



تولیدات دامی

دوره ۲۱ ■ شماره ۲ ■ تابستان ۱۳۹۸

صفحه‌های ۱۸۰-۱۶۵

شناسایی LncRNA های مرتبط با رشد عضله سینه مرغ به وسیله روش RNA-seq

سید نادر آلبوشوکه^{۱*}، محمدرضا بختیاری زاده^۲

۱. استادیار، گروه علوم دامی، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان خوزستان، سازمان تحقیقات، آموزش و ترویج کشاورزی، اهواز، ایران.
۲. استادیار، گروه علوم دام و طیور، پردیس ابوریحان، دانشگاه تهران، پاکدشت، ایران.

تاریخ پذیرش مقاله: ۱۳۹۷/۱۲/۲۶

تاریخ دریافت مقاله: ۱۳۹۷/۱۰/۱۶

چکیده

هدف از این آزمایش بررسی و شناسایی RNA های غیررمزکننده بلند (lncRNAs) مرتبط با عضله اسکلتی مرغ بومی اصفهان و جوجه تجاری راس ۷۰۸ بود. پس از استخراج RNA از چهار نمونه عضله سینه در سن ۲۸ روزگی، توالی‌یابی جفتی با استفاده از پلاتفرم illumine Hiseq 2000 انجام شد. برای هم‌ترازی خوانش‌ها با ژنوم مرجع مرغ اهلی، از نرم‌افزار Hisat2 و جهت سرهم‌بندی رونوشت‌ها از بسته نرم‌افزاری Stringtie استفاده شد. در مجموع ۱۰۹۷ lncRNA شناسایی شد که از این تعداد ۹۲۵ ژن و رونوشت بین‌ژنی (ایتروژنی) و ۱۷۲ ژن و رونوشت ایتروژنی بودند. همچنین تعداد lncRNA های جدید شناسایی شده در دو گروه بین‌ژنی و ایتروژنی به ترتیب ۴۳۲ و ۱۲۸ بود. آنالیز بیان افتراقی ژن منجر به شناسایی ۱۹ ژن و ۲۰ رونوشت با تفاوت بیان معنادار بین دو گروه شد. بررسی جایگاه‌های ژنی lncRNA های با تفاوت معنادار، نشان داد که این ژن‌ها در مجاورت ۴۵ ژن رمزکننده پروتئینی قرار دارند. از این تعداد بیان پنج ژن رمزکننده پروتئینی (ژن *SCD* در جوجه تجاری و ژن‌های *GALNT15*، *USP7*، *KLHDC4* و *ASB1* در مرغ بومی) - که روند بیان آنها هم‌سو با بیان lncRNA های هم‌جوارشان بود - بین دو نژاد تفاوت معنادار داشتند. بررسی عملکردی این ژن‌ها نشان داد که همگی در رشد عضله اسکلتی مؤثر هستند. نتایج حاصل از این پژوهش نشان داد، lncRNA های شناسایی شده، احتمالاً پتانسیل تنظیم ژن‌های درگیر در رشد عضله اسکلتی را داشته و می‌توانند بخشی از تفاوت در سرعت رشد مشاهده شده بین دو نژاد مرغ مورد بررسی را توجیه نمایند.

کلیدواژه‌ها: رشد، عضله سینه، مرغ بومی اصفهان، lncRNA، RNA-seq.

مقدمه

اثر *linc-MYH*، *lincRNA-p21* و *gga-lnc-0181* بر رشد ماهیچه‌ای گزارش شده است.

نتایج برخی پژوهش‌های اخیر نشان داده‌اند، *lncRNA*ها در جوجه گوشتی تنظیم‌کننده رشد و نمو و تمایز سلول‌های ماهیچه اسکلتی می‌باشند. در مقایسه ساختار بیان ژنی عضله اسکلتی در جوجه آرورا کرز و یک نژاد مرغ بومی چینی، دو *lncRNA* خاص ماهیچه‌ای (*TCONS-00069348* و *TCONS-00064133*) گزارش شد [۱۷]. بررسی ساختار بیان ژنی عضله سینه جوجه تجاری راک سفید و مقایسه آن با جوجه ژینگوای (*Xinghua*) بومی چینی نشان داد که *lncRNA-six1* در فاصله ۴۳۲ نوکلئوتیدی بالادست ژن *six1* قرار دارد و به‌طور مستقیم میزان بیان این ژن را تنظیم می‌کند [۲].

بر خلاف پژوهش‌های متعدد صورت‌گرفته، تا به حال پژوهش اندکی در این خصوص روی نژادهای مرغ بومی ایران انجام شده است. بر همین اساس در پژوهش حاضر به‌منظور بررسی تفاوت بیان *lncRNA*های مرتبط با رشد و نمو عضله سینه از دو نژاد مرغ با سرعت رشد‌های متفاوت و فناوری توالی‌یابی *RNA* استفاده شد. همچنین با استفاده از روش‌های محاسباتی، *lncRNA*های جدیدی در مرغ شناسایی شد.

مواد و روش‌ها

این آزمایش در سالن مرغداری مرکز تحقیقات کشاورزی و منابع طبیعی صفی‌آباد دزفول انجام شد. به این منظور از جوجه راس ۷۰۸ به‌عنوان نژاد با سرعت رشد بالا و از مرغ بومی اصفهانی به‌عنوان نژاد با سرعت رشد پایین استفاده شد. ۲۰۰ قطعه جوجه از هر کدام از دو نژاد خریداری و تا پایان ۲۸ روزگی با شرایط یکسان و مشابه از نظر مدیریت پرورش (دما، رطوبت و نور) نگهداری شدند. دما در شروع دوره ۳۲ درجه سانتی‌گراد بود و پس

مرغ بومی اصفهان جزو مرغان دومنظوره تخمگذار-گوشتی است و در مقایسه با جوجه‌های گوشتی تجاری نظیر جوجه راس ۷۰۸ سرعت رشد پایین‌تری دارد. رشد و نمو عضله اسکلتی متأثر از ژنتیک و محیط است. پیشرفت‌های ژنتیکی ناشی از راهبردهای تلاقی و به‌گزینی، موجب ایجاد سویه‌هایی با سرعت رشد بالا شده است. خصوصیات زیستی نه تنها در سطح *DNA* بلکه در سطح *mRNA* قبل و پس از رونویسی به دقت کنترل و تنظیم می‌گردد. به‌طوری‌که بخش اعظم ژنوم (بیش از ۹۰ درصد) می‌تواند مورد رونویسی قرار گیرد. با این حال فقط بخش کوچکی از آن (یک تا دو درصد) *RNA*های رمزکننده پروتئین بوده و باقیمانده، شامل ده‌ها هزار مکان بین‌ژنی، مرتبط با *RNA*های غیررمزکننده پروتئینی می‌باشند.

*lncRNA*ها رونوشت‌هایی غیررمزکننده و بدون چهارچوب خوانش باز (*Open reading frame*)، با طولی بین ۲۰۰ تا افزون بر ۱۰۰ هزار نوکلئوتید بوده که اکثراً توسط آنزیم *RNA* پلی‌مراز II رونویسی شده و همچنین بیشتر آنها دارای دم پلی A و کلاهک (*Cap*) در سمت 5' خود می‌باشند. معمولاً سطح بیان *lncRNA* کمتر از سطح بیان *mRNA*ها بوده و برخی از آنها دارای بیان اختصاصی در یک بافت یا سلول می‌باشند. گزارش‌های متعددی مرتبط با نقش تنظیمی این ژن‌ها در مراحل مختلف زیستی وجود دارد. گزارش شده است که در موش و در طی ماهیچه‌زایی *lncMyoD* که پایین‌دست ژن *MyoD* قرار گرفته و توسط آن فعال می‌گردد، می‌تواند تمایز سلول‌های عضله اسکلتی و میوزن را تنظیم نماید [۵]. همچنین گزارش شده که *linc-YY1* می‌تواند تمایز میوزنی و بازسازی ماهیچه‌ای را از طریق اثر متقابل با فاکتور رونویسی *YY1* افزایش دهد [۲۵]. به همین ترتیب

تولیدات دامی

رفرنس ژنومی مرغ *Gallus_gallus* (نسخه ۵) و فایل حاشیه‌نویسی GTF (نسخه ۹۴) مربوط به مرغ خانگی از پایگاه داده به آدرس تارنمای (<http://asia.ensembl.org/info/data/ftp/index.html>) دریافت شد. جفت خوانش‌های مربوط به هر نمونه که فیلترهای کیفی را طی کرده بودند با استفاده از نرم‌افزار HISAT2 [۹] با ژنوم مرجع هم‌ردیف شدند.

خوانش‌های هم‌تراز شده برای هر نمونه سپس به‌طور جداگانه و با استفاده از نرم‌افزار Stringtie [۱۶] به رونوشت‌ها سرهم‌بندی شدند. پس از اتمام فرایند سرهم‌بندی خوانش‌ها و ایجاد رونوشت‌ها، با استفاده از ابزار Stringtie رونوشت‌های تمام نمونه‌ها ادغام شد تا مجموعه رونوشتی واحدی ایجاد شود.

برای شناسایی ماهیت رونوشت‌ها از نرم‌افزار Cuffcompare [۲۰] و فایل رفرنس‌دهی ژنوم مرغ (فایل GTF) استفاده شد. به این ترتیب رونوشت‌های با حروف i به‌عنوان ایترونی، u به‌عنوان بین‌ژنی و x به‌عنوان رفرنس‌دهی شده (ژن‌های شناخته‌نشده) گروه‌بندی شدند. رونوشت‌های باقیمانده با استفاده از معیارهای زیر حذف شدند:

- رونوشت‌های تک‌اگزونی که طولی بیش از ۱۰۰۰۰ نوکلئوتید داشتند حذف شدند.
- تمامی رونوشت‌های مربوط به ژن‌های رمزکننده، ژن‌های کاذب، پیش‌سازهای tRNA، miRNA، rRNA و snoRNA حذف شدند.
- رونوشت‌های دارای یک اگزون همپوشان با تکرارهای ساده حذف شدند (بر اساس فایل UCSC Repeat Masker).
- رونوشت‌های دارای حداقل یک مورد معنی‌دار در سطح ۰/۰۰۰۱ مقایسه‌شده با پایگاه‌های داده miRbase، (به‌وسیله ابزار BLASTx)، UniprotKB

از آن روزانه نیم درجه سانتی‌گراد کاهش داده شد تا در ۲۶ روزگی به دمای ۲۴ درجه سانتی‌گراد برسد. دسترسی به آب آشامیدنی و خوراک آزاد بود. در پایان ۲۸ روزگی از هر یک از دو گروه ۱۰ قطعه جوجه خروس به‌طور تصادفی انتخاب، توزین و کشتار شد. از هر جوجه از یک مکان مشابه در عضله بالادست سینه (Pectoralis major muscle) نمونه‌گیری صورت گرفت (دو گرم) و در کریوتیوب‌های استریل و فاقد RNase و DNase قرار داده شد. نمونه‌های گرفته‌شده بلافاصله به تانک ازت مایع منتقل و سپس تا زمان استخراج RNA در فریزر دمای -۸۰ درجه سانتی‌گراد نگهداری شدند.

با استفاده از کیت تریزول شرکت اینویترژن (Invitrogen, CA, USA) و بر اساس دستورالعمل شرکت (Cat, No. 15596-018)، از نمونه‌ها RNA استخراج گردید. خلوص RNA با استفاده از اسپکتروفوتومتری (NanoDrop™ 2000c, Thermo Scientific, USA) بررسی شد. نسبت جذب ۲۶۰ به ۲۸۰ نانومتر در نمونه‌های RNA در محدوده ۱/۸ تا دو بود. سه نمونه RNA مربوط به هر گروه به نسبت مساوی با یکدیگر مخلوط شده تا در کل و به‌ازای هر یک از دو گروه آزمایشی، دو تکرار از نمونه‌های ادغام‌شده (Pooled) به‌دست آید. چهار نمونه حاصل با دستگاه (Illumina Hiseq 2000, USA) و مطابق دستورالعمل مربوطه در شرکت BGI چین توالی‌یابی شد. قبل از انجام هم‌ترازی (Alignment) کیفیت توالی خوانش‌های خام با استفاده از نرم‌افزار Fastqc (<http://www.bioinformatics.babraham.ac.uk/projects>) بررسی شد. برای حذف خوانش‌های با کیفیت پایین و یا خوانش‌های خام حاوی توالی‌های مربوط به آداپتور و پرایمر و Poly-N از نرم‌افزار Trimmomatic (نسخه 0.36) استفاده شد. خوانش‌های پیرایش‌شده جهت تأیید کیفیت، دوباره با نرم‌افزار Fastqc بررسی شدند.

تولیدات دایمی

NONCODE database (نسخه ۵) به آدرس تارنمای <http://www.noncode.org/> دریافت شد. سپس از رویه Blastn جهت جستجوی هومولوژی LncRNAهای کاندیدا با LncRNAهای گزارش شده در مرغ اهلی استفاده شد. رونوشت‌های با همولوژی معنادار (۰/۰۰۰۵) با توالی‌های مورد بررسی، به‌عنوان LncRNAهای شناخته‌شده و باقیمانده رونوشت‌ها به‌عنوان LncRNAهای جدید گزارش شدند.

بیان رونوشت‌ها با استفاده از نرم‌افزار Stringtie اندازه‌گیری و به‌صورت FPKM (Fragments Per Kilobase of exon per Million mapped reads) نمایش داده شد. به‌طوری‌که خطاهای ناشی از تفاوت عمق خوانش و تفاوت اندازه ژنی رونوشت‌ها در نظر گرفته شد. برای شناسایی ژن‌های با بیان متفاوت بین مرغ بومی و جوجه تجاری از برنامه Cuffdiff استفاده شد. رونوشت‌های با تفاوت بیان معنی‌دار شده در سطح (FDR \leq ۰/۰۵) و مقدار بیان چند برابری بیشتر از ۰/۸ به‌عنوان ژن‌های با تفاوت بیان معنی‌دار معرفی شدند.

پژوهش‌های اخیر نشان داده‌اند LncRNAها می‌توانند ژن‌های رمزکننده پروتئینی مجاور خود (در فاصله ۱۰ کیلوبازی بالادست یا پایین دست) را تنظیم کنند [۱۵]. بر همین اساس ژن‌های مجاور (حداکثر تا فاصله ۱۰ هزار بازی) LncRNAهای شناسایی‌شده در پژوهش حاضر بررسی شدند. بررسی عملکردی ژن‌های کاندیدای هدف از طریق آنالیز عملکردی و آنالیز مسیرهای زیستی ژن‌ها (KEGG) با استفاده از نرم‌افزار DAVID [۳] انجام شد. نتایج بررسی عملکردی برای فرایند زیستی با سطح معنی‌دار تصحیح شده برای خطای چند آزمونی با روش FDR (P-adjusted \leq 0.05) مورد بررسی قرار گرفت.

به‌منظور بررسی وجود توالی‌های همولوگ در دیگر گونه‌ها برای LncRNAهای شناسایی‌شده در پژوهش حاضر

(نسخه ۲۱ و با ابزار BLASTn) و Rfam (به‌وسیله ابزار BLASTn) حذف شدند.

- رونوشت‌های دارای دامین‌های پروتئینی شناخته‌شده در پایگاه داده Pfam، با استفاده از hmmscan مربوط به بسته نرم‌افزاری HMMER (نسخه 3.1b) حذف شدند.

- توان رمزکنندگی پروتئینی رونوشت‌ها با پنج ابزار مختلف شامل CPC2 [۸] (با نمره بالاتر از نیم)، CNCI [۱۹] (با نمره بالاتر از صفر)، CPAT [۲۱] (با اسکور بالاتر از صفر) و FEEInc [۲۳] (با فراسنجه‌های پیش‌فرض) اندازه‌گیری شده و آن دسته از رونوشت‌هایی که توسط حداقل سه ابزار از پنج ابزار فوق، واجد پتانسیل رمزکنندگی پروتئینی تشخیص داده شدند، از مجموعه داده حذف شدند.

- رونوشت‌های بین‌ژنی که در فاصله کمتر از هزار نوکلئوتیدی از ژن‌های رمزکننده پروتئینی شناخته شده قرار داشتند، حذف گردیدند.

- با استفاده از ابزار TransDecoder (آدرس تارنمای <https://transdecoder.github.io>) چهارچوب خوانش باز (ORF) تخمین زده شد و رونوشت‌های دارای ORF \leq 300 برای مرحله بعدی آنالیز در نظر گرفته شدند.

- رونوشت‌های با مقدار (یک \geq FPKM) که در حداقل دو نمونه بیان شده بودند حفظ شدند.

در نهایت مجموعه رونوشت‌های باقیمانده به‌عنوان LncRNAهای جدید و کاندیدا در دو گروه بین‌ژنی و ایترونی در نظر گرفته شدند.

با توجه به این‌که تا به حال تعداد زیادی از LncRNAها در ژنوم مرغ در بانک اطلاعاتی NONCODE گزارش شده ولی در فایل رفرنس‌دهی ژنوم وارد نشده است، LncRNAهای مرغ خانگی از پایگاه داده

تولیدات دایمی

بیش از ۱۳۱ میلیون خوانش جفتی ۱۵۰ نوکلئوتیدی ایجاد نمود که پس از کنترل کیفیت ۵۱۴۷۷۱۴۳ خوانش بدون آلودگی حاصل شد. میانگین نمره کیفی ۳۸ و میانگین درصد CG بین ۵۲ تا ۵۴ بود. این موضوع نشان دهنده کیفیت بالای خوانش‌ها و مناسب بودن داده‌ها جهت آنالیزهای بعدی می‌باشد. با استفاده از نرم‌افزار HISAT2 بیش از ۷۲ درصد خوانش‌ها با ژنوم مرغ هم‌ردیف شدند. در مرحله بعدی و با استفاده از نرم‌افزارهای Stringtie و Cuffcompare و فایل حاشیه نویسی مرغ اهلی، تجزیه و تحلیل داده‌های RNA-seq نشان داد در مجموع ۲۰۴۹۹ ژن و رونوشت، ۳۷۶۶۳ ژن و رونوشت شناخته شده و ۹۳۴۳ ژن و رونوشت جدید در نمونه‌های ماهیچه سینه بیان شده است.

برای شناسایی LncRNA های مرغ، معیارهای متعددی در نظر گرفته شد تا رونوشت‌های با پتانسیل بالای کدکنندگی حذف گردند. پس از اعمال معیارهای بیان‌شده در بخش مواد و روش‌ها، ۹۴۰ ژن و ۱۷۰۶ رونوشت شناسایی شد. از این تعداد ۷۳۲ ژن و ۱۴۶۰ رونوشت در جوجه تجاری و ۷۵۲ ژن و ۱۵۳۶ رونوشت در مرغ بومی و ۶۳۸ ژن و ۱۲۹۲ رونوشت مشابه نیز در هر دو گروه بیان شده بودند. همچنین ۹۷ ژن و ۱۶۷ رونوشت فقط در جوجه تجاری و ۱۲۰ ژن و ۲۴۴ رونوشت نیز فقط در مرغ بومی بیان شده بودند.

با مقایسه ۲۶۴۶ ژن و رونوشت با LncRNA های گزارش شده در فایل رفرنس‌دهی ژنوم مرغ (بر اساس بانک اطلاعاتی Ensembl)، ۴۱۹ ژن و ۱۱۳۱ رونوشت مشابه یافت شد که با پیشنهاد Annotated نام‌گذاری شدند. باقیمانده LncRNA ها با LncRNA های گزارش شده مرتبط با مرغ در بانک اطلاعاتی Noncode بلست شدند که منجر به شناسایی ۲۴۷ ژن و ۲۸۹ رونوشت گردید. این گروه به‌عنوان شناخته‌شده (Known) و ۲۷۵ ژن و ۲۸۵

از نرم‌افزار BLASTn (نسخه 2.2.32+) استفاده شد. به این منظور ابتدا رونوشت‌های IncRNAs انسان (GRCh38) و موش (GRCm38) از پایگاه داده Ensembl (نسخه ۹۴) دریافت شد و رونوشت‌هایی که دارای مشابهت معنادار با این توالی‌ها بودند، به‌عنوان رونوشت‌های حفاظت‌شده معرفی شدند. معیارهای مورد استفاده برای مقایسه، داشتن حداقل ۵۰ درصد پوشش هم‌ترازی (Querycoverage) با توالی پایگاه داده و حداقل ۷۰ درصد یکسانی (Identity percentage) در توالی بود.

هدف از انجام این آنالیز شناسایی LncRNA هایی بود که دارای ژن رمزکننده مجاور (بالادست و پایین دست) یکسان در گونه‌های دیگر بودند. گزارش شده است که علی‌رغم حفاظت‌شدگی پایین توالی‌ها در دیگر گونه‌ها، ژن‌های بالادست و پایین‌دست آن‌ها در گونه‌های مختلف حفاظت‌شده می‌باشد. به این منظور ابتدا ژن‌های مجاور Noncode های موجود در بانک اطلاعاتی (نسخه ۵) برای گونه‌های انسان بر اساس فایل رفرنس‌دهی (نسخه ۹۴) بانک اطلاعاتی Ensembl ژنوم GRCh38 و گاو بر اساس فایل رفرنس‌دهی (نسخه ۹۴) بانک اطلاعاتی Ensembl ژنوم UMD3.1 استخراج شد. چنانچه هر کدام از LncRNA های شناسایی شده در این تحقیق، با LncRNA های انسانی یا گاوی از نظر ترتیب و جهت قرارگیری دو ژن رمزکننده هم‌جوارش شباهت داشتند، به‌عنوان ژن‌های با حفاظت هم‌جواری گزارش شد. در این خصوص از نظر فاصله مابین LncRNA ها و ژن‌های رمزکننده پروتئینی هم‌جوار محدودیتی در نظر گرفته نشد. با این حال بیشترین فاصله مشاهده شده ۴۲۱۸۴ نوکلئوتید بود.

نتایج و بحث

نتایج توالی‌یابی RNA در مجموع و به ازای هر چهار نمونه

تولیدات دایمی

میانگین میزان CG نواحی بین‌ژنی بیشتر از نواحی ایترونی می‌باشند. توزیع lncRNAها روی کروموزوم‌های مختلف در شکل ۱ نمایش داده شده است.

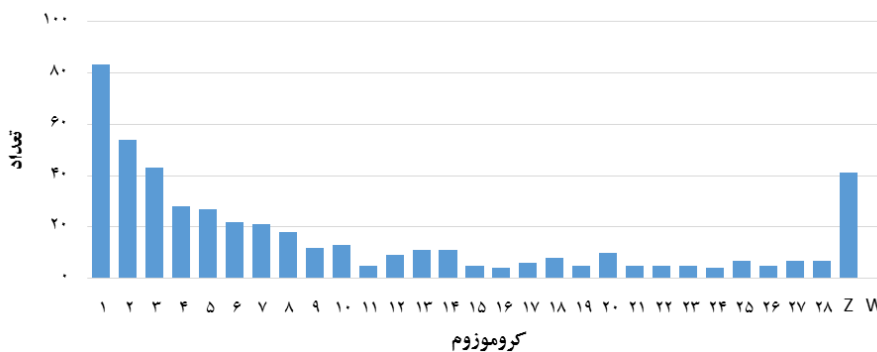
توزیع lncRNAهای شناسایی شده روی ژنوم نشان داد که آن‌ها در تمام کروموزوم‌های مرغ وجود داشته و بیشترین تعداد آن‌ها بر روی کروموزوم شماره یک قرار داشتند. بالاترین میزان بیان lncRNA در هر دو نژاد مورد بررسی متعلق به سه lncRNA بین‌ژنی شناخته شده (MSTRG.11691، MSTRG.10160 و MSTRG.1563) بود. همچنین در گروه lncRNAهای ایترونی نیز MSTRG.7711 بالاترین میزان بیان را در هر دو نژاد نشان داد. بیان بالای این ژن‌ها احتمالاً مؤید اهمیت و نقش آن‌ها در ماهیچه اسکلتی مرغ است.

رونوشت باقیمانده که با هیچ‌کدام از فایل‌های مرجع پایگاه‌های داده شناسایی نشده بودند به‌عنوان جدید (New) نامگذاری شدند (جدول ۱).

میانگین طول lncRNAها در نمونه‌های بررسی شده ۱۲۵۵/۶ نوکلئوتید بود. این مقدار بزرگ‌تر از میانگین طول lncRNA در انسان (۱۰۰۰ نوکلئوتید) و موش (۵۵۰ نوکلئوتید) می‌باشد. وجود مقادیر بالاتر از میزان CG در نواحی اگزونی نسبت به ایترونی، مشخصه توالی‌های رمزکننده بوده و به معنای رونویسی، اسپلایسینگ و ترجمه کارآمدتر این نواحی تفسیر می‌شود. مشابه چنین روندی در lncRNAهای نواحی اگزونی نسبت به ایترونی مشاهده شده است [۶]. نتایج این پژوهش نیز نشان داد، از نظر تعداد ژن و رونوشت و نیز تعداد اگزون و همین‌طور

جدول ۱. خصوصیات رونوشت‌ها و ژن‌های lncRNA در نمونه بافت ماهیچه سینه مرغ خانگی

خصوصیت	ایترونی جدید	بین‌ژنی جدید	ایترونی شناخته شده	بین‌ژنی شناخته شده
تعداد ژن	۶۳	۲۱۲	۲۲	۲۲۵
تعداد رونوشت	۶۵	۲۲۰	۲۲	۲۶۷
میانگین طول رونوشت	۹۹۸/۳۷	۹۵۱/۰۸	۱۳۱۶/۸۶	۱۷۵۵/۹۰
بیشینه و کمینه طول رونوشت	۲۱۸ - ۵۵۶۵	۲۰۶ - ۵۹۰۵	۲۲۱ - ۷۱۲۹	۲۰۱ - ۸۴۴۷
میانگین تعداد اگزون	۱/۳۸۴۶	۱/۳۹۵۵	۱/۷۷۲۷	۲/۸۷۶۴
بیشینه و کمینه طول اگزون	۵ - ۱	۵ - ۱	۴ - ۱	۱۶ - ۱
میانگین میزان گوانین - سیتوزین (CG)	۰/۴۳۱۴	۰/۴۳۹۰۱۶	۰/۴۹۳۸۹۹	۰/۵۰۴۷۱۶



شکل ۱. توزیع کروموزومی lncRNAها در ژنوم مرغ در این پژوهش

تولیات دامی

منجر به افزایش شکست پروتئین ها و افزایش گلوکونئوزن می شود. پژوهش های انجام شده در مرغ خانگی همچنین نشان داده اند که تیمار گلوکوکورتیکوئیدی می تواند منجر به مقاومت انسولینی و یا به عبارتی هایپرگلیسمی و هایپرانسولینی شوند. مقاومت به انسولین یکی دیگر از مسیرهای معنی دار شده در مرغ بومی بود و این موضوع می تواند به احتمال زیاد نشان دهنده وجود استرس در مرغ بومی در مقایسه با جوجه تجاری باشد. در طول آزمایش نیز فنوتیپ پر جنب و جوش جوجه های بومی در مقایسه با ظاهر آرام و کم تحرک جوجه تجاری موید این موضوع بود. به هر حال عارضه مقاومت به انسولینی همبستگی بالا و مثبتی را با کاهش رشد و نمو عضله اسکلتی طیور نشان داده است [۱۳].

پژوهش های متعددی پیشنهاد کرده اند، LncRNA ها بیان ژن های هم جوار خود را تنظیم می نمایند. نتایج حاصل از بررسی ژن های مجاور LncRNA های بین ژنی، منجر به شناسایی ۹۶۰ ژن هدف شد. از این تعداد ۴۷۹ ژن بالادست و ۴۸۱ ژن پایین دست ژن ها قرار داشت. در جدول ۲ مهم ترین مسیرهای عملکردی معنادار مربوط به این ژن ها ارائه شده است.

در مرغ بومی بالاترین سطح معنی داری مسیرهای عملکردی مربوط به مسیر پیام رسانی انسولین (Insulin signaling pathway) بود. این مسیر نقش اصلی را در هموستاز گلوکز و انرژی و نیز رشد ایفا می کند. استرس با تأثیر بر محور آدرنال هیپوفیز هیپوتالاموسی می تواند منجر به ترشح سطوح بالای گلوکوکورتیکوئیدی شده و این

جدول ۲. مهم ترین مسیرهای غنی شده مرتبط با ژن های هدف LncRNA ها در مرغ بومی و جوجه تجاری

P-adjust	ژن های هدف در مجاور LncRNA های بین ژنی	نام مسیر
		مسیرهای غنی شده در مرغ بومی
1.13E-05	ENSGALG0000006670, ENSGALG0000002656, ENSGALG0000009435, ENSGALG0000004998, ENSGALG0000008081, ENSGALG0000007636, ENSGALG0000004959, ENSGALG00000011295, ENSGALG00000012755, ENSGALG0000007307,	Insulin signaling pathway
0.002434	ENSGALG0000004998, ENSGALG0000008081, ENSGALG0000007636, ENSGALG0000004959, ENSGALG00000010638, ENSGALG0000002656,	FoxO signaling pathway
0.002635	ENSGALG0000004998, ENSGALG0000002556, ENSGALG00000016781, ENSGALG00000014645, ENSGALG00000015596	MAPK signaling pathway
0.005197	ENSGALG0000006670, ENSGALG0000009435, ENSGALG0000008081, ENSGALG0000007636, ENSGALG0000004959	Insulin resistance
0.043583	ENSGALG00000024485, ENSGALG00000011673, ENSGALG0000000579	SNARE interactions in vesicular transport
		مسیرهای غنی شده در جوجه تجاری
4.3E-07	ENSGALG0000003814, ENSGALG0000002229, ENSGALG0000005733, ENSGALG00000017283, ENSGALG00000011109, ENSGALG0000000220, ENSGALG00000015596, ENSGALG00000017242	Wnt signaling pathway
0.000609	ENSGALG00000012015, ENSGALG0000001371, ENSGALG0000006757, ENSGALG00000012010	Glycosaminoglycan biosynthesis - heparan sulfate / heparin
0.002076	ENSGALG00000011109, ENSGALG00000017186, ENSGALG00000015596, ENSGALG0000002786	NOD-like receptor signaling pathway

بالاترین سطح معنی‌داری را داشت. مسیر Wnt بر رشد هایپرتروفی و تمایز میوتیوب‌ها تأثیر مثبت دارد. مشخص شده است که این مسیر می‌تواند رشد و نمو ماهیچه اسکلتی را در مراحل مختلف زندگی تحت تأثیر قرار دهد. مسیر گلایکوزآمینوگلیکان-هپارین سولفات نیز با داشتن چهار ژن معنادار شد. پروتئوگلیکان‌های سولفات هپاران در ساختمان بافت همبند، کولاژن‌های ماتریس خارج سلولی و غشای پایه ماهیچه‌ای شرکت داشته و می‌تواند با باندشدن با فاکتورهای رشد فیروبلستی و لیگاند‌های پروتئینی در تنظیم طیف وسیعی از فرایندهای بیولوژیکی و از جمله رشد و نمو، مدل‌دهی مجدد ماتریس خارج سلولی و رگزایی (آنژیوژنز) شرکت نمایند. مسیر پیام‌رسان گیرنده‌های شبه NOD مرتبط با فرایندهای ایمنی بوده و در واکنش به تنش حرارتی نیز فعال می‌گردد. همچنین گزارش شده بیان این مسیر با فنوتیپ چاقی همبستگی بالایی داشته و در آدیپوسیت‌های انسانی فعال شدن این مسیر مرتبط با التهاب ناشی از افزایش چربی بافتی عنوان شده است [۲۴].

در مجموع بیان ۱۹ ژن و ۲۰ رونوشت بین دو گروه مرغ با هم تفاوت معنادار داشتند ($FDR \leq 0.05$). از این تعداد هشت ژن و شش رونوشت در جوجه تجاری و ۱۳ ژن و ۱۲ رونوشت در مرغ بومی بیش بیان شده بودند (شکل ۲ و جدول ۳).

در مرحله بعدی lncRNAهای با تفاوت بیان معنادار بین دو نژاد مورد بررسی قرار گرفته و مشخص شد در نزدیکی این ۳۹ lncRNA، ۴۵ ژن رمزکننده پروتئینی قرار دارند. از این تعداد ۳۹ ژن در فاصله کمتر از ۱۰ هزار نوکلئوتید و مابقی در فاصله کمتر از صد هزار نوکلئوتیدی lncRNAها قرار داشتند. در میان این ۴۵ ژن، بیان پنج ژن بین دو نژاد اختلاف معنی‌دار داشت. روند بیان این پنج ژن هم‌سو با بیان lncRNAهای مربوطه به آنها بود.

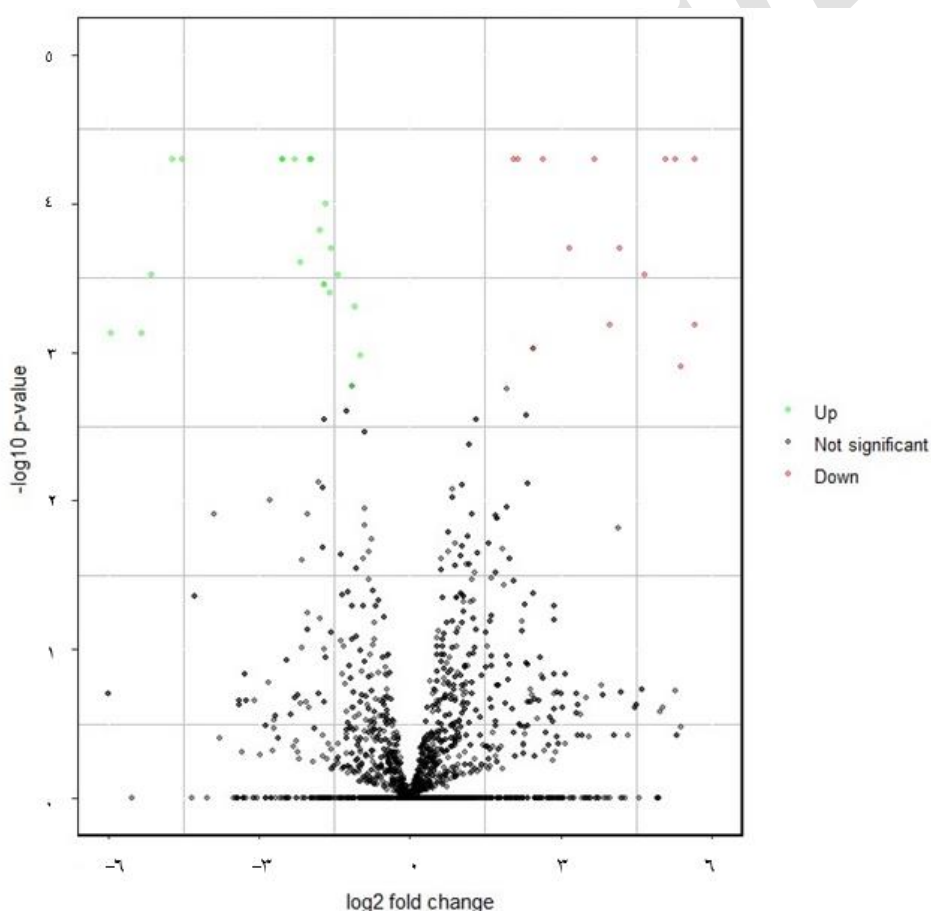
یکی دیگر از مسیرهای مرتبط با این موضوع مسیرهای پیامی Foxo بود. فعال شدن این مسیر در مرغ بومی می‌تواند منجر به افزایش بیان لیگازهای یوبیکویتینی E3 شده که آتروفی ماهیچه‌ای را در پی دارد. سامانه پروتئوزوم-یوبیکویتین عمده مسیر شکست غریلوزوم پروتئین‌های سلولی می‌باشد. در این مسیر یوبیکویتین به‌صورت کووالانسی به پروتئین‌های هدف متصل شده و در نتیجه به پروتئوزوم‌های 26S اجازه شکست پروتئینی و آتروفی سلولی را می‌دهد. یوبیکویتیناسیون شامل یکسری آنزیم‌های چندتایی، آنزیم فعال‌کننده یوبیکوتین (E1)، آنزیم اتصال‌دهنده (E2) و لیگاز یوبیکویتینی (E3) می‌باشد. لیگازهای E3 با شکست پروتئین‌ها و به‌خصوص اجزای فیبرهای ماهیچه‌ای موجب تحلیل رفتن ماهیچه می‌شوند. تاکنون تعداد زیادی از لیگازهای یوبیکویتینی E3 شناسایی شده‌اند که هر کدام به‌صورت اختصاصی پروتئین‌های خاصی را هدف قرار می‌دهند. برای شروع آتروفی در عضله اسکلتی، وجود FoxO3 شرط لازم و کافی به‌شمار آمده و افزایش بیان آن موجب افزایش رونویسی مجموعه‌ای از ژن‌های رمزکننده لیگازهای E3 همانند FBXO32 و FBXO30 می‌شود. عمل برخی از لیگازهای یوبیکویتینی E3 به شکل کمپلکس‌های چندتایی انجام شده و پروتئین‌های Cullin یکی از زیرواحدهای فعال‌کننده آن می‌باشند. بر همین اساس و با توجه به افزایش بیان ژن‌های FoxO3، FBXO32 و FBXO30 در مرغ بومی در مقایسه با جوجه تجاری می‌توان نتیجه‌گیری نمود که به احتمال زیاد این موضوع یکی از علل پایین‌تر بودن میزان رشد مرغ بومی در مقایسه با مرغ‌های تجاری است.

در جوجه تجاری تقریباً هر سه مسیر زیستی معنادار، با رشد و نمو ماهیچه اسکلتی مرتبط بودند. مهم‌ترین آن‌ها مسیر پیام‌رسان Wnt بود که با داشتن هشت ژن

تولیدات دامی

(peroxisome proliferator-activated receptors) شرکت می‌نماید. گزارش‌ها حاکی از آن است که افزایش بیان *SCD* سبب افزایش سوخت‌وساز چربی‌ها شده و ممانعت در بیان آن و یا بروز جهش‌های غیرفعال‌کننده، سبب کاهش وزن و لاغری می‌گردد [۱۴]. بنابراین با توجه به این‌که سوخت‌وساز چربی‌ها در جوجه‌های با سرعت رشد بالا، بیشتر از جوجه‌های با سرعت رشد پایین است، میزان بیان این ژن و بیان *IncRNA* آن احتمالاً مرتبط به یکدیگر می‌باشد.

نخستین ژن *SCD* بود که در فاصله ۲۰۸ نوکلئوتیدی پایین‌دست *IncRNA* مربوطه‌اش (*ENSGALG00000043779*) قرار دارد و هر دو در جوجه تجاری بیان بالاتر نسبت به نژاد بومی نشان دادند. گزارش شده ژن *SCD* در هموستاز انرژی، سوخت‌وساز لیپیدها و کنترل وزن بدن مرغ نقش دارد. برخی از ژن‌ها نظیر *SCD* دارای نقش‌های چندگانه هستند. این ژن هم در زیست تولید اسیدهای چرب غیراشباع و هم در سوخت‌وساز اسیدهای چرب و پیام‌دهی *PPAR*



شکل ۲. نمایش پراکندگی بیان *lncRNA* ها با استفاده از نمودار *Volcano*. مقدار بیان چند برابری (*Fold Change*) در مقیاس منفی لگاریتم پایه دو بعنوان معیار عرضی نمودار و سطح معنی‌داری (*P-value*) مربوط به تفاوت بیان ژن در مقیاس منفی لگاریتم پایه ۱۰ بعنوان معیار عمودی نمایش داده شده است. *lncRNA* های با تفاوت بیان معنی‌دار در در جوجه تجاری به رنگ قرمز و در مرغ بومی به رنگ سبز و ایزوفرم‌های با بیان مشابه به رنگ مشکی نمایش داده شده‌اند.

تولیدات دائمی

جدول ۳. خصوصیات lncRNAهای با تفاوت بیان معنی‌دار در دو مرغ بومی و تجاری

<i>P-adjust</i>	<i>-log2 FC</i>	گروه‌بندی	نام ژن / رونوشت
0.01199	7.476	ژن‌های آنوتیت‌شده	<i>ENSGALG00000034160</i>
0.02802	2.457	ژن‌های آنوتیت‌شده	<i>ENSGALG00000043779</i>
0.02566	19.655	رونوشت‌های آنوتیت‌شده	<i>ENSGALT00000051744</i>
0.00593	-2.285	رونوشت‌های آنوتیت‌شده	<i>ENSGALT00000056849</i>
0.01069	-1.667	رونوشت‌های آنوتیت‌شده	<i>ENSGALT00000066371</i>
0.03790	-1.078	رونوشت‌های آنوتیت‌شده	<i>ENSGALT00000071755</i>
0.00593	2.650	رونوشت‌های آنوتیت‌شده	<i>ENSGALT00000083946</i>
0.00593	-1.957	رونوشت‌های آنوتیت‌شده	<i>ENSGALT00000087952</i>
0.00277	-2.526	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.11671</i>
0.00277	-2.525	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.6274</i>
0.01049	-2.173	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.14159</i>
0.00277	-1.983	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.1275</i>
0.01369	-1.698	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.13905</i>
0.01199	-1.434	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.13282</i>
0.04273	-1.158	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.3636</i>
0.03026	-0.968	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.4044</i>
0.00278	8.163	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.12126</i>
0.02165	7.422	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.11722</i>
0.03429	8.103	ژن‌های بین‌ژنی شناخته‌شده	<i>MSTRG.11740</i>
0.00593	-2.525	رونوشت‌های بین‌ژنی شناخته‌شده	<i>MSTRG.11671.1</i>
0.00593	-1.982	رونوشت‌های بین‌ژنی شناخته‌شده	<i>MSTRG.1275.1</i>
0.01504	-1.781	رونوشت‌های بین‌ژنی شناخته‌شده	<i>MSTRG.6274.1</i>
0.02885	-1.698	رونوشت‌های بین‌ژنی شناخته‌شده	<i>MSTRG.13905.1</i>
0.04519	20.228	رونوشت‌های بین‌ژنی شناخته‌شده	<i>MSTRG.11722.1</i>
0.00593	20.969	رونوشت‌های بین‌ژنی شناخته‌شده	<i>MSTRG.12126.1</i>
0.008654	-1.571	ژن‌های بین‌ژنی جدید	<i>MSTRG.2323</i>
0.01199	-1.434	ژن‌های بین‌ژنی جدید	<i>MSTRG.13282</i>
0.00277	2.065	ژن‌های بین‌ژنی جدید	<i>MSTRG.14163</i>
0.00277	-6.785	ژن‌های بین‌ژنی جدید	<i>MSTRG.5539</i>
0.02367	-4.598	ژن‌های بین‌ژنی جدید	<i>MSTRG.5531</i>
0.01199	-4.431	ژن‌های بین‌ژنی جدید	<i>MSTRG.5414</i>
0.00593	4.891	ژن‌های بین‌ژنی جدید	<i>MSTRG.5721</i>
0.01881	4.340	ژن‌های بین‌ژنی جدید	<i>MSTRG.555</i>
0.02566	-4.431	رونوشت‌های بین‌ژنی جدید	<i>MSTRG.5414.1</i>
0.04970	-4.598	رونوشت‌های بین‌ژنی جدید	<i>MSTRG.5531.1</i>
0.00593	2.141	رونوشت‌های بین‌ژنی جدید	<i>MSTRG.14163.1</i>
0.00593	-6.785	رونوشت‌های بین‌ژنی جدید	<i>MSTRG.5539.1</i>
0.03215	-1.581	رونوشت‌های بین‌ژنی جدید	<i>MSTRG.2323.1</i>
0.00593	4.891	رونوشت‌های بین‌ژنی جدید	<i>MSTRG.5721.1</i>

تولیدات دائمی

قسمت مرکزی سیستم پروتئوزوم یوبیکوتین E3 قرار می‌گیرد. معلوم شده است که Cullin-3 در ناحیه z دیسک سارکومر میوفیبریل‌های ماهیچه اسکلتی قرار دارد و در شکست پروتئین‌های ماهیچه شرکت می‌نماید [۱۱]. بنابراین افزایش بیان *CUL3* مرتبط با افزایش بیان ژن *KLHDC5* بوده و این امکان وجود دارد که بیان *KLHDC5* به واسطه بیان IncRNA هم‌جوارش تنظیم شود.

گزارش شده است که پروتئین‌های خانواده ASB با کمک پروتئین‌های Cullin و Rbx نیز می‌توانند موجب شکل‌گیری و فعال‌شدن لیگازهای یوبیکوتینی E3 گردند [۱۰]. در واقع ژن *ASB1* یکی دیگر از ژن‌های هم‌جوار IncRNA های با تفاوت معنی‌دار بین دو گروه بومی و تجاری بود که در فاصله ۲۷۶۲ نوکلئوتیدی پایین دست MSTRG.1222 قرار داشت. در مرغ بومی بیان MSTRG.1222 ($\log_2FC = -1.25863$) و ژن *ASB1* ($\log_2FC = -0.85925$) افزایش معنی‌داری نسبت به جوجه تجاری داشت. این موضوع مؤید آن است که ژن *ASB1* می‌تواند از طریق فعال‌سازی و ساخت لیگازهای یوبیکوتینی، در فرایند آتروفی ماهیچه‌ای دخالت داشته باشد. بر همین اساس در مرغ بومی افزایش بیان ژن‌های *Cul3*، *KLHDC4* و *ASB1* احتمالاً مرتبط با آتروفی و کاهش رشد ناشی از افزایش بیان ژن‌های *FoxO3*، *FBXO30* و *FBXO32* بوده است.

یکی دیگر از ژن‌های مرتبط با این موضوع، *USP7* است که روی کروموزوم ۱۴ و در فاصله ۵۹۹ نوکلئوتیدی پایین دست MSTRG.3636 قرار دارد. این ژن و IncRNA مربوط به آن در مرغ بومی دارای افزایش بیان معنی‌داری نسبت به جوجه تجاری بودند. پروتئین محصول *USP7*، جزو پروتئازهای خاص یوبیکوتین بوده و سبب جداسازی یوبیکوتین از پیش‌ماده شده و به این ترتیب

دومین ژن *GALNT15* بود که در فاصله ۱۲۸۱۷ نوکلئوتیدی پایین دست MSTRG.5539 قرار دارد. این ژن و IncRNA مربوط به آن در مرغ بومی دارای بیان بالاتر و معنادار نسبت به جوجه تجاری بودند. ژن *GALNT15* در فرایندهای گلیکوزیلاسیون پروتئین‌ها دخالت دارد. در بسیاری از جنبه‌های رشد و نمو، تغییرات ژنتیکی و اپی‌ژنتیکی در ژن‌های رمزکننده گلیکوزیل ترانسفرازها می‌تواند منجر به تغییرات پاتولوژیکی گردد [۷]. قندها با اتصال O و N دو شکل عمده از گلیکوزیلاسیون پروتئین‌های سطح سلولی و ترشحی می‌باشند. قندها با اتصال N در افزایش پایداری، ترشح و پیچ‌خوردگی پروتئینی دخالت دارند. چنانچه فرایندهای N گلیکوزیلاسیون مختل گردند واکنش‌های مرتبط به تنش شبکه آندوپلاسمیکی و فرایندهای آپوپتوز سلولی افزایش می‌یابند. بنابراین می‌توان افزایش بیان این ژن را مرتبط با توان مرغ‌های بومی در واکنش به تنش‌های محیطی دانست.

ژن *KLHDC4* در فاصله ۳۷۷۲ نوکلئوتیدی پایین دست ژن IncRNA مربوطه‌اش (MSTRG.2659) قرار داشته و هر دو در مرغ بومی بیان بالاتر و معنادار نسبت به جوجه تجاری داشتند. گزارش شده در انسان عدم تنظیم بیان پروتئین‌های kelch توأم با بروز برخی بیماری‌ها نظیر سرطان، کاهش رشد و نمو ماهیچه اسکلتی و بروز ناهنجاری‌های عصبی می‌باشد. نتایج برخی پژوهش‌های اخیر ثابت نموده‌اند پروتئین‌های kelch در ارتباط با یک نوع پیچیده از لیگازهای یوبیکوتینی E3 بنام لیگاز یوبیکوتینی Culling-RING بوده که به‌عنوان پیش‌ماده جهت فعال‌سازی این مجموعه نقش مهمی ایفا می‌نمایند [۱]. در مرغ بومی ژن *CUL3* افزایش بیان معنی‌داری نسبت به جوجه تجاری داشت. این ژن پروتئینی به نام Cullin-3 را کدهی می‌کند که در

تولیدات دامی

Evf-2, XIST, HOTAIR و Air دارای مشابهت بسیار وضعیفی در سطح توالی بین گونه‌ها می‌باشند. بنابراین در گونه‌های جانوری مختلف، lncRNAها می‌توانند علی‌رغم عملکرد مشابه، بدون مشابهت و یا دارای مشابهت کمی در توالی‌های خود باشند. ژن‌های lncRNAها در مقایسه با ژن‌های رمزکننده پروتئین درجات بالاتری از بیان اختصاصی در بافت را نشان می‌دهند. این ویژگی در مطالعات متعددی گزارش شده و نشان‌دهنده نقش‌های خاص تنظیمی آن‌ها در فرایندهای مهم سلولی می‌باشد [۲۲]. در واقع اگر lncRNAها صرفاً در اثر اشتباهات رونویسی به‌وجود می‌آیند، باید تنوع بسیار کمتری را در میزان بیان بین بافتی نشان می‌دادند. با توجه به این مباحث و نیز بیان اختصاصی در بافت، می‌توان دریافت که شناسایی کامل نقش تنظیمی lncRNAها نیازمند مطالعات بیشتری است.

نتیجه آنالیز این بخش نشان داد از مجموع ۲۲۵ ژن و ۲۶۷ رونوشت بین‌ژنی شناخته‌شده و ۲۱۲ ژن و ۲۲۰ رونوشت بین‌ژنی جدید، ۱۰۸ lncRNA بین‌ژنی (۵۹ شناخته‌شده و ۴۹ جدید) از نظر ترتیب و جهت قرارگیری ژن‌های هم‌جوار مرغ با انسان مطابقت داشتند. این مقدار برای lncRNAهای گاو با مرغ برابر با ۸۵ (۳۴ شناخته‌شده و ۵۱ جدید) بود. این موضوع می‌تواند مؤید عملکرد مشابه lncRNAها و ژن‌های هم‌جوارشان در مسیرهای سوخت‌وسازی مشترک در این گونه‌ها باشد. میزان بیان ۱۰ lncRNAs با بالاترین بیان با حفاظت هم‌جواری بین انسان و مرغ در جدول ۴ ارائه شده است. همان‌طور که در این جدول مشاهده می‌شود، بالاترین مقدار در هر دو مرغ مربوط به MSTRG.11186 بود. بنابراین شاید بتوان نتیجه گرفت این lncRNAها اولاً بر بیان ژن‌های عضله اسکلتی نقش مؤثری داشته و ثانیاً در انسان و مرغ وظیفه یکسان و مشابهی را انجام می‌دهند.

موجب پایداری بیشتر پروتئین‌های پیش‌ماده و تنظیم فعالیت‌های یوبیکوئیتیناسیون و شکست پروتئینی می‌گردد. گزارش شده است که USP7 با دی‌یوبیکوئیتیناسیون FoxO موجب خروج آن از هسته به سیتوپلاسم و غیرفعال شدن توان رونویسی و فعال‌سازی لیگازهای E3 از این طریق می‌گردد. به این ترتیب افزایش بیان ژن USP7 در مرغ بومی، احتمالاً از طریق تأثیر بر FoxO3، می‌تواند سبب کاهش نسبی میزان شکست پروتئین‌های ماهیچه‌ای و تنظیم فرآیند بازگردش (Turn Over) پروتئینی ماهیچه مرتبط به آن شده باشد.

بنابراین با توجه به اثر ژن‌های گفته شده بر میزان رشد هایپرتروفی ماهیچه‌ای و نیز با توجه به فاصله هم‌جواری و میزان بیان این پنج ژن با lncRNAهای مربوطه و نیز هم‌سو بودن مقادیر بیان هر ژن با lncRNA هم‌جوار آن، می‌توان نتیجه گرفت که به احتمال زیاد lncRNAهای مذکور بر بیان ژن‌های هم‌جوارشان مؤثر بوده‌اند.

با استفاده از مقایسه‌های جفتی تمامی توالی‌ها و با معیارهای تعیین‌شده جهت شناسایی حفاظت توالی، شباهت هیچ‌یک از lncRNAهای بین‌ژنی شناخته‌شده و جدید بین مرغ با انسان و مرغ با موش معنی‌دار نبود. بالاترین مقدار هم‌ترازی برای lncRNAهای بین‌ژنی شناخته‌شده مربوط به MSTRG.13772.1 (با ۳۰۹ نوکلئوتید) و ENST00000623789.1 انسانی با بیش از ۱۳ درصد (۸۴ درصد شباهت توالی) و برای lncRNAهای بین‌ژنی جدید، مربوط به MSTRG.11970.1 (با ۳۳۰ نوکلئوتید) و ENST00000507794.2 انسانی با بیش از ۱۱ درصد (۸۵ درصد شباهت توالی) بود. این نتایج در توافق با گزارش‌هایی است که نشان داده‌اند با افزایش فاصله تکاملی بین گونه‌ها، تعداد lncRNAهای با توالی حفاظت‌شده کاهش می‌یابد [۴]. به‌عنوان مثال برخی از شناخته‌شده‌ترین lncRNAهای انسانی نظیر Gomafu،

تولیدات دامی

افزایش پروتئوگلیکین‌ها است که بیانگر نقش احتمالی آن‌ها در فرآیندهای ترشحی است. بنابراین افزایش بیان این lncRNA در مرغ بومی می‌تواند به احتمال زیاد مرتبط با افزایش بیان ژن‌های مجاور آن‌ها باشد.

در این پژوهش معلوم شد روش توالی‌یابی RNA-seq بخوبی قادر به شناسایی lncRNAs می‌باشد. همچنین مشخص شد lncRNAs در تنظیم بیان ژن‌های هم‌جواری مؤثر می‌باشند. بررسی عملکردی ژن‌های هم‌جوار lncRNAs شناسایی شده، نشان داد که این ژن‌ها در جوجه تجاری در مسیرهای مرتبط با رشد و نمو ماهیچه اسکلتی و در مرغ بومی در مسیرهای تنظیم‌کننده انرژی دخالت معنادار دارند.

بین میزان لگاریتم بیان چند برابری lncRNAها با ژن پایین‌دست همبستگی مثبت و معنی‌دار برآورد شد ($r=0/351$) درحالی‌که برای ژن بالادست رابطه معناداری مشاهده نشد. در جدول ۵، ۱۰ ژن lncRNAs با حفاظت هم‌جواری بین انسان و مرغ که بالاترین مقادیر لگاریتم بیان چند برابری را داشتند، نمایش داده شده است. به‌عنوان مثال بالاترین مقدار بیان چند برابری در مرغ بومی مربوط به MSTRG.5539 بود و GALNT15 ژن بالادست و DPH3 ژن پایین‌دست آن نیز بیان بالاتری نسبت به جوجه تجاری داشتند. در انسان GALNT15 در سوخت‌وساز لیپوپروتئین‌ها و چربی‌گذاری درون ماهیچه‌ای و در نتیجه در مسیرهای انرژی دخالت دارد [۱۸]. کاهش بیان DPH3 نیز توأم با

جدول ۴. بالاترین مقادیر بیان lncRNAهای با حفاظت هم‌جواری در مرغ و انسان

نام lncRNA	میزان بیان در جوجه تجاری ^۱	میزان بیان در جوجه بومی ^۱	ژن بالادست	ژن پایین‌دست
MSTRG.11125	35.0685	12.8052	TSPAN18	CD82
MSTRG.11186	37.9562	31.7963	OIP5	CHP1
MSTRG.11920	27.3474	18.2967	HECTD2	PCGF5
MSTRG.11923	6.60974	9.74959	TNKS2	PPPIR3C
MSTRG.12647	5.06537	13.4687	GPD2	NR4A2
MSTRG.1954	10.8691	21.8294	P2RY2	FCHSD2
MSTRG.358	17.453	5.14037	POC1B	DUSP6
MSTRG.6156	5.71883	6.89991	CHD7	RAB2A
MSTRG.6362	7.22939	9.59545	HAS2	SNTB1
MSTRG.8612	8.09165	15.8355	AFTPH	LGALS1

۱. مقادیر بیان بر اساس معیار FPKM هستند.

جدول ۵. بالاترین مقادیر بیان نسبی lncRNAهای با حفاظت هم‌جواری در مرغ و انسان

نام lncRNA	Log2FC ^۱	ژن بالادست	Log2FC	ژن پایین‌دست	Log2FC
MSTRG.5539	-23.2584	DPH3	-0.55197	GALNT15	-1.2521
MSTRG.8625	-21.0462	WDR92	-0.75819	CID	-0.3012
MSTRG.6887	-21.0216	H6PD	-0.34781	GPR157	-1.3239
MSTRG.5414	-20.9045	GTPBP10	-0.56703	FAM237B	-
MSTRG.8812	-3.28181	ZFP36L2	-	HAAO	-1.0353
MSTRG.9471	24.6023	FAM84A	-0.03667	NBAS	-0.4111
MSTRG.356	23.8685	TMTC2	-	SLC6A15	0.22688
MSTRG.3021	22.6735	C3orf14	4.20808	FEZF2	-0.7477
MSTRG.12126	20.9693	CHST15	-0.32524	OAT	0.47467
MSTRG.10484	4.62663	BMPRI1B	0.3877	UNC5C	0.32934

۱. در پنج ردیف نخست lncRNAها در مرغ بومی و در پنج ردیف بعدی در جوجه تجاری بیان بالاتر داشتند.

تولیدات دائمی

منابع

- Bennett, E. J., Rush, J., Gygi, S. P. & Harper, J. W. (2010). Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. *Cell*, 143(6): 951-965.
- Cai, B., Li, Z., Ma, M., Wang, Z., Han, P., Abdalla, B. A., ... Zhang, X. (2017). LncRNA-Six1 Encodes a Micropeptide to Activate Six1 in Cis and Is Involved in Cell Proliferation and Muscle Growth. *Frontiers in Physiology*, 8(230). doi:10.3389/fphys.2017.00230
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C. & Lempicki, R. A. (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome biology*, 4(9): 1.
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., ... Knowles, D. G. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome research*, 22(9): 1775-1789.
- Gong, C., Li, Z., Ramanujan, K., Clay, I., Zhang, Y., Lemire-Brachat, S. & Glass, David J. (2015). A Long Non-coding RNA, LncMyoD, Regulates Skeletal Muscle Differentiation by Blocking IMP2-Mediated mRNA Translation. *Developmental Cell*, 34(2): 181-191. doi:https://doi.org/10.1016/j.devcel.2015.05.009
- Haerty, W. & Ponting, C. P. (2015). Unexpected selection to retain high GC content and splicing enhancers within exons of multiexonic lncRNA loci. *Rna*.
- Hussain, M. R. M., Nasir, J. & Al-Aama, J. Y. (2014). Clinically significant missense variants in human GALNT3, GALNT8, GALNT12, and GALNT13 genes: intriguing in silico findings. *J Cell Biochem*, 115(2): 313-327. doi:10.1002/jcb.24666
- Kang, Y.-J., Yang, D.-C., Kong, L., Hou, M., Meng, Y.-Q., Wei, L. & Gao, G. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic acids research*, 45(W1): W12-W16.
- Kim, D., Langmead, B. & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature methods*, 12(4): 357-360.
- Kohroki, J., Nishiyama, T., Nakamura, T. & Masuho, Y. (2005). ASB proteins interact with Cullin5 and Rbx2 to form E3 ubiquitin ligase complexes. *FEBS Lett*, 579(30): 6796-6802. doi:10.1016/j.febslet.2005.11.016
- Lange, S., Perera, S., Teh, P. & Chen, J. (2012). Obscurin and KCTD6 regulate cullin-dependent small ankyrin-1 (sAnk1. 5) protein turnover. *Molecular biology of the cell*, 23(13): 2490-2504.
- Li, A., Zhang, J. & Zhou, Z. (2014). PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC bioinformatics*, 15(1): 311.
- Liu, W. & Zhao, J. (2014). Insights into the molecular mechanism of glucose metabolism regulation under stress in chicken skeletal muscle tissues. *Saudi journal of biological sciences*, 21(3): 197-203.
- Ntambi, J. M., Miyazaki, M., Stoehr, J. P., Lan, H., Kendziorski, C. M., Yandell, B. S., ... Attie, A. D. (2002). Loss of stearoyl-CoA desaturase-1 function protects mice against adiposity. *Proceedings of the National Academy of Sciences*, 99(17): 11482-11486.
- Ørom, U. A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., ... Huang, Q. (2010). Long noncoding RNAs with enhancer-like function in human cells. *Cell*, 143(1): 46-58.
- Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature protocols*, 11(9): 1650.
- Ren, T., Li, Z., Zhou, Y., Liu, X., Han, R., Wang, Y., ... Kang, X. (2018). Sequencing and characterization of lncRNAs in the breast muscle of Gushi and Arbor Acres chickens. *Genome*, 61(5): 337-347.
- Schjoldager, K. T., Vester-Christensen, M. B., Bennett, E. P., Levery, S. B., Schwientek, T., Yin, W., ... Clausen, H. (2010). O-glycosylation modulates proprotein convertase activation of angiotensin-like protein 3: possible role of polypeptide GalNAc-transferase-2 in regulation of concentrations of plasma lipids. *J Biol Chem*, 285(47): 36293-36303. doi:10.1074/jbc.M110.156950
- Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., ... Zhao, Y. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic acids research*, 41(17): e166-e166.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., ... Pachter, L. (2012).

- Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*, 7(3): 562.
21. Wang, L., Park, H. J., Dasari, S., Wang, S., Kocher, J.-P. & Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic acids research*, 41(6): e74-e74.
 22. Wu, S. C., Kallin, E. M. & Zhang, Y. (2010). Role of H3K27 methylation in the regulation of lncRNA expression. *Cell Res*, 20(10): 1109-1116. doi:10.1038/cr.2010.114
 23. Wucher, V., Legeai, F., Hedan, B., Rizk, G., Lagoutte, L., Leeb, T., . . . Lohi, H. (2017). FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic acids research*, 45(8): e57-e57.
 24. Yin, Z., Deng, T., Peterson, L. E., Yu, R., Lin, J., Hamilton, D. J., . . . Zhan, M. (2014). Transcriptome analysis of human adipocytes implicates the NOD-like receptor pathway in obesity-induced adipose inflammation. *Molecular and cellular endocrinology*, 394(1-2): 80-87.
 25. Zhou, L., Sun, K., Zhao, Y., Zhang, S., Wang, X., Li, Y., . . . Bao, X. (2015). Linc-YY1 promotes myogenic differentiation and muscle regeneration through an interaction with the transcription factor YY1. *Nature communications*, 6: 10026.



Animal Production

(College of Abouraihan – University of Tehran)

Vol. 21 ■ No. 2 ■ Summer 2019

Identification of LncRNAs related to growth of chicken's breast muscle using RNA-seq

Seyed Nader Albooshoke^{1*}, Mohammad Reza Bakhtiarizadeh²

1. Assistant Professor, Animal Science Department, Khuzestan Agricultural and Natural Resources Research and Education Center, AREEO, Ahwaz, Iran.
2. Assistant Professor, Department of Animal Science, Abouraihan Campus, University of Tehran, Pakdasht, Iran.

Received: January 6, 2019

Accepted: March 17, 2019

Abstract

This experiment was carried out to identify lncRNAs associated with skeletal muscle of the Isfahan native chicken and the Ross 708 commercial broiler chicken. To do this, after extraction of RNA from the breast muscle samples at the age of 28 days, paired-end sequencing was performed using the Illumina HiSeq 2000 platform. Hisat2 software was used to align the clean reads to the chicken reference genome and the Stringtie software was used to assemble the transcripts. A total of 1097 lncRNAs were identified as 925 of which were intergenic and 172 were intronic. Also, the number of novel lncRNAs in intergenic and intronic groups were 432 and 128, respectively. Differential gene expression analysis led to the identification of 19 genes and 20 transcripts differentially expressed lncRNAs between two groups. Syntenic analysis showed that differentially expressed lncRNAs are located near by 45 protein encoding genes. Of these, the expression of five gene coding proteins (SCD gene in commercial chickens and GALNT15, KLHDC4, USP7 and ASB1 genes in native chicken) - whose expression was consistent with the expression of their lncRNA - were significantly expressed between two breeds. Functional enrichment analysis of these genes showed that all of them are involved in the skeletal muscle growth. The results of this study showed that the identified lncRNAs probably have the potential to regulate the genes involved in skeletal muscle growth. In this regard, they possibly cause the differences in growth rates between the two chicken breeds.

Keywords: Breast muscle, growth, Isfahan native chicken, lncRNA, RNA-seq.