



Study of differential gene expression of Holstein cow's fertility in corpus luteum using RNA-Seq data

Ghorban Elyasi Zarringhabaie ¹, Mostafa Sadeghi ², Seyed Reza Miraie Ashtiani ³

1. Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: Gh.elyasi@ut.ac.ir

2. Corresponding Author, Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: Sadeghimos@ut.ac.ir

3. Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran. E-mail: Ashtiani@ut.ac.ir

Article Info	ABSTRACT
Article type: Research Article	The fertility of female cows is the main factor for the survival of dairy farming. This study was performed to determine the differentially expressed genes in fertility in Holstein cows using phenotypic data and next-generation sequencing (NGS) technology. Corpus luteum samples from 18 Holstein cows at up to two parturitions were used in two high and low fertility groups, three samples from each group were pooled together and a total of three biological samples from each were submitted for whole genome sequencing. The preparation and analysis steps were performed on the Galaxy 22.01 platform. DAVID (2021) database were used for gene ontology. The results of differential gene expression analysis in luteal tissue showed that 19 genes out of 13049 expressed transcripts had a significant difference in expression between high and low-fertility cows with an FDR-adjusted p-value (q-value) of less than 0.05. UBE3B, NIF3L1 and ORC2 genes were three functional marker genes that were more highly expressed in the corpus luteum of high-fertility cows. The UBE3B gene is involved in biological processes as a protein involved in the catabolic process of ubiquitin-dependent protein. NIF3L1 is involved in the biological process of neuronal differentiation, positive transcriptional regulation, DNA pattern, and negative transcriptional regulation by nucleic acid pattern. The ORC2 gene plays a role in the biological process of DNA replication. The genes KRT8, PHLDB3, PPT2, LOC787628, PPYR1, TOX, TP73, DHX8, KCNN1, CLEC6A, PXMP4, LRRC26, SLC34A3, and OR13C7 have been shown to be more highly expressed in the corpus luteum in low fertility cows. Among these genes, three genes PPT2, PPYR1 and CLEC6A were functional indicators. The loci UBE3B, NIF3L1 and ORC2, which were expressed more strongly in the corpus luteum of highly fertile cows, can be used as functional indicator genes in the genetic selection of dairy cows.
Article history: Received: 26 June 2022 Received: 29 November 2022 Accepted: 3 December 2022 Published online: 23 September 2023	
Keywords: <i>Alignment,</i> <i>Galaxy, Holstein,</i> <i>Ontology,</i> <i>Transcriptome.</i>	

Cite this article: Zarringhabaie, Gh. E., Sadeghi, M., & Miraie Ashtiani, S. R (2023). Study of differential gene expression of Holstein cow's fertility in corpus luteum using RNA-Seq data. *Iranian Journal of Animal Science*, 54 (3), 239-251. DOI: <http://doi.org/10.22059/IJAS.2022.344531.653893>



© The Author(s).

DOI: <http://doi.org/10.22059/IJAS.2022.344531.653893>

Publisher: The University of Tehran Press.

Extended Abstract

Objective

The fertility of female cows is the main factor for the survival of dairy farming. The interval between calvings and the response to the first insemination are used as the most important fertility traits. Infertility and low fertility are important and widespread problems in domestic animals, and the main limitation of reproductive efficiency in mammalian species is embryonic mortality. This study was performed to determine the differentially expressed genes in fertility in Holstein cows using phenotypic data and next-generation sequencing (NGS) technology.

Materials and Methods

Corpus luteum samples from 18 Holstein cows at up to two parturitions were used in two high and low

fertility groups, three samples from each group were pooled together and a total of three biological samples from each were submitted for whole genome sequencing. RNA sequencing was performed with Illumina HiSeq 2500 using paired-end reads, each read being 75 bp in length. The preparation and analysis steps were performed on the Galaxy platform. Quality control was performed with Galaxy FastQC. For low-quality samples, the Trimmomatic software was used to remove substandard samples. HISAT2 software was used to match the transcriptome reads to the cow reference genome. After mapping the transcripts to the reference genome, the expression level of each transcript was counted using the featureCounts software. Differential gene expression analysis was used for two treatment groups using DESeq2. Gene ontology using the DAVID database was performed for genes that showed significantly different expressions between the two treatments.

Results

Results from the Trimmomatic software showed that only 0.87% of the data was deleted and approximately 99.13% of the raw data was used for mapping to the reference genome. The results of the alignment showed that 93.48% of the transcriptome sequence mapped to the reference genome, 90.58% to one site, and only 2.90% to more than one site. The results of differential gene expression analysis in luteal tissue showed that 19 genes out of 13049 expressed transcripts had a significant difference in expression between high and low-fertility cows with an FDR-adjusted p-value (q-value) of less than 0.05. UBE3B, NIF3L1 and ORC2 genes were three functional marker genes that were more highly expressed in the corpus luteum of high-fertility cows. The UBE3B gene is involved in biological processes as a protein involved in the catabolic process of ubiquitin-dependent protein. NIF3L1 is involved in the biological process of neuronal differentiation, positive transcriptional regulation, DNA pattern, and negative transcriptional regulation by nucleic acid pattern. The ORC2 gene plays a role in the biological process of DNA replication. The genes KRT8, PHLDB3, PPT2, LOC787628, PPYR1, TOX, TP73, DHX8, KCNN1, CLEC6A, PXMP4, LRRC26, SLC34A3, and OR13C7 have been shown to be more highly expressed in the corpus luteum in low fertility cows. Among these genes, three genes PPT2, PPYR1 and CLEC6A were functional indicators. The PPT2 gene is responsible for the removal of palmitoyl groups from a macromolecule, the PPYR1 gene is involved in the biological process of cell surface receptor signaling pathway, neuropeptide signaling pathway, nutritional behavior and blood circulation and the CLC34A3 gene plays a role in the biological process of cellular phosphate ion homeostasis.

Conclusion

the loci UBE3B, NIF3L1 and ORC2, which were expressed more strongly in the corpus luteum of highly fertile cows, can be used as functional indicator genes in the genetic selection of dairy cows.



مطالعه بیان افتراقی ژن‌های مرتبط با باروری در بافت جسم زرد گاو نژاد هلشتاین با استفاده از داده‌های RNA-Seq

قربان الیاسی زرین قبایی^۱ | مصطفی صادقی^۲ | سیدرضا میرائی آشتیانی^۵

۱. گروه علوم دامی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانه Gh.elyasi@ut.ac.ir
 ۲. نویسنده مسئول، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانه Sadeghimos@ut.ac.ir
 ۳. گروه علوم دامی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران. رایانه Ashtiani@ut.ac.ir

اطلاعات مقاله	چکیده
<p>نوع مقاله: مقاله پژوهشی</p> <p>تاریخ دریافت: ۱۴۰۱/۰۴/۰۵</p> <p>تاریخ بازنگری: ۱۴۰۱/۰۹/۰۱</p> <p>تاریخ پذیرش: ۱۴۰۱/۰۹/۱۲</p> <p>تاریخ انتشار: ۱۴۰۲/۰۷/۰۱</p> <p>کلیدواژه‌ها: ترانسکریپتوم، گالاکسی، هستی‌شناسی، هلشتاین، هم‌ترازی.</p>	<p>صفت باروری بعنوان عامل اصلی برای تداوم سودآوری در مزارع گاو شیرده محسوب می‌شود. پژوهش حاضر، به منظور تعیین ژن‌های بیان شده افتراقی در باروری گاو نژاد هلشتاین با استفاده از داده‌های ترانسکریپتوم صورت گرفت. بدین منظور، در مجموع نمونه بافت جسم زرد ۱۸ رأس گاو نژاد هلشتاین با نوبت زایش ۲ به بالا که به‌صورت ترکیب شد (Pooling) در دو گروه با باروری زیاد و باروری کم هر کدام در سه تکرار بیولوژیک تقسیم‌بندی شده بودند استفاده شد. تمامی مراحل آماده‌سازی توالی‌های خام و تجزیه و تحلیل آن‌ها در پلتفرم گالاکسی نسخه ۲۲/۰۱ انجام شد و متعاقباً، هستی‌شناسی ژن با استفاده از پایگاه داده DAVID نسخه ۲۰۲۱ صورت پذیرفت. در نهایت، نتایج آنالیز بیان افتراقی ژن در بافت جسم زرد نشان داد که از ۱۳۰۴۹ رونوشت ترانسکریپتومی بیان شده، ۱۹ ژن به‌عنوان ترانسکریپتوم‌های شاخص عمل کرده‌اند ($q\text{-value} \leq 0.05$). همچنین، ژن‌های UBE3B، NIF3L1، ORC2 سه ژن شاخص عملکردی بودند که مقدار بیان بیشتری در جسم زرد گاوهای با باروری زیاد داشتند. ژن UBE3B در فرآیندهای بیولوژیک به‌عنوان پروتئین درگیر در فرآیند کاتابولیک پروتئین وابسته به یوئیکوتیتین است. در جهت تفسیر بیشتر نتایج، ژن NIF3L1 در فرآیند بیولوژیک تمایز نورو، تنظیم مثبت رونویسی، الگوی DNA، تنظیم منفی رونویسی با الگوی اسید نوکلئیک، نقش دارد. همچنین، ژن ORC2 در فرآیند بیولوژیک همانندسازی DNA نقش بازی می‌کند. ژن‌های KRT8، PHLDB3، PPT2، LOC787628، PPYR1، TOX، TP73، DHX8، KCNN1، CLEC6A، PXMP4، LRRC26، SLC34A3 و OR13C7 در گاوهای با باروری کم، مقدار بیان بیشتری در جسم زرد نشان داده است. سه ژن PPYR1، PPT2 و CLEC6A جزو ژن‌های شاخص محسوب می‌شوند. به‌عنوان نتیجه گیری کلی، جایگاه‌های UBE3B، NIF3L1، ORC2 که مقدار بیان بیشتری در جسم زرد گاوهای با باروری زیاد داشتند می‌توانند بعنوان ژن‌های شاخص عملکردی در انتخاب ژنتیکی گاوهای شیری مورد استفاده قرار گیرند.</p>

استناد: الیاسی زرین قبایی، قربان؛ صادقی، مصطفی؛ و میرائی آشتیانی، سیدرضا (۱۴۰۲). مطالعه بیان افتراقی ژن‌های مرتبط با باروری در بافت جسم زرد گاو نژاد هلشتاین با استفاده از داده‌های RNA-Seq. *نشریه علوم دامی ایران*، ۵۴ (۳)، ۲۵۱-۲۳۹. DOI: <http://doi.org/10.22059/IJAS.2022.344531.653893>



© نویسندگان.

DOI: <http://doi.org/10.22059/IJAS.2022.344531.653893>

ناشر: مؤسسه انشارات دانشگاه تهران.

مقدمه

در ارزیابی‌های ژنتیکی گاوهای شیرده، توان باروری و صفات مربوط به بقا در مدل حیوانی چند صفتی با همدیگر ارزیابی می‌گردند که شامل چند صفت ترکیبی بوده و به‌عنوان نشانگر خوبی برای صفات تولید می‌باشند (Olori *et al.* 2003). فاصله گوساله‌زایی بعنوان صفت اصلی باروری مورد استفاده قرار می‌گیرد و همراه با پاسخ به تلقیح اول، تعداد تلقیح و سن در اولین زایش به کار گرفته می‌شود. استفاده از فاصله گوساله‌زایی به‌عنوان تعریفی از صفت باروری در بسیاری از کشورها بعنوان سیستم‌های زایش در طول سال مطرح می‌باشد؛ بنابراین، بهبود باروری بازده تولید گاوهای شیری را افزایش خواهد داد و توسعه عملکرد تولیدمثلی در گاو می‌تواند ماندگاری یک گاو، کاهش فاصله گوساله‌زایی و کاهش هزینه‌های دامپزشکی را به دنبال داشته باشد (Shahabi *et al.* 2019)؛ بنابراین، باروری گاوها فاکتور اصلی برای پایداری یک گاو‌داری می‌باشد هرچند از سال‌ها پیش باروری در گاو کاهش یافته است و تبدیل به یکی از دلایل اصلی حذف و جایگزینی گاوها شده است (Barbat *et al.* 2010). بدین معنی که در پرورش گاو شیرده صفات مرتبط با تولید شیر با افزایش قابل توجهی مواجه شده و به‌طور هم‌زمان، صفات باروری در گاو کاهش یافته است که در نتیجه همبستگی منفی تولید و باروری در گاو شیرده است (Cai *et al.* 2019). ناباروری و کم باروری از مشکلات مهم و فراگیر در حیوانات اهلی و انسان هست و بزرگ‌ترین محدودیت برای کارایی تولیدمثل در گونه‌های پستانداران، مرگومیر رویان است. زنده‌مانی رویان، عامل اصلی مؤثر بر تولید و بازده اقتصادی در کلیه سیستم‌های تولید گوشت و شیر توسط نشخوارکنندگان است (Minten *et al.* 2013). در گاو، تکامل رویان قبل از لانه‌گزینی در مقایسه با سایر گونه‌های پستانداران طولانی‌تر می‌باشد. در طول این مرحله، رویان سالم تحت تقسیم سریع میتوزی قرار می‌گیرد در سه مرحله حیاتی رشد رویان (تراکم، تشکیل بلاستوسیست و طویل شدن) تکامل می‌یابد و به شدت به انواع عوامل بیرونی از جمله محیط خارج سلولی رحم وابسته است (Killeen *et al.* 2016). علاوه بر این، تغییرات در مراحل مختلف چرخه فحلی به دلیل مکمل‌های غذایی، تغییرات در سطوح پروژسترون، عفونت و آبستنی رخ می‌دهد (Killeen *et al.* 2016). مسیرهای بیولوژیک متعددی در فرآیندهای مرتبط با باروری در گاو شیری وجود دارد بنابراین، باروری به‌عنوان یک صفت پیچیده و مرکب تحت تأثیر بسیاری از ژن‌ها و واریانت‌ها است که هر کدام به‌طور طبیعی دارای اثرات کوچکی هستند و این بدان معنی است که قدرت یافتن واریانت‌هایی ژنتیکی فردی با اثرات اصلی بالقوه پایین است (Snelling *et al.* 2013).

پیشینه پژوهش

مطالعات GWAS^۵ بعنوان، یکی از تکنیک‌های ارتباط ژن با عملکردهای بیولوژیک صورت گرفته و ژن FAM181A به‌عنوان ژن مرتبط با زایش در اولین سرویس آبستنی و روزهای باز شناخته شده است (Nayeri *et al.* 2016). روش تجزیه و تحلیل ریزآرایه ۴۳۰ ژن متفاوت بیان شده را بین حیوانات با باروری زیاد و باروری کم شناسایی کرد. تجزیه و تحلیل مسیر نبوغ غنی‌سازی برای میزبانی از مسیرهای بیولوژیکی از جمله متابولیسم لیپید، حمل‌ونقل مولکولی، پاسخ ایمنی، مورفولوژی و توسعه سلولی و رشد و تکثیر سلولی را نشان داد ژن متفاوت بیان شده مهم شامل ALB، BMRP2، CCL28، COL4A3/4، FADS1، ITGA6، LDLR، PLCB3، PPARG، PTGS2 و SLC27A4 بود؛ و ژن‌های متفاوت بیان شده در هر دو روز هفتم و چهاردهم آبستنی شامل PCCB، SLC25A24، DAP و COL4A4 است (Killeen *et al.* 2016). با این وجود،

1Calving Interval (CIV)

2Calving to First Service

3Number of Services (NS)

4Age at First Calving (AFC)

5Genome-wide association studies (GWAS)

6Differentially Expressed Genes (DEGs)

فن آوری‌های ریزآرایه برای مطالعات سطح RNA در مقیاس بزرگ، مانند شناسایی ژن‌های بیان شده متفاوت بین مراحل رشد یا بین گروه‌های سالم و بیمار استفاده شده است. با این حال، ماهیت مبتنی بر هیبریداسیون آن، توانایی فهرست‌نویسی و کمی سازی مولکول‌های RNA بیان شده در شرایط مختلف را محدود می‌کند. RNA-seq به دلیل وضوح بهتر و تکرارپذیری بالاتر، به سرعت جایگزین فناوری ریزآرایه شده است (Mortazavi *et al.* 2008). این روش، می‌تواند برای گسترش دانش ما در مورد رویدادهای پیوند جایگزین، ژن‌ها و رونوشت‌های جدید و رونوشت‌های همجوشی استفاده شود؛ بنابراین، با توسعه فناوری‌های تحلیل مولکولی، توالی‌یابی و بیوانفورماتیک، توالی‌یابی عمیق RNA با توان عملیاتی بالا (RNA-seq) بستری را برای اندازه‌گیری الگوهای بیان ژن در مقیاس بزرگ فراهم می‌کند. اخیراً، به منظور شناسایی ژن‌های با بیان متفاوت (DEGs) و واحدهای رونوشت جدید، RNA-seq به طور گسترده در حیوانات اهلی استفاده شده است (Chen *et al.* 2015). برنامه‌های بیوانفورماتیک متعددی برای تجزیه و تحلیل داده‌های RNA-seq توسعه یافته است. حتی ابزارهایی که برای یک هدف توسعه یافته‌اند بر اساس رویکردهای متمایز با استفاده از الگوریتم‌ها و مدل‌های مختلف هستند. تنوع روش این امکان را فراهم می‌کند تا پروتکل‌های تجزیه و تحلیل را با انتخاب برنامه‌ای که بهترین تناسب را برای هر هدف خاص ارائه می‌دهد، سفارشی‌سازی گردد (Yang and Kim, 2015).

در یک مطالعه، تفاوت در بیان ژن بین گوسفند سیاه Qira و تخمدان گوسفند Hetian با استفاده از فناوری RNA-Seq تجزیه و تحلیل شد. گوسفند سیاه قیرا و گوسفند هتیان هم سن و هم جثه بودند و از یک منطقه آمده بودند، اما باروری آن‌ها بسیار متفاوت بود (متوسط نرخ بره زایی گوسفند سیاه قیرا ۲۱۵/۵ درصد و گوسفند هتیان ۱۰۲/۵ درصد بود). در مجموع ۱۲۵۲ ژن به طور قابل توجهی متفاوت بیان شد. تجزیه و تحلیل مسیر KEGG نشان داد که برخی از این ژن‌ها در بیوسنتز استروئید، بیوسنتز هورمون استروئیدی، TGF β ، انسولین، Wnt، Notch و سایر مسیرهای سیگنال دهی نقش دارند. داده‌های ارائه شده نشان داد که تفاوت‌های زیادی در الگوهای بیان ژن تخمدان بین گوسفند سیاه Qira و گوسفند Hetian وجود دارد (Chen *et al.* 2015). علاوه بر این، Minten *et al.* (۲۰۱۳) تنوع ژنتیکی ذاتی احتمالی را بین تلیسه‌های بارور و نابارور/نابارور با علائم بیان آندومتر تغییر یافته و ارتباط‌های ژنتیکی متوسط توصیف کردند (Minten *et al.* 2013). هدف از پژوهش حاضر تعیین ژن‌های با بیان افتراقی در فرایند باروری گاوهای هلشتاین با استفاده از داده‌های RNA-Seq است.

روش‌شناسی پژوهش

در پژوهش حاضر، از داده‌های RNA-Seq مرتبط با ۱۸ رأس گاو هلشتاین گاو‌داری هلدینگ شیرین عسل استفاده شد که دارای نوبت زایش ۲ به بالا بودند که به دو گروه مجزا با باروری زیاد و باروری کم تقسیم‌بندی شدند گاوهایی که دارای باروری زیاد بودند روزهای باز کمتر از ۷۰ روز، فاصله گوساله‌زایی کمتر از ۳۵۰ روز و تعداد تلقیح منجر به آبستنی کمتر از ۱/۳ داشتند و در مقابل گاوهایی که دارای باروری کم بودند روزهای باز بیشتر از ۱۵۰ روز، فاصله گوساله‌زایی بیشتر از ۴۵۰ روز و تعداد تلقیح منجر به آبستنی بیشتر از ۳ داشتند. برای کاهش اثر تعداد نوبت زایش و کاهش هزینه‌های تعیین توالی کل ژنوم نمونه بافت هر سه نمونه با هم با رعایت غلظت ادغام حجمی شده و در مجموع، از هر گروه سه نمونه بیولوژیک برای تعیین توالی کل ژنوم به شرکت BGI^۴ چین ارسال شد. تعیین توالی RNA با استفاده از illumina Hiseq 2500 به روش خوانش دو طرفه تولید شد که طول هر خوانش ۷۵ جفت باز بود. داده‌های خام در فرمت Fastq به حجم ۱۰ گیگابایت و در حدود ۴۰

1 Alternative splicing events
2 Fusion transcripts
3 Pool
4 Beijing Genomics Institute
5 Paired-End

میلیون جفت باز دریافت شد و تمامی مراحل آماده‌سازی و تجزیه و تحلیل آن‌ها در پلتفرم گالاکسی نسخه ۲۲/۰۱ انجام شد (Blankenberg *et al.* 2010). سنجش کیفیت داده‌ها با استفاده از FastQC نسخه ۰/۷۳ نرم‌افزار گالاکسی صورت گرفت (Andrews 2010). برای نمونه‌هایی که دارای کیفیت پایینی بودند از نرم‌افزار Trimmomatic نسخه ۰/۳۸/۰ نرم‌افزار گالاکسی برای حذف توالی‌هایی با کیفیت کم و همچنین حذف آداپتورهایی که در هنگام تعیین توالی به ابتدای خوانشها اضافه شده بود استفاده شد (Bolger *et al.* 2014). برای هم‌ردیفی خوانش‌های ترانسکریپتوم با ژنوم مرجع گاو (Merchant *et al.* 2014) از نرم‌افزار HISAT2 نسخه ۲/۲/۱ نسخه گالاکسی استفاده شد (Kim and Langmead 2015). در ادامه، پس از نقشه یابی رونوشت‌ها بر روی ژنوم مرجع، شمارش میزان بیان هر کدام از ترانسکریپت‌ها با استفاده از نرم‌افزار featureCounts نسخه ۲/۰۱ گالاکسی (Liao *et al.* 2014) صورت گرفت. آنالیز بیان افتراقی برای دو گروه تیماری با استفاده از DESeq2 نسخه ۲/۱۱/۴۰/۷ گالاکسی استفاده شد (Love *et al.* 2014). پس از انجام آنالیز بیان افتراقی ژن‌های بیان شده، برای ژن‌هایی که بین دو تیمار به‌صورت معنی‌داری بیان متفاوت به‌دست آمده بود هستی‌شناسی ژن با استفاده از پایگاه داده DAVID نسخه ۲۰۲۱ صورت پذیرفت (Huang *et al.* 2008).

یافته‌های پژوهش

سنجش کیفیت داده‌ها برای خوانش‌های دو طرفه به طور جداگانه با ۱۰ آزمون مختلف (کیفیت توالی به ازای هر باز ، کیفیت توالی به ازای هر کاشی ، امتیاز کیفیت به ازای توالی ، محتوای توالی به ازای هر باز ، محتوای GC به ازای توالی ، محتوای بازهای نوکلئوتیدی خوانش نشده (N) به ازای هر باز ، توزیع طول توالی‌های نوکلئوتیدی ، سطوح تکراری توالی‌های نوکلئوتیدی ، توالی‌های نوکلئوتیدی موجود بیش از حد ، محتوای آداپتوری) صورت می‌گیرد که نتایج هر آزمون بطور جداگانه و به صورت گرافیکی ارائه می‌گردد. وضعیت قبول، رد یا سطح هشدار برای آزمون‌های ده‌گانه ملاک سنجش کیفیت داده‌ها می‌باشد ولی پارامتر کیفیت توالی نوکلئوتیدی به ازای هر باز از اهمیت بیشتری برخوردار هست؛ که بعنوان مثال نتیجه یکی از تکرارهای باروری زیاد در شکل ۱ آورده شده است.

1 <https://usegalaxy.org/22.01>

2 Gene Ontology

3 Per base sequence quality

4 Per tile sequence quality

5 Per sequence quality scores

6 Per base sequence content

7 Per sequence GC content

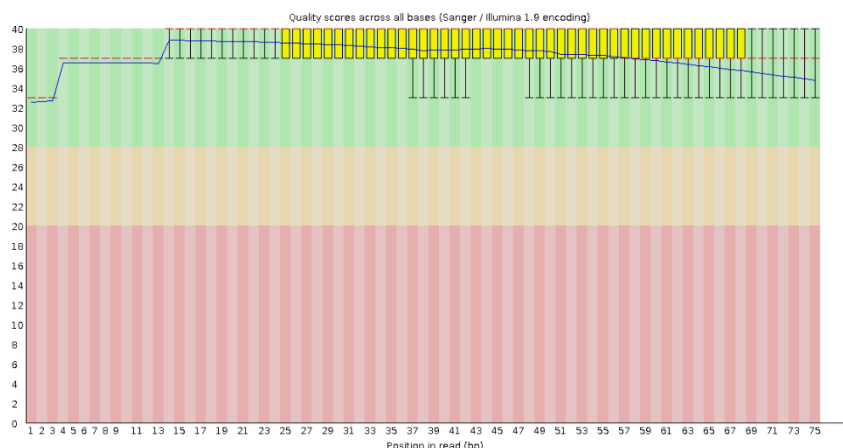
8 Per base N content

9 Sequence Length Distribution

10 Sequence Duplication Levels

11 Overrepresented sequences

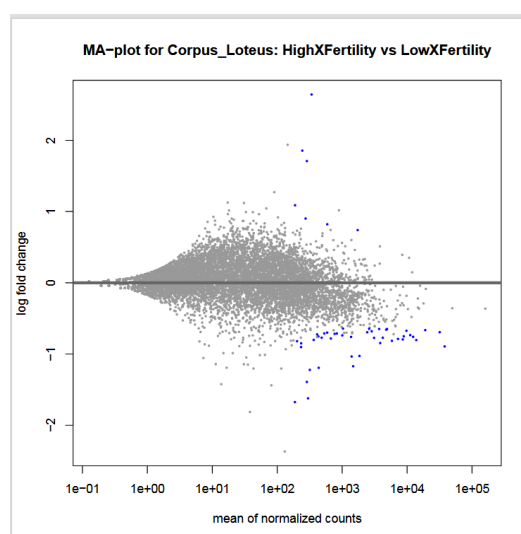
12 adapter Content



شکل ۱. مشاهده خروجی امتیاز کیفیت به ازای هر توالی بازی

با توجه به کیفیت بالای داده‌های حاصل از تعیین توالی به روش NGS^۱ و با توجه به اینکه باقیمانده آداپتورهای الصاقی تعیین توالی نیز در داده‌های خام وجود نداشت نتایج مرحله بعد از ویرایش با نرم‌افزار Trimmomatic (Bolger *et al.* 2014) نشان داد که تنها ۰/۸۷ درصد داده‌ها حذف و در حدود ۹۹/۱۳ درصد داده‌های خام برای هم‌ترازی با ژنوم مرجع استفاده شده است این در حالی است که حذف داده‌های با کیفیت پایین می‌تواند به مراتب بیشتر از این مقدار باشد (Attari *et al.* 2019). نتایج حاصل از هم‌ترازی با ژنوم مرجع نشان داد که بطور میانگین ۹۳/۴۸ درصد از توالی ترانسکریپتومی بر روی ژنوم مرجع نقشه یابی شدند که از این میزان ۹۰/۵۸ درصد بصورت اختصاصی بر روی یک ناحیه و تنها ۲/۹۰ به بیش از ۱ جایگاه اتصال پیدا کردند. لذا، از کل داده‌های RNA-Seq تنها ۶/۵۲ درصد داده‌ها توسط نرم‌افزار HISAT2 نسخه ۲/۲/۱ نسخه گالاکسی (Kim *et al.* 2015) بر روی ژنوم مرجع شناسایی نشد؛ که آنالیز میزان بیان با استفاده از نرم‌افزار FeatureCounts نسخه ۲/۰/۱ گالاکسی (Liao *et al.* 2014) بر روی خوانش‌های هم‌تراز شده با ژنوم مرجع صورت پذیرفت.

نتایج آنالیز بیان افتراقی ژن در بافت جسم زرد نشان داد که ۱۳۰۴۹ رونوشت ترانسکریپتومی بیان شده بود (شکل ۲) که از این تعداد ۳۶۵ ژن با سطح معنی‌داری ۰/۰۵ تفاوت بیان در گاوهایی با باروری زیاد و باروری کم نشان دادند از آنجایی که ارزش P در داده‌های حاصل از RNA-Seq به تنهایی قابل اعتماد نیست از مقدار P همسان‌سازی شده با نرخ کشف اشتباه (q-value) استفاده شد که ۱۹ ژن با سطح معنی‌داری کمتر ۰/۰۵ بعنوان ترانسکریپتوم‌های شاخص شناسایی شدند (جدول ۱). از این تعداد ژن برای ۲ عدد ترانسکریپتوم معنی‌دار (ENSBTAG00000048513 و ENSBTAG00000050175) اطلاعات نام و مشخصات ژن در پایگاه‌های داده بدست نیامد همچنین برای ژن‌های OR13C7، PHLDB3 و TOX نیز اطلاعات حاشیه‌نویسی بدست نیامد ولی تعداد ۹ عدد از ۱۶ ژن شناخته شده (۵۶ درصد) در فرآیندهای بیولوژیکی، ۱۳ عدد (۸۱ درصد) در محتوای سلولی، ۱۲ عدد (۷۵ درصد) در عملکرد مولکولی و ۹ عدد (۵۶ درصد) در مسیرهای زیستی ایفای نقش می‌کنند (شکل شماره ۳). با توجه به اطلاعات شکل مشخص می‌گردد که ژن‌های PPYR1، ORC2، UBE3B، TP73، PPT2 و DHX8 در تمامی فعالیت‌های حاشیه‌نویسی مشارکت دارند.

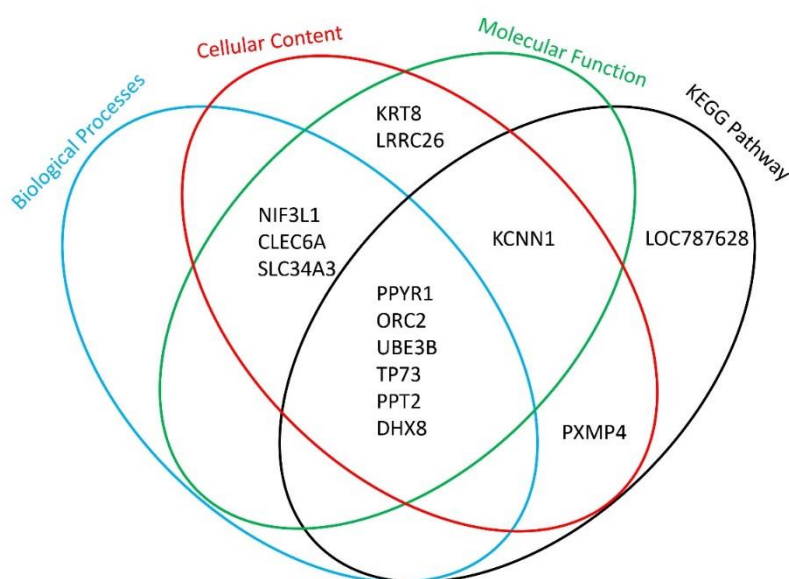


شکل ۲. توزیع ژن‌های بیان شده در جسم زرد بر اساس لگاریتم پایه دوم تغییرات چند برابر

جدول ۱- مشخصات ژن‌های بیان شده افتراقی در جسم زرد

کد شناسایی ژن	نام ژن	کروموزوم	تعداد بیان متوسط	لگاریت ۲ (تغییرات)	خطای استاندارد	ارزش P	ارزش Q
ENSBTAG00000002698	Ubiquitin protein ligase E3B(UBE3B)	۱۷	۳۳۹	۲/۶۴	۰/۳۲	۰/۰۰	۰/۰۰
ENSBTAG00000018282	NGG1 interacting factor 3 like 1(NIF3L1)	۲	۲۴۳	۱/۸۶	۰/۲۷	۰/۰۰	۰/۰۰
ENSBTAG00000031473	Origin recognition complex subunit 2(ORC2)	۲	۲۸۶	۱/۷۱	۰/۲۵	۰/۰۰	۰/۰۰
ENSBTAG00000004436	Palmitoyl-protein thioesterase 2(PPT2)	۲۳	۱۸۷	-۱/۶۷	۰/۳۱	۰/۰۰	۰/۰۰
ENSBTAG00000004760	Pancreatic polypeptide receptor 1(PPYR1)	۲۸	۲۹۸	-۱/۶۲	۰/۲۷	۰/۰۰	۰/۰۰
ENSBTAG00000023635	C-type lectin domain family 6 member A(CLEC6A)	۵	۲۸۵	-۱/۳۹	۰/۲۴	۰/۰۰	۰/۰۰
ENSBTAG00000048513	epithelial splicing regulatory protein 1(LOC788414)	۲	۳۱۷	-۱/۲۲	۰/۳۱	۰/۰۰	۰/۰۱
ENSBTAG00000045734	Leucine rich repeat containing 26 (LRRC26)	۱۱	۴۳۵	-۱/۱۹	۰/۲۴	۰/۰۰	۰/۰۰
ENSBTAG00000052921	Olfactory receptor family 13 subfamily C member 7 (OR13C7)	۸	۱۴۷۸	-۱/۱۷	۰/۲۲	۰/۰۰	۰/۰۰
ENSBTAG00000046379	Solute carrier family 34 member 3(SLC34A3)	۱۱	۱۳۹۹	-۱/۰۴	۰/۲۹	۰/۰۰	۰/۰۴
ENSBTAG00000005812	Tumor protein p73(TP73)	۱۶	۱۸۴۷	-۱/۰۳	۰/۲۳	۰/۰۰	۰/۰۰
ENSBTAG00000038738	Peroxisomal membrane protein 4(PXMP4)	۱۳	۲۳۴	-۰/۸۵	۰/۲۴	۰/۰۰	۰/۰۵
ENSBTAG00000004377	Pleckstrin homology like domain family B member 3(PHLDB3)	۱۸	۵۸۳۵	-۰/۸۱	۰/۲۳	۰/۰۰	۰/۰۴
ENSBTAG00000004596	Cytochrome P450 11B1, mitochondrial(LOC787628)	۱۴	۱۳۸۴۶	-۰/۸۰	۰/۲۳	۰/۰۰	۰/۰۵
ENSBTAG00000004954	Thymocyte selection associated high mobility group box (TOX)	۱۴	۸۵۷۵	-۰/۸۰	۰/۲۳	۰/۰۰	۰/۰۵
ENSBTAG00000008626	Potassium calcium-activated channel subfamily N member 1(KCNN1)	۷	۷۲۴۲	۰/۷۹	۰/۲۲	۰/۰۰	۰/۰۴
ENSBTAG00000050175	vomer nasal type-1 receptor 4-like(LOC107131477)	۱۸	۴۲۲۳	-۰/۷۷	۰/۱۹	۰/۰۰	۰/۰۱
ENSBTAG00000008339	DEAH-box helicase 8(DHX8)	۱۹	۳۱۰۱	-۰/۷۷	۰/۲۲	۰/۰۰	۰/۰۵
ENSBTAG00000000836	Keratin 8(KRT8)	۵	۱۳۷۳	-۰/۷۶	۰/۲۲	۰/۰۰	۰/۰۵

همان‌گونه که در جدول ۱ قابل مشاهده است ژن‌های UBE3B، NIF3L1 و ORC2 سه ژن شاخص عملکردی هستند که میزان بیان بیشتری در گاوهای با باروری زیاد دارند مطالعات هستی‌شناسی با استفاده از پایگاه DAVID نسخه ۲۰۲۱ نشان داد که ژن UBE3B در فرایندهای بیولوژیکی به‌عنوان پروتئین درگیر در فرآیند کاتابولیک پروتئین وابسته به یوبیکوئیتین است که بر روی کروموزوم شماره ۱۷ گاو تعیین نقشه شده است. NIF3L1 در فرآیند بیولوژیکی تمایز نورون، تنظیم مثبت رونویسی، الگوی DNA، تنظیم منفی رونویسی با الگوی اسید نوکلئیک، نقش دارد همچنین ژن ORC2 در فرآیند بیولوژیکی همانندسازی DNA نقش بازی می‌کند. از نظر محتوای سلولی ژن UBE3B در هسته و سیتوپلاسم وجود دارد. ژن NIF3L1 جزو هسته، سیتوپلاسم و میتوکندری بوده و ژن ORC2 جزو بافت‌های کروموزوم هسته‌ای، ناحیه تلومر، هتروکروماتین، کینتوکور داخلی کروموزوم متراکم، مجتمع تشخیص همانندسازی منشأ هسته‌ای، سانتروزوم و غشاء می‌باشد.



شکل ۳. نمودار ون برای حاشیه‌نویسی ژن‌های با بیان متفاوت افتراقی

بحث

در تحقیق حاضر ۱۳۰۴۹ عدد رونوشت ترانسکریپتومی در جسم زرد بیان گردید که تنها ۳۶۵ ژن با سطح معنی‌داری ۰/۰۵ تفاوت بیان در گاوهای با باروری زیاد و باروری کم نشان دادند. (Ghafouri et al (2022) تعداد ۱۳۸۷ ژن را در روز ۱۳ چرخه فحلی در بافت جسم زرد در گاوهای با باروری زیاد و کم بصورت بیان متفاوت معنی‌دار ($p \leq 0.000001$, Fold change ≥ 0.5) گزارش نمودند (Ghafouri et al. 2022). در یک مطالعه دیگر، تعداد ژن‌های با بیان افتراقی در بافت جسم زرد بین گاوهای بارور و کم بارور ۵۶۰ عدد به شرط سطح معنی‌داری اصلاح شده کمتر از ۰/۰۵ گزارش شده است (Moore et al. 2016).

ژنوم مرجع گاو نسخه Bos_taurus_UMD_3.1.1 (bostau8) نتیجه نمونه ترکیب شده خون گاوهای نر و ماده نژاد هرфорд است که با عمق خوانش 9X با استفاده از تکنولوژی سانجر در تاریخ ۲۵ نوامبر ۲۰۱۴ معرفی شده است دارای ۲,۶۴۰,۱۲۳,۳۱۰ جفت باز است که بر روی ۳۰ کروموزوم تعیین نقشه شده و شامل ۳۰,۸۱۱ ژن است که از این تعداد ۲۱,۰۸۹

ژن کد کننده و ۵,۵۲۰ ژن غیر کد کننده می‌باشد (Merchant et al. 2014). در مطالعه حاضر، میزان همترازی با ژنوم مرجع بطور میانگین ۹۳/۴۸ بدست آمد که در مقایسه با مطالعات صورت گرفته ۷۰-۶۰ درصد (Attari et al. 2019) و ۷۱-۶۵ درصد (Moore et al. 2016) مقدار قابل توجهی است. مطالعات غفوری و همکاران (۲۰۲۲) نشان داد که ژن‌های COL4A3, NT5DC3, ALPK3, OAS1X, MASP1, JSP.1 در بافت جسم زرد گاوهایی با باروری زیاد نسبت به گاوهای با باروری کم کاهش بیان داشته است. در مطالعه دیگری، ۴۰۳ ژن با بیان متفاوت در بافت آندومتریم گاوهایی با باروری زیاد و کم شناسایی شدند که ۱۲۳ عدد از این ژن‌ها در سه شبکه فیزیولوژیکی اکتین و اجزای اسکلت سلولی، عملکرد ایمنی و حمل‌ونقل یون فعالیت داشتند (Moran et al. 2017). علاوه بر آن، مطالعات مختلفی بر روی بیان ژن در خصوص باروری در بافت‌های مختلف انجام شده و ژن‌های دخیل در این خصوص گزارش شده است (Gonella-Diaza et al. 2017). ژن UBE3B دارای عملکرد مولکولی و فعالیت ترانسفرازی یوبیکوئیتین-پروتئین و فعالیت لیگازی است، ژن NIF3L1 وظیفه مولکولی اتصال به فاکتور رونویسی و همچنین اتصال پروتئینی یکسان را بر عهده دارد و در این میان عملکرد مولکولی ژن ORC2 اتصال به منشأ تکثیر DNA است. در مسیرهای بیولوژیکی دایره المعارف KEGG ژن UBE3B در مسیر فرایندی پروتئولیز با واسطه یوبیکوئیتین نقش دارد اما برای NIF3L1 مسیری مشخص نشده است و لیکن ژن ORC2 بر اساس داده‌های ثبت شده در این پایگاه در چرخه سلولی نقش دارد. در مطالعات پیشین، میزان بیان ژن‌های SYNM, PARM1, NXPE2, COL12A1, ADAMDEC1, SERPINA14, S100A9, PI16, MSTN, CD83, CA2, C2, SAA3, C5 در گاوهای با باروری زیاد نسبت به گاوهای با باروری کم بصورت مشترک در بافت آندومتریم و بافت جسم زرد بیشتر بوده است که ژن NXPE2 که در کروموزوم ۱۵ گاو تعیین نقشه شده است به عنوان ژن کاندیدا برای باروری شناخته شده است (Ghafouri et al. 2022).

در تحقیق حاضر، ژن‌های KRT8, PHLDB3, PPT2, LOC787628, PPYR1, FOX, TP73, DHX8, KCNN1, CLEC6A, PXMP4, LRR26, SLC34A3 و OR13C7 در گاوهای با باروری کم میزان بیان بیشتری نشان داده است. از میان این ژن‌ها سه ژن PPT2, PPYR1 و CLEC6A جزو ژن‌های شاخص محسوب می‌گردند. ژن PPT2 در فرایند زیستی دپالمتویلاسیون ماکرومولکول حذف گروه‌های پالمیتول از یک ماکرومولکول را بر عهده دارد. ژن PPT2 جزو ساختار سلولی لیزوزوم، اندامک محدود به غشای داخل سلولی و آگزوزوم خارج سلولی است. بر اساس اطلاعات حاشیه‌نویسی در خصوص عملکرد مولکولی این ژن فعالیت هیدرولاز پالمیتویل (پروتئین) را بر عهده دارد که در حضور مولکول آب ترکیب پالمیتویل-پروتئین را به پروتئین و پالمیتات تجزیه می‌کند. در مسیرهای KEGG این ژن در چرخه افزایش طول اسید چرب، مسیرهای متابولیک، متابولیسم اسیدهای چرب و لیزوزوم نقش ایفا می‌کند (Dennis et al. 2003). ژن PPYR1 در فرایند زیستی مسیر سیگنال دهی گیرنده سطح سلول، مسیر سیگنال دهی نوروپپتید، رفتار تغذیه و گردش خون نقش دارد. PPYR1 جزء جدایی‌ناپذیر غشای پلاسمایی و متشکل از محصولات ژنی و کمپلکس‌های پروتئینی است که حداقل بخشی از توالی پپتیدی آن‌ها در ناحیه آب‌گریز غشا جاسازی شده است. از نظر عملکرد مولکولی این ژن فعالیت گیرنده پپتید YY و همچنین، فعالیت گیرنده پلی پپتیدی پانکراس کنترل می‌کند. ژن PPYR1 در مسیر زیستی برهمکنش لیگاند-گیرنده عصبی ایفای نقش می‌کند (Dennis et al. 2003). ژن CLEC6A جزء جدایی‌ناپذیر غشاء سلولی است بدین صورت که جزء یک غشاء

1Protein Coding

2None Coding

3Molecular Function

4KEGG Pathway

5Cell Cycle

6Macromolecule depalmitoylation

7Palmitoyl-(protein) hydrolase activity

8Neuroactive ligand-receptor interaction

مشکل از محصولات ژنی و کمپلکس‌های پروتئینی است که حداقل بخشی از توالی پپتیدی آن‌ها در ناحیه آب‌گریز غشا جاسازی شده است؛ که در فرآیندهای بیولوژیکی پاسخ ایمنی تطبیقی، تنظیم مثبت سیگنال دهی I-kappaB کیناز NF-kappaB، پاسخ ایمنی ذاتی، تنظیم مثبت ترشح سیتوکین و پاسخ دفاعی به قارچ نقش بازی می‌کند و وظیفه مولکولی اتصال کربوهیدرات را بر عهده دارد. اتصال به یک کربوهیدرات شامل: مونوساکاریدها، الیگوساکاریدها و پلی ساکاریدها و همچنین مواد به دست آمده از مونوساکاریدها از طریق احیای گروه کربونیل (آلدیتول‌ها)، با اکسیداسیون یک یا چند گروه هیدروکسی برای تولید آلدئیدها، کتون‌ها، یا اسیدهای کربوکسیلیک مربوطه، یا با جایگزینی یک یا چند گروه هیدروکسی با یک اتم هیدروژن است (Dennis et al. 2003). ژن CLC34A3 با ایفای نقش در فرایند بیولوژیکی هموستاز یون فسفات سلولی در هر فرآیندی که در حفظ حالت پایدار داخلی یون‌های فسفات در سطح یک سلول دخیل می‌باشد. CLC34A3 در سیتوپلاسم، دیواره Brush، جزء جدایی‌ناپذیر غشاء، غشای پلاسمایی آپیکال و وزیکول وجود دارد. فعالیت ناقل گذرنده فسفات غیر آلی، فعالیت سمپولتر سدیم؛ فسفات و فعالیت ناقل گذرنده فسفات وابسته به سدیم تحت نظارت این ژن می‌باشد (Dennis et al. 2003).

نتیجه‌گیری

به عنوان نتیجه‌گیری کلی، نتایج حاصل از آنالیز داده‌های RNA-Seq با استفاده از پلتفرم گالاکسی نشان داد که حدود ۹۹/۱۳ درصد از داده‌های خام دارای کیفیت مناسب بوده و از این مقدار در حدود ۹۳/۴۸ درصد داده‌ها بر روی ژنوم مرجع گاوی هم‌تراز شده است. در بیان افتراقی ژن ۱۳۰۴۹ عدد ژن در جسم زرد بیان شده است که از این تعداد ۳۶۵ عدد در گاوهایی با باروری زیاد و باروری کم اختلاف بیان در سطح ۰/۰۵ نشان دادند و در صورتی که سطح معنی‌داری اصلاح شده بر اساس نرخ کشف اشتباه مدنظر قرار گیرد تعداد ۱۹ رونوشت اختلاف بیان بین تیمارها نشان دادند. ژن‌های UBE3B، NIF3L1 و ORC2 سه ژن شاخص عملکردی هستند که میزان بیان بیشتری در گاوهای باروری زیاد دارند که در فعالیت‌های بیولوژیکی به‌عنوان پروتئین درگیر در فرآیند کاتابولیک پروتئین وابسته به یوبیکوئیتین، تمایز نورون، تنظیم مثبت رونویسی، الگوی DNA، تنظیم منفی رونویسی با الگوی اسید نوکلئیک و همانندسازی DNA نقش بازی می‌کنند. ژن‌های PPT2، PHLDB3، KRT8، LOC787628، PPYR1، TOX، TP73، DHX8، KCNN1، CLEC6A، PXMP4، LRRC26، SLC34A3 و OR13C7 در گاوهای با باروری کم میزان بیان بیشتری نشان داده است ژن‌های شاخص این گروه در فعالیت‌های بیولوژیکی دپالمتویلاسیون ماکرومولکول، مسیر سیگنال دهی گیرنده سطح سلول، مسیر سیگنال دهی نوروپپتید، رفتار تغذیه، گردش خون و هموستاز یون فسفات سلولی دخالت دارند. به عنوان نتیجه‌گیری کلی، جایگاه‌های UBE3B، NIF3L1 و ORC2 که مقدار بیان بیشتری در جسم زرد گاوهای با باروری زیاد داشتند می‌توانند بعنوان ژن‌های شاخص عملکردی در انتخاب ژنتیکی گاوهای شیری مورد استفاده قرار گیرند.

REFERENCES

- Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Attari, M., Moradi Shahrababak, H., Nehzati Paghale, G., Banabazi, M. H. & Hashemi, M. (2019). Study of differential gene expression in queen, drone and worker honey bee using RNA-seq data. *Iranian Journal of Animal Science*, 50(2), 103-113. doi: 10.22059/ijas.2019.257296.653635 (In Persian)
- Barbat, A., Le Mézec, P., Ducrocq, V., Mattalia, S., Fritz, S., Boichard, D. & Humblot, P. (2010). Female fertility in French dairy breeds: current situation and strategies for improvement. *Journal of Reproduction and Development*, 56(S), S15-S21.

- <https://doi.org/10.1262/jrd.1056S15>
- Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., Nekrutenko, A. & Team, G. (2010). Manipulation of FASTQ data with Galaxy. *Bioinformatics*, 26(14), 1783-1785. <https://doi.org/10.1093/bioinformatics/btq281>
- Bolger, A. M., Lohse, M. & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Cai, Z., Guldbandsen, B., Lund, M. S. & Sahana, G. (2019). Prioritizing candidate genes for fertility in dairy cows using gene-based analysis, functional annotation and differential gene expression. *BMC Genomics*, 20(1), 1-9. <https://doi.org/10.1186/s12864-019-5638-9>
- Chen, H. Y., Shen, H., Jia, B., Zhang, Y. S., Wang, X. H. & Zeng, X. C. (2015). Differential gene expression in ovaries of Qira black sheep and Hetian sheep using RNA-Seq technique. *PLoS One*, 10(3), e0120170. <https://doi.org/10.1371/journal.pone.0120170>
- 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), R60. <https://doi.org/10.1186/gb-2003-4-9-r60>
- Ghafouri, F., Sadeghi, M., Bahrami, A., Abdollahi-Arpanahi, R., Javanmard, A. & Miraei-Ashtiani, S. R. (2022). Comparison of differential expression profiles of candidate genes related to fertility traits using transcriptome perspective based on RNA-Seq in Holstein dairy cows. *Iranian Journal of Animal Science*, 52(4), 217-229. <https://doi.org/10.22059/ijas.2021.322630.653824> (In Persian)
- Gonella-Diaza, A. M., da Silva Andrade, S. C., Sponchiado, M., Pugliesi, G., Mesquita, F. S., Van Hoeck, V. & Binelli, M. (2017). Oviductal transcriptional profiling of a bovine fertility model by next-generation sequencing. *Genomics Data*, 13, 27-29. <https://doi.org/10.1016/j.gdata.2017.06.004>
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44. <https://doi.org/10.1038/nprot.2008.211>
- Killeen, A. P., Diskin, M. G., Morris, D. G., Kenny, D. A. & Waters, S. M. (2016). Endometrial gene expression in high-and low-fertility heifers in the late luteal phase of the estrous cycle and a comparison with midluteal gene expression. *Physiological Genomics*, 48(4), 306-319. <https://doi.org/10.1152/physiolgenomics.00042.2015>
- Kim, D., Langmead, B. & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), 357-360. <https://doi.org/10.1038/nmeth.3317>
- Liao, Y., Smyth, G. K. & Shi, W. (2014). FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923-930. <https://doi.org/10.1093/bioinformatics/btt656>
- Love, M. I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1-21. <https://doi.org/10.1186/s13059-014-0550-8>
- Merchant, S., Wood, D. E. & Salzberg, S. L. (2014). Unexpected cross-species contamination in genome sequencing projects. *PeerJ*, 2, e675. <https://doi.org/10.7717/peerj.675>
- Minten, M. A., Bilby, T. R., Bruno, R. G., Allen, C. C., Madsen, C. A., Wang, Z. & Geary, T. W. (2013). Effects of fertility on gene expression and function of the bovine endometrium. *PLoS One*, 8(8), e69444. <https://doi.org/10.1371/journal.pone.0069444>
- Moore, S. G., Pryce, J. E., Hayes, B. J., Chamberlain, A. J., Kemper, K. E., Berry, D. P. & Fair, T. (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. *Biology of Reproduction*, 94(1), 1-11. <https://doi.org/10.1095/biolreprod.115.132951>
- Moran, B., Butler, S. T., Moore, S. G., MacHugh, D. E. & Creevey, C. J. (2017). Differential gene expression in the endometrium reveals cytoskeletal and immunological genes in lactating dairy cows genetically divergent for fertility traits. *Reproduction, Fertility and Development*, 29(2),

- 274-282. <https://doi.org/10.1071/RD15128>
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 5(7), 621-628. <https://doi.org/10.1071/RD15128>
- Nayeri, S., Sargolzaei, M., Abo-Ismael, M. K., May, N., Miller, S. P., Schenkel, F. & Stothard, P. (2016). Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genetics*, 17(1), 1-11. <https://doi.org/10.1186/s12863-016-0386-1>
- Olori, V., Pool, M., Calus, M., Cromie, A. & Veerkamp, R. (2003). Joint evaluation of survival and fertility in dairy cattle with a linear model. *Interbull Bulletin* (30), 20.
- Shahabi, A., Tahmoorespour, M. & Kazemipour, 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. DOI: [10.22103/jab.2019.13510.1112](https://doi.org/10.22103/jab.2019.13510.1112)
- Snelling, W., Cushman, R., Keele, J., Maltecca, C., Thomas, M., Fortes, M. & Reverter, A. (2013). Breeding and genetics symposium: Networks and pathways to guide genomic selection. *Journal of Animal Science*, 91(2), 537-552. <https://doi.org/10.2527/jas.2012-5784>
- Yang, I. S. & Kim, S. (2015). Analysis of whole transcriptome sequencing data: workflow and software. *Genomics and Informatics*, 13(4), 119. DOI: [10.5808/GI.2015.13.4.119](https://doi.org/10.5808/GI.2015.13.4.119)