



"مقاله پژوهشی"

پوشش کل ژنوم جهت شناسایی ژن‌ها و مسیرهای زیستی مرتبط با صفات تیپ
در گاو بر پایه آنالیز مسیر

حسین محمدی^۱ و ابودر نجفی^۲

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

۲- استادیار گروه علوم دام و طیور، پردیس ابوریحان، دانشگاه تهران، پاکدشت، ایران

تاریخ دریافت: ۱۴۰۰/۹/۷ تاریخ پذیرش: ۱۴۰۰/۱۱/۱۳

صفحه: ۱۳۹ تا ۱۴۷

چکیده مبسوط

مقدمه و هدف: صفات تیپ توصیف کننده وضعیت اسکلتی دام بوده و دارای همبستگی ژنتیکی متوسط تا بالا با صفات مهم اقتصادی شامل باروری، طول عمر اقتصادی و خصوصیات لاشه در گاو می‌باشد. پژوهش حاضر با هدف مطالعه پوشش کل ژنوم بر اساس آنالیز مجموعه‌های ژنی برای شناسایی جایگاه‌های ژنی مؤثر بر صفات تیپ با استفاده از آرایه‌های ژنومی 50K بود.

مواد و روش‌ها: برای این منظور از اطلاعات فنوتیپی و ژنوتیپی صفات طول بدن (BL)، ارتفاع قد از جدوگاه (WH)، دور سینه (CG) و عمق سینه (CD) ۲۸۶ نمونه از گاوهای آمیخته استفاده شد. ابتدا آنالیز پوشش کل ژنومی برای صفات تیپ در برنامه GEMMA نسخه ۰/۹۸ انجام شد. سپس با استفاده از بسته نرم افزاری biomaR2 ژن‌های معنی‌داری که در داخل و یا ۲۵ کیلوباز بالا و پایین دست نشانگرهای معنی‌دار قرار داشتند، شناسایی شدند. در نهایت تفسیر مجموعه ژنی با بسته نرم افزاری goseq برنامه R با هدف شناسایی عملکردهای بیولوژیکی ژن‌های نزدیک به مناطق انتخابی و ژن‌های کاندیدا از طریق پایگاه‌های GO، KEGG، DAVID و PANTHER انجام شد.

یافته‌ها: در این پژوهش تعداد ۱۲ نشانگر تک نوکلئوتیدی واقع روی کروموزوم‌های ۱، ۳، ۵، ۷، ۸، ۱۰، ۱۳، ۱۶، ۱۷، ۲۲، ۲۳ و ۲۵ شناسایی شدند که با ژن‌های MYOG، CAPN3 و (CG) HMGA1، GLP1R، (WH) TRAK1، SYN3، TGFBR3، (BL) PLCB4، TMEM130، PTBP1، PLCB1، MYO18B (CD) مرتبط بودند. برخی از این ژن‌های شناسایی شده در مناطق ژنومی معنی‌دار با مطالعات قبلی هم خوانی داشت. در تفسیر مجموعه ژنی تعداد ۱۷ مسیر هستی شناسی ژنی و بیوشیمیایی با صفات تیپ شناسایی شد ($P < 0.05$). از این بین، مسیرهای توسعه ساختار بافت عضله، تشکیل ساختار آناتومیکی در برگزیده مورفوژن، هموستازی ساختار آناتومیکی، تمایز و تفرق استئوکلاست‌ها، تمایز و تفرق سلول‌های عضلات اسکلتی و مسیر سیگنال‌دهی کلسیم عملکردهای مهمی را در ارتباط با رشد و توسعه عضلات اسکلتی، هموستازی گلوکز، فرآیند استخوان‌سازی، تفرق سلول‌های عضلات اسکلتی، تنظیم یون کلسیم و فعال‌سازی مسیر سیگنال‌دهی کلسیم بر عهده داشتند.

نتیجه‌گیری: یافته‌های این تحقیق ضمن تأیید نتایج مطالعات قبلی در زمینه پوشش ژنومی صفات مرتبط با تیپ و همچنین شناسایی مناطق ژنومی جدید، می‌تواند در انتخاب ژنتیکی برنامه‌های اصلاح نژادی گاو مفید باشد.

واژه‌های کلیدی: آنالیز مسیر، پوشش ژنومی، صفات تیپ، گاو

مقدمه

صفات تیپ به طور مستقیم و غیرمستقیم در تصمیم‌گیری برای حذف حیوان مؤثر هستند (۵). مطالعات صورت گرفته نشان‌دهنده مفید بودن صفات تیپ به عنوان صفات همبسته با صفات باروری، خوراک مصرفی و تولید گوشت جهت انتخاب می‌باشد (۷، ۲۴).

وجود فاصله نسلی طولانی و نبود امکانات برای رکوردبرداری برخی از صفات سلامتی و باروری باعث توجه به صفات دیگری شده است که علاوه بر رکوردبرداری آسان‌تر و وراثت‌پذیری بالاتر، در ابتدای عمر حیوان قابل اندازه‌گیری بوده و با صفات مهم اقتصادی همچون سلامتی و باروری نیز همبستگی ژنتیکی متوسط تا بالایی دارند. صفات تیپ از جمله صفات مهم در این مورد می‌باشند (۱۰).

همبستگی ژنتیکی بین صفات عرض سینه و عمق بدن با طول عمر اقتصادی در گاوهای هلشتاین مجارستانی مثبت گزارش شده است، بطوریکه گاوهای با خصوصیات عرض سینه و عمق بدن بیشتر، در معرض خطر حذف بالاتری از گله می‌باشند (۲۶). همچنین همبستگی ژنتیکی مثبت بین صفات تیپ (عمق بدن و طول بدن) با صفات تولید شیر، چربی و پروتئین گزارش شده است (۳۳). ارتباط معنی‌داری بین صفات

تیپ با عملکرد تولیدمثلی نیز گزارش شده است، بطوریکه صفات عرض سینه از لحاظ ژنتیکی با برخی صفات باروری همبستگی ژنتیکی داشته و عرض سینه کمتر منجر به فاصله گوساله‌زایی بیشتر و سخت‌زایی بیشتر می‌شود (۱۰).

در مطالعات پوشش ژنومی ارتباط بین یک نشانگر SNP و یک صفت با استفاده از نشانگرهای با تراکم بالا در سطح ژنوم شناسایی می‌شود. این اطلاعات می‌تواند برای انتخاب به کمک نشانگر مفید واقع شده و به درک بهتر بیولوژی مولکولی صفات مورد مطالعه کمک نماید (۱۴). یکی از ایرادات تحقیقات مطالعات پوشش 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 (۲۰).

یک جایگزین مناسب برای حل این مشکل، انجام مطالعات پوشش کل ژنومی بر مبنای روش آنالیز مسیر^۱ با تجزیه و تحلیل مجموعه‌های ژنی است. در واقع در این روش به جای انجام تجزیه برای یک SNP یا یک ژن، ارتباط بین صفت مورد مطالعه و واریانت‌های ژنتیکی در یک دسته یا گروه ژنی که به طور عملکردی با هم مرتبط هستند بررسی می‌شود. به عبارت دیگر پیوستگی بین یک مجموعه ژنی معنی‌دار با فنوتیپ، مورد

¹ Pathway-based analysis

<https://datadryad.org/stash/share/qFZOgMFnj57>
 (0ee02psNIaTao70PjL_bnMb6sxfScK8) استفاده گردید.

جهت اطمینان از کیفیت داده‌های تعیین ژنوتیپ مراحل مختلف کنترل کیفیت روی داده‌های اولیه تعیین ژنوتیپ شده قبل از آنالیز پیوستگی انجام شد.

برای فیلتراسیون داده‌های تعیین ژنوتیپ شده، ابتدا نمونه‌هایی که فراوانی نرخ تعیین ژنوتیپ آنها کمتر از ۹۰٪ بود، شناسایی و حذف شد (۳۰). در مرحله بعد نشانگرهایی که حداقل فراوانی آللی در آنها کمتر از ۱٪ بود حذف شدند. سپس نشانگرهایی که نرخ تعیین ژنوتیپ آنها در نمونه‌ها کمتر از ۹۵٪ بود شناسایی و حذف شدند. در نهایت برای SNP‌های باقیمانده آنهایی که در تعادل هاردی-واینبرگ قرار نداشتند به عنوان معیاری از خطای تعیین ژنوتیپ کنار گذاشته شدند (۳۰). سطح احتمال 10^{-6} برای تعادل هاردی-واینبرگ در نظر گرفته شد که با استفاده از تصحیح بنفرونی به دست آمد. مراحل مختلف فیلتراسیون با استفاده از نرم‌افزار PLINK انجام شد (۲۱). از مجموع ۴۷۸۴۳ نشانگر به کار رفته در این تحقیق، ۳۸۳۲۲ نشانگر و ۲۷۹ دام توانستند مراحل مختلف کنترل کیفیت را بگذرانند.

برای بررسی وجود یا عدم وجود لایه‌بندی جمعیتی، از روش کنترل ژنومیک با استفاده از فاکتور تورم کنترل جمعیتی (۸) در نرم افزار PLINK استفاده شد. جهت ارتباط فنوتیپ‌ها با ژنوتیپ‌ها از نرم‌افزار GEMMA (۳۴) نسخه ۰/۹۸ استفاده گردید.

اساساً آنالیز پوشش ژنومی بر پایه تجزیه و تحلیل مجموعه‌های ژنی در سه مرحله انجام می‌گردد: ۱) تعیین مکان SNP‌های معنی‌دار که مقدار *P*-value آنها کمتر از ۰/۰۵ بود با استفاده از بسته نرم افزاری *biomart2* (۱۲) در محیط R و با استفاده از ژنوم مرجع گاو نسخه (*ARS UCD1.2*) به ژن‌هایی که نشانگر SNP مورد نظر در داخل آن ژن و یا kb 25 بالادست یا پایین دست (۱۹) آن ژن قرار داشت، ارتباط داده شدند (۲) ارتباط ژن‌ها به طبقات عملکردی و مسیرهای بیوشیمیایی از پایگاه‌های اطلاعاتی شامل هستی شناسی ژن (GO, <http://www.geneontology.org>), مسیرهای بیوشیمیایی (<http://www.genome.jp/kegg>), (KEGG), Panther (<http://www.pantherdb.org>) و Reactome (<http://www.reactome.org>) (۳) پوشش کل ژنومی بر پایه آنالیز مسیر با استفاده از توزیع فوق هندسی^۱ و آماره Fisher's exact test مورد آزمون قرار گرفت. تجزیه غنی‌سازی مجموعه ژنی با استفاده از بسته نرم افزاری *goseq* (۳۱) در محیط نرم‌افزار R انجام گردید. برای تفسیر بهتر عملکرد ژن‌های به دست آمده از پایگاه‌های اطلاعاتی آنلاین GeneCards (<http://www.genecards.org>) و UniProtKB (<http://www.uniprot.org>) استفاده شد.

نتایج و بحث

آمار توصیفی صفات مورد بررسی در مطالعات پوشش کل ژنومی در جدول ۱ نشان داده شده است. لایه‌بندی جمعیتی که

آزمون قرار می‌گیرد (۲۷). در این روش به دنبال ژن‌هایی که به تنهایی اثر آنها بر صفت مورد نظر معنی‌دار نشده و دارای اثرات متوسطی می‌باشند، ولی اثر تجمعی آنها روی صفت دارای اثر معنی‌دار است دنبال می‌شوند. برای اینکه بتوان تفسیر درستی از کنار هم قرار دادن این ژن‌ها حاصل شود، از مسیرهای زیستی به عنوان بستری معنی‌دار که عملکرد مجموع ژن‌ها در آنها یک یا چند هدف واحد را پیگیری می‌کند استفاده می‌شود (۱۸).

در مطالعه‌ی پوشش کل ژنومی در گاوهای هلشتاین با هدف شناسایی ژن‌های کاندیدای مرتبط با صفات ساختار دست و پا انجام شده بود و ژن‌های کاندیدای *INPP4A*, *ADIPOR2*, *CADPS* و *XKR4*, *PCDH7*, *ALDH1A2*, *DNMT3A* گزارش شده است (۱).

برای اولین بار پناگریکانو و همکاران (۱۹) نشان دادند که تجزیه و تحلیل پوشش 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 پوشش ژنومی بر مبنای مسیر با استفاده از آنالیز غنی‌سازی مجموعه‌های ژنی روی خصوصیات لخته‌شدگی شیر، تولید پنیر و استحکام دل‌مه 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 (۸).

در یک جمعیت گاو گوشتی نژاد نلور نشان داده شده که روش پوشش ژنومی بر مبنای مسیر کارایی بالاتری برای یافتن مناطق ژنومی و درک بهتری از معماری ژنتیکی صفات مرتبط با کیفیت گوشت نسبت به آنالیز پوشش ژنومی بر پایه تک نشانگری دارد (۶).

هدف از انجام پژوهش حاضر، تجزیه و تحلیل مجموعه‌های ژنی و مسیرهای زیستی مرتبط با صفات تیپ در گاوهای آمیخته با استفا‌ه‌های 50K و براساس پوشش کل ژنوم بر مبنای مسیر می‌باشد.

مواد و روش‌ها

در این پژوهش ابتدا فایل‌های تعیین ژنوتیپ با استفاده از تراشه‌های ژنومی همراه با رکوردهای فنوتیپی مرتبط با صفات تیپ که مربوط به گونه گاو بوده و در پایگاه‌های ذخیره‌ای عمومی مختلف داده‌های ژنومی (*Figshare*, *Dryad*, *Zenodo*, *Animal Genome*, *Frontiersin*) ذخیره شده بودند، استخراج و بر حسب اطلاعات مفید گروه بندی انجام شد. در نهایت، پس از جمع‌بندی و کنترل داده‌های تعیین ژنوتیپ و رکوردهای فنوتیپی، از اطلاعات صفات تیپ گاوهای گوشتی آمیخته حاصل از تلاقی بین نژادهای *Baoule* (نژاد با جثه بزرگ) از گله جثه کوچک) با نژاد *Girolando* (نژاد با جثه بزرگ) از گله تحت رکورد برداری کمپانی *INERA* تحت نظارت دانشگاه علوم زیستی و منابع طبیعی وین اتریش در بورکینافاسو بود، استفاده شد (۳۰). برای دستیابی به اطلاعات رکوردهای فنوتیپی و ژنوتیپی از سایت

معیار عدم تعادل لینکاژی) بیش از ۰/۵ (دستور -indep- -- pairwise 50 5 0.5) با یکدیگر از مجموعه داده‌ها حذف شدند (۲۵). همچنین برای اطمینان بیشتر فاکتور لامبدا برای هر ۴ صفت مورد مطالعه محاسبه گردید که برای صفت طول بدن ۱/۰۱۴، صفت ارتفاع قد از جدوگاه ۰/۹۸۲، صفت دور سینه ۱/۰۰۸ و صفت عمق سینه ۱/۰۳۱ به‌دست آمد. آماره لامبدا که از تقسیم میانه مقادیر کای مربع مشاهده شده بر میانه مورد انتظار (۰/۴۵۶) حاصل می‌شود، فاکتور تورم می‌باشد. اگر مقدار آماره لامبدا کمتر یا مساوی یک باشد، نشان دهنده عدم وجود اثر لایه بندی جمعیتی می‌باشد (۹).

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 مثبت کاذب شود. در نتیجه برای کنترل اختلاط و ارتباط ژنتیکی بین این حیوانات، SNP‌هایی که در حالت عدم تعادل لینکاژی بالایی با یکدیگر قرار داشتند، حذف شدند. بدین منظور در پنجره‌هایی شامل SNP ۵۰ و با حرکت SNP ۵ رو به جلو در هر مرحله، SNP‌های دارای r^2

جدول ۱- آمار توصیفی صفات تیپ بررسی شده در گاو

Table 1. Descriptive statistics of studied type traits in cattle

صفات (واحد)	تعداد	میانگین	انحراف معیار	حداقل	حداکثر
ارتفاع قد از جدوگاه (cm)	۲۸۶	۱۲۷/۵۲	۱۰/۲۵	۱۰۲/۰۰	۱۴۵/۰۰
طول بدن (cm)	۲۸۶	۱۱۵/۹۸	۸/۳۷	۸۴/۰۰	۱۴۱/۰۰
عمق سینه (cm)	۲۸۶	۵۹/۷۰	۵/۳۶	۴۸/۰۰	۱۰۲/۰۰
دور سینه (cm)	۲۸۶	۱۲۶/۸۰	۱۱/۷۰	۸۶/۰۰	۱۵۲/۰۰

ضروری می‌باشد و نقش محوری در بسیاری از سیگنال‌های خارج سلولی دارند (<https://www.genecards.org>). از دیگر مسیرهای اصلی معنی‌دار مرتبط با طول بدن می‌توان به مسیر هموستازی ساختار آناتومیکی^۲ و ورود یون کلسیم به سیتوزول^۳ اشاره کرد. از میان ژن‌های کاندیدای موجود، ژن‌های کاندیدای *PTBP1* و *TMEM130* دارای بیشترین ارتباط معنی‌دار با طول بدن بودند. ژن *PTBP1* نقش اساسی در تنظیم آزاد سازی و ترشح انسولین در بدن داشته و همچنین و تمایز سلول‌های عصبی اثرگذار می‌باشد (۱۵).

در مطالعه‌ی پویش ژنومی با هدف شناسایی مناطق ژنومی و ژن‌های کاندیدای مرتبط با اندازه بدن در گاو گوشتی سمینتال، ژن کاندیدای *PTBP1* گزارش شده است (۳). علاوه بر این، با مطالعه پویش کل ژنومی در گاوهای هلشتاین چینی، ژن کاندیدای *TMEM130* مرتبط با اندازه بدن گزارش شده است (۲۲).

از مسیرهای بیوشیمیایی معنی‌دار KEGG مرتبط با طول بدن می‌توان به مسیر سیگنال‌دهی کلسیم^۴ اشاره کرد (شکل ۲). تحلیل مسیرهای زیستی KEGG نشان‌دهنده این است که ژن *PLCB4* به طور معنی‌داری با مسیر سیگنال‌دهی کلسیم در ارتباط می‌باشد. این مسیر زیستی نقش مهمی در رشد و توسعه استخوان دارد. با مطالعه پویش کل ژنومی در نژادهای آمیخته Landrace×Yorkshire با هدف شناسایی ژن‌های کاندیدای مرتبط با صفات تیپ، ژن کاندیدای *PLCB4* در ارتباط با اندازه بدن گزارش شده است (۳۲).

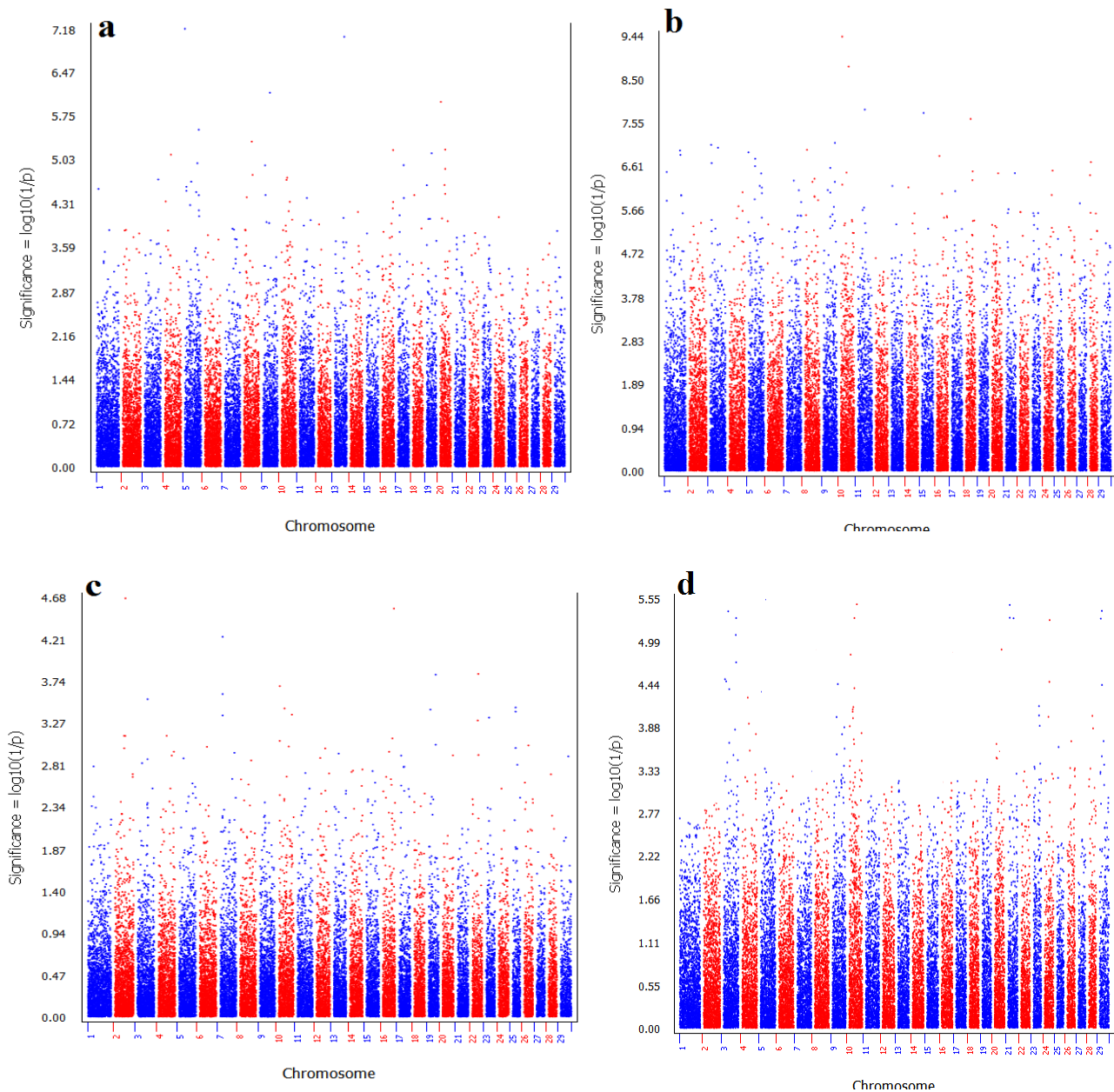
نتایج حاصل از پویش کل ژنومی برای چهار صفت مورد مطالعه در (شکل ۱) نشان داده شده است. تعداد مجموعه‌های ژنی حاصل از پایگاه‌های داده‌ای مختلف شامل ۱۸۳ طبقات هستی‌شناسی (فرایند زیستی، عملکرد مولکولی و اجزای سلولی) و ۴۱ مسیر بیوشیمیایی KEGG و PANTHER بود. همانطور که در (جدول ۲) مشاهده می‌شود، طبقات عملکردی در هستی‌شناسی فرایندهای زیستی، عملکرد مولکولی، اجزای سلولی و مسیرهای KEGG با صفات تیپ دارای ارتباط هستند ($P < 0.05$). مسیرهای بیوشیمیایی که شامل بیش از ۳ ژن و کمتر از ۲۰۰ ژن داشتند شناسایی شده‌اند.

مسیر تشکیل ساختار آناتومیکی در برگبرنده مورفوژن^۱ یکی از مهمترین مسیرهای شناسایی شده مؤثر بر طول بدن می‌باشد. از بین ژن‌های معنی‌دار در این مسیر مرتبط با طول بدن می‌توان ژن *PLCB1* را نام برد.

در مطالعه پویش کل ژنومی در گاوهای هلشتاین مرتبط با صفات تیپ و عملکرد گوساله‌زایی، ژن کاندیدای *PLCB1* را گزارش کردند (۲). همچنین در مطالعه‌ی پویش کل ژنومی با هدف شناسایی ژن‌های کاندیدای مرتبط با طول بدن و ارتفاع قد خوک‌های Large White، ژن کاندیدای *PLCB1* به عنوان ژن کاندیدای مرتبط با طول بدن گزارش شده است (۱۶). ژن‌های کاندیدای *PLCB1* و *PLCB4* جزوی از خانواده پروتئینی فسفولیپازها بوده که نقش کاتالیزوری در تشکیل اینوزیتول تری فسفات (IP3) از فسفات دی‌اینوزیتول بی‌فسفات (PIP2) دارند. حضور یون کلسیم جهت انجام این واکنش

1-Anatomical structure formation involved in morphogenesis
3- Calcium ion import into cytosol

2- Anatomical structure homeostasis
4- Calcium signaling pathway



شکل ۱- پلات منهن برای صفات ارتفاع قد از جدوگاه (a)، طول بدن (b)، دور سینه (c) و عمق سینه (d)
 Figure 1. Manhattan plots for body conformation traits withers height (a) body length (b) chest girth (c) and chest depth (d)

ماهچه‌های موجب بازسازی ماتریکس خارج سلولی فیبرهای ماهیچه‌ای شوند (<https://www.genecards.org>). مطالعه پوشش کل ژنومی در گاوهای گوشتی Wagyu با هدف شناسایی ژن‌های کاندیدای مرتبط با صفت ارتفاع بدن، ژن کاندیدای *SYN3* را گزارش کردند (۴). همچنین در پژوهش حاضر ناحیه معنی‌دار مؤثر روی ارتفاع قد از جدوگاه در ناحیه ۵۱ مگابازی کروموزوم ۳ بدست آمد با مطالعه قبلی مشترک بود (۳۰).

مسیرهای فرآیند زیستی معنی‌دار مرتبط با ارتفاع قد از جدوگاه می‌توان به مسیرهای انتقالی یون کلسیم^۱ و افزایش رشد دربرگیرنده مورفوژن^۲ اشاره کرد. از بین ژن‌های معنی‌دار در این مسیر ژن‌های کاندیدای *TRAK1* و *SYN3* در مطالعات قبلی در ارتباط با صفات تیپ گزارش شده‌اند. ژن‌های خانواده *SYN* از جمله *SYN3* در هموستازی فیبرهای ماهیچه‌ای و ماتریکس خارج سلولی نقش مهمی ایفاء می‌نمایند. همچنین گزارش شده است ژن‌های خانواده *SYN* می‌توانند با تخریب بافت‌های همبند و کلاژنی اطراف فیبرهای

۱ Calcium ion transport

2- Developmental growth involved in morphogenesis

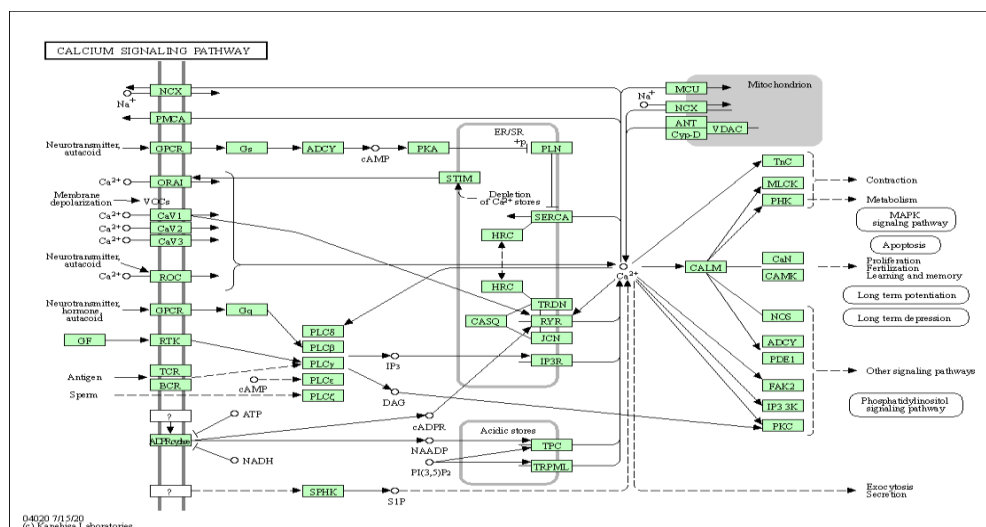
جدول ۲- تجزیه و تحلیل غنی‌سازی مجموعه‌های ژنی معنی‌دار ($p < 0.05$) مرتبط با صفات تیپ در گاوTable 2. Gene set enrichment analysis significantly ($p < 0.05$) associated with type traits in cattle

دسته هستی‌شناسی (صفت)	نام مسیر (GO)	ژن‌های معنی‌دار	ژن‌های کاندیدای مرتبط	p-adjust
فرایند زیستی GO_BP (BL)	GO:0048646: Anatomical structure formation involved in morphogenesis	۷	<i>APOH, CIB1, FBN2, GDF3, GNPAT, PLCB1, PAPPB, RB1, SMG7, TRAF6, CIB2, CTGF, MKS1, PTBP1, PTPN1, RBP4, STRAP, SMC5, TEPI, PRKCQ, MC4R, USH2A, NPHP4</i>	۰/۰۰۳
	GO:0060249: anatomical structure homeostasis	۱۶	<i>FOS, NOX3, TAB1, TAB2, TRAF2, TMEM130, TYROBP, CAMK4, CSF1R, GRB2, ILIR1, MAPK8, PIK3R1, PIK3R3, SIRPB1, TGFB2, AKT3, FOSL2, FCGR2A, FCGR2B, FNI, ITGB1, SYN3, OGN, PLXNA1, PLXNA2, SMAD2, SMAD9, TAB1, FBN2, NLK, STRAP, TGFB3, TRIM33, ZEB1, ZFH3</i>	۰/۰۰۴
	GO:0051209 calcium ion import into cytosol	۲۰	<i>FOS, NOX3, TAB1, TAB2, TRAF2, TMEM130, TYROBP, CAMK4, CSF1R, GRB2, ILIR1, MAPK8, PIK3R1, PIK3R3, SIRPB1, TGFB2, AKT3, FOSL2, FCGR2A, FCGR2B, FNI, ITGB1, SYN3, OGN, PLXNA1, PLXNA2, SMAD2, SMAD9, TAB1, FBN2, NLK, STRAP, TGFB3, TRIM33, ZEB1, ZFH3</i>	۰/۰۰۲۸
GO_BP (WH)	GO:0048589: Developmental growth involved in morphogenesis	۱۶	<i>FOS, NOX3, TAB1, TAB2, TRAF2, TMEM130, TYROBP, CAMK4, CSF1R, GRB2, ILIR1, MAPK8, PIK3R1, PIK3R3, SIRPB1, TGFB2, AKT3, FOSL2, FCGR2A, FCGR2B, FNI, ITGB1, SYN3, OGN, PLXNA1, PLXNA2, SMAD2, SMAD9, TAB1, FBN2, NLK, STRAP, TGFB3, TRIM33, ZEB1, ZFH3</i>	۰/۰۳۷
	GO:0006816: Calcium ion transport	۷	<i>HTR1D, HTR1E, ABL1, CACNA1G, CAPN3, TRAK1, TRPC1</i>	۰/۰۱۹
GO_BP (CG)	GO:0007204: Positive regulation of cytosolic calcium ion concentration	۶	<i>CAPN3, FFAR4, PTGFR, PTPN6, CASQ2, HMGAI</i>	۰/۰۲۳
	GO:0055074: Calcium ion homeostasis	۱۱	<i>BAK1, VAPB, ERO1A, ESR1, GALR1, GLP1R, F2RL3, PTGER4, SPPL3, SMDT1, GLP1R</i>	۰/۰۳۷
GO_BP (CD)	GO:0061061: muscle structure development	۹	<i>EGLN1, ERBB4, GPCPD1, GDF3, MEGF10, MYO18B, CHL3A1, DLL4, EGLN1</i>	۰/۰۱۶
	GO:0035914: Skeletal muscle cell differentiation	۸	<i>RB1, SAP30, CAPN3, FOXN2, MYOG, NR4A1, VAX1, BTG2</i>	۰/۰۴۲
اجزای سلولی GO_CC	GO:0034704: Calcium channel complex	۱۱	<i>AKAP6, SESTD1, ANXA2, ANXA5, CACNA1C, CACNA1B, CACNA1G, CACNA1H, CCDC109B, SMDT1, TRPC4</i>	۰/۰۰۱۷
مسیرهای بیوشیمیایی KEGG Pathway	hsa04380: Osteoclast differentiation	۲۰	<i>FOS, NOX3, TAB1, TAB2, TRAF2, TGFB3, TYROBP, CAMK4, CSF1R, GRB2, ILIR1, MAPK8, PIK3R1, PIK3R3, SIRPB1, TGFB2, AKT3, FOSL2, FCGR2A, FCGR2B, ADCY2, PLCB4, CALM3, HR2H, ITPR3, LHCGR, MYLK, PHKB, PTK2B, PPP3CA, P2RX4, P2RX7, SPHK1, SPHK2, TACR2, PDGFRA, SLC25A4</i>	۰/۰۰۵۱
KEGG Pathway	hsa04020: Calcium signaling pathway	۱۷	<i>FOS, NOX3, TAB1, TAB2, TRAF2, TGFB3, TYROBP, CAMK4, CSF1R, GRB2, ILIR1, MAPK8, PIK3R1, PIK3R3, SIRPB1, TGFB2, AKT3, FOSL2, FCGR2A, FCGR2B, ADCY2, PLCB4, CALM3, HR2H, ITPR3, LHCGR, MYLK, PHKB, PTK2B, PPP3CA, P2RX4, P2RX7, SPHK1, SPHK2, TACR2, PDGFRA, SLC25A4</i>	۰/۰۱۲

و همچنین نقش کلیدی در فعالیت گیرنده‌های سلول‌های پروتئین‌های مورفوزئیک استخوانی (BMPs) مخصوصاً BMP2 دارد (۳۹).

در مطالعه پویس کل ژنومی مرتبط با صفات تیپ مرتبط با دست و پا در خوک، ژن کاندیدای *TGFBR3* مرتبط با اندازه بدن گزارش شده است (۳۳).

از مسیرهای بیوشیمیایی معنی‌دار KEGG مرتبط با ارتفاع قد از جدوگانه می‌توان به مسیر تمایز استئوکلاست‌ها اشاره کرد (شکل ۳). از بین ۲۰ ژن معنی‌دار در این مسیر ژن کاندیدای *TGFBR3* در مطالعات قبلی در ارتباط با صفات تیپ مختلف ارتباط معنی‌داری گزارش شده است. ژن *TGFBR3* به عنوان یک مدیاتور اصلی در مسیر سیگنال‌دهی $TGF-\beta$ نقش داشته



شکل ۲- مسیر سیگنال‌دهی کلسیم و ژن‌های کاندیدای مرتبط با رشد و توسعه استخوان (پایگاه داده KEGG)
Figure 2. Calcium signaling pathway and candidate genes related to bone development (KEGG database)

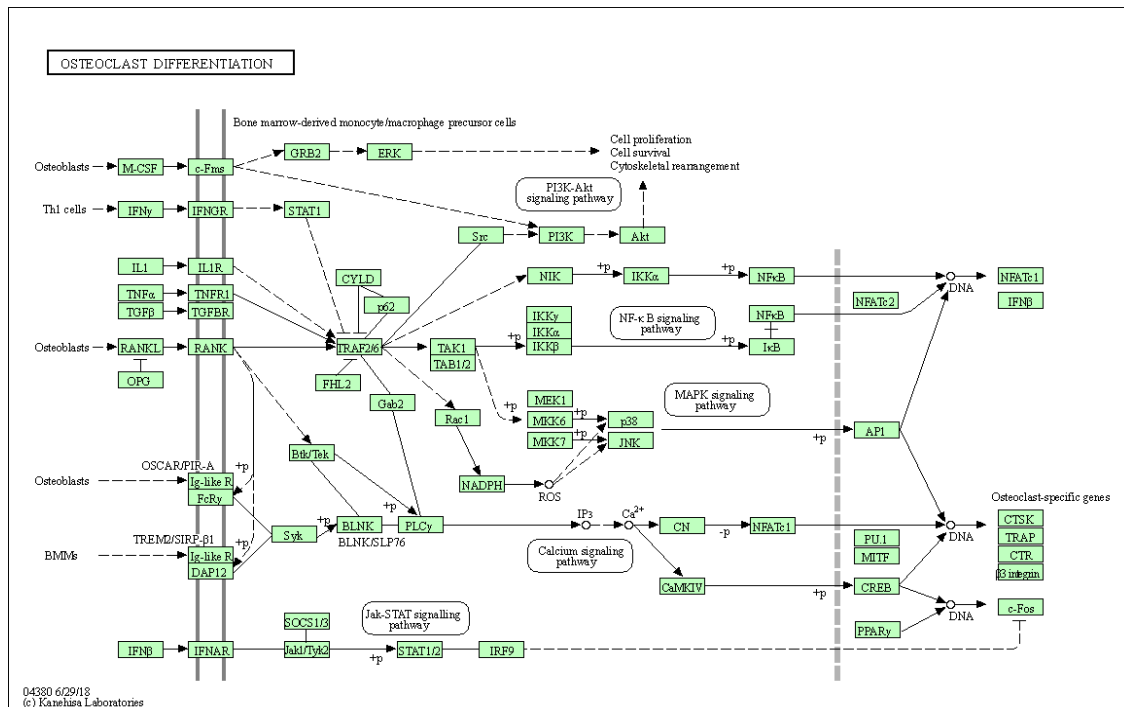
با هدف شناسایی مناطق ژنومی مرتبط با صفات اسکلتی در گاوهای گوشتی، ژن کاندیدای *HMGAI* مرتبط با عرض سینه را گزارش کردند (۱۱).

از مسیرهای هستی‌شناسی معنی‌دار مرتبط با عمق سینه، مسیر تمایز سلول‌های عضلات اسکلتی^۳ بدست آمد که جزء فرآیندهای زیستی می‌باشد. از میان ژن‌های کاندیدای موجود در این مسیر، ژن‌های *CAPN3* و *MYOG* بیشترین ارتباط را با عمق سینه نشان دادند.

کالپاین در ارتباط با کالپاستاتین سیستم پروتئازی وابسته به یون کلسیم به نام سیستم کالپاین-کالپستاتین تشکیل می‌دهند. این سیستم نقش اساسی در فرآیندهای فیزیولوژیکی وابسته به یون کلسیم از قبیل توسعه و تفرق سلول‌های عضلانی و اسکلتی دارد (<https://www.genecards.org>). با بررسی چند شکلی ژن کاندیدای *CAPN3* با صفات مرتبط با رشد در بره‌های نژاد باریکی، ارتباط معنی‌داری گزارش شده است (۱۳).

از مسیرهای مهم و معنی‌دار مرتبط با دور سینه می‌توان به مسیر هموستازی یون کلسیم^۱ اشاره کرد که از بین ژن‌های معنی‌دار در این مسیر، ارتباط معنی‌داری بین ژن کاندیدای *GLPIR* در مطالعات قبلی با وزن بدن گزارش شده است. در مطالعه پوشش کل ژنومی در نژادهای مختلف گاوهای گوشتی با هدف شناسایی مناطق ژنومی مرتبط با رشد عضله، ژن کاندیدای *GLPIR* در نژاد هر فوردها گزارش شده است (۱۱). ژن *GLPIR* جزوی از خانواده ژنی گیرنده گلوکان بوده که نقش کلیدی در تنظیم ترشح انسولین را برعهده دارند و همچنین در سلول‌های مغزی بیان شده و وظیفه کنترل اشتها را دارد (۱۷).

مسیر معنی‌دار دیگر مرتبط با صفت دور سینه مسیر تنظیم مثبت غلظت یون کلسیم سیتوزولی^۲ بود و شامل ژن کاندیدای *HMGAI* بود. ژن *HMGAI* در فسفریله کردن گلیکوپروتئین‌ها و تشکیل استخوان در انسان و موش از طریق تفرق سلول‌های استئوبلاست نقش کلیدی دارد (<https://www.genecards.org>). مطالعه پوشش کل ژنومی



شکل ۳- مسیر تمایز استئوکلاست و ژن‌های کاندیدای مرتبط با صفات استخوان (پایگاه داده KEGG)

Figure 3. Osteoclast differentiation pathway and candidate genes related to bone traits (KEGG database)

می‌توان به مسیر توسعه ساختار عضلانی^۲ نام برد که از بین ۱۰ ژن، ژن کاندیدای *MYO18B* در متابولسم لیپید و اسیدهای چرب در بدن نقش دارند (۲۸، ۲۲).

نتیجه‌گیری کلی

با بررسی منابع انجام شده، پژوهش حاضر اولین مطالعه پوشش کل ژنومی برپایه آنالیز مسیر مرتبط با صفات تیپ در گاوهای آمیخته بوده است، لذا در تجزیه و تحلیل‌ها تلاش گردید از آزمون‌های پر کاربرد و سختگیرانه برای شناسایی مسیرهای زیستی معنی‌دار استفاده گردد. بررسی مناطق ژنومی

MYOG پروتئینی است که نقش عمده‌ای در تنظیم عضلانی دارد و به عنوان یک فاکتور رونویسی به طور مستقیم بیان و تفرق سلول‌های عضلانی را تنظیم می‌کند. این ژن به عنوان یک عامل نظارتی در فرآیند میوژنیز^۱ در سطوح اولیه بیان می‌شود (<https://www.genecards.org>). در مطالعه قبلی پوشش کل ژنومی صفات تیپ مرتبط با سینه که براساس مدل خطی مختلط تک متغیره انجام شده بود، در ناحیه ۵۵/۸ Mb روی کروموزوم ۱۶ در ارتباط با عرض سینه معنی‌دار گزارش شده بود (۳۰) که با منطقه شناسایی شده در پژوهش حاضر همخوانی داشت. مسیر دیگر معنی‌دار مرتبط با عمق سینه را

1 Calcium ion homeostasis

3- Skeletal muscle cell differentiation

2- Positive regulation of cytosolic calcium ion concentration

4- Myogenesis

5- Muscle structure development

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. از طرفی در این تحقیق به دلیل عدم دسترسی به رکوردهای فنوتیپی و اطلاعات ژنوتیپی مرتبط با صفات مورد مطالعه تحقیق حاضر در دام‌های بومی کشور، از اطلاعات فنوتیپی و ژنوتیپی کشور اتریش استفاده شد. لذا استفاده از نتایج این تحقیق در جمعیت‌های گاوهای کشور نیاز به مطالعات بیشتر دارد تا در این جمعیت‌ها نیز تأیید شوند.

شناسایی شده روی کروموزوم‌های ۳، ۵، ۷، ۱۳، ۱۶، ۱۷، ۲۲، ۲۳ و ۲۵ با استفاده از پایگاه داده نشان داد که این مناطق با صفات تیپ مرتبط می‌باشند. البته باید توجه کرد که یکی از محدودیت‌های اجرای مطالعات پویش 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. با توجه به تعداد کم نمونه مورد استفاده در پژوهش حاضر نیز باید در استفاده از ژن‌های کاندیدای شناسایی شده مرتبط با

منابع

1. Abdalla, I.M., X. Lu, M. Nazar, Y. Fan, Z. Zhang, X. Wu, T. Xu and Z. Yang. 2021. Genome-wide association study on reproduction-related body-shape traits of Chinese holstein cows. *Animals (Basel)* 11(7): 1927.
2. Abo-Ismael, M.K., L.F. Brito, S.P. Miller, M. Sargolzaei, D.A. Grossi, S.S. Moore, G. Plastow, P. Stothard, S. Nayeri and F.S. Schenkel. 2017. Genome-wide association studies and genomic prediction of breeding values for calving performance and body conformation traits in Holstein cattle. *Genetic Selection Evolution*, 49(1): 82.
3. An, B., L. Xu, J. Xia, X. Wang, J. Miao, T. Chang, M. Song, J. Ni, J. Li and H. Gao. 2020. Multiple association analysis of loci and candidate genes that regulate body size at three growth stages in Simmental beef cattle. *BMC Genetics*, 21(1): 32.
4. An, B., J. Xia, T. Chang, X. Wang, L. Xu, L. Zhang, X. Gao, Y. Chen and H. Gao. 2019. Genome-wide association study reveals candidate genes associated with body measurement traits in Chinese Wagyu beef cattle. *Animal Genetics*, 50(4): 386-390.
5. Bani Saadat, H., S. Varkoohi and S. Razagh Zadeh. 2018. Investigation of genetic, phenotypic and environmental trends for biometric traits in Makuie sheep. *Research on Animal Production*, 8(18): 139-146 (In Persian).
6. Braz, C.U., J.F. Taylor and T. Bresolin. 2019. Sliding window haplotype approaches overcome single SNP analysis limitations in identifying genes for meat tenderness in Nelore cattle. *BMC Genetics*, 20: 8.
7. Crowley, J., R. Evans, N. Mc Hugh, T. Pabiou, D. Kenny and M. McGee. 2011. Genetic associations between feed efficiency measured in a performance test station and performance of growing cattle in commercial beef herds. *Journal of Animal Science*, 89: 3382-3393.
8. Dadousis, C., S. Pegolo, G.J.M. Rosa, D. Gianola, G. Bittante and A. Cecchinato. 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.
9. Devlin, B., S.A. Bacanu and K. Roeder. 2004. Genomic control to the extreme. *Nature Genetics*, 36: 1129-1130.
10. Doyle, J.L., D.P. Berry, R.F. Veerkamp, T.R. Carthy, SW. Walsh and D.C. Purfield. 2020a. Genomic regions associated with skeletal type traits in beef and dairy cattle are common to regions associated with carcass traits, feed intake and calving difficulty. *Frontiers Genetic*, 11: 20.
11. Doyle, J.L., D.P. Berry, R.F. Veerkamp, T.R. Carthy, SW. Walsh and D.C. Purfield. 2020b. Genomic regions associated with muscularity in beef cattle differ in five contrasting cattle breeds. *Genetic Selection Evolution*, 52(1): 2.
12. Durinck, S., P.T. Spellman, E. Birney and W. Huber. 2009. Mapping identifiers for the integration of genomic datasets with the R/bioconductor package biomaRt. *Nature Protocols*, 4: 1184-1191.
13. Ghanem, N., M. Zayed, I. Mohamed, M. Mohammady and M. Shehata. 2021. Co-expression of candidate genes regulating growth performance and carcass traits of Barki lambs in Egypt. *Agriculture science*, 5: 202-216.
14. Gholizadeh, M., Gh. Rahimi Mianji and A. Nejati Javaremi. 2014. Linkage disequilibrium estimation and haplotype based genome-wide association to detect QTLs affecting twinning rate in Baluchi sheep. *Research on Animal Production*, 10(5): 166-178 (In Persian).
15. Hansen, T.H., H. Vestergaard, T. Jørgensen, M.E. Jørgensen, T. Lauritzen, I. Brandslund, C. Christensen, O. Pedersen, T. Hansen and A.P. Gjesing. 2015. Impact of PTBP1 rs11085226 on glucose-stimulated insulin release in adult Danes. *BMC Medicine Genetics*, 16:17.
16. Hong, Y., J. Ye, L. Dong, Y. Li, L. Yan, G. Cai, D. Liu, C. Tan and Z. Wu. 2021. Genome-wide association study for body length, body height, and total teat number in large white pigs. *Frontiers Genetics*, 12: 650370.
17. Li, P., H.K. Tiwari, W.Y. Lin, D.B. Allison, W.K. Chung, R.L. Leibel, N. Yi and N. Liu. 2014. Genetic association analysis of 30 genes related to obesity in a European American population. *International Journal Obesity (Lond)*, (5): 724-729.

18. Mooney, M.A. and B. Wilmot. 2015. Gene set analysis: A step-by-step guide. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 168: 517-527.
19. Peñagaricano, F., K.A. Weigel, G.J. Rosa and H. Khatib. 2013. Inferring quantitative trait pathways associated with bull fertility from a genome-wide association study. *Frontiers Genetics*, 3:307-314.
20. Peng, G., L. Luo and H. Siu. 2010. Gene and pathway-based second wave analysis of genome-wide association studies. *European Journal of Human Genetics*, 18:111-117.
21. Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M.A.R. Ferreira and D. Bender. 2007. PLINK: a toolset for whole-genome association and population-based linkage analysis. *The American Journal of Human Genetics*, 81: 559-575.
22. Resnyk, C.W., C. Chen, H. Huang, C.H. Wu J. Simon, E. Le Bihan-Duval, M.J. Duclos and L.A. Coghburn. 2015. RNA-seq analysis of abdominal fat in genetically fat and lean chickens highlights a divergence in expression of genes controlling adiposity, hemostasis, and lipid metabolism. *PLoS One*, 9(10): e0139549.
23. Reimer, C., C.J. Rubin, A.R. Sharifi, N.T. Ha, S. Weigend, K.H. Waldmann, O. Distl, S.D. Pant, M. Schlather and H. Simianer. 2018. Analysis of porcine body size variation using re-sequencing data of miniature and large pigs. *BMC Genomics*, 19(1): 687.
24. Ring, S., A. Twomey, H. Byrne, M. Kelleher, T. Pabiou and M. Doherty. 2018. Genetic selection for hoof health traits and cow mobility scores can accelerate the rate of genetic gain in producer-scored lameness in dairy cows. *Journal of Dairy Science*, 101: 10034-10047.
25. Tao, L., X.Y. He, L.X. Pan, J.W. Wang, S.Q. Gan and Chu, M.X. 2020. Genome-wide association study of body weight and conformation traits in neonatal sheep. *Animal Genetics*, 51(2): 336-340.
26. Török, E., I. Komlósi, V. Szónyi, B. Béri, G. Mészáros and J. Posta. 2021. Combinations of linear type traits affecting the longevity in Hungarian holstein-friesian cows. *Animals (Basel)*, 11(11): 3065.
27. Wang, L., P. Jia and R.D. Wolfinger. 2011. Gene set analysis of genome-wide association studies: Methodological issues and perspectives. *Genomics*, 98: 1-8.
28. Wu, X., M. Fang, L. Liu, S. Wang, J. Liu, X. Ding, S. Zhang, Q. Zhang, G. Su and D. Sun. 2013. Genome wide association studies for body conformation traits in the Chinese Holstein cattle population. *BMC Genomics*, 17(14): 897.
29. Xiong, D.H., X.G. Liu, Y.F. Guo, L.J. Tan, L. Wang, B.Y. Sha, Y.W. Lundberg, R.R. Recker, Y.Z. Liu, Y.J. Liu, J.M. Zmuda and H.W. Deng. 2009. Genome-wide association and follow-up replication studies identified ADAMTS18 and TGFBR3 as bone mass candidate genes in different ethnic groups. *American Journal of Human Genetics*, 84(3): 388-98.
30. Yougbare, B., A. Soudre, D. Ouedraogo, B.L. Zoma, A.S.R. Tapsoba, M. Sanou, S. Ouedraogo-Kone, P.A. Burger, M. Wurzinger, N. Khayatzadeh, H.H. Tamboura, O.A. Mwai, A. Traore, J. Solkner and G. Meszaros. 2021. Genome-wide association study of trypanosome prevalence and morphometric traits in purebred and crossbred Baoule cattle of Burkina Faso. *PLoS ONE*, 16(8): e0255089.
31. Young, M.D., M.J. Wakefield, G.K. Smyth and A. Oshlack. 2010. Method gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biology*, 11: 14-23.
32. Zhang, H., Z. Zhuang, M. Yang, R. R. Ding, J. Quan, S. Zhou, T. G. Z. Xu, E. Zheng, J. Yang and Z. Wu. 2021. Genome-wide detection of genetic loci and candidate genes for body conformation traits in Duroc × Landrace × Yorkshire crossbred pigs. *Frontiers Genetic*, 12: 664343.
33. Zink, V., L. Zavadilová, J. Lassen, M. Štipkova, J. Vacek and L. Štolc. 2014. Analyses of genetic relationships between linear type traits, fat-to-protein ratio, milk production traits, and somatic cell count in first-parity Czech Holstein cows. *Czech Journal of Animal Science*, 59: 539-547.
34. Zhou, X. and M. Stephens. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44: 821.

Genome-wide Association Study to Identify Genes and Biological Pathways Associated with Type Traits in Cattle using Pathway Analysis

Hossein Mohammadi¹ and Abozar Najafi²

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

2- Assistant Professor, Department of Animal and Poultry Science, College of Aburaihan, University of Tehran, Tehran, Iran

Received: 28 November, 2021 Accepted: 2 February, 2022

Extended Abstract

Introduction and Objective: Type traits describing the skeletal characteristics of an animal are moderately to strongly genetically correlate with other economically important traits in cattle including fertility, longevity and carcass traits. The present study aimed to conduct a genome wide association studies (GWAS) based on gene-set enrichment analysis for identifying the loci associated with conformation traits using the 50K arrays.

Material and Methods: Genome wide association study was performed with biometric traits using GEMMA software v. 0.98. Using the biomaRt2 R package R the SNP were assigned to genes if they were within the genomic sequence of the gene or within a flanking region of 25 kb up- and downstream of the gene. For the assignment of the genes to functional categories, the GO, KEGG, DAVID and PANTHER databases were used.

Results: In this research, 12 SNP markers on chromosomes 1, 3, 5, 7, 8, 10, 13, 16, 17, 22, 23 and 25 located in PLCB1, PTBP1, TMEM130, PLCB4 (BL), TGFBR3, SYN3, TRAK1 (WH), GLP1R, HMGA1 (CG) and CAPN3, MYOG and MYO18B (CD) genes were identified. Some of the genes that were found are consistent with some previous studies related biometric traits. According to pathway analysis, 17 pathways from gene ontology and biological pathways were associated with biometric traits ($p < 0.05$). Among these pathways, muscle structure development, anatomical structure formation involved in morphogenesis, anatomical structure homeostasis, Osteoclast differentiation, skeletal muscle cell differentiation and calcium signaling pathway have important functions in development of skeletal muscle, glucose homeostasis, osteogenesis process, regulation of ion calcium and activation.

Conclusion: The finding of this research confirms the results of previous studies from genomic scanning related to type traits, also revealed additional genomic regions, could be useful for genetic selection in the cattle breeding programs.

Keywords: Cattle, Genome scan, Pathway analysis, Type trait