دانشگاه علوم کشاورزی و منابع طبیعی ساری
پژوهشهای تولیدات دامی

بررسی عملکرد پروتئین‌های پیش‌بینی شده از داده‌های RNA-Seq در گاوهای نژاد هلشتاین و کلیستانی

مصطفی قادریزفره‌ای[۱]، آزاده ترابی[۲]، محمد اسماعیل‌پور[۳]، مینا سلیم‌پور[۴] و محمد حسین بناءبازی[۵]

۱- دانشیار، گروه علوم دامی، دانشگاه یاسوج، یاسوج، ایران، (نویسنده مسوول: mghaderi@yu.ac.ir, mosmos741@yahoo.com)
۲- استادیار، گروه کشاورزی، دانشگاه پیام نور، تهران، ایران
۳- استادیار، دانشکده کشاورزی و منابع طبیعی اهر، دانشگاه تبریز، اهر، ایران
۴- دانشجوی دکترای ژنتیک و اصلاح نژاد دام، گروه علوم دامی، دانشکده کشاورزی و منابع طبیعی، دانشگاه گیلان، رشت، ایران
۵- استادیار پژوهشی، بخش پژوهش‌های بیوتکنولوژی، موسسه تحقیقات علوم دامی کشور، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران

تاریخ دریافت: ۱۳۹۸/۰۹/۲۷ تاریخ پذیرش: ۱۳۹۹/۰۲/۳۰
صفحه: ۱۲۱ تا ۱۳۵

چکیده

پژوهش حاضر به‌منظور تعیین نیم‌رخ بیانی عددی ژن‌های متفاوت بیان شده در نژادهای هلشتاین و کلیستانی و همچنین بررسی عملکرد پروتئین‌های پیش‌بینی شده از ژن‌های متفاوت بیان شده بین دو نژاد مذکور با استفاده از داده‌های RNA-Seq انجام شد. در این مطالعه توالی کل mRNA خون گاو ماده هلشتاین آمریکا و گاو ماده کلیستانی پاکستان از طریق هم‌ردیفی و مکان‌یابی خوانش‌های RNA-Seq روی ژنوم مرجع گاو بدست آمد و سپس تعیین نیم‌رخ بیانی عددی ژن‌های متفاوت بیان شده انجام شد. تعداد ۲٤٦١٦ ژن و ۲٦۷۱٦ ایزوفرم روی ترانسکریپتوم این دو نژاد شناسایی شدند که از این میان تعداد ٤١ ژن شناسایی شدند که تفاوت بیانی بسیار بالا و معنی‌داری را نشان دادند (p<۰۰۰۰۱۵). حدود یک سوم ژن‌هایی که عملکردشان بین دو نژاد متغیر بود، کدکننده آنزیم‌های هیدرولاز و پنج مورد از پروتئین‌های پیش بینی شده از جمله پروتئین‌های عامل رونویسی FosB/JunD، انکوژن FOS ، ساختار elonginb، زیرواحد عامل رونویسی AP-1، کمپلکس روی zif268 و عامل تنظیمی-رونویسی U2AF هستند که در تنظیم رونویسی و بیان ژن‌ها و ویرایش RNA نقش دارند. بررسی میان‌کنش شبکه‌ای بین پروتئین‌های پیش‌بینی شده از ژن‌های متفاوت بیان شده نشان داد که پروتئین‌های پیش‌بینی شده در مسیرهای مختلفی مانند مسیر پیام‌دهی TNF، عفونت سالمونلایی، مسیر پیام‌دهی MAPK و آرتریت روماتوئید دخالت دارند. همچنین بررسی‌ها نشان داد نژاد هلشتاین و کلیستانی به‌علت قرار گرفتن در شرایط محیطی متفاوت و فاصله تکاملی، توانسته‌اند با شرایط پیرامونی به‌خوبی سازگار شوند و این روند در سطح مولکولی از میزان بیان ژن‌های اختصاصی شناسایی شده در این پژوهش کاملاً قابل فهم است. از دلایل تأییدکننده این پیشنهاد می‌توان به نقش بیشتر سامانه‌ی ایمنی در نژاد کلیستانی اشاره نمود که به‌علت قرار گرفتن در محیطی با آلودگی بالاتر در مقایسه با نژاد هلشتاین، میزان بیان پروتئین‌های دارای خاصیت باکتریوسایدی (Cathelicidin 1) در آن بالا رفته و در نتیجه فعالیت سامانه‌ی ایمنی در آن افزایش یافته است. همچنین در این پژوهش مشخص شد که ژن IL1B بیشترین درجه تعامل ژن-دارو را با داروی Canakinumab دارد. نتایج این پژوهش نشان می‌دهد که در مقایسه‌های بین نژادی، نگاه دقیق به فعالیت پروتئین‌هایی که توسط ژن‌های متفاوت بیان می‌شوند می‌توانند تفاوت نژادی در سطح ملکولی را بهتر توضیح دهند.

واژه‌های کلیدی: گاو هلشتاین، RNA-Seq، ایزوفرم، شبکه برهم کنش پروتئینی

مقدمه

بیان ژن سازوکار زیستی پیچیده‌ای دارد که سبب بروز ژنوتیپ به‌صورت فنوتیپ می‌شود و توسط محیطی که در آن رشد می‌کند کنترل می‌شود و در صورت عدم نیاز به فرآورده ژن، آن ژن به‌صورت خاموش و غیرفعال باقی خواهد ماند (۴،۲۳). امروزه با استفاده از فن‌آوری‌های توالی‌یابی میلیون‌ها توالی از، قطعات شکسته شده‌ی ژنوم با اندازه‌های مختلف جهت اتصال به آداپتور تولید و در پروژه‌های ژنومی، فراژنومی و بیان ژنی استفاده می‌شوند (۷،۲۶). ظهور نسل جدیدی از سامانه‌های پربازده تعیین توالی امکان دستیابی به آزمایش‌های تعیین توالی اگزونی را فراهم کرده است که هدف چنین آزمایش‌هایی به‌طور خاص تعیین توالی زیر مجموعه‌هایی از ژنوم‌های موجودات است که پروتئین‌ها را کدگذاری می‌کنند و پیدایش چنین سامانه‌هایی در کاهش هزینه‌ها و به‌دنبال آن افزایش سریع در بازده توالی‌یابی نقش به‌سزایی داشته است. نتایج پروژه دانش‌نامه‌ی اجزای DNA انسانی نیز نشان داده است که ترانسکریپتوم در میان سطوح مختلف داده‌های امیکس از جامعیت بیشتری برخوردار است. حدود ۸۰ درصد ژنوم رونویسی شده، ولی تمامی این رونویسی‌ها ترجمه نمی‌شوند (۲۷). بخش زیادی از RNAهای رونویسی شده به‌صورت نارمزگر باقی می‌مانند که احتمال می‌رود نقش تنظیمی روی سایر بخش‌های ژنوم داشته باشند (۵،۲۸،۳۳). جامعیت ترانسکریپتوم به حدی است که محققین پروژه ENCODE بر این باورند که ترانسکریپتوم می‌تواند فاصله‌ی بین ژنوتیپ و فنوتیپ واقعی را پر کرده و پس زمینه‌های آن را آشکار سازد. این جامعیت موجب شده تا به‌جای ژن، ترانسکریپت یا به‌عبارتی رونوشت ژن به‌عنوان واحد بنیادی ژنوم و واحد توارث پیشنهاد شود (۳). فراهم آوردن داده‌های ترانسکریپتومی با اشکال و ابزارهای مختلف از جمله لکه‌گذاری سادرن، واکنش زنجیره‌ای پلیمراز در لحظه،

[Downloaded from rap.sanru.ac.ir on 2025-07-04]
[DOI: 10.52547/rap.11.28.121]

ریزآرایه‌های بیان ژن و توالی‌یابی RNA امکان‌پذیر است. در این میان RNA-Seq که داده‌هایی دقیق‌تر و معتبرتر از انواع دیگر ارایه می‌دهد، یک تکنیک ابتکاری است که مطالعات ترانسکریپتوم را بر اساس تکنولوژی‌های نسل جدید توالی‌یابی انجام می‌دهد و بر بسیاری از مشکلات ریز ارایه غلبه می‌کند (۱۰،۲۶،۳۳). تکنیک RNA-Seq جهت بررسی مواردی از جمله شناسایی ژن‌های جدید یا ایزوفرم‌های آن در ژنوم، مکان‌هایی که اسپلایسینگ به‌طور متناوب در آن‌ها رخ می‌دهد، استفاده می شود (۳۳،۱۲،۲۷). روش RNA-seq بر اساس برش mRNAهای رونویسی شده (یا به‌عبارت دیگر کل ترانسکریپتوم یا کل اگزوم) به قطعات کوچک‌تر و توالی‌یابی این قطعات با استفاده از چارچوب‌های فنی متنوعی از جمله Illumina /Solexa و Roch454 و ABI Solid است (۲۲،۱۵) که تفاوت عمده بین این چارچوب‌ها در اندازه قطعات RNAی توالی‌یابی شده است (۲۱). در روش RNA-seq تفاوت میزان رونویسی در بین قطعات کوچک شده RNA به‌صورت تفاوت در تعداد دفعات خوانش قطعات به‌هنگام توالی‌یابی مشخص می‌شود (۱۴،۳۵). از این‌رو، علاوه بر فراهم نمودن توالی نوکلئوتیدی قطعات کوچک حاصل از کل ترانسکریپتوم، میزان رونویسی شدن هر یک از آن‌ها به‌صورت کمی و با یک معیار عددی دقیق به‌نام FPKM یا شاخص محاسبه بیان ژن به‌دست می‌آید. دقت اندازه‌گیری بیان ژن با این روش را آنالیز دیجیتالی بیان ژن می‌نامند (۲۴).

گاو از جنس Bos بوده و اکثر نژادهای گاو گوشتی از گونه‌ی*Bos taurus* می‌باشند. بنابراین اتوزوم‌های گاو به‌صورت مخفف BTA1 تا BTA29 و کروموزوم‌های جنسی X و Y به‌ترتیب به‌صورت BTX و BTY نام‌گذاری شده‌اند (۶،۱). ژنوم گاو از ۳۰ جفت کروموزوم اتوزوم و یک جفت کروموزوم جنسی تشکیل شده است. از دو نوع گاو اهلی موجود، گاوهای بوس تائوروس از اروپا و شمال امریکا منشا گرفته‌اند و مناسب مناطق معتدل هستند، ولی گاوهای بوس ایندیکوس کوهان‌دار بوده و از جنوب آسیا منشا یافته و پراکنش آن‌ها بیشتر در مناطق حاره‌ای است (۱۳). تفاوت فنوتیپی قابل توجهی بین این دو گونه وجود دارد و همچنین از نظر مقابله با تنش حرارتی و نیز مقاومت در مقابل انگل‌ها و بیماری‌ها متفاوت هستند (۱۷). نژاد کلیستانی (Bos indicus) شباهت بسیار زیادی با نژاد سیستانی دارد و به تحمل حرارت و مقاومت در برابر کنه و بیماری مشهور است (۱۱). در مقابل گاو هلشتاین جزو نژادهای غالب در صنعت گاو شیری است. علت انتخاب این دو نژاد تفاوت فنوتیپی قابل توجه بین این دو نژاد است. یکی از اقدامات مهم در حیوانات اهلی مطالعه ژن‌ها و پروتئین‌های مرتبط با صفات اقتصادی و بررسی آن‌ها در سطح سلولی یا کروموزومی است (۲۰،۲۲).

Huang و همکاران (۲۰۱۲) ویژگی‌های ترانسکریپتوم گلبول‌های سفید خون را در دو نژاد گاو تائورین (هلشتاین و جرزی) و یک نژاد ایندیسین (کلیستانی) تعیین و مقایسه نمودند. تحلیل ژن‌های متفاوت بیان شده، شباهت زیاد بین هلشتاین و جرزی و تفاوت زیاد بین نژادهای تائورین و ایندیسین را نشان داد. به‌غیر از تحقیق مذکور تاکنون تحقیقات بسیار اندکی مبنی بر تعیین پروفایل بیان برای دو نژاد به‌صورت مجزا انجام نشده است. بسیاری از پژوهش‌هایی که تاکنون در حوزه مختلف داده‌های امیکس در حیوانات مزرعه‌ای بالاخص گاوهای شیری صورت گرفته، بدون در نظر گرفتن تعامل ژن–دارو بوده است. بنابراین هدف از انجام پژوهش حاضر بررسی ژن‌های متفاوت بیان شده بین این دو نژاد بود. هچنین علاوه بر این تلاش شد که شبکه برهم‌کنشی پروتئینی پیش بینی شده حاصل از ژن‌های متفاوت بیان شده به‌دست آید و فهرستی از داروهای کاندید برای ژن‌های متفاوت بیان شده با استفاده از تعامل ژن–دارو استخراج شود.

مواد و روش‌ها

در این پژوهش، از داده‌های Huang و همکاران (۱۸) که مربوط به گاوهای نژاد هلشتاین و کلیستانی است استفاده شد (۳۰). این دو سری از داده‌ها از ادغام ۷۵ نمونه خون گاوهای ماده هلشتاین و ۷۰ نمونه خون گاو ماده کلیستانی تولید شدند. پس از اخذ داده‌ها، تمامی مراحل آماده‌سازی و تجزیه و تحلیل داده‌ها در محیط گرافیکی (GNU) سیستم عامل لینوکس اوبونتو نسخه ۱۴،۰۴ انجام شد. مراحل آماده‌سازی داده‌ها شامل تبدیل فرمت sra به فرمت fastq با نرم‌افزار Sratoolkits نسخه ۲،۵،۲ بود. سپس سنجش کیفیت داده‌ها با نرم‌افزار FastQC نسخه ۳،۱۱،۰ (۱) انجام شد. سنجش کیفیت داده‌ها برای خوانش‌های جفتی یا دوسویه (رفت و برگشت) به‌طور جداگانه با ۱۰ آزمون مختلف در محیط جاوا انجام شد. این آزمون‌ها شامل: کیفیت توالی به‌ازای tile، امتیاز کیفیت به‌ازای توالی، محتوای نوکلئوتیدی توالی به‌ازای هر باز، محتوای GC به‌ازای توالی، محتوای باز خوانده نشده (N) به‌ازای هر باز، توزیع طول توالی‌ها، تکراری بودن توالی، توالی‌هایی که بیش از حد وجود دارند، محتوای آداپتوری و محتوای Kmer بودند. نرم‌افزار نتیجه هر آزمون را جداگانه کنترل کرد و رد، قبول و هشدار به‌ترتیب با رنگ‌های قرمز، سبز و نارنجی مشخص شد و نتایج ۱۰ آزمون به‌صورت مجزا جهت چگونگی انجام ویرایش استفاده شد. در پژوهش حاضر از Illumina 1.5 جهت استاندارد کدگذاری کیفیت برای سامانه ایلومینای Genome Analyzer IIx استفاده شد. پس از ویرایش داده‌ها مجدداً بررسی کیفیت با آزمون سنجش کیفیت انجام شد. ویرایش خوانش‌ها، با نرم‌افزار Trimmomatic (۵) نسخه ۳۳،۰ صورت گرفت. این مرحله شامل حذف آداپتورها و حذف یا ویرایش (کوتاه نمودن) خوانش‌های بی‌کیفیت بود. مرحله هم‌ردیفی خوانش‌های کوتاه روی ژنوم مرجع با استفاده از نرم‌افزار TopHat2 انجام شد که از Bowtie2 به‌عنوان هم‌ردیف کننده استفاده شد. به‌منظور تشکیل ترانسکریپتوم و مکان‌یابی خوانش‌ها روی ژنوم مرجع از ژنوم مرجع UMD3.1 (ویرایش ۸۲) و اطلاعات حاشیه‌نویسی نسخه‌ی ۸۴ به فرمت‌های gtf و gff استفاده شد. لازم به ذکر است دسترسی به اطلاعات مربوط به مراحل تهیه کتابخانه‌ها، اطلاعات فنی دستگاه توالی‌یابی و

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

آزمایشگاه مربوطه وجود نداشت، لذا مقادیر مربوط به اندازه فاصله درون کتابخانه‌ای و انحراف معیار با دستورات TopHat2 و نمونه‌گیری از داده‌های هر نژاد با نرم‌افزار bwa برآورد شد. ژنوم مرجع ایندکس و فایل BAM با استفاده از دستورات بسته SAMtools (۲۱)، مرتب و ایندکس شدند. شناسایی و شمارش رونوشت‌ها یا به‌عبارتی آنالیز بیان افتراقی ژن‌ها توسط پکیج Cufflinks نسخه ۲،۲،۰ انجام شد. این بسته نرم‌افزاری دارای سه ابزار متفاوت: Cufflinks, Cuffmerge, Cuffdiff برای بررسی بیان ژنی است که هم‌ردیفی ترانسکریپتوم به‌وسیله خوانش‌های جداگانه برای هر جمعیت، با استفاده Cufflinks و تلفیق دو ترانسکریپتوم با Cuffmerge انجام شد. بین دو نژاد هلشتاین و کلیستانی مورد پژوهش آنالیز بیان افتراقی ژن‌ها با Cuffdiff محاسبه شد. به‌منظور مصورسازی نتایج نمودارهای قابل فهم، از بسته CummeRbund در نرم‌افزار R استفاده شد و تجزیه و تحلیل هستی‌شناسی ژن‌ها انجام گرفت.

استخراج ساختار پروتئینی

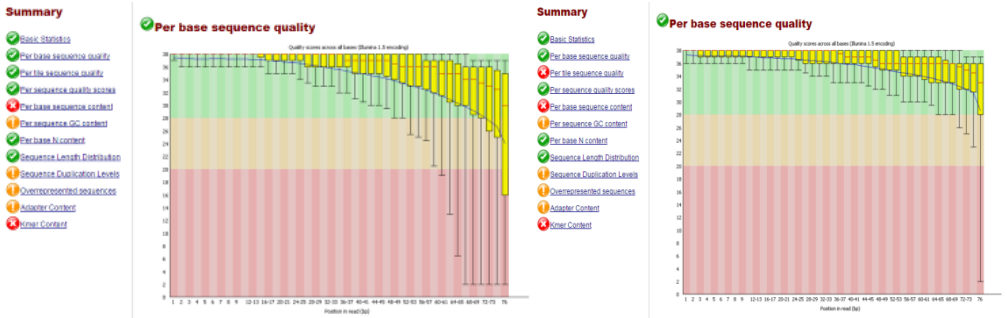
ابتدا نام ژن‌های متفاوت بیان شده به پایگاه Uniprot (قسمت UniprotKB) که شامل ساختار اول پروتئین‌ها است، وارد شدند و از این طریق ساختار اول و عملکرد پروتئین‌ها مشخص شدند. داده‌ها با فرمت Fasta بارگذاری و ذخیره شدند. سپس ساختارهای اول به سایت Blastp داده شد. ساختار اول پروتئین‌های مورد نظر با ساختار اول همه پروتئین‌هایی که ساختار سوم آن‌ها معلوم بود، مقایسه شدند. پروتئین‌هایی مشابه با ساختار پروتئین مد نظر یافت شد و کد pdb پروتئین‌های مشابه ثبت شدند. ثابت شده است در صورتی‌که حداقل ۴۰ درصد توالی‌های اسید آمینه‌ای دو پروتئین یکسان باشند، ساختار سوم و عملکرد آن دو پروتئین نیز یکسان خواهند بود (۲۹). به این ترتیب با استفاده از سایت Blastp کد PDB پروتئین مورد بررسی یا پروتئین‌های مشابه آن به‌دست آمد (جدول۲). سپس با مراجعه به پایگاه داده همکاری پژوهش بیوانفورماتیک ساختاری RCSB، ساختار سوم پروتئینی ژن‌های متفاوت بیان‌شده استخراج شد تا عملکرد زیستی این پروتئین‌ها مورد کنکاش بیشتر قرار گیرد. با بررسی عملکرد پروتئین‌ها مشخص خواهد شد که تفاوت در عملکرد کدام پروتئین‌ها باعث متفاوت شدن گونه هلشتاین و کلیستانی در سطح پروتئینی شده است.

استخراج کاندیدهای دارویی

فهرست ژن‌های متفاوت بیان شده برای بررسی تعامل آن‌ها با داروها مورد استفاده قرار گرفت. برای این‌کار از تارنمای http://dgidb.org/ و https://www.drugbank.c/ استفاده شد.

نتایج و بحث

کیفیت خوانش‌ها در نژاد هلشتاین و کلیستانی نسبتاً مطلوب ارزیابی شد. لازم به ذکر است که کیفیت خوانش در ۱۰ تا ۱۵ نوکلئوتید در انتهای رشته‌ها پایین بود که معمولاً این مورد در تمام روش‌های کلاسیک و پربرون‌داد توالی‌یابی متداول است (شکل ۱).



شکل ۱- آزمون‌های ده گانه سنجش کیفیت و آزمون کیفیت انفرادی نوکلئوتیدهای خوانش شده برای نژاد هلشتاین(چپ) و کلیستانی(راست)

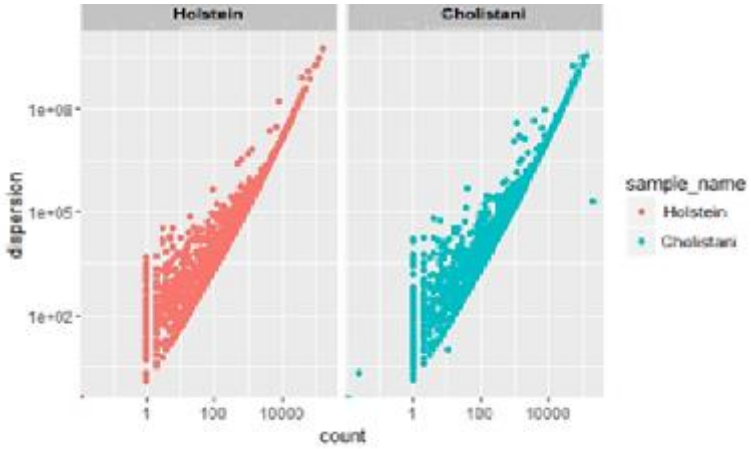
Figure 1. Ten quality assessment tests and unique quality test for sequenced nucleotides for Holstein (left) and Cholistani (right) breeds

حداقل طول قابل قبول برای خوانش‌ها پس از ویرایش ۵۰ جفت باز در نظر گرفته شد. در مرحله ویرایش داده‌ها تعداد خوانش‌های بی‌کیفیت برای دو نژاد حذف شد. در نژاد کلیستانی از مجموع ۲۰،۹۴۰،۰۶۳ خوانش اولیه پس از ویرایش تعداد ۱۹،۳۷۹،۴۸۷ خوانش کیفیت مناسب داشته و باقی ماندند. به‌عبارت دیگر در نژاد کلیستانی ۱،۵۶۰،۵۷۶ خوانش (حدود ۷/۵ درصد از کل خوانش‌ها) از آنالیز حذف شد در نژاد هلشتاین از مجموع ۲۱،۰۷۸،۴۷۷ خوانش اولیه پس از ویرایش، تعداد ۱۹،۳۴۶،۱۷۲ خوانش کیفیت مناسب داشته و باقی ماندند و ۱،۷۳۲،۳۰۵ خوانش (حدود ۸ درصد از کل خوانش‌ها) از آنالیز کنار گذاشته شدند. از ۱۹،۳۷۹،۴۸۷ خوانش پیش‌رو و پس‌رویِ ویرایش شده برای نژاد کلیستانی به‌ترتیب ۱۵،۶۴۷،۱۲۱ (۸۰/۷٪) و ۱۵،۳۶۵،۰۵۷ (۷۹/۳٪) خوانش مکان‌یابی شدند. بیش از شش درصد معادل نهصد هزار خوانش با چند مکان هم‌ردیف شدند که هم‌ردیفی چندگانه داشتند. از این میان بیش از ۲۲۰۰۰ خوانش با بیش از ۲۰ مکان ژنومی همردیف شدند. نرخ مکان‌یابی کل برای نژاد کلیستانی ۸۰ درصد به‌دست آمد. علاوه بر این، تعداد ۱۳،۸۵۵،۲۶۰ جفت‌خوانش (دوسویه) برای نژاد کلیستانی همردیف گردیدند که از این تعداد ۵/۹ درصد (معادل ۵۷۸،۸۱۸ جفت‌خوانش) هم‌ردیفی چندگانه و ۳ درصد (معادل ۴۱۶۰۰۰ جفت‌خوانش) هم‌ردیفی ناجور داشتند.

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

نرخ همردیفی جفتی جور شده ۶۹/۳ درصد به‌دست آمد و نرخ مکان‌یابی برای خوانش‌های پیش‌رو و پس‌رو در نژاد هلشتاین به‌ترتیب ۶۵/۶ و ۵۴/۸ درصد برآورد شد. همچنین نرخ همردیفی چندگانه به‌ترتیب ۴/۶ و ۴/۴ درصد و تعداد خوانش‌های هم‌ردیفی روی بیش از ۲۰ موقعیت ژنومی حدود هشت هزار بود. نرخ همردیفی کل ۶۰/۲ درصد و نرخ هم‌ردیفی جفتی جور شده ۵۰/۸ درصد به‌دست آمد. در مجموع و با توجه به پژوهش‌های قبلی، نرخ هم‌ردیفی در حد مطلوبی ارزیابی شد (شکل ۲).



شکل ۲- نمودار توزیع و پراکنش خوانش‌های به‌دست آمده در این پژوهش بر حسب نژادهای مورد بررسی
Figure 2. Plot of reads distribution in this study based upon investiged breeds

باتوجه به اطلاعات به‌دست آمده از فایل خروجی Cuffdiff، ۲۴۶۱۶ ژن و ۲۶۷۱۶ ایزوفرم روی ترانسکریپتوم نمونه جمعیت گاو هلشتاین آمریکا و کلیستانی پاکستان یافت شد که در این میان ۴۱ ژن شناسایی شد (جدول ۱) که بیان متفاوتی داشتند (p<0/000015). یک ژن حاوی دو ایزوفرم است که می‌تواند پروتئین‌های مختلفی را تولید نماید. بررسی این ژن‌ها در سطوح ایزوفرم‌های آن‌ها به‌منظور درک بهتر پیچیدگی‌های ترانسکریپتوم انجام شد. ژن ENSBTAG00000014332 در موقعیت کروموزوم ۵ از نوکلئوتید ۹۹۶۴۷۱۲۲ تا ۹۹۷۰۱۶۳۰، از نظر عملکردی مشابه ژن‌های NKG2A و NKG2B (پروتئین‌های پیوسته غشایی نوع ۲) است. نام دو ایزوفرم آن به‌صورت ENSBTAT00000056514 و ENSBTAT00000039039 (حاوی ۷۱۱ جفت باز و ۶ اگزون) ثبت شده است که ایزوفرم ENSBTAT00000056514 دارای یک اگزون بیشتر (۷۲۹ جفت باز) نسبت به دیگری می‌باشد. بیش‌ترین مقدار بیان در ژن‌های کروموزوم x مشاهده شد که احتمالاً به‌علت وجود ژن‌های مربوط به سامانه ایمنی است که روی این کروموزوم قرار دارند و این کروموزوم با طول ۱۴۸۸۲۳۸۹۹ جفت‌باز یکی از طولانی‌ترین کروموزوم در ژنوم گاو است. نتایج حاصل از تفاوت بیان ژن بین دو جمعیت را می‌توان برای بررسی ارتباط چندین صفت، تعریف شبکه‌هایی از ژن‌ها که بر صفات همبسته موثرند و نیز به‌منظور معرفی ژن‌های کاندیدای مرتبط با صفات مورد نظر استفاده نمود. اسلینگ و همکاران (۲۰۱۳) بیان کردند چنین شبکه‌هایی را می‌توان از نتایج پژوهش‌های بیان ژن افتراقی (که با استفاده از ریزآرایه‌ها یا توالی‌یابی RNA صورت می‌گیرند و به فهرستی از ژن‌های متفاوت بیان‌شده ختم می‌گردند) به‌دست آورد. این شبکه‌ها با حاشیه‌نویسی، آنالیز ماهیت و شناسایی مسیرهای مولکولی تکمیل و تایید می‌شوند (۳۲).

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

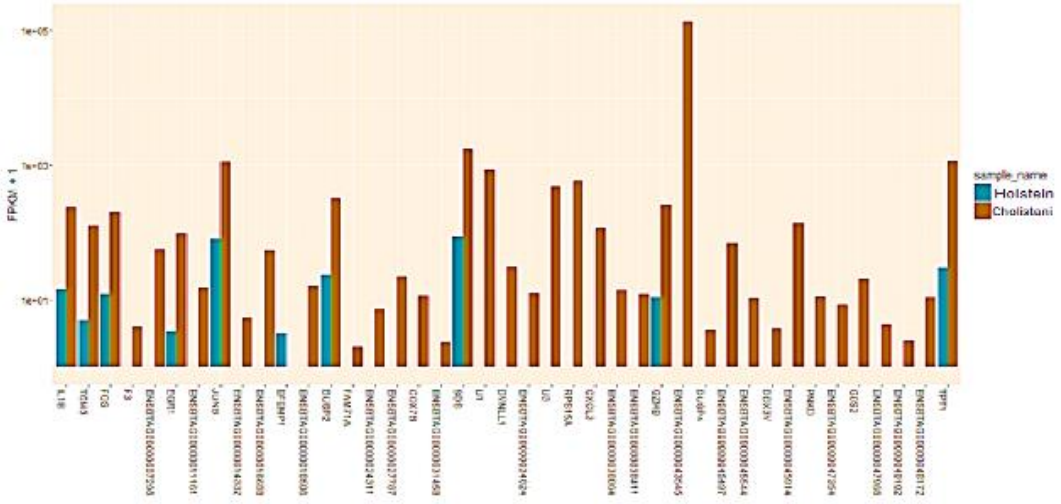
[DOI: 10.52547/rap.11.28.121]

جدول ۱- فهرستی از ژن‌های با بیان متفاوت بین در گاوهای هلشتاین و کلیستانی

Table 1. List of genes with different expression in Cholistani and Holstein cattle

کد ژن های حاشیه نویسی	نام ژن	جایگاه ژنی	P-Values
ENSBTAG00000001785	TGM3	13:53261634-53304213	0.00005
ENSBTAG00000004322	FOS	10:86883738-86887169	0.00005
ENSBTAG00000007101	F3	3:49110155-49121034	0.0001
ENSBTAG00000007296	-	1:82549641-82550013	0.00005
ENSBTAG00000010069	EGR1	7:51438726-51442500	0.0001
ENSBTAG00000011161	-	8:38683732-38684116	0.0001
ENSBTAG00000012046	JUNB	7:13851130-13852933	0.00005
ENSBTAG00000014332	–	5:99647122-99701630	0.00005
ENSBTAG00000016688	-	X:66722398-66723397	0.00005
ENSBTAG00000017448	EFEMP1	11:38338743-38408288	0.0001
ENSBTAG00000018506	-	5:58013469-58014035	0.00005
ENSBTAG00000020350	DUSP2	11:2250454-2252492	0.00005
ENSBTAG00000021166	FAM71A	16:72811910-72814133	0.0001
ENSBTAG00000027787	-	18:65621141-65621509	0.00005
ENSBTAG00000030735	COX7B	7:42303402-42303827	0.00005
ENSBTAG00000031458	-	4:101019787-101021837	0.00015
ENSBTAG00000031814	SDS	17:63302946-63311098	0.00015
ENSBTAG00000032308	U1	3:21433393-21433557	0.00005
ENSBTAG00000034170	DYNLL1	20:11746000-11746270	0.00005
ENSBTAG00000034824	-	20:2680573-3054892	0.0001
ENSBTAG00000037147	U3	20:64476312-64476463	0.00005
ENSBTAG00000037600	RPS15A	8:84843153-84843546	0.00005
ENSBTAG00000037778	CXCL3	6:90811061-90813079	0.00005
ENSBTAG00000038064	-	2:11700813-11701755	0.00005
ENSBTAG00000038411	-	8:62722029-62722587	0.00005
ENSBTAG00000039813	GZMB	21:35135770-35137852	0.00005
ENSBTAG00000043545	-	MT:5686-7299	0.00005
ENSBTAG00000044208	DUSP4	27:24815380-24828237	0.00005
ENSBTAG00000045497	–	21:66764060-66764393	0.0001
ENSBTAG00000045544	EIF2S3Y	X:144159055-144234222	0.00005
ENSBTAG00000045750	DDX3Y	X:143364825-143375658	0.00005
ENSBTAG00000045914	–	7:21247238-21247649	0.00005
ENSBTAG00000047277	PNKD	18:46516556-46516985	0.00005
ENSBTAG00000047354	-	5:32529438-32529956	0.00015
ENSBTAG00000047766	G0S2	16:75539495-75540404	0.00015
ENSBTAG00000047966	-	16:6155902-6156793	0.00005
ENSBTAG00000048102	-	X:143736737-143800954	0.00015
ENSBTAG00000048172	ZRSR2Y	X:142664999-142690152	0.00005
ENSBTAG00000048229	TPT1	25:1394200-1436118	0.00005

نتایج تفاوت بیان ژن در شکل ۳ نمایش داده شده است. در نمودار نقطه‌ای میزان انحراف هر نقطه از خط در واقع نشانگر میزان بیان می‌باشد. نقاطی که بالای خط قرار دارند بیان بیشتر و نقاطی که پایین خط قرار دارند بیان کمتری دارند. همچنین نقاطی که روی خط قرار گرفتند که بیانشان تغییری نکرده است. کنکاش حاشیه‌نویسی ژنومی الگوی بیانی ژن‌های متفاوت بیان شده در دو نژاد نشان داد که ژن‌های درگیر در مسیرهای متابولیکی مرتبط با سامانه ایمنی (IL1B interleukin 1 beta) در نژاد کلیستانی میزان خوانش بالایی داشتند؛ این در حالی است که در نژاد هلشتاین بیان کم یا بیان صفر داشتند (شکل ۳) بیشترین میزان بیان برای ژن میتوکندریایی ENSBTAG00000043545 با مقدار ۱۳۵۰۳۱FPKM در نمونه جمعیت کلیستانی بوده است. این ژن در موش ژن کد کننده tRNA بوده و در سنتز اسیدآمینه فنیل‌آلانین نقش دارد که یکی از بیست اسید آمینه‌ی اصلی یاخته‌های زنده است و در گاو ژنی با عملکرد ناشناخته محسوب می‌شود و اطلاعات زیادی در فایل حاشیه‌نویسی گاو برای آن ثبت نشده است (۱۶).



شکل ۳- تفاوت بیان ژن‌ها در هر یک از نژادهای مورد بررسی

Figure 3. Diferential gene expression in each of investigated breeds

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

آنالیز ماهیت ژن (GO) و مسیرهای درگیر نشان داد ژن‌ها در ۲۰ مسیر درگیر می‌باشند؛ به‌طوری که تعداد زیادی از ژن‌ها در مسیرهای منجر به پاسخ ایمنی، زنجیره انتقال الکترون و مسیرهای ترجمه قرار گرفتند. این مسیرها ممکن است در سطوح مختلف مقاومت به بیماری و تنش حرارتی در نژادهای تائورین و ایندیسین نقش داشته باشند. از جمله نتایج جالب این پژوهش، قرار گرفتن تعداد زیادی از این ژن‌ها در مسیرهای مربوط به ایمنی در نژاد کلیستانی است. می‌توان نتیجه گرفت دو زیرگونه ایندیسین و تائورین در بیان ژن‌های مرتبط با تحمل حرارتی و مقاومت به بیماری تفاوت آشکاری را نشان می‌دهند این امر دلیلی بر اهمیت پرورش نژادهای بومی در شرایط محیطی نامساعد است.

ساختار پروتئینی ژن‌های پیش‌بینی شده

در جدول ۲ کد PDB و درصد یکسان بودن ساختار اول پروتئین‌های یافت شده از سایت Uniprot با پروتئین‌هایی که دارای ساختار سوم هستند (یعنی کد PDB آن‌ها ذکر شده است) و همچنین عملکرد پروتئین‌ها نشان داده شده است. با توجه به اطلاعات حاصل از سایت Uniprot، عملکرد پروتئین‌های حاصل از ژن‌های بیان شده در جدول ۲ ذکر شده است. همان‌طور که مشخص است در اغلب موارد عملکرد ذکر شده برای ساختار اول در سایت Uniport با عملکرد ذکر شده در سایت PDB برای پروتئین‌های مشابه یکسان که به‌علت شباهت زیاد بین آن‌ها است. در مواردی که عملکردی برای پروتئین در سایت Uniprot پیشگویی نشده بود با مقایسه پروتئین مشابه در بانک PDB می‌توان به عملکرد پروتئین پی‌برد. به‌عنوان مثال در سایت Uniprot ساختار اول عملکرد ژن‌ها یا پروتئین‌های ۴، ۲۴، ۸، ۳۴، ۳۶ و ۱۳ مشخص نشده است، ولی عملکرد پروتئین‌های مشابه در سایت PDB تعیین شده است. همان‌طور که از جدول ۲ مشخص است، هشت مورد از این پروتئین‌های یافت شده در خانواده هیدرولازها قرار می‌گیرند که به پروتئین‌های ۲۴، ۲۶، ۲۸، ۲۹، ۱۱، ۲۷، ۲۲و ۳ مربوط می‌باشند. به‌عبارت دیگر تقریباً یک سوم ژن‌هایی که عملکردشان در این دو نژاد تغییر می‌کند مربوط به آنزیم‌های هیدرولاز هستند. همچنین پنج پروتئین ۲، ۱۳، ۶، ۱۲ و ۵ در تنظیم رونویسی و بیان ژن‌ها دخالت داشته و پروتئین شماره ۳۶ در ویرایش RNA دخیل می‌باشد. بقیه ژن‌ها هم مربوط به عملکردهای مختلف می‌باشند.

جدول ۲- شناسایی پروتئین‌های یافت شده در این پژوهش

Table 2. Identification of proteins found in this study

Number	PDB code	% Ident	function PDB	Classification	Function Uniprot
30	4hga	99	histone H3.3-H4 heterodimer in complex with its chaperone DAXX	CHAPERONE / APOPTOSIS	BOVIN Histone H3
21	1qnk	82	truncated human GROB [5-73]	CHEMOKINE	C-X-C motif chemokine
4	1a38	98	PROTEIN ZETA BOUND TO R18 PEPTIDE	COMPLEX (SIGNAL TRANSDUCTION / PEPTIDE)	Uncharacterized protein
17	1f3c	100	8kda dynein light chain (dlc8)	CONTRACTILE PROTEIN	Dynein light chain 1, cytoplasmic
14	2y69	100	Bovine heart cytochrome c oxidase	ELECTRON TRANSPORT	Cytochrome c oxidase subunit 7B, mitochondrial
24	1euf	100	bovine duodenase (new serine protease)	HYDROLASE	Uncharacterized protein
26	3ezz	99	human MKP-2 Dual specificity protein phosphatase 4	HYDROLASE	Dual specificity protein phosphatase
28	5koy	99	Mouse pgp 34 linker deleted bound with ATP Multidrug resistance protein 1A	HYDROLASE	BOVIN Eif2s3y (Fragment) eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked
29	5e7i	97	active catalytic core of the human DEAD-box protein DDX3	HYDROLASE	BOVIN DEAD box polypeptide 3 Y-linked short isoform
11	1m3g	95	catalytic domain of mapk phosphatase pac-1	HYDROLASE	Dual specificity protein phosphatase
27	4rnd	54	subunit DF-assembly of the eukaryotic V-ATPase	HYDROLASE	Uncharacterized protein
22	4arz	50	Gtr1p-Gtr2p complexed with GTP-GDP	HYDROLASE	Uncharacterized protein
3	1wqv	75	Human Factor Viia-Tissue Factor	HYDROLASE	TF_BOVIN Tissue factor
1	3o4o	67	Interleukin-1 receptor complex	IMMUNE SYSTEM	Interleukin-1 beta precursor
8	3bdw	52	Human CD94/NKG2A	IMMUNE SYSTEM RECEPTOR	Uncharacterized protein
34	5vlq	60	TTLL3 Glycylase	LIGASE	Uncharacterized protein
15	1pwe	78	Rat Liver L-Serine Dehydratase Apo Enzyme	LYASE	L-serine dehydratase/L-threonine deaminase
10	1emn	46	fibrillin ca2+ binding epidermal growth factor-like domains	MATRIX PROTEIN	EGF-containing fibulin-like extracellular matrix protein 1
20	2zkq	100	Ribosomal 40S subunit	Ribosomal	40S ribosomal protein S15a
36	4yh8	41	yeast U2AF complex	SPLICING	Uncharacterized protein
2	2wt7	100	FOS Fos proto-oncogene, AP-1 transcription factor subunit	TRANSCRIPTION	FOS_BOVIN Proto-oncogene c-Fos
13	1vcb	97	The vhl-elonginc-elonginb structure	TRANSCRIPTION	Uncharacterized protein
6	5vpa	81	Transcription factor FosB/JunD bZIP domain	TRANSCRIPTION	Transcription factor jun-B
12	3rd3	44	Pseudomonas aeruginosa transcriptional regulator PA2196	TRANSCRIPTION	BOVIN Family with sequence similarity 71 member A
5	1aay	98	ZIF268 ZINC FINGER-DNA COMPLEX	TRANSCRIPTION / DNA	Early growth response protein
37	1yz1	99	human translationally controlled tumour associated protein	UNKNOWN FUNCTION	Translationally-controlled tumor protein

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

نتایج بررسی پروتئین‌ها نشان می‌دهد عملکرد این پروتئین‌ها به‌طور دقیق و آزمایشگاهی مشخص نشده ولی بیشتر آن‌ها دارای شواهدی در سطح بیان ژن هستند. از طرف دیگر بر اساس نظریه تکامل داروین که بیانگر قرابت دور یا نزدیک موجودات و پروتئین‌های موجود در آن‌ها است، می‌توان انتظار داشت که با بررسی توالی پروتئین‌ها از پایگاه داده‌های پروتئینی، به اقوام دور یا نزدیک موجودات پی برد و یا به پروتئین‌هایی دست یافت که به‌دنبال مضاعف‌شدگی درون ژنوم موجود زنده دچار یک تکامل موازی شده‌اند. بنابراین برای تعیین خویشاوندی بین پروتئین‌هایی که در پایگاه داده Uniprot فاقد نتایج تایید شده بودند، BLASTP انجام شد و نتایج قابل‌توجهی از این بررسی به‌دست آمد. به‌عنوان مثال پروتئین G3MYI2 (Uniprot code) که ژن مشخصی برای آن تعیین نشده و نامی هم برای آن پیشنهاد نشده است، بعد از BLAST مشخص شد که ۹۵ درصد اسیدهای‌آمینه‌ی آن با پروتئین Oral-facial-digital syndrome 1 Y-linkedisoform 1 دارای همسانی می‌باشد. این میزان همسانی بیانگر شباهت بسیار زیاد بین این دو پروتئین می‌باشد که از این‌رو می‌توان انتظار داشت پروتئین G3MYI2 در جایی دیگر از ژنوم دچار مضاعف‌شدگی شده است و در اثر جهش‌های مستقل تفاوت بین آن‌ها ایجاد شده است. نکته قابل توجه دیگر این که این پروتئین در حالت عادی در سلول‌های مورد بررسی بیان نمی‌شود، بلکه در اثر شرایط محیطی اعمال شده بیان شده‌اند. پروتئین G3N3S4، فاقد بررسی‌های آزمایشگاهی است، بعد از انجام BLAST مشخص شد که از نظر توالی اسیدهای‌آمینه‌ای با نوع پیش بینی شده‌ی پروتئین Tubulin monoglycylase TTLL3 از ۹۹ درصد همسانی برخوردار است. مورد دیگر، پروتئین A0A0N4STN1 می‌باشد که هیچ ژن یا پروتئین مشخصی برای آن گزارش نشده است، اما بعد از انجام BLAST مشخص شد این پروتئین با همسانی ۱۰۰ درصدی (Coverage=100% & Indentity=100%) پروتئین Cathelicidin 1 گاوی می‌باشد. در پژوهشی نشان داده شد که این پروتئین روی باکتری‌های *E.coli* و *S. aureus* اثر کشندگی دارد (۱۷). تفاوت بیان این پروتئین در نژاد هلشتاین و کلیستانی می‌تواند ناشی از شرایط متفاوت محیطی برای دو گونه‌ی مورد بررسی از نظر میزان آلودگی‌های باکتریایی باشد. پروتئین E1DBS9، فاقد اطلاعات ژن بود و پس از انجام BLAST مشخص شد که این پروتئین جزئی از پروتئین های ابرخانواده 3-3-14 بوده و با همسانی ۱۰۰ درصدی به‌طور حتم پروتئین- TPA: tyrosine 3/tryptophan 5 monooxygenase activation protein, zeta polypeptide-like می‌باشد و همانند سایر پروتئین‌های ابرخانواده 3-3-14 انتظار می‌رود این پروتئین (E1DBS9) در بازآرایش اسکلت سلولی و مهاجرت سلولی و همچنین در فعال یا غیرفعال کردن چرخه سلولی نقش داشته باشد (۸). پروتئین E1BGP6 در پایگاه داده Uiprot فاقد اطلاعات ژن و پروتئین بود اما در سطح بیان پروتئینی شامل اطلاعات آزمایشگاهی بود. نتایج BLAST پروتئین نشان داد که جزیی از ابرخانواده P-Loop_NTPase است و به‌نظر می‌رسد بر اساس همسانی ۱۰۰ درصدی و پوشش کامل بین توالی ورودی و استخراج شده، همان پروتئین پیش‌بینی شده‌ی Ras-related GTP-binding protein A باشد. پروتئین C0KTL0 نیز در پایگاه داده Uniprot فاقد اطلاعات ژن می‌باشد ولی برای پروتئین نام Eif2s3y پیشنهاد شده است و رونوشت‌های این پروتئین در سطح آزمایشگاهی تایید شده است. نتایج BLAST نشان داد این پروتئین به ابرخانواده SelB وابسته است و با همسانی ۱۰۰ درصدی و پوشش کامل توالی، احتمالا به پروتئین پیش‌بینی شده Eukaryotic translation initiation factor 2 subunit 3, Y-linked در سلول‌های گاوی شباهت دارد. از آن‌جایی که این توالی با همسانی ۹۹ درصدی با پروتئین Eukaryotic translation initiation factor 2 subunit 3, Y-linked-like isoform X1 & X2 در سلول‌های گوزن شباهت دارد، در نتیجه با احتمالی نزدیک به اطمینان می‌توان این پروتئین را نوع گاوی پروتئین موجود در گوزن دانست. البته این موضوع نیاز به بررسی بیشتری دارد. با بررسی مثال‌های گفته شده می‌توان انتظار داشت پروتئین‌های فاقد تأیید آزمایشگاهی دارای عملکردهایی مشابه با عملکرد اقوام مشابه‌شان باشند که در اثر شرایط محیطی متفاوت، فرصت بیان شدن پیدا کرده‌اند و یا به‌علت متفاوت بودن شرایط محیطی، دارای عملکردهای دیگری در شبکه میان‌کنشی سلولی شده‌اند. بررسی Gene Antology پروتئین‌های به‌دست آمده نشان داد که از مجموع ۴۱ پروتئین، ۲۳ پروتئین در پایگاه DAVID دارای سوابق ثبت شده می‌باشند. بررسی میان‌کنش شبکه‌ای بین آن‌ها نشان می‌دهد این پروتئین‌ها می‌توانند در مسیرهایی مانند مسیر سیگنال‌دهی TNF، عفونت سالمونلایی، مسیر سیگنال‌دهی MAPK و آرتریت روماتوئید دارای عملکرد باشند (جدول ۳).

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

جدول ۳- نتایج GO استخراج شده از DAVID برای شبکه میان‌کنشی رسم شده

Table 3. GO results extracted from DAVID for the interactive network plotted

Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
bta04668:TNF signaling pathway	4	18.18181818	0.001128568	Q2KIE5, Q0VBZ5, P09428, O77628	16	108	7583	17.55324074	0.086376208	0.086376208	1.193299329
bta05132:Salmonella infection	3	13.63636364	0.011331628	Q2KIE5, P09428, O77628	16	83	7583	17.13027108	0.598161753	0.366092872	11.41043831
bta04010:MAPK signaling pathway	4	18.18181818	0.012940233	F1MGU0, F1MM08, P09428, O77628	16	257	7583	7.376459144	0.64724269	0.293424338	12.93085218
bta05323:Rheumatoid arthritis	3	13.63636364	0.014666117	P39942, P09428, O77628	16	95	7583	14.96644737	0.693328518	0.255836529	14.53579446
bta04380:Osteoclast differentiation	3	13.63636364	0.027999594	Q0VBZ5, P09428, O77628	16	134	7583	10.61054104	0.896886459	0.36516167	26.05991067
bta05168:Herpes simplex infection	3	13.63636364	0.052943972	F1MGU0, P09428, O77628	16	190	7583	7.483223684	0.987115908	0.515817631	43.91510511
bta05020:Prion diseases	2	9.090909091	0.061518541	P09428, Q29W20	16	32	7583	29.62109375	0.993776198	0.515977992	49.08414157
bta05332:Graft-versus-host disease	2	9.090909091	0.080006548	P80219, P09428	16	42	7583	22.56845238	0.998732993	0.565642462	58.79123477
bta05166:HTLV-I infection	3	13.63636364	0.097212151	Q29W20, P27674, O77628	16	269	7583	5.285548327	0.999720194	0.59709165	66.28490716
bta04940:Type I diabetes mellitus	2	9.090909091	0.098154176	P80219, P09428	16	52	7583	18.22836538	0.999742615	0.562419069	66.65704241

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

همچنین این پروتئین‌ها در فرایندهای زیستی از قبیل تنظیم مرگ سلولی، پاسخ ایمنی و تمایز سلول‌های ماهیچه‌ای اسکلتی نقش دارند (جدول ۴). فعال شدن یا غیر فعال شدن این مسیرها می‌تواند بر درگیری سامانه‌ی ایمنی بدن به‌دلیل شرایط نامساعد محیطی دلالت نماید که به تبع آن بر میزان تولید، رشد و تمایز سلول‌های ماهیچه‌ی اسکلتی تاثیرگذار خواهد بود.

جدول ۴- نتایج Biological Process(BP) حاصل از بررسی انجام شده در DAVID

Table 4. BP results from the DAVID survey

Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GO:0010941~regulation of cell death	2	9.090909091	0.011518538	P63103, Q0VBZ5	21	8	13824	164.5714286	0.770383457	0.770383457	12.57216289
GO:0006955~immune response	3	13.63636364	0.038063792	Q2KIE5, P80219, P09428	21	215	13824	9.18538206	0.9927627	0.914927679	36.24029073
GO:0071277~cellular response to calcium ion	2	9.090909091	0.042548164	Q0VBZ5, O77628	21	30	13824	43.88571429	0.996001925	0.841285366	39.60346811
GO:0035914~skeletal muscle cell differentiation	2	9.090909091	0.06317173	Q29W20, O77628	21	45	13824	29.25714286	0.999748338	0.874048132	53.08203001
GO:0042493~response to drug	2	9.090909091	0.095294104	Q0VBZ5, O77628	21	69	13824	19.08074534	0.999997005	0.921424259	68.69497873

شاخص‌های GO مرتبط با عملکرد مولکولی این پروتئین‌ها نشان داد که پروتئین‌های مورد بررسی به‌عنوان فاکتور رونویسی و تنظیم‌کننده‌ی بیان ژن کنش دارند (جدول ۵).

جدول ۵- نتایج عملکردهای مولکولی استخراج شده از DAVID

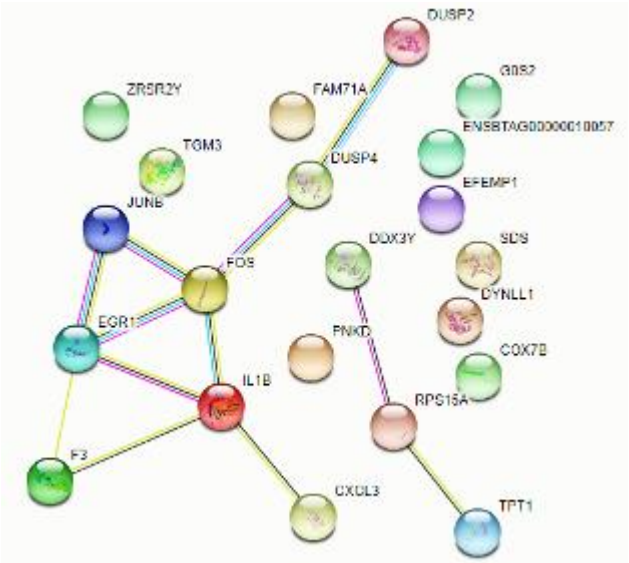
Table 5. Results of the Molecular Functions Derived from DAVID

Term	Count	%	Pvalue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GO:0008134~transcription factor binding	3	13.63636364	0.005752542	P63103, Q0VBZ5, O77628	21	75	13095	24.94285714	0.3007083 4	0.3007083 4	5.655707013
GO:0001077~transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding	3	13.63636364	0.032558731	Q0VBZ5, Q29W20, O77628	21	187	13095	10.00381971	0.871552463	0.641604218	28.39708418

برای درک بهتر از نحوه میان‌کنش بین این پروتئین‌ها در شبکه میان‌کنشی موجود در سلول، پروتئین‌های مورد بررسی در پایگاه داده String بارگذاری شدند، چرا که این پایگاه داده دارای بیشترین تعداد میان‌کنش ثبت شده بوده و از بیشترین پوشش سازواره‌ایی برخوردار می‌باشد؛ با انتخاب گونه مورد نظر در این نرم‌افزار، شبکه میان‌کنشی بین پروتئین‌ها بر اساس سازواره Taurus Bos ترسیم شد. از بین پروتئین‌های بارگذاری شده دو پروتئین «ELOB وOFD1Y» شناسایی نشدند که با بررسی بیشتر در پایگاه داده Uniprot مشخص شد که فاقد اطلاعات تایید شده در سازواره (*Bos taurus*) می‌باشند و فقط اطلاعاتی در سطح بیان ژنی از آن‌ها موجود است که می‌تواند در نتیجه تغییرات بیان ژنی در شرایط مختلف و در نتیجه بیان فاکتورهای رونویسی متفاوت و رونوشت‌برداری از ژن‌های متفاوت باشد. شبکه میان‌کنشی رسم شده بین پروتئین‌های استخراج شده نشان می‌دهد که ارتباط میان‌کنشی اندکی بین این پروتئین‌ها وجود دارد که می‌تواند به‌دلیل تفاوت‌های زیاد بین مسیرهای دربرگیرنده آن‌ها باشد و یا به‌علت شباهت بسیار زیاد بین این دو گونه، تنها پروتئین‌هایی باشند که ارتباط میان‌کنشی زیادی با یکدیگر ندارند. با این وجود دو مجموعه میان‌کنشی بین این پروتئین‌ها شناسایی شد: گروه اول: شبکه میان‌کنشی بین پروتئین‌های DDX3Y, RPS15A, TPT1 و EIF2S3Y و گروه دوم: شبکه میان‌کنشی بین پروتئین‌های DUSP2, DUSP4, FOS, IL1B, JUNB,CXCL3, F3 و EGR1 (شکل ۴).

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]



شکل ۴- شبکه تعامل پروتئین‌های پیش‌بینی شده حاصل از ژن‌های متفاوت بیان شده توسط پایگاه داده String
Figure 4. The network of interactions of the predicted proteins derived from differentially expressed genes by String database

با بررسی مجموع اعضای هر شبکه در پایگاه DAVID مشخص شد پروتئین‌های زیرشبکه گروه اول از نظر فرایند زیستی، عملکرد مولکولی و جای‌گیری درون‌سلولی با یکدیگر ارتباطی ندارند. پروتئین‌های شبکه گروه دوم از نظر فرایندهای زیستی (BP) در فرایندهایی نظیر شکل‌گیری اندودرم، پاسخ سلولی به یون کلسیم، تنظیم مثبت رونویسی توسط RNA پلیمراز II، پاسخ به دارو و تمایز سلول‌های ماهیچه اسکلتی نقش دارند و همچنین از نظر عملکردهای مولکولی (MF) در فرآیندهایی نظیر فعال‌سازی رونویسی، دخالت در اتصال اختصاصی توالی DNA به مناطق مرکزی پیشران پروکسیمال RNA پلیمراز II و فعالیت فسفاتازی سرین/ترئونین/تیروزین کینازی MAP نقش دارند. همگی موارد ذکر شده بیان‌گر نقش این پروتئین‌ها در فرایندهای رونویسی و تقسیم سلولی می‌باشد که می‌تواند ناشی از افزایش نیاز سلول به سنتز mRNA پروتئین‌های درون‌سلولی باشد که شرایط محیطی و سازگاری‌های ژنتیکی باعث افزایش بیان در این ژن‌ها شده است.

فهرست ژن‌های متفاوت بیان شده از جدول ۱ برای بررسی تعامل آن‌ها با داروهای احتمالی کاندید مورد استفاده قرار گرفت. جدول ۶ نتیجه حاصل از تعامل ژن – دارو را نشان می‌دهد. جدول ۶ نشان می‌دهد که برای تمامی ژن‌های متفاوت بیان شده در این پژوهش، ترکیب یا داروهای کاندید یافت نشد. به‌طوریکه برای تعدای از ژن‌های متفاوت بیان شده مثل DUSP2، FAM71A، COX7B، SDS، RNU1-1، DYNLL1 و RPS15A هیچگونه داروی کاندیدی یافت نشد. همچنین، همانطور که پیشتر بحث گردید، اصولا آندسته از ژن‌هایی بین دو نژاد متفاوت بیان شدند که– بر اساس حاشیه‌نویسی ژنومی آن‌ها– در ایمنی درگیر بودند.

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

جدول ۶– فهرست داروهای کاندید احتمالی برای ژن‌های متفاوت بیان شده در این پژوهش (آندسته از داروهایی که تاییده سازمان غذا و داروی امریکا را گرفته‌اند، برجسته نشان داده شده‌اند)

Table 6. List of possible candidate drugs for differentially expressed genes in this study (some of the drugs that have been approved by the US Food and Drug Administration are boldy shown)

Score	PMIDs	Sources	Interaction Type	Drug	Gene
7	19169963	MyCancerGenome TdgClinical Trial ChemblInteractions TEND DrugBank TTD	inhibitor, binder, antibody	CANAKINUMAB	IL1B
5	23319019 23553601	ChemblInteractions DrugBank TTD	inhibitor, binder	RILONACEPT	IL1B
3	16122880	DrugBank TTD	antagonist, inhibitor	GALLIUM NITRATE	IL1B
3	None found	TdgClinicalTrial ChemblInteractions DrugBank	n/a	GEVOKIZUMAB	IL1B
3	2954576 1431212	NCI	n/a	ACITRETIN	IL1B
3	2350191 1331350	NCI	n/a	NIMUSTINE HYDROCHLORIDE	IL1B
2	None found	DrugBank TTD	Inhibitor	IBUDILAST	IL1B
2	8048000	NCI	n/a	PENTOXIFYLLINE	IL1B
2	1888883	NCI	n/a	NICARDIPINE	IL1B
2	8370344	NCI	n/a	PENTAMIDINE	IL1B
2	2788696	NCI	n/a	THYROGLOBULIN	IL1B
2	3260587	NCI	n/a	CEFACLOR	IL1B
2	8615653	NCI	n/a	PHORBOL MYRISTATE ACETATE	IL1B
2	16389574	NCI	n/a	RESVERATROL	IL1B
2	3260587	NCI	n/a	OFLOXACIN	IL1B
2	12773123	NCI	n/a	RALOXIFENE	IL1B
2	3497116	NCI	n/a	MAFOSFAMIDE	IL1B
2	2534682	NCI	n/a	ERYTHROMYCIN	IL1B
2	7523795	NCI	n/a	CYTARABINE	IL1B
2	16815316	NCI	n/a	LANSOPRAZOLE	IL1B
2	7589278	NCI	n/a	HYDROQUINONE	IL1B
2	11943316	NCI	n/a	FLUTICASONE PROPIONATE	IL1B
2	8077674	NCI	n/a	MELATONIN	IL1B
2	8724378	NCI	n/a	SODIUM beta-NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE	IL1B
2	2162889	NCI	n/a	HYDROCORTISONE	IL1B
2	9342951	NCI	n/a	LITHIUM	IL1B
2	9176529	NCI	n/a	BECLOMETHASONE DIPROPIONATE	IL1B
2	2686646	NCI	n/a	VERAPAMIL	IL1B
1	None found	TTD	Antibody	AMG-108	IL1B
1	None found	DrugBank	n/a	VX-702	IL1B
1	*None found*	TTD	n/a	CELASTROL	IL1B
1	None found	PharmGKB	n/a	MEDRONIC ACID	IL1B
1	None found	TdgClinicalTrial	n/a	DIACEREIN	IL1B
1	None found	PharmGKB	n/a	CLODRONATE DISODIUM	IL1B
1	None found	DrugBank	n/a	BELNACASAN	IL1B
1	None found	DrugBank	n/a	TALMAPIMOD	IL1B
1	None found	DrugBank	n/a	ETIPREDNOL DICLOACETATE	IL1B
3	17139284 17016423	DrugBank	n/a	CHEMBL384759	TGM3
3	17139284 17016423	DrugBank	n/a	CHEMBL1204628	TGM3
3	17139284 17016423	DrugBank	n/a	CHEMBL283807	TGM3
3	19261492 19757196	DrugBank	Inhibitor	NADROPARIN CALCIUM	FOS
2	2119237	NCI	n/a	THROMBIN	FOS
2	9661997	NCI	n/a	PILOCARPINE	FOS
2	11301212	NCI	n/a	BACLOFEN	FOS
2	9242432	NCI	n/a	PHORBOL MYRISTATE ACETATE	FOS
2	11336444	NCI	n/a	MAGNESIUM SULFATE	FOS
2	15773551	NCI	n/a	ANTIBIOTIC	FOS
2	2491749	NCI	n/a	INTERFERON BETA	FOS
2	8677009	NCI	n/a	INTERLEUKIN 1BETA	FOS
2	9588740	NCI	n/a	PACLITAXEL	FOS
2	16750178	NCI	n/a	ALCOHOL	FOS
2	1517537	NCI	n/a	BROMOACETIC ACID	FOS
2	8593857	NCI	n/a	DAIDZEIN	FOS
2	9187311	NCI	n/a	PHENOBARBITAL	FOS
2	11033072	NCI	n/a	URSODIOL	FOS
2	11680511	NCI	n/a	BROMOCRIPTINE	FOS
2	8301097	NCI	n/a	NIMODIPINE	FOS
2	10592235	DrugBank	n/a	CHEMBL377956	F3
2	10592235	DrugBank	n/a	CHEMBL209022	F3
2	10592235	DrugBank	n/a	CHEMBL204997	F3
2	9024781	NCI	n/a	DEXAMETHASONE	JUNB
2	10592235	DrugBank	n/a	CHOLIC ACID	COX7B
6	14596599 15155761 14646100 16580895 15689518	DrugBank	Cofactor	PYRIDOXAL PHOSPHATE	SDS
6	4377655 14688104 17139284 17016423 500557	DrugBank	n/a	SERINE	SDS
2	12210718	NCI	n/a	DEXAMETHASONE	SDS
1	*None found*	Guide To Pharmacology Interactions	Inhibitor	CHEMBL357080	GZMB
2	15262975	NCI	n/a	FLUOROURACIL	TPT1

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

در جدول ۶ مشاهده می‌شود که ژن IL1B– که از ژن‌های مهم درگیر در سیستم ایمنی به‌شمار می‌رود در تعامل دارویی با تعدادی از ترکیبات و داروها ظاهر شده است. این نشان می‌دهد محصولات نهایی این ژن در فعالیت‌های زیستی و در اشکال گوناگون درگیر هستند و دارای اهمیت دارویی ویژه‌ای هستند. در این میان داروی Canakinumab از بالاترین ارتباط با IL1B برخوردار بود. همچنین این دارو هیچگونه واکنش تداخلی با سایر اعضای خانواده اینترلوکین ۱ از جمله اینترلوکین ۱ آلفا را ندارد. فهرست داروهایی که در جدول ۶ ارائه شده است بر اساس کلیت ارتباط دارو با ژن و بدون فیلتراسیون همزمان و بر اساس استاندارد سازمان غذا و داروی آمریکا FDA، Antineoplastic و Immunotherapies بوده است. با اعمال فیلتراسیون همزمان، داروی Canakinumab در صدر داروهای کاندید برای ژن IL1B قرار گرفت اگرچه برای تعداد زیادی از ژن‌ها بعد از اعمال فیلتراسیون همزمان این سه معیار، هیچ دارویی کاندید نشد. این موضوع در جدول ۶ نشان داده شده است. در این پژوهش تفسیر نتایج حاصل برای ژنوم گاو می‌تواند صرفاً از دیدگاه هومولوژی ژن‌ها بین ژنوم انسان و گاو ارزشمند باشد به این علت که در پایگاه‌های داده‌ای مثل http://dgidb.org از حاشیه‌نویسی ژنومی انسان استفاده شده است. از نگاه تعامل ژن–دارو، داده‌ها و پژوهش‌های کافی در حوزه علوم دامی وجود ندارد و در پژوهش حاضر جهت بررسی این مهم از پایگاه داده http://www.ecodrug.org استفاده شد. نتایج برای داده‌های مورد بررسی در این پژوهش نشان داد هیچگونه دارویی که بر اساس ژنوم حیوانات مزرعه‌ای ساخته شده باشد با استفاده از مفهوم تعامل ژن–دارو وجود ندارد. این پژوهش می‌تواند آغازی جهت انجام پژوهش‌های آتی در این زمینه باشد. در جدول ۶ فهرستی از داروهای مختلف در تعامل با یک ژن ارایه شده است. این فهرست دارویی برای یک ژن می‌تواند در پژوهش‌های آینده جهت کنکاش دارویی ژن مورد نظر بیشتر مورد بررسی قرار گیرد. چرا که به‌عنوان مثال هنگامی‌که یک دارو به‌همراه داروی دیگری مصرف می‌شود، تأثیر یک دارو در بدن تغییر می‌کند. تعامل دارو–دارو می‌تواند باعث تأخیر، کاهش یا افزایش جذب هر دو دارو شود. همچنین می‌تواند باعث کاهش یا افزایش عملکرد هر یک از داروها یا هر دو دارو با هم شود و یا عوارض جانبی ایجاد کند.

از مجموع نتایج به‌دست آمده می‌توان استنباط کرد که این دو نژاد گاو به‌علت قرار گرفتن در شرایط محیطی متفاوت و فاصله تکاملی، توانسته‌اند با شرایط پیرامونی به‌خوبی هماهنگ شوند و این هماهنگی در سطح مولکولی از میزان بیان شدن ژن‌های اختصاصی یافت شده در این مطالعه قابل مشاهده است. از دلایل تأییدکننده می‌توان به درگیری بیشتر سامانه‌ی ایمنی در نژاد کلیستانی اشاره نمود که به‌علت قرار گرفتن در محیطی با آلودگی بالاتر در مقایسه با نژاد هلشتاین، میزان بیان پروتئین‌های دارای خاصیت باکتریوسایدی (Cathelicidin 1) در آن بالا رفته است و همچنین فعالیت سامانه‌ی ایمنی در آن افزایش یافته است. این افزایش تحریک و فعالیت سامانه‌ی ایمنی می‌تواند باعث بروز افزایش بیان ژن‌های تولیدکننده انرژی (ENSBTAG00000043545) در میتوکندری، افزایش فعالیت RNA پلیمراز II و فعال شدن مسیرهای سیگنال‌دهی افزاینده‌ی تقسیم سلولی (MAPK) (به‌ویژه در سلول‌های سامانه‌ی ایمنی) در نژاد کلیستانی شود. در نتیجه می‌توان برخی از ویژگی‌های ریخت‌شناسی متمایز کننده این دو نژاد را به چنین تغییراتی نسبت داد. نتایج این پژوهش می‌تواند در بالا بردن سامانه ایمنی در گاو هلشتاین– که انتظار داریم از نرم ایمنی کمتری نسبت به گاو کلیستانی برخوردار باشد، مورد استفاده قرار بگیرد. دارهای پیش‌بینی شده در این پژوهش بایستی در عمل مورد بررسی قرار بگیرند تا اهمیت درستی پیش‌بینی در این پژوهش به سنجه ترازو درآید.

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

منابع

1. Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
2. Bae, J.S., H.S. Cheong, L.H. Kim, S. NamGung, T.J. Park, J.Y. Chun and H.D. Shin. 2010. Identification of copy number variations and common deletion polymorphisms in cattle. BMC genomics, 11(1): 1.
3. Banabazi, M.H., M. Ghaderi-Zefrehei, I. Imumorin and S. Peters. 2013. Whole Transcriptome Value Index (WTVI): A Methodology for Integrating Functional Sequences from RNA-Seq Data into Animal Selection. 21st International Conference on Plant and Animal Genome, 12-16 pp, San Diego, United States.
4. Banabazi, M.H., S.R. Miraei ashtiani and M.N. Khalil. 2013. Regulation Gene expression, cell cycle in Saccharomyces by weight correlation. 3rd Iranian Agricultural Biotechnology Conference, .305-308 pp ,Mashhad, Iran.
5. Bolger, A.M., M. Lohse and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15): 2114-2120.
6. Brandwein, D. and Z. Wang. 2017. Interaction between Rho GTPases and 14-3-3 Proteins. International journal of molecular sciences, 18(10): 2148.
7. Cahais, V., P. Gayral, G. Tsagkogeorga, J. Melo-Ferreira, M. Ballenghien, L. Weinert and N. Galtier. 2012. Reference-free transcriptome assembly in non-model animals from next-generation sequencing data. Molecular ecology resources, 12(5): 834-845.
8. Esmaeelizadeh, A. and M.R. Mohammad Abadi. 2009. A molecular genome scan to map quantitative trait loci affecting bovine carcass weight. Agricultural Biotechnology Journal, 1(1): 117-130
9. Farooq, U., H.A. Samad, F. Sher, M. Asim and M.A. Khan. 2010. Cholistan and Cholistani breed of cattle. Pakistan Veterinary Journal, 30(2).
10. Flintoft, L. 2008. Transcriptomics: digging deep with RNA-Seq. Nature Reviews Genetics, 9(8): 568-568.
11. Fries, R. and A. Ruvinsky. 1999. The Genetics of Cattle. New York: CABI Publising.
12. Gan, Q., I. Chepelev, G. Wei, L. Tarayrah, K. Cui, K. Zhao and X. Chen. 2010. Dynamic regulation of alternative splicing and chromatin structure in Drosophila gonads revealed by RNA-seq. Cell research, 20(7): 763-783.
13. Guttman, M., M. Garber, J.Z. Levin, J. Donaghey, J. Robinson, X. Adiconis and J.L. Rinn. 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nature biotechnology, 28(5): 503-510.
14. Haas, B.J. and M.C. Zody. 2010. Advancing RNA-seq analysis. Nature biotechnology, 28(5): 421.
15. Hansen, P.J. 2004. Physiological and cellular adaptations of zebu cattle to thermal stress. Animal Reproduction Science, 82: 349-360.
16. https://bovinegenome.elsiklab.missouri.edu/
17. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5285380
18. Huang, W., A. Nadeem, B. Zhang, M. Babar, M. Soller and H. Khatib. 2012. Characterization and comparison of the leukocyte transcriptomes of three cattle breeds. PLoS One, 7(1): e30244.
19. Jafari Darehdor, AH., M.R. Mohammadabadi, A.K. Esmailizadeh, A. Riahi Madvar. 2016. Investigating expression of CIB4 gene in different tissues of Kermani Sheep using Real Time qPCR. Small Ruminant Research, 4: 119-132.
20. Levin, J., X. Adiconis, M. Yassour, D. Thompson, M. Guttman, M. Berger and A. Regev. 2010. Development and evaluation of RNA-Seq methods. Genome biology, 11(Suppl 1): 1-1.
21. Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics, 25(14): 1754-1760.
22. Marguerat, S. and J. Bähler. 2010. RNA-seq: from technology to biology. Cellular and Molecular Life Sciences, 67(4): 569-579.
23. Mohammadabadi, MR., A.H.D. Jafari and F. Bordbar. 2017. Molecular analysis of CIB4 gene and protein in Kermani sheep. Brazilian Journal of Medical and Biological Research, 50: e6177.
24. Mortazavi, A., B.A. Williams, K. McCue, L. Schaeffer and B. Wold. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature Methods, 5(7): 621-628.
25. Oshlack, A., M.D. Robinson and M.D. Young. 2010. From RNA-seq reads to differential expression results. Genome biol, 11(12): 220.
26. Pauerstein, P. 2011. RNA-Seq: Current Methods and Potential Applications.
27. Pennisi, E. 2012. ENCODE project writes eulogy for junk DNA. Genomics Science, 337(1159): 61.
28. Ponting, C.P., P.L. Oliver and W. Reik. 2009. Evolution and functions of long noncoding RNAs. Cell, 136(4): 629-641.
29. Rost, B. 1999. Twilight zone of protein sequence alignments. Protein Engineering, 12: 85–94.
30. Salimpour, M., S.R. Miraei-Ashtiani and M.H. Banabazi .2019. Differential gene expression of two bovine Bos taurus (Holstein) and Bos indicus (Cholistani) sub-species using RNA-Seq data. Iranian Journal of Animal science, 50(1): 47-55.
31. Saxena, A.K., D. Singh and G. Singh. 2009. Structural interaction between drug - DNA and protein- A novel approach for bioinformatics in medicine. Biomedical Research, 20(1): 28-34.
32. Snelling, W.M., R.A. Cushman, J.W. Keele, C. Maltecca, M.G. Thomas, M.R.S. Fortes and A. Reverter. 2013. BREEDING AND GENETICS SYMPOSIUM: Networks and pathways to guide genomic selection1–3. Journal of Animal Science, 91(2): 537-552.
33. Sultan, M., M.H. Schulz, H. Richard, A. Magen, A. Klingenhoff, M. Scherf and D. Schmidt. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science, 321(5891): 956-960.

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

34. Wang, K.C. and H.Y. Chang. 2011. Molecular mechanisms of long noncoding RNAs. Molecular cell, 43(6): 904-914.
35. Wilhelm, B.T. and J.R. Landry. 2009. RNA-Seq, quantitative measurement of expression through massively parallel RNA-sequencing. Methods, 48(3): 249-257.
36. Wang, K.C. and H.Y. Chang. 2011. Molecular mechanisms of long noncoding RNAs. Molecular cell, 43(6): 904-914.
37. Wilhelm, B.T. and J.R. Landry. 2009. RNA-Seq, quantitative measurement of expression through massively parallel RNA-sequencing. Methods, 48(3): 249-257.

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]

Investigating the Function of Predicted Proteins from RNA-Seq Data in Holstein and Cholistani Cattle Breeds

Mostafa Ghaderi-Zefrehei[1], Azadeh Torabi[2], Mohammad Esmaeilpour[3], Mina Salimpour[4] and Mohammad Hossein Banabazi[5]

1- Associate Professor, Department of Animal Science, Yasouj University, Yasouj, Iran
(Corresponding Author: mosmos741@yahoo.com, mghaderi@yu.ac.ir)
2- Assistant professor, Department of Agriculture, Payame Noor University, Tehran, Iran
3- Assistant Professor, Ahar Faculty of Agriculture and Natural Resources, University of Tabriz, Ahar, Iran
4- Ph.D candidate, Department of Animal Science, Faculty of Agriculture and Natural Resources, University of Guilan, Rasht, IRAN
5- Assistant Professor, Department of Biotechnology, Animal Science Research Institute of IRAN (ASRI), Agricultural Research ,Education & Extension Organization (AREEO), Karaj, IRAN
Received: December 18, 2019 Accepted: May 19, 2020

Abstract

This study was performed to determine the digital expression profile of different genes expressed in Holstein and Cholistani breeds as well as to evaluate the performance of predicted proteins derived from differentially expressed genes between these two breeds using RNA-Seq data. For this purpose, the whole mRNA sequence for a blood sample of American Holstein and Pakistani Cholistani cattle populations was obtained and by sequencing and locating RNA-Seq reads on the bovine reference genome and determining the digital expression profile, the differentially expressed genes were obtained. The results of this study showed that there were 24616 genes and 26716 isoforms on the transcriptome of these two breeds, out of which, 41 genes were identified with substantial and significant differential expression ($P < 0.000015$). It was also found that approximately one-third of genes whose functions is altered accros two breeds, encode hydrolase enzymes, five of the predicted proteins, FOS Fos proto-oncogene proteins, AP-1 transcription factor subunit, the vhl-elonginc-elonginb structure, transcription factor FosB/ JunD bZIP domain, T yeast U2AF complex pseudomonas aeruginosa transcriptional regulator PA2196 and zif268 zinc finger-dna complex that are involved in transcriptional regulation and RNA editing. Investigation of network interactions between predicted proteins from differentially expressed genes showed that predicted proteins are involved in different pathways such as TNF signaling pathway, Salmonella infection, MAPK signaling pathway and rheumatoid arthritis. Overall, these two breeds were found to be well-adapted to environmental conditions due to different environmental conditions and evolutionary distance and this coordination at the molecular level of the expression of specific genes was found in this study. One of the reasons supporting this is the greater role of the immune system in the Cholistani breed due to its higher exposition to contamination than the Holstein breed that led the expression of bactericidal proteins (Cathelicidin 1) was up regulaged, as a result, the activity of the immune system might be improved. The study also found that the IL1B gene had the highest degree of gene-drug interaction with Canakinumab drug.The results of this study indicate that in breed comparisons, a close look at the activity of proteins produced by different genes could better explain breed differences at the molecular level.

Keywords: Holstein cattle, RNA-Seq, Isoform, Protein interaction network

[Downloaded from rap.sanru.ac.ir on 2025-07-04]

[DOI: 10.52547/rap.11.28.121]