R., & Kalcheim, C. (2005). Lineage analysis of the avian dermomyotome sheet reveals the existence of single cells with both dermal and muscle progenitor fates. *Development*, 132(4), 689-701.

Biressi, S., & Rando, T. A. (2010, October). Heterogeneity in the muscle satellite cell population. In *Seminars in Cell & Developmental Biology,* (Vol. 21, No. 8, pp. 845-854). Academic Press.

[Downloaded from rap.sanru.ac.ir on 2026-10-22]

[DOI: 10.61882/rap.2026.1565]

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23(19), 2633-2635.

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., ... & Stanton, L. W. (2004). Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation. *Nature Biotechnology*, 22(6), 707-716.

Buckingham, M., & Relaix, F. (2007). The role of Pax genes in the development of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell functions. *Annual Review of Cell and Developmental Biology*, 23(1), 645-673.

Chen, M., Cho, J., & Zhao, H. (2011). Incorporating biological pathways via a Markov random field model in genome-wide association studies. *PLoS Genetics*, 7(4), e1001353.

Chimusa, E. R., Mbiyavanga, M., Mazandu, G. K., & Mulder, N. J. (2016). ancGWAS: a post genome-wide association study method for interaction, pathway and ancestry analysis in homogeneous and admixed populations. *Bioinformatics*, 32(4), 549-556.

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(2), 1223-1231.

Diao, Y., Guo, X., Li, Y., Sun, K., Lu, L., Jiang, L., ... & Wu, Z. (2012). Pax3/7BP is a Pax7-and Pax3-binding protein that regulates the proliferation of muscle precursor cells by an epigenetic mechanism. *Cell Stem Cell*, 11(2), 231-241.

Esmaeili-Fard, S. M., Gholizadeh, M., Hafezian, S. H., & Abdollahi-Arpanahi, R. (2021). Genes and pathways affecting sheep productivity traits: Genetic parameters, genome-wide association mapping, and pathway enrichment analysis. *Frontiers in Genetics*, 12, 710613.

Ge, F., Jia, C., Bao, P., Wu, X., Liang, C., & Yan, P. (2020). Accuracies of genomic prediction for growth traits at weaning and yearling ages in yak. *Animals*, *10*(10), 1793.

Gill, R. A., Ahmar, S., Ali, B., Saleem, M. H., Khan, M. U., Zhou, W., & Liu, S. (2021). The role of membrane transporters in plant growth and development, and abiotic stress tolerance. *International Journal of Molecular Sciences*, 22(23), 12792.

Gholizadeh, M., Rahimi Mianji, G., & Nejati Javaremi, A. (2015). Linkage disequilibrium estimation and haplotype based genome-wide association to detect QTLs affecting twinning rate in Baluchi sheep. *Research on Animal Production*, 5(10), 166-178. [In Persian]

Ha, N. T., Gross, J. J., van Dorland, A., Tetens, J., Thaller, G., Schlather, M., ... & Simianer, H. (2015). Gene-based mapping and pathway analysis of metabolic traits in dairy cows. *PloS one*, 10(3), e0122325.

Hayes, B. (2013). Overview of statistical methods for genome-wide association studies (GWAS). *Genome-Wide Association Studies and Genomic Prediction*, 149-169.

Hu, Y. H., Warnatz, H. J., Vanhecke, D., Wagner, F., Fiebitz, A., Thamm, S., ... & Janitz, M. (2006). Cell array-based intracellular localization screening reveals novel functional features of human chromosome 21 proteins. *BMC Genomics*, 7(1), 155.

Jia, C., Li, C., Fu, D., Chu, M., Zan, L., Wang, H., ... & Yan, P. (2020). Identification of genetic loci associated with growth traits at weaning in yak through a genome-wide association study. *Animal Genetics*, 51(2), 300-305.

Jiang, H., Chai, Z. X., Cao, H. W., Zhang, C. F., Zhu, Y., Zhang, Q., & Xin, J. W. (2022). Genome-wide identification of SNPs associated with body weight in yak. *BMC Genomics*, 23(1), 833.

Liu, W. B., Liu, J., Liang, C. N., Guo, X., Bao, P. J., Chu, M., ... & Yan, P. (2014). Associations of single nucleotide polymorphisms in candidate genes with the polled trait in D atong domestic yaks. *Animal Genetics*, 45(1), 138-141.

Mackay, T. F., Stone, E. A., & Ayroles, J. F. (2009). The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics*, 10(8), 565-577.

Mani, M., Lee, S. Y., Lucast, L., Cremona, O., Di Paolo, G., De Camilli, P., & Ryan, T. A. (2007). The dual phosphatase activity of synaptojanin1 is required for both efficient synaptic vesicle endocytosis and reavailability at nerve terminals. *Neuron*, 56(6), 1004-1018.

Neibergs, H. L., Mutch, J. L., Neupane, M., Seabury, C. M., Taylor, J. F., Garrick, D. J., ... & Johnson, K. A. (2016). Gene set enrichment analysis for feed efficiency in beef cattle.

McNerney, M. P., & Styczynski, M. P. (2018). Small molecule signaling, regulation, and potential applications in cellular therapeutics. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 10(2), e1405.

Pasandideh, M., Rahimi-Mianji, G., & Gholizadeh, M. (2018). A genome scan for quantitative trait loci affecting average daily gain and Kleiber ratio in Baluchi Sheep. *Journal of Genetics*, 97(2), 493-503.

Peng, G., Luo, L., Siu, H., Zhu, Y., Hu, P., Hong, S., ... & Xiong, M. (2010). Gene and pathway-based second-wave analysis of genome-wide association studies. *European Journal of Human Genetics*, 18(1), 111-117.

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.

[Downloaded from rap.sanru.ac.ir on 2026-10-22]

[DOI: 10.61882/rap.2026.1565]

Puri, P. L., & Sartorelli, V. (2000). Regulation of muscle regulatory factors by DNA-binding, interacting proteins, and post-transcriptional modifications. *Journal of Cellular Physiology*, 185(2), 155-173.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, 47(W1), W191-W198.

Sakisaka, T., Itoh, T., Miura, K., & Takenawa, T. (1997). Phosphatidylinositol 4, 5-bisphosphate phosphatase regulates the rearrangement of actin filaments. *Molecular and Cellular Biology*, 17(7), 3841-3849.

Turner, S. D. (2014). qqman: an R package for visualizing GWAS results using QQ and manhattan plots. *Biorxiv*, 005165.

Wang, L., Jia, P., Wolfinger, R. D., Chen, X., & Zhao, Z. (2011). Gene set analysis of genome-wide association studies: methodological issues and perspectives. *Genomics*, 98(1), 1-8.

Wang, J., Liu, Y., & Tian, B. (2024). Protein-small molecule binding site prediction based on a pre-trained protein language model with contrastive learning. *Journal of Cheminformatics*, 16(1), 125.

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., ... & Liu, X. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, Proteomics & Bioinformatics*, 19(4), 619-628.

Zhang, Y., Zhang, Y., Sun, K., Meng, Z., & Chen, L. (2019). The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. *Journal of Molecular Cell Biology*, 11(1), 1-13.

[Downloaded from rap.sanru.ac.ir on 2026-10-22]

[DOI: 10.61882/rap.2026.1565]