



Reconstruction, analysis and comparison of gene networks topology based on RNA-Seq data involved in reproductive and fertility complex traits.

Amin Shahabi

PhD. Student, Department of Animal Science, Ferdowsi University of Mashhad, Mashhad, Iran.

Tel: +989125233634, Email: shahabi.amin@mail.um.ac.ir

Mojtaba Tahmoorespur

* Corresponding Author. Professor, Department of Animal Science, Ferdowsi University of Mashhad, Mashhad, Iran. Email: tahmoores@um.ac.ir

Ali Kazemi-Pour

Assistant Professor, Shahid Bahonar University of Kerman, Iran.

Abstract

Objective

In spite of the importance of fertility in different species, there has been little success dissecting the levels of OMICS. To better understand the molecular basis of fertility, we study transcriptome profiling of different tissues.

Materials and methods

liver, muscle, endometrium and corpus luteum tissues between cows with either good or poor genetic merit for fertility using RNA-Seq data sets. We first compiled a master list of genes related to corpus luteum that change with level of fertility and then reconstructed the network.

Results

A few genes were identified in liver, muscle and endometrium between high and low fertile cows but in corpus luteum circumstance was different. 264 genes and 6 key modules were disclosed through clustering for mRNA master list for corpus luteum. All these genes, being involved in at least one of the biological process, namely proteolysis, actin cytoskeleton organization, immune system process, biological adhesion, cell differentiation and lipid metabolic process, have an overexpression pattern ($P < 0.01$).

Conclusions

Finally, the identification of genes and the construction of their regulatory networks may give new insights into biological procedures. As well as, this study increases our understanding of the contribution of different tissues transcriptome to phenotypic fertility in dairy cattle.

Keywords: Fertility, Transcriptome, Co-expression gene network, Reconstruction of gene networks

Citation: Shahabi A, Tahmoorespur M, Kazemi-Pour A (2019) Reconstruction, analysis and comparison of gene networks topology based on RNA-Seq data involved in reproductive and fertility complex traits. *Agricultural Biotechnology Journal* 11 (2), 57-78.

Agricultural Biotechnology Journal 11 (2), 57-78.

DOI: 10.22103/jab.2019.13510.1112

Received: March 19, 2019; Accepted: June 16, 2019

© Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society

بازسازی، آنالیز و مقایسه توپولوژی شبکه‌های ژنی مبتنی بر داده‌های RNA-Seq دخیل در

صفات چند فاکتوره تولید مثلی و باروری

امین شهبابی

دانشجوی دکترا، بخش علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران. ایمیل:

shahabi.amin@mail.um.ac.ir

مجتبی طهمورث پور

*نویسنده مسئول، استاد، بخش علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران. ایمیل:

tahmoores@um.ac.ir

علی کاظمی پور

استادیار، گروه زراعت و اصلاح نباتات، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان، کرمان، ایران

تاریخ دریافت: ۱۳۹۷/۱۲/۲۸، تاریخ پذیرش: ۱۳۹۸/۰۳/۲۶

چکیده

هدف: علیرغم اهمیت باروری در گونه‌های مختلف، موفقیت‌های کمی در فهم اساس باروری از منظر لایه‌های مختلف OMICS به دست آمده است. برای فهم بهتر اساس مولکولی باروری، نیمرخ ترانسکریپتوم بافت‌های مختلف در گاو مورد بررسی قرار گرفت. **مواد و روش‌ها:** بافت‌های کبد، عضله، اندومتريوم و جسم زرد در گاوهایی با شایستگی ژنتیکی بالا و پایین در صفات تولیدمثلی با استفاده از داده‌های RNA-Seq مورد پژوهش قرار گرفت. در ابتدا فهرست ژن‌های مربوط به جسم زرد که تفاوت بیانی داشتند تهیه و در ادامه شبکه هم بیان ژنی برای این فهرست ژنی ایجاد شد.

نتایج: تعداد ژن‌های محدودی در بافت‌های کبد، عضله و اندومتريوم تفاوت بیان نشان دادند ولی در مقابل تعداد قابل توجهی ژن در جسم زرد تفاوت بیانی از خود نشان دادند. بر این اساس، تعداد ۲۶۴ ژن و ۶ ماژول عملکردی بعد از طبقه بندی ژن‌ها برای جسم زرد شناسایی شد. همه این ژن‌ها حداقل در یکی از فرآیندهای زیستی از قبیل پروتئولایسیز، سازماندهی آکتین، پاسخ ایمنی، چسبندگی سلولی، تمایز سلولی و متابولیک چربی نقش داشتند ($p < 0.01$).

نتیجه‌گیری: شناسایی ژن‌ها و بازسازی شبکه‌های تنظیمی می‌تواند افق جدیدی در راستای فهم چگونگی سازو کار فرآیندهای زیستی باز کند. همچنین این پژوهش فهم ما را در نقش بافت‌های مختلف روی باروری در گاوهای شیری را نشان می‌دهد.

کلید واژه : باروری، ترانسکریپتوم، شبکه هم بیان ژن، ایجاد شبکه

مقدمه

بهبود باروری، بازده تولید گاوهای شیری را افزایش خواهد داد. توسعه عملکرد تولید مثلی در گاو می تواند ماندگاری یک گاو، کاهش فاصله گوساله زایی و کاهش هزینه های دامپزشکی را به دنبال داشته باشد. باروری گله ها، فاکتور اصلی برای پایداری یک گاوداری می باشد. هر چند از سال های پیش، باروری در گاو کاهش یافته است و تبدیل به یکی از دلایل اصلی حذف و جایگزینی گاوها شده است (Barbat 2010). مطالعات ناچیزی در دهه های اخیر در جهت بررسی اساس ژنتیکی این چالش انجام شده است. بنابراین، یک دانش کاربردی در سازو کار مولکولی در جهت کمک در شناسایی ژن های کلیدی و بررسی اثرات متقابل بین در صفات اقتصادی مهم مورد انتظار می باشد. هر دوی برنامه های نقشه یابی QTL و ژن های مهم و مطالعات ژنومیک برای یافتن ژن های کلیدی مرتبط با باروری و تولید مثل انجام گرفته است. متأسفانه هیچ گونه تصویر مشخصی در بررسی لایه های مختلف OMICS وجود ندارد. در نتیجه موفقیت تولید مثلی در گاو به نظر می رسد که در نتیجه اثرات متقابل پیچیده بین لایه های مختلف و فاکتورهای مرتبط باشد. از سوی دیگر، گاو گونه ای است که بالقوه و بالفعل به عنوان گونه مدل برای مطالعه اساس باروری در دیگر پستانداران مورد توجه قرار گرفته است. باروری گاو از ارزش اقتصادی بالایی برخوردار است (Walsh, 2011). نگره داری و پایداری آبستینی شامل یک تعامل پیچیده بین اندومترיום تخمدان، جنین و جسم زرد می باشد. اندومترיום با ترشح همومون هایی به توسعه جنین کمک می کند (Forde et al. 2013; Forde et al. 2014a; Forde et al. 2014b). بعد از تخمک ریزی، سازماندهی سلولی فولیکول تخمک ریزی کرده برای پایداری جسم زرد لازم و ضروریست (Robinson et al. 2014). بنابراین، اندومترיום و جسم زرد اهداف مشخصی برای مطالعات بیان ژن برای شناسایی ژن هایی که تفاوت بیانی نشان می دهند می باشد (Robinson et al. 2010). از سویی به یک رویکرد جامع برای رمزگشایی حجم وسیع اطلاعات تولید شده توسط فناوری های نوین زیستی نیاز است. تولید اطلاعات با خروجی بالا در مطالعات OMICS ما را قادر به تجزیه و تحلیل هم زمان و بررسی اجزاء سلولی در هر شرایطی می کند. همچنین، الگوریتم ها و ابزارهای محاسباتی بسیاری برای شناسایی ماژول های زیستی در شبکه های زیستی توسعه یافته اند (Chen et al. 2011). در نتیجه استراتژی مطالعات لایه های مختلف OMICS می تواند در شناسایی و بازسازی فرآیندهای زیستی و مسیرهای جدید کمک شایانی نماید (Bugrim et al. 2004). پیشرفت های اخیر در زیست شناسی مولکولی نیاز ما را به این رویکرد افزایش داده است. بنابراین، بازسازی، تجزیه و تحلیل و بررسی شبکه های زیستی و فرآیندهای زیستی شامل اثرات متقابل بین پروتئینی (PPI)، مسیرهای متابولیکی، شبکه های تنظیمی ژنی (GRV) و مسیرهای انتقال پیام ابزار ارزشمندی در درک مفهوم زیستی می باشند (Kann 2007). تجزیه و تحلیل مسیرهای مولکولی پتانسیل خوبی برای نشان دادن سازو کار مولکولی صفات، همچنین پاسخ به تیمارهای مختلف دارد. در ایران حدود ۷/۹ میلیون راس گاو وجود دارد که از این تعداد ۴۵/۹ درصد بومی، ۴۳/۶ درصد دورگ و ۱۰/۵۱ درصد رجیستر شده (اکثراً هلشتاین)

هستند. سهم حیوانات اهلی در اقتصاد ملی ۴ درصد از درآمد ناخالص ملی (GDP) است. بر اساس گزارش فائو (FAO 1993)، تولیدات گاو شیری در طی دو دهه گذشته تغییرات ساختاری قابل توجه و مهمی داشته است و گله‌های بزرگتری از گاو ایجاد شده است. تلقیح مصنوعی در مناطق مختلفی از جهان، به ویژه در ایران به طور قابل توجهی توسعه یافته و مورد استفاده قرار گرفته است (Ebrahimi et al. 2015a). پرورش دهندگان گاو شیری در ایران چهار گروه از نرها را معمولاً مورد استفاده قرار می‌دهند که براساس منشاشان عبارتند از گاوهای نر آمریکایی، کانادایی، اروپایی و ایرانی. از طرف دیگر، این چهار گروه همچنین می‌توانند به دو گروه طبقه بندی شوند: نرهای summarized و نرهای sampling. در گروه اول آزمون نتاج و برآورد توانایی تولید دختران آنها امکان پذیر است. گروه دوم برای صفات تولید بالا یا صفات همبسته با آنها براساس شجره‌شان انتخاب می‌شوند. لذا، شناسایی بسیار دقیق نرهایی که حامل صفات تولیدی بالا و صفات وابسته به آنها برای انتقال به فرزندان هستند عدم تمایل بسیاری از تولیدکنندگان گاوهای شیری ایرانی برای استفاده از نرهای sampling این واقعیت است که عملکرد دختران آنها تا حدودی غیر قابل پیش بینی است. اما، بعضی از تولیدکنندگان گاوهای شیری ترجیح می‌دهند تا از نرهای sampling استفاده کنند تا نقش گاوهای خارجی را تعدیل کنند یا پایین آورند (Ebrahimi et al. 2015b). به علاوه، طبق آخرین آمار رسمی وزارت جهاد کشاورزی، تعداد ۱۸۸۳۰ واحد صنعتی گاو داری با ظرفیت ۲۰۴۸۵۶۳ راس گاو شیرده در کشور مشغول فعالیت هستند. طی سال های ۱۳۸۳ تا ۱۳۸۷ تولید شیر دارای یک روند رو به رشد بوده است (Kharrati Koopaei et al. 2011). با وجود روند افزایشی تولید شیر در کشور اما هنوز سرانه مصرف شیر از حد استاندارد جهانی پایین تر است (Kharrati Koopaei et al. 2012a). سرانه مصرف شیر در کشور برای هرنفر برابر با ۹۵ کیلوگرم می باشد، در حالی که سرانه مصرف شیر در جهان برابر با ۱۶۹ کیلوگرم و در اروپا برابر با ۳۵۰ کیلوگرم در سال است (Kharrati Koopaei et al. 2011). با توجه به آمار و اطلاعات موجود می‌توان دریافت که اهداف اصلاح نژادی در ایران بایستی برای افزایش تولید در کشور برنامه ریزی شود. لذا مطالعه و بررسی نژادهای مختلف گاو از جنبه‌های مختلف در ایران اهمیتی دو چندان می یابد (Kharrati Koopaei et al. 2012b). در اواخر دهه ۸۰ میلادی مطالعات و بررسی های به عمل آمده روشن نمود که مکانیسم‌های مولکولی در زمره مهم‌ترین فرایندهای ژنتیکی (مشمول بر همانندسازی DNA، رونویسی، ترجمه و حتی نحوه تنظیم ژن ها) هستند (Mohammadabadi and Tohidinejad 2017). ماده ژنتیکی یک سلول دارای تعداد زیادی ژن می باشد که هیچ گاه به طور همزمان بیان نمی‌شوند و در یک زمان خاص فقط تعداد کمی از آنها بیان شده و پروتئین یا آنزیم مورد نیاز سلول را تولید می‌نمایند (Tohidi nezhad et al. 2015). نیاز به بیان ژن توسط محیطی که در آن رشد می‌کند کنترل می‌شود و در صورت عدم نیاز به فرآورده ژن، آن ژن به صورت خاموش و غیرفعال باقی خواهد ماند. ساز و کار بیان ژن اولین بار در باکتری E.coli کشف شد. بیان ژن های یوکاریوتی تحت کنترل موقت و چندبعدی می باشد. تنها یک مجموعه نسبتاً کوچک از تمام ژنوم در هر یک از انواع بافت ها بیان می شود و نیز بیان ژن‌ها به مرحله نمو بستگی دارد. بنابراین، بیان ژن در یوکاریوت‌ها برای هر بافت اختصاصی است. همچنین مقدار محصولات ژن که در همان بافت و نیز در سایر بافت‌هایی که آن محصول را می‌سازند، ساخته شده سبب تنظیم بیان آن ژن می‌-

شود (Mohammadabadi et al. 2017). یکی از اقدامات اساسی در حیوانات اهلی مطالعه ژن‌ها و پروتئین‌های مرتبط با صفات اقتصادی و مطالعه آنها در سطح سلولی یا کروموزومی است (Jafari Darehdor et al. 2016). کاوش در تفاوت بیان ژن‌ها در بافت‌های مختلف به روشن‌سازی نتایج ژنتیکی و فیزیولوژیکی انتخاب برای عملکرد، همچنین ممکن است به فهم بهتر شبکه‌های تنظیمی و مسیرها که برای صفات تولید مثلی مهم هستند کمک کند. از سوی دیگر سودمندترین رویکرد در جهت نگاه کلی به شبکه‌ها و مسیرها بازسازی شبکه‌های تنظیم بیان ژنی با استفاده از الگوریتم‌های مختلف یا پایگاه داده‌ای اثرات متقابل می‌باشد (Blais and Dynlacht 2005; Bansal et al. 2007; Emmert-Streib et al. 2014). شبکه‌های تنظیم ژنی اطلاعات ارزشمندی در مورد ژن‌های اصلی تولید می‌کند و ما را قادر به شناسایی ژن‌های کلیدی در یک بافت خاص می‌کند (Basso et al. 2005; Emmert-Streib et al. 2014). اگر چه پژوهش‌های زیادی روی گاوهای ایران انجام شده است (Mohammadabadi et al. 2004; Mohammadabadi et al. 2005; Sulimova et al. 2007; Alinaghizadeh et al., 2007; Mohammadabadi and Mohammadi 2010; Ruzina et al. 2010; Ghasemi et al. 2010; Mohammadabadi et al. 2011; Pasandideh et al. 2015; Barazandeh et al. 2011; Askari et al. 2016; Pasandideh et al. 2016) ولی تاکنون پروفایل ترانسکریپتوم کبد، عضله، اندومتر و جسم زرد با استفاده از داده‌های RNA-Seq مورد مطالعه قرار نگرفته است. بنابراین، هدف این مطالعه تهیه لیست ژن‌های مربوط به ترانسکریپتوم جسم زرد بین گاوهای بارور و نابارور، و ادغام اثرات متقابل پروتئینی و شبکه‌های تنظیم ژنی پ و تجزیه و تحلیل توپولوژیکی روی این شبکه بود.

مواد و روش‌ها

بازسازی شبکه نقش کلیدی در مطالعات OMICS دارد. از این رو فرآیند بازسازی به طور گسترده‌ای با دانش ما در رابطه با باروری در ارتباط است. ما سه دسته داده با خروجی بالا در این مطالعه استفاده نمودیم که تمامی داده‌های مربوط به داده‌های RNA-Seq بودند. دسته داده‌ها به طور جداگانه مورد تجزیه و تحلیل قرار گرفت. در همه حالات برای شناسایی ژن‌ها و پروتئین‌ها از HGNC (The HUGO Gene Nomenclature Committee) استفاده شد.

بیان ژن داده‌های RNA-Seq: دسته داده‌ها برای بیان ژن در بافت‌های مختلف متعلق به گاوهای بارور در مقایسه با گاوهای غیر بارور با جست‌جو در پایگاه داده‌ای GEO (Gene Expression Omnibus) انجام شد. برای افزایش تعداد ژن‌ها همچنین برای افزایش داده‌های مورد استفاده پایگاه‌های NCBI (Barrett et al. 2013) و Array Express (Kolesnikov et al. 2015) نیز برای گونه‌های مورد جست‌جو قرار گرفت. در مجموع ۱۳۴ نمونه انتخاب شد. ۲۴ نمونه مربوط به کد دسترسی GSE74076 (Moore et al. 2016) و ۹۶ نمونه از کد دسترسی GSE62159 (Moran et al. 2016) و ۱۴ نمونه مربوط به کد دسترسی GSE52438 (Moran et al. 2015) انتخاب شدند.

مراحل کار داده‌های RNA-Seq: ابتدا برای کنترل کیفیت داده‌های موجود از نرم‌افزار کنترل کیفیت FastQC و برای تبدیل فرمت فایل‌ها به فرمت سنکر و ایلومینا از FastQ Groomer استفاده شد. پس از این مرحله داده‌ها نیاز به پیرایش آدپتورها داشتند که برای این منظور نرم‌افزار Trimmomatic (Bolger et al. 2014) به کار گرفته شد. در مرحله بعدی که مربوط به نقشه یابی قطعات خوانده شده روی ژنوم مرجع گاوی بود از نرم‌افزار TopHat2 (Kim et al. 2013) بهره گرفته شد. سپس با استفاده از نرم‌افزارهای CuffLink و CuffCompare (Trapnell et al. 2010) خوانش‌های انجام گرفته ادغام و ترانسکرپتوم واحد تشکیل شد و در پایان میزان تفاوت در بیان ژن‌ها با استفاده از نرم‌افزار CuffDiff (Trapnell et al. 2010) مورد ارزیابی قرار گرفت.

حاشیه نویسی ژن‌ها: برای حاشیه نویسی ژن‌ها پایگاه‌های مختلفی از جمله DAVID، g:Profiler، (<http://biit.cs.ut.ee/gprofiler/>) و PANTHER مورد استفاده قرار گرفتند.

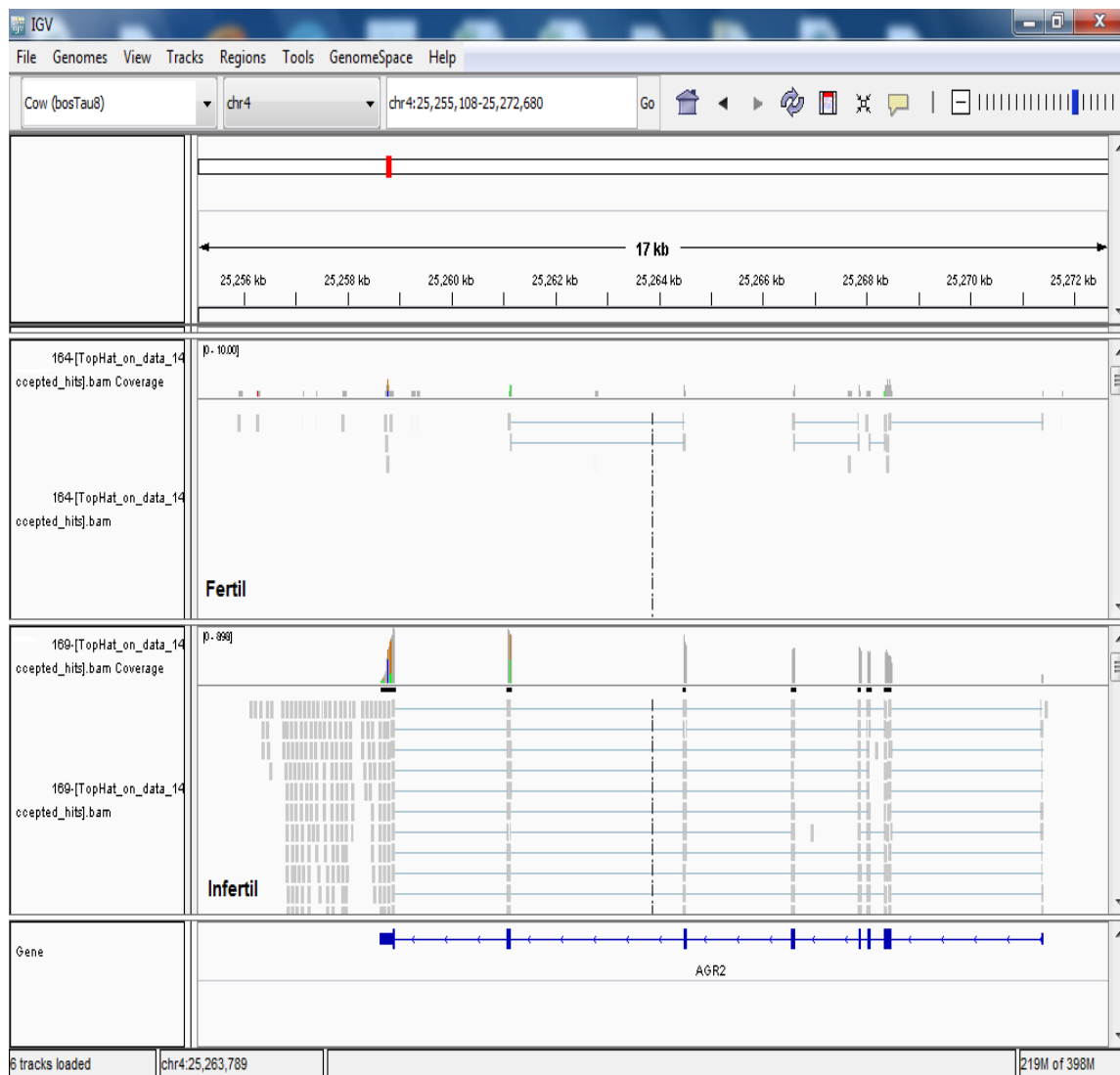
بازسازی شبکه هم بیانی ژنی: شبکه هم بیان ژنی بر اساس اثرات متقابل و لیست ژنی مربوط بازسازی شد. اثرات متقابل پروتئینی (PPI)، BIND (Biomolecular Interaction Network Database) MIPS (Mammalian Biological General Repository for Protein-Protein Interactions Database) و BioGRID (Interaction Datasets) در بازسازی شبکه استفاده شد (ChatrAryamontri et al. 2013). علاوه بر این داده‌های مربوط به اثرات متقابل از پایگاه‌هایی از قبیل GeneMANIA (Mostafavi et al. 2008) و STRING (Search Tool for the Retrieval of Interacting Genes or Proteins) مورد استفاده قرار گرفت. در نهایت برای بازسازی شبکه استخراج، ترکیب، تصویر سازی و تجزیه و تحلیل از نرم افزار Cytoscape استفاده شد. Species مولکولی در شبکه بازسازی شده نمایانگر گره‌ها و اثرات متقابل بین آن‌ها نمایانگر یا لهای بین آنها بودند. برخی از اثرات متقابل به دست آمده به وسیله پایگاه‌های داده ای از قبیل PANTHER (Protein ANalysis THrough Evolutionary Relationships) Reactome و KEGG (Kyoto Encyclopedia of Genes and Genomes) چک شدند.

طبقه بندی شبکه هم بیان ژنی: مشخصات توپولوژیکی شبکه هم بیان ژنی با نرم افزار Cytoscape مورد ارزیابی قرار گرفت و برای طبقه بندی از نرم افزارهای MCODE (Bader & Hogue 2003) و ClusterOne (Nepusz et al. 2012) استفاده شد.

نتایج و بحث

لیست ژنی کاندید: براساس گردش کار توضیح داده شده در بخش مواد و روش‌ها در مجموع ۲۸۳ ژن (۳ و ۶ و ۱۰ و ۲۶۴ ژن به ترتیب در بافت‌های عضله، کبد، اندومتر و جسم زرد) برای داده‌های RNA-seq شناسایی شدند. این ژن‌ها بر

اساس پایگاه داده ای مربوطه حاشیه نویسی شدند. برای اعتبار سنجی نقشه یابی صحیح یکی از مهم ترین ژن ها با نرم افزار IGV تصویرسازی شد که در شکل ۱ نشان داده شده است. ژن های تفاوت بیان نشان داده در گاوهای بارور در مقایسه با گاوهای نابارور ($FDR < 0.01$, Fold change > 1.5) بر اساس سلسله مراتب طبقه بندی ژن ها تعداد ۱۵۵ ژن به طور معنی داری افزایش بیان و ۱۰۹ ژن کاهش بیان ($FDR < 0.01$, Fold change < -1.5) در جسم زرد از خود نشان دادند. در تجزیه تحلیل فرآیند های زیستی و مسیرهای مربوط به ژن های تغییر بیان نشان داده، فرآیندهای پرتئولایسیز، سازمان دهی آکتین، پاسخ ایمنی، چسبندگی سلول، تمایز سلولی و فرآیند متابولیسم چربی شناسایی شدند.



شکل ۱. چگونگی خوانش ها در ژن AGR2 با استفاده از نرم افزار IGV

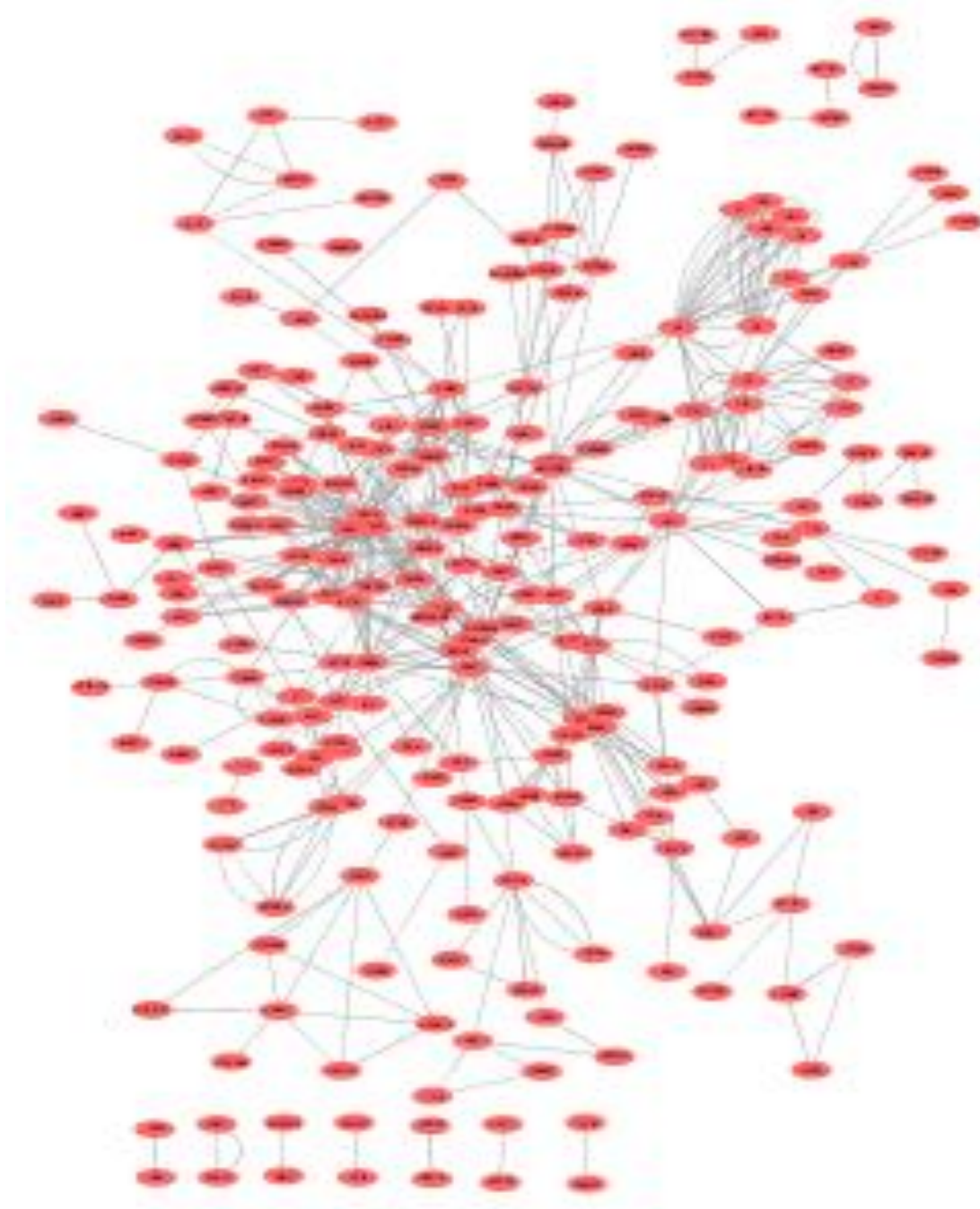
Figure 1. How to read the AGR2 gene using the IGV software

نتایج این مطالعه نشان داد که ادغام داده های ترانسکریپتوم مربوط به دسته داده های مختلف در یک بافت مشخص از بافت های کلیدی تولید مثلی یک رویکرد نوین و مفیدی برای روشن کردن سازو کار دخیل در عملکرد تولید مثل گاوها پیش روی ما قرار داد. تجزیه و تحلیل تفرق بیانی بین گاوها با باروری بالا و پایین تعداد ۳ و ۶ و ۱۰ و ۲۶۴ ژن به ترتیب در بافت های عضله، کبد، اندومتریم و جسم زرد را شناسایی نمود. ژن های مربوط به بافت های عضله و کبد همانطور که انتظار می رفت به طور مشخصی در هیچ مسیر زیستی خاصی حضور نداشته اما ژن های مربوط به بافت اندومتریم همگی در فرآیند سیستم ایمنی نقش داشتند. در مقابل ژن های شناسایی شده در جسم زرد در فرآیندهای زیستی مختلفی از قبیل پرتولایسیز، سازماندهی آکتین، پاسخ ایمنی، چسبندگی سلولی، تمایز سلول و فرآیند متابولیسم چربی دخیل بودند. روش های محاسباتی ماژول های تنظیمی ژن ها به فهم فرآیندهای پیچیده زیستی کمک خواهند نمود. شبکه هم بیانی ژنی تعریف شده در این تحقیق شامل ژن ها و شرایط مرتبط در باروری می باشد. این شبکه می تواند شانس جدیدی برای اعمال تیمارها روی صفات ژنتیکی از قبیل فرآیند تولید مثلی ایجاد نماید. از سوی دیگر فهم سازو کار مولکولی باروری برای حل پیچیدگی فرآیند تولید مثلی نیاز است. هر چند چندین مطالعه (Bao & Garverick et al. 1998; Vitt et al. 2000; Hayashi et al. 2010; Bahrami et al. 2017) الگوی متفاوت بیان ژن را در فرایندهای مختلف تولید مثل نشان داده اند. در این تحقیق شبکه هم بیان ژن مرتبط با جسم زرد مطالعه شد (شکل ۲).

در این راستا با داده های مختلف که از جسم زرد به دست آمده بود ادغام شدند. بنابراین به دلیل تعداد بالای نمونه های جمع آوری شده می توان گفت که این تحقیق از لحاظ زیستی قابل اتکا و تکرار پذیر است. در واقع ما یک رویکرد تلفیقی در این تحقیق استفاده نمودیم تا تمامی ژن ها و اثرات متقابل بین آن ها را شناسایی کنیم. نتایج ۶ ماژول کلیدی دخیل در باروری را شناسایی نمود. در این ماژول ها تعداد ۴۴ ژن در همه ماژول ها بدون تکرار وجود داشتند که به ترتیب در ماژول ۱، ۲، ۳، ۴، ۵ و ۶ شامل ۳۲، ۲۰، ۱۶، ۶۵، ۲۸ و ۱۶ ژن وجود داشت. در ماژول ۱، مهم ترین ژن ها شامل CAPN3, PLS1, CNN1, ACTA1 و KLHL17 کاهش بیان از خود نشان دادند. همه ژن های این ماژول در سازماندهی آکتین نقش داشتند (شکل ۳).

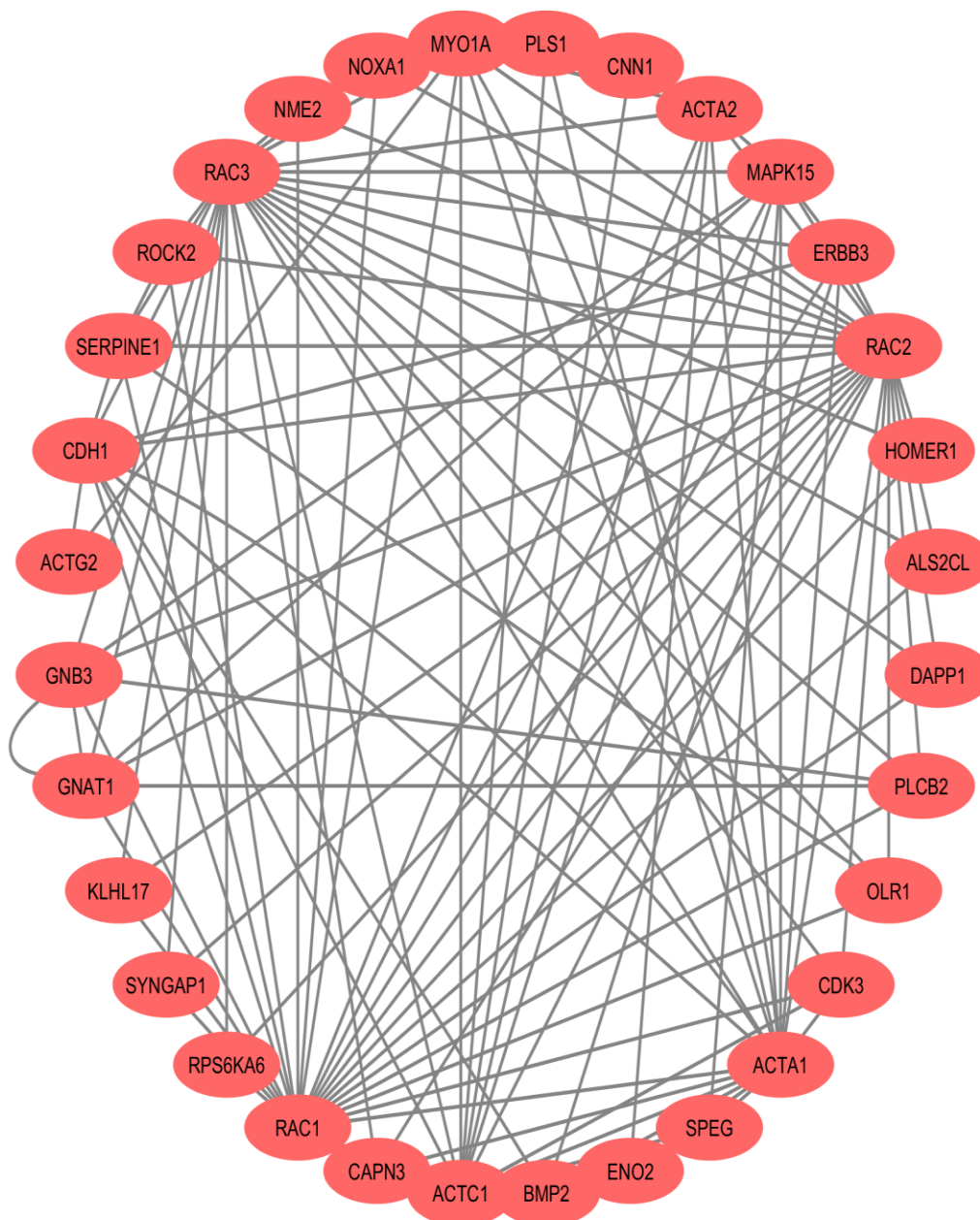
در ماژول ۲، ژن های C3, SERPINE1, CFB, C5 و CFI به عنوان مهم ترین ژن شناسایی شدند که همگی آن ها افزایش بیان داشتند. فرآیند بیولوژیکی کلیدی در این ماژول که تحت تاثیر این ژن ها قرار داشت فرآیند پروتولایسیز بود (شکل ۴).

در ماژول ۳، ژن های کلیدی که در فرآیند سیستم ایمنی دخیل بودند شامل C3, C4BPB, SERPINE1, CFB و MASP2 و C4BPA بودند که ژن های C3, C4BPB, SERPINE1, CFB و C4BPA افزایش بیان و ژن MASP2 کاهش بیان از خود نشان دادند (شکل ۵).



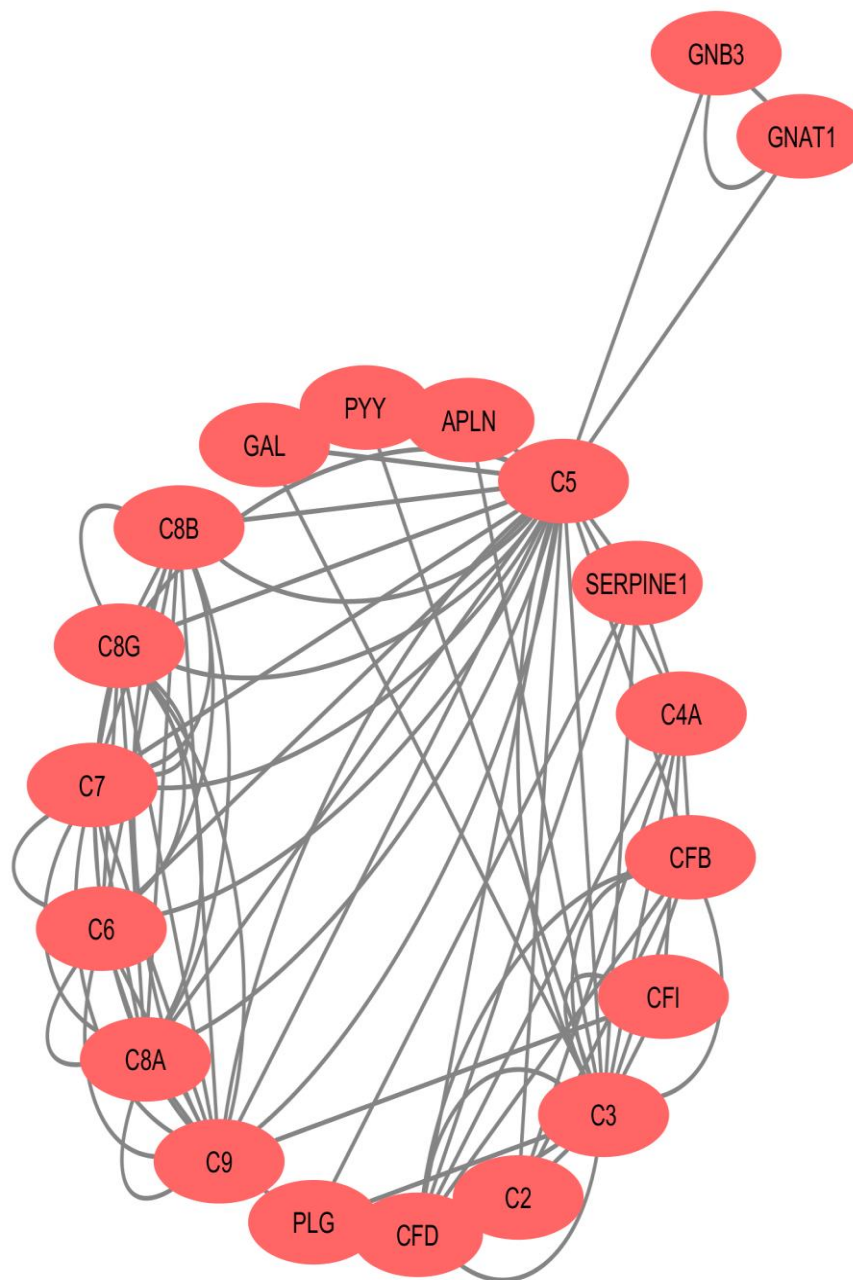
شکل ۲. شبکه‌ی ژنی مبتنی بر داده‌های RNA-Seq دخیل در صفات چند فاکتوره تولید مثلی و باروری

Figure 2. Gene networks based on RNA-Seq data involved in multiple reproductive and reproductive traits



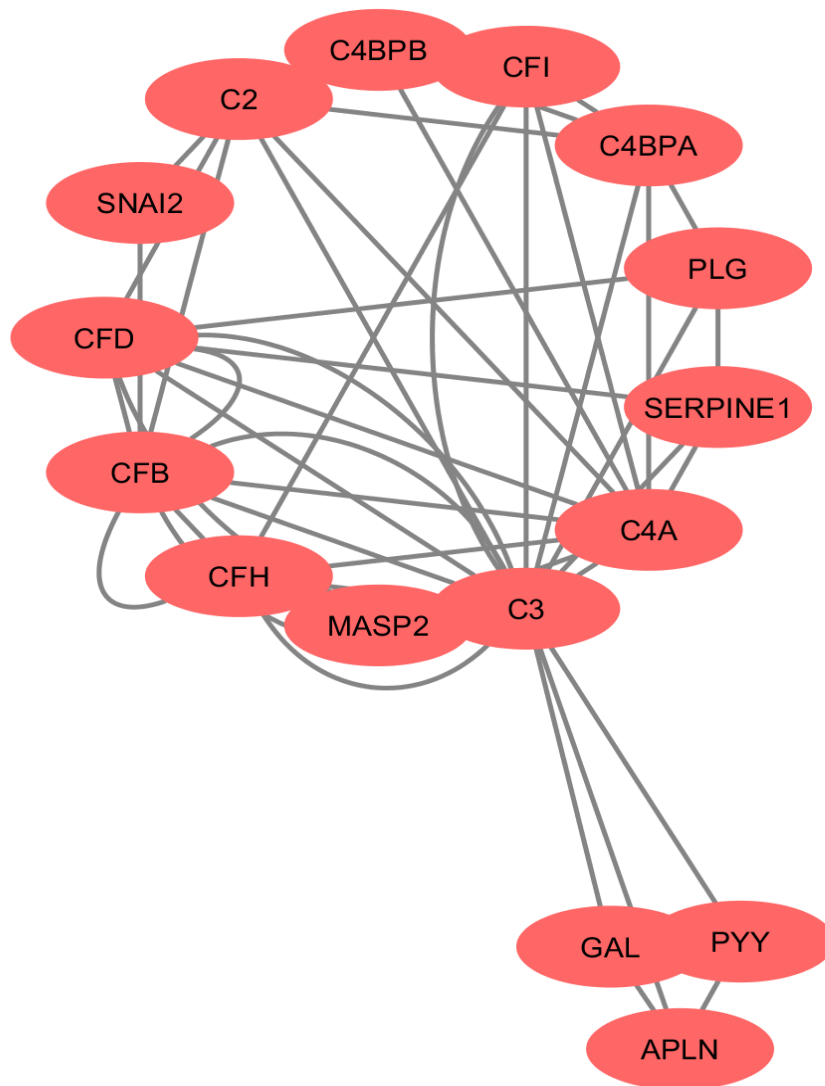
شکل ۳. ماژول شماره یک که نشان دهنده تعداد ژن‌ها و اثرات متقابل بین آن‌هاست. اشکال بیضی (نودها) بیانگر ژن‌ها و یال‌های بین آن‌ها بیانگر اثر متقابل بین ژنی می‌باشد

Figure 3. The module number one, which represents the number of genes and interactions between them. Oval shapes (nodes) representing genes and edges between them represent the interactions between the genes.



شکل ۴. ماژول شماره دو که نشان دهنده تعداد ژن‌ها و اثرات متقابل بین آن‌هاست. اشکال بیضی (نودها) بیانگر ژن‌ها و یال‌های بین آن‌ها بیانگر اثر متقابل بین ژنی می‌باشد

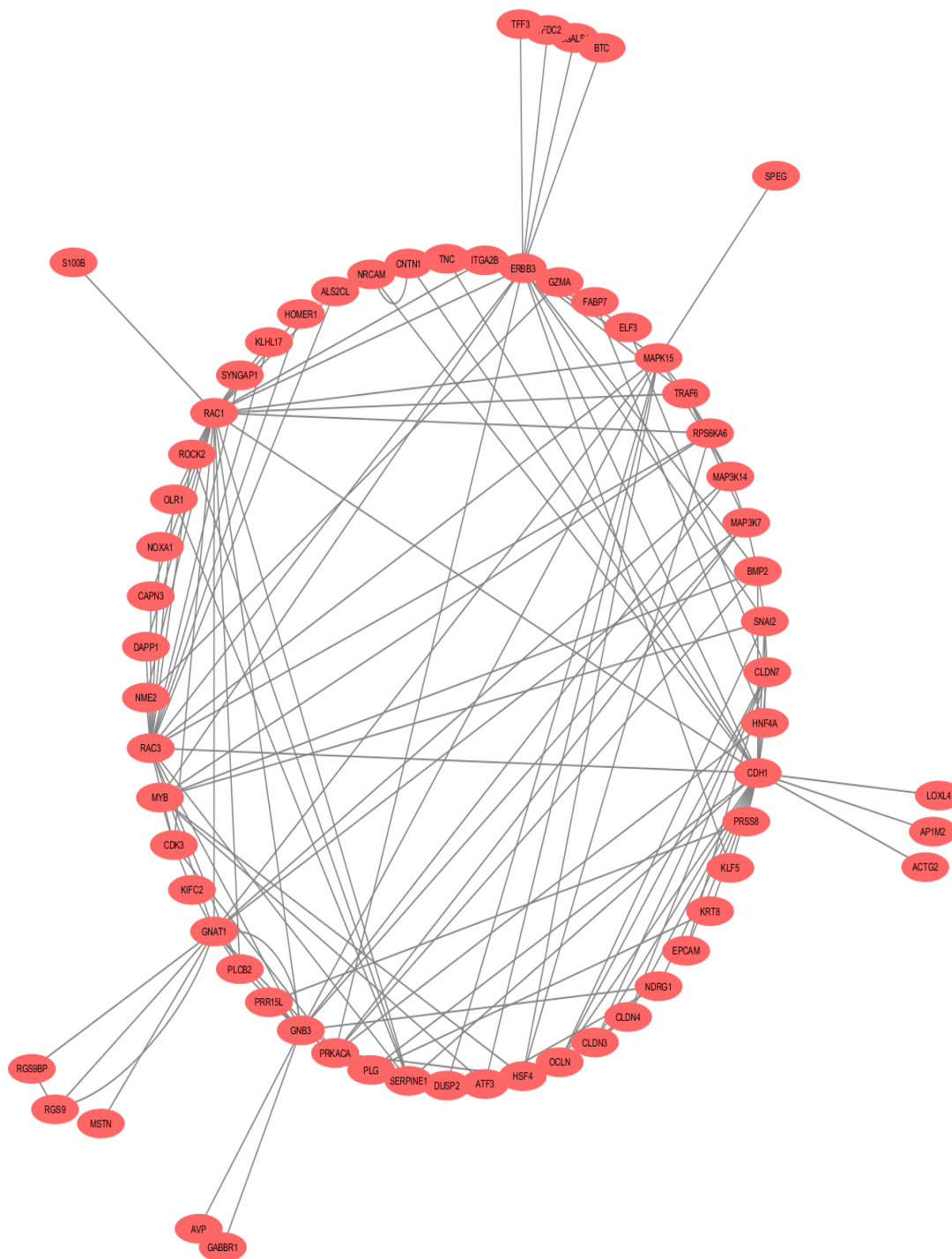
Figure 4. Module 2, which shows the number of genes and interactions between them. Oval shapes (nodes) representing the genes and edges between them represent the interactions between the genes.



شکل ۵. ماژول شماره سه که نشان دهنده تعداد ژن‌ها و اثرات متقابل بین آنهاست. اشکال بیضی (نودها) بیانگر ژن‌ها و یال‌های بین آنها بیانگر اثر متقابل بین ژنی می‌باشد

Figure 5. Module 3, which shows the number of genes and interactions between them. Oval shapes (nodes) representing the genes and edges between them represent the interactions between the genes.

در ماژول ۴، مهم‌ترین ژن‌ها شامل ITGA2B, CDH1, SERPINE1, BMP2, TNC, NRCAM, ERBB3 و CLDN3, CLDN4, MYB بودند که ژن‌های CDH1, SERPINE1, BMP2, TNC, NRCAM, ERBB3 و CLDN3, CLDN4, MYB افزایش بیان و ژن ITGA2B کاهش بیان از خود نشان داد. همه ژن‌های این ماژول در فرآیند چسبندگی سلول نقش داشتند (شکل ۶).



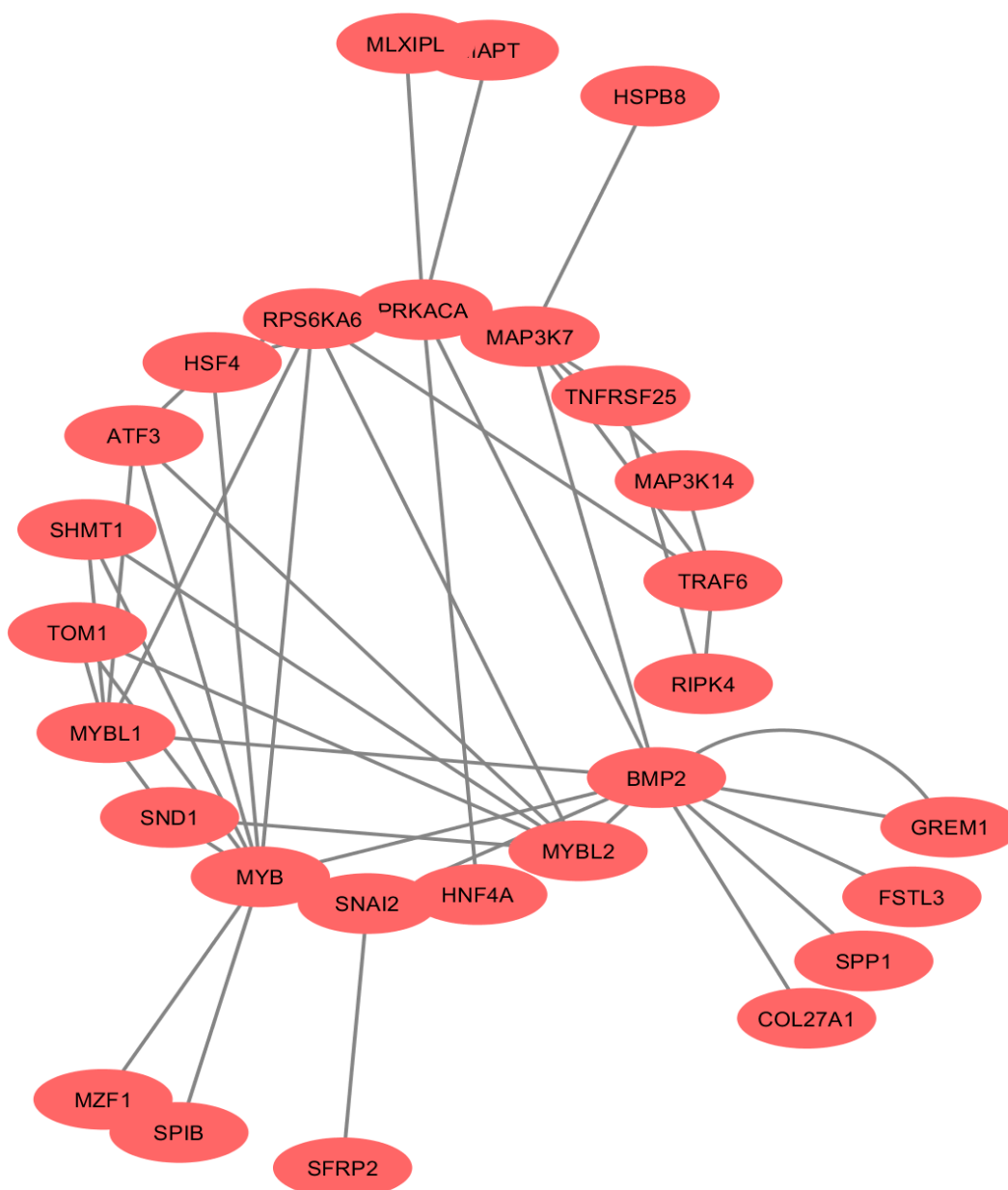
شکل ۶. ماژول شماره چهار که نشان دهنده تعداد ژن‌ها و اثرات متقابل بین آن‌هاست. اشکال بیضی (نودها) بیانگر ژن‌ها و یال‌های بین آن‌ها بیانگر اثر متقابل بین ژنی می‌باشد

Figure 6. Module 4, which shows the number of genes and interactions between them. Oval shapes (nodes) representing the genes and edges between them represent the interactions between the genes.

در ماژول ۵، ژن های BMP2 و ATF3, HSF4, COL27A1, MYB, MAPT, SPIB, SFRP2, GREM1

شناسایی شدند که مهم ترین ژن های این ماژول شامل BMP2 و ATF3, MYB, SPIB, SFRP2, GREM1 افزایش

بیان و ژن های HSF4, COL27A1 و MAPT کاهش بیان از خود نشان دادند (شکل ۷).

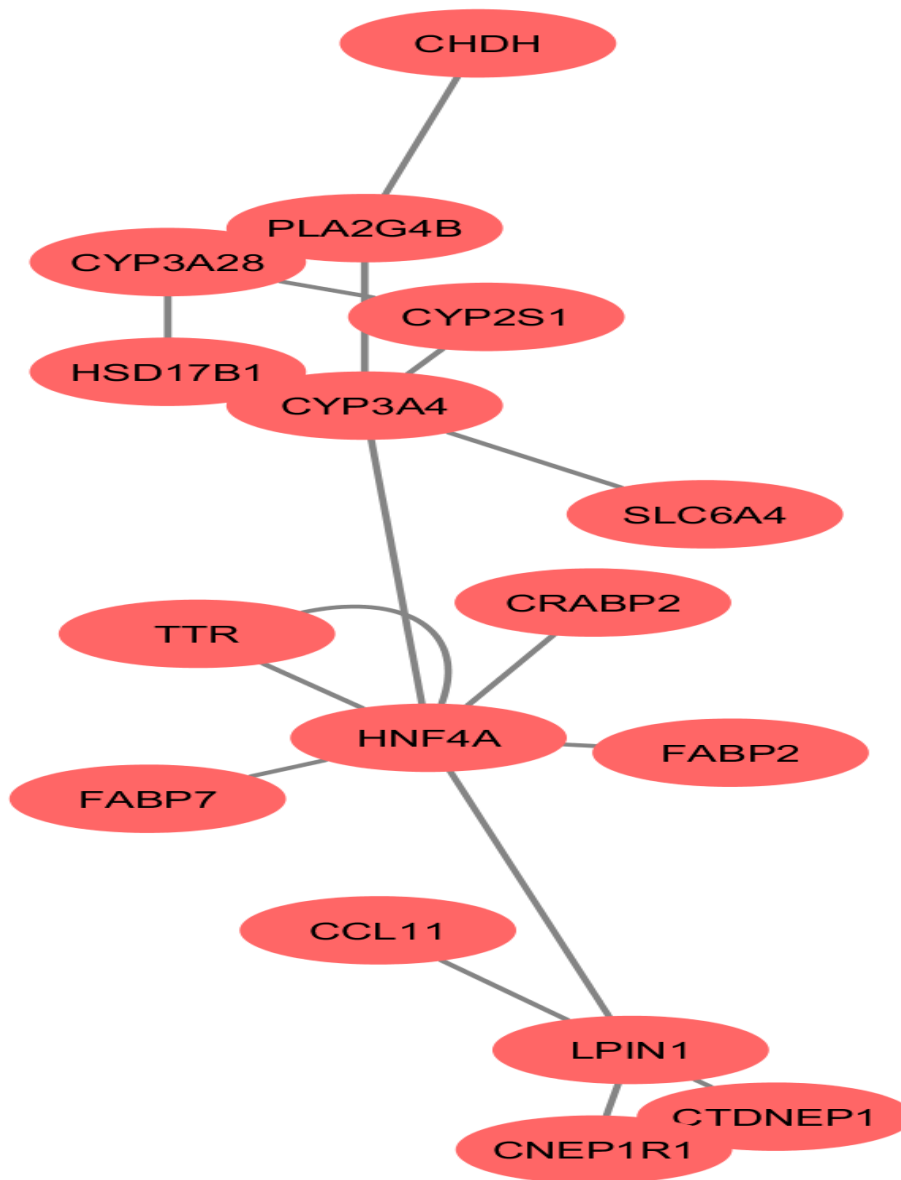


شکل ۷. ماژول شماره پنج که نشان دهنده تعداد ژن ها و اثرات متقابل بین آنهاست. اشکال بیضی (نودها)

بیانگر ژن ها و یال های بین آنها بیانگر اثر متقابل بین ژنی می باشد

Figure 7. Module 5, which shows the number of genes and interactions between them. Oval shapes (nodes) representing the genes and edges between them represent the interactions between the genes.

در ماژول ۶، مهم ترین ژن ها شامل LPIN1, HSD17B1, PLA2G4B, CYP2S1, HNF4A, TTR و CRABP2 بودند که ژن های HSD17B1, CYP2S1, HNF4A و CRABP2 افزایش و ژن های LPIN1 و PLA2G4B کاهش بیان از خود نشان دادند. این ژن ها در فرآیند متابولیسم چربی نقش خود را ایفا می کردند (شکل ۸).



شکل ۸. ماژول شماره شش که نشان دهنده تعداد ژن ها و اثرات متقابل بین آنهاست. اشکال بیضی (نودها) بیانگر ژن ها و یال های بین آنها بیانگر اثر متقابل بین ژنی می باشد.

Figure 8. Module 6, which shows the number of genes and interactions between them. Oval shapes (nodes) representing the genes and edges between them represent the interactions between the genes.

نتیجه گیری

در این تحقیق با استفاده از روش محاسباتی جدید به بازسازی شبکه هم بیان ژنی پرداخته شد که این شبکه بر اساس استفاده از پروفایل بیان ژن ها بود. این تحقیق می تواند افق جدیدی پیش روی ما برای شناسایی فرآیندهای زیستی در سطح مولکول ارائه نماید. اطلاعاتی که در این تحقیق گزارش شده است ممکن است برای شناسایی و فهم سازو کار مولکول باروری در گاو مفید باشد. هر چند آزمایشات بیشتری برای فهم نقش عملکردی ماژول های ژنی در مراحل مختلف صفات مرتبط با باروری نیاز است.

منابع

- پسندیده مجید، خراتی کوپایی حامد، محمدآبادی محمدرضا، اسماعیلی زاده کشکوئیه علی (۱۳۹۵) ارتباط آلل های ژن های OPN و PPARGC1A با تعداد سلول های بدنی شیر در گاوهای هلشتاین ایران. مجله ژنتیک نوین ۱۱(۳)، ۳۵۷-۳۶۵.
- توحیدی نژاد فاطمه، محمدآبادی محمدرضا، اسماعیلی زاده کشکوئیه علی، نجمی نوری عذرا (۱۳۹۳) مقایسه سطوح مختلف بیان ژن Rheb در بافت های مختلف بز کرکی راینی. مجله بیوتکنولوژی کشاورزی ۶(۴)، ۳۵-۵۰.
- جعفری دره در امیر حسین، محمدآبادی محمدرضا، اسماعیلی زاده کشکوئیه علی، ریاحی مدوار علی (۱۳۹۵) بررسی بیان ژن CIB4 در بافت های مختلف گوسفند کرمانی با استفاده از Real Time qPCR. مجله پژوهش در نشخوارکنندگان ۴(۴)، ۱۳۲-۱۱۹.
- خراتی کوپایی حامد، محمدآبادی محمدرضا، انصاری مهباری سعید، اسماعیلی زاده کشکوئیه علی، ترنگ علیرضا، نیکبختی مهدی (۱۳۹۰) بررسی تنوع ژنتیکی جایگاه ژن DGAT1 و میزان ارتباط آن با تولید شیر در جمعیت گاو های هلشتاین ایران. مجله پژوهش های علوم دامی ایران ۳(۲)، ۱۸۵-۱۹۲.
- خراتی کوپایی حامد، محمدآبادی محمدرضا، ترنگ علیرضا، خراتی کوپایی محمود، اسماعیلی زاده کشکوئیه علی (۱۳۹۱) بررسی ارتباط چند شکلی آلی ژن DGAT با بیماری ورم پستان در جمعیت گاوهای هلشتاین ایران. مجله ژنتیک نوین ۷(۱)، ۱۰۴-۱۰۱.
- محمدآبادی محمدرضا، محمدی اکرم (۱۳۸۹) بررسی ژنوتیپ های بتالاکتوگلوبولین در گاوهای بومی و هلشتاین استان کرمان. مجله تولیدات دامی ۱۲(۲)، ۶۱-۶۷.

Reference

- Alinaghizadeh R, Mohammad Abadi MR, Moradnasab Badrabadi S (2007) Kappa-casein gene study in Iranian Sistani cattle breed (*Bos indicus*) using PCR-RFLP. *Pak J Biol Sci*, 10, 4291-4294.
- Askari N, Mohammadabadi MR, Baghizadeh A (2011) ISSR markers for assessing DNA polymorphism and genetic characterization of cattle, goat and sheep populations. *Iran J Biotechnol* 9, 222-229.
- Bader GD, Hogue CW (2003) An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinform* 4, 1-10.
- Bahrami A, Miraie-Ashtiani SR, Sadeghi M, Najafi A (2017) miRNA-mRNA network involved in folliculogenesis interactome: systems biology approach. *Reproduction* 154, 51-65.
- Bansal M, Belcastro V, Ambesi-Impombato A, Di Bernardo D (2007) How to infer gene networks from expression profiles. *Mol System Biol* 3, 78-87.
- Bao B, Garverick HA (1998) Expression of steroidogenic enzyme and gonadotropin receptor genes in bovine follicles during ovarian follicular waves: a review. *J Anim Sci* 76, 1903–1921.
- Barazandeh A, Mohammadabadi MR, Ghaderi M, Nezamabadipour H (2016) Predicting CpG Islands and Their Relationship with Genomic Feature in Cattle by Hidden Markov Model Algorithm. *Iran J App Anim Sci* 6, 571-579.
- Barbat, A, Le Mézec P, Ducrocq S, Humblot P (2010) Female fertility in French dairy breeds: Current situation and strategies for improvement. *J Reprod and Devel* 56, S15–S21.
- Barrett T, Wilhite SE, Ledoux P et al. (2013) NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research* 41, 991–995.
- Basso K, Saito M, Sumazin P et al. (2010) Integrated biochemical and computational approach identifies *BCL6* direct target genes controlling multiple pathways in normal germinal center B cells. *Blood*, 115, 975e984.
- Blais A, Dynlacht BD (2005) Constructing transcriptional regulatory networks. *Genes and Development*, 19, 1499e1511.
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Bugrim A, Nikolskaya T, Nikolsky Y (2004) Early prediction of drug metabolism and toxicity: systems biology approach and modeling. *Drug Dis Tod* 9, 127–135.
- Chatr-Aryamontri A, Breitkreutz BJ, Heinicke S et al. (2013) The BioGRID interaction database: 2013 update. *Nucleic Acids Research* 41, 816–823.

- Chen L, Li W, Zhang L et al. (2011) Disease gene interaction pathways: a potential framework for how disease genes associate by disease-risk modules. *PLoS One* 6, e24495.
- Ebrahimi Z, Mohammadabadi MR, Esmailizadeh AK et al. (2015a) Association of PIT1 gene with milk fat percentage in Holstein cattle. *Iran J Appl Anim Sci* 5, 575-582.
- Ebrahimi Z, Mohammadabadi MR, Esmailizadeh AK, Khezri A (2015b) Association of PIT1 gene and milk protein percentage in Holstein cattle. *J Livest Sci Technol* 3, 41-49.
- Emmert-Streib F, de Matos Simoes R, Mullan P et al. (2014) The gene regulatory network for breast cancer: integrated regulatory landscape of cancer hallmarks. *Frontiers in Genet* 5, 15.
- Forde N, Mehta J, McGettigan P et al. (2013) Alterations in expression of endometrial genes coding for proteins secreted into the uterine lumen during conceptus elongation in cattle. *BMC Genom* 14, 321.
- Forde N, McGettigan PA, Mehta JP et al. (2014) Proteomic analysis of uterine fluid during the pre-implantation period of pregnancy in cattle. *Reproduction* 147, 575–587.
- Forde N, Simintiras CA, Sturmey R et al. (2014) Amino acids in the uterine luminal fluid reflects the temporal changes in transporter expression in the endometrium and conceptus during early pregnancy in cattle. *PLoS One* 9, e100010.
- Ghasemi M, Baghizadeh A, Abadi MRM (2010) Determination of genetic polymorphism in Kerman Holstein and Jersey cattle population using ISSR markers. *Aust J Basic Appl Sci* 4, 5758-5760.
- Hayashi K-G, Ushizawa K, Hosoe M, Takahashi T (2010) Differential genome wide gene expression profiling of bovine largest and second-largest follicles: identification of genes associated with growth of dominant follicles. *Reprod Biol Endo* 8, 11.
- Jafari Darehdor AH, Mohammadabadi MR, Esmailizadeh AK, Riahi Madvar A (2016) Investigating expression of CIB4 gene in different tissues of Kermani Sheep using Real Time qPCR. *J Rumin Res* 4, 119-132 (In Persian).
- Javanmard A, Mohammadabadi MR, Zarrigabayi GE et al. (2008) Polymorphism within the intron region of the bovine leptin gene in Iranian Sarabi cattle (Iranian *Bos taurus*). *Russ J Genet* 44, 495-497.
- Kann MG (2007) Protein interactions and disease: computational approaches to uncover the etiology of diseases. *Brief Bioinformatics* 8, 333–346.
- Kolesnikov N, Hastings E, Keays M et al. (2015) Array Express update—simplifying data submissions. *Nucleic Acids Res* 43, 1113–1116.

- Kharrati Koopaei H, Mohammad Abadi MR, Ansari Mahyari S et al. (2012a) Effect of DGAT1 variants on milk composition traits in Iranian Holstein cattle population. *Anim Sci Papers Report* 30, 231-240.
- Kharrati Koopaei H, Mohammadabadi MR, Ansari Mehryari S et al. (2011) Genetic Variation of DGAT1 Gene and its Association with Milk Production in Iranian Holstein Cattle Breed Population. *Iran J Anim Sci Res* 3, 185-192 (In Persian).
- Kharrati koopaei H, Mohammadabadi MR, Tarang A et al. (2012b) Study of the association between the allelic variations in DGAT1 gene with mastitis in Iranian Holstein cattle. *Modern Genet J* 7, 101-104 (In Persian).
- Kim D, Pertea G, Trapnell C et al. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Gen Biol* 14, 36.
- Lucy MC (2001) Reproductive loss in high-producing dairy cattle: Where will it end?. *J Dairy Sci* 84, 1277-1293.
- Mohammad Abadi MR, Mohammadi A (2010) Study of beta-lactoglobulin genotypes in native and Holstein cattle of Kerman province. *J Anim Prod* 12, 61-67 (in Persian).
- Mohammadabadi M R, Shaikhaev GO, Sulimova GE et al. (2004) Detection of bovine leukemia virus proviral DNA in Yaroslavl, Mongolian and black pied cattle by PCR. *Cell Mol Biol Letter* 9, 766-768.
- Mohammadabadi MR, Jafari AHD, Bordbar F (2017) Molecular analysis of CIB4 gene and protein in Kermani sheep. *Brazil J Med Biol Res* 50, e6177.
- Mohammadabadi MR, Kovalenko TA, Nasiri MR, Sulimova GE (2005) inter-Simple-Sequence Repeat (ISSR)-PCR for the identification of polymorphism in some native cattle breeds. *Proceeding of the 3rd Moscow international congress: Biotechnology: state of the art and prospects of development, Moscow, Russia.*
- Mohammadabadi MR, Soflaei M, Mostafavi H, Honarmand M (2011) Using PCR for early diagnosis of bovine leukemia virus infection in some native cattle. *Genet Mol Res* 10, 2658-2663.
- Mohammadabadi MR, Tohidinejad F (2017) Characteristics determination of Rheb gene and protein in Raini Cashmere goat. *Iran J Appl Anim Sci* 7, 289-295.
- Moore SG, Pryce JE, Hayes BJ, Chamberlain AJ (2016) Differentially Expressed Genes in Endometrium and Corpus Luteum of Holstein Cows Selected for High and Low Fertility Are Enriched for Sequence Variants Associated with Fertility. *Biol Reprod* 94(1), 19.

- Moran B, Cummins SB, Creevey CJ, Butler ST (2016) Transcriptomics of liver and muscle in Holstein cows genetically divergent for fertility highlight differences in nutrient partitioning and inflammation processes. *BMC Gen* 17, 603.
- Moran B, Butler ST, Moore SG, MacHugh DE (2015) Differential gene expression in the endometrium reveals cytoskeletal and immunological genes in lactating dairy cows genetically divergent for fertility traits. *Reprod Fert Devel* 25.
- Mostafavi S, Ray D, Warde-Farley D et al. (2008) GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Gen Biol* 1, 4.
- Nepusz T, Yu H, Paccanaro A (2012) Detecting overlapping protein complexes in protein-protein interaction networks. *Nat Meth* 9, 471–472.
- Pasandideh M, Kharrati Koopae H, Mohammad Abadi MR, Esmailzadeh Koshkoiyeh A (2016) Association of the OPN and PPARGC1A Alleles with Milk Somatic Cell Count in Iranian Holstein cattle. *Modern Genet J* 11, 357-365 (In Persian).
- Pasandideh M, Mohammadabadi MR, Esmailzadeh AK, Tarang A (2015) Association of bovine PPARGC1A and OPN genes with milk production and composition in Holstein cattle. *Czech J Anim Sci* 60, 97-104.
- Robinson RS, Woad KJ, Hunter MG et al. (2014) Corpus luteum development and angiogenesis. In: Juengel JL, Miyamoto A, Price C, Reynolds LP, Smith MF, Webb R (eds.), *Reproduction in Domestic Ruminants VIII*. London: Soc Reprod Fert 1, 327–343.
- Robinson M, McCarthy D, Smyth G (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.
- Ruzina MN, Shtyfurko TA, Mohammadabadi MR et al. (2010) Polymorphism of the BoLA-DRB3 gene in the Mongolian, Kalmyk, and Yakut cattle breeds. *Russ J genet* 46, 456-463.
- Sulimova GE, Abani MA, Rostamzadeh J et al. (2007) Allelic polymorphism of kappa-casein gene (CSN3) in Russian cattle breeds and its informative value as a genetic marker. *Russ J Genet* 43, 88-95.
- Tohidi nezhad F, Mohammadabadi MR, Esmailzadeh AK, Najmi Noori A (2015) Comparison of different levels of Rheb gene expression in different tissues of Raini Cashmir goat. *Agric Biotechnol J* 6, 35-50 (In Persian).
- Trapnell C, Williams BA, Pertea G et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotech* 28, 511–515.

Vitt U, Hayashi M, Klein C, Hsueh A (2000) Growth differentiation factor-9 stimulates proliferation but suppresses the follicle-stimulating hormone-induced differentiation of cultured granulosa cells from small antral and preovulatory rat follicles. *Biol Reprod* 62, 370–377.

Walsh SW, Williams EJ, Evans ACO (2011) A review of the causes of poor fertility in high milk producing dairy cows. *Anim Reprod Sci* 123, 127–138.