خوشه‌بندی مبتنی بر آنتولوژی ژن‌های هدف ریزRNAهای مؤثر بر تولید شیر

- **سعیده اسکندری‌نسب:** گروه علوم دامی، دانشکده کشاورزی، دانشگاه یاسوج، یاسوج، ایران
- **زهرا رودباری*:** گروه علوم دامی، دانشکده کشاورزی، دانشگاه جیرفت، جیرفت، ایران
- **محمدرضا بحرینی‌بهزادی:** گروه علوم دامی، دانشکده کشاورزی، دانشگاه یاسوج، یاسوج، ایران

تاریخ دریافت: آذر ۱۳۹۸ تاریخ پذیرش: اسفند ۱۳۹۸

چکیده

شیر گاو مایعی بسیار مغذی است و بخش مهمی از رژیم غذایی یک فرد می‌باشد و تمامی مواد مورد نیاز بدن انسان را در خود دارد. استراتژی‌هایی که از طریق آن می‌توان تولید شیر را تحت تأثیر قرار داد، شناسایی ژن‌هایی که روی تولید و ترکیب شیر اثر می‌گذارند و استراتژی دیگر شناسایی ریزRNAهایی‌که بیان ژن‌های مؤثر را تحت تأثیر قرار می‌دهند. در این مطالعه ابتدا داده‌های ریزRNAهای مربوط به غده پستانی گاو شیری با کد شناسایی E-GEOD-61227 از پایگاه GEO پیاده‌سازی شدند. بعد از مشخص شدن میزان بیان ریزRNAها، شناسایی ژن‌های هدف ریزRNAها به کمک پایگاه داده‌ای miRwalk انجام گردید. سپس برای فرآیند خوشه‌بندی ژن‌ها بر مبنای ساختار درختی آنتولوژی ژن که شامل سه زیر آنتولوژی فرآیند بیولوژیکی، عملکرد ملکولی و جزء سلولی است از نرم‌افزار AgriGO استفاده گردید. در مطالعه حاضر براساس خوشه‌بندی مبتنی بر آنتولوژی ژن‌های هدف نتایج نشان داد که فعالیت‌های بیولوژیکی مورفوژنز بافت پستانی و تکثیر سلول‌های اپیتلیال در سطح بسیار معنی‌داری می‌باشندکه این فرآیندها در تکامل و توسعه غدد پستان نقش بسیار مهمی دارند ونتایج براساس فعالیت مولکولی نشان داد که سیگنال‌دهی با استفاده از پروتئین‌های گیرنده کینازی بیش‌ترین سطح معنی‌داری به خود اختصاص داده است. هم‌چنین نتایج مربوط به بخش اجزا سلولی نشان داد بیش‌ترین شمار ژن‌های هدف در بخش داخلی غشا سلولی ارگانل و غشا پلاسمایی قرار دارند. بنابراین، ژن‌های هدفی که تنظیم کننده این فرآیندهای معنی‌دار می‌باشند پتانسیل ایفای نقش قابل توجه جهت یافتن راهکاری هدفمند در راستای بهبود صفت تولید شیر را دارند.

کلمات کلیدی: آنتولوژی ژن، ریزRNA، خوشه‌بندی ژن، تولید شیر

* پست الکترونیکی نویسنده مسئول: roudbari.zahra@ujiroft.ac.ir

(DOI): 10.22034/aej.2020.118071

مقدمه

در سال‌های اخیر، با ظهور فناوری‌های زیستی پیشرفته، حجم داده‌های زیستی به‌طور قابل ملاحظه‌ای افزایش یافته است. این مسئله منجر به ایجاد یک مبحث بین‌رشته‌ای به نام بیوانفورماتیک شده است. بیوانفورماتیک علم نوینی است که با تلفیق علوم زیست‌شناسی، کامپیوتر و ریاضیات (به‌ویژه آمار)، تلاش می‌کند به مسائل زیستی در زمینه‌های سلولی و ملکولی پاسخ دهد (Goujon و همکاران، ۲۰۱۰). در حال حاضر درک شبکه‌های تنظیم‌کننده ژنتیکی و فهم فرآیندهای تنظیمی در یک سلول در سطح ژنی، یک هدف مهم در بیولوژی محاسباتی می‌باشد. از آن‌جاکه درک شهودی رفتار سیستم‌های تنظیمی مشکل است، استفاده از شبیه‌سازی و مدل‌سازی کامپیوتری این شبکه‌ها غیر قابل اجتناب است (Bar-Joseph، ۲۰۰۴). یکی از مسائل کلیدی در درک رفتار سلولی کشف فعل و انفعالات بین ژن‌ها است که به کمک استنتاج شبکه‌های ژنی ارتباط بین ژن‌ها کشف می‌شود و این شبکه‌های ژنی مجموعه‌ای از ژن‌ها می‌باشند که در سلول به‌صورت غیرمستقیم با هم تعامل دارند (Niemi و همکاران، ۲۰۰۷). تحلیل شناسایی شبکه‌های تنظیم‌کننده ژنی مشخص می‌کند که ژن‌های کلیدی چگونه، کی و کجا سایر ژن‌ها را تحت تأثیر قرار می‌دهند، هم‌چنین این شبکه‌ها اطلاعاتی در مسیرهای ژنی فراهم می‌کنند که می‌توانند کاربردهای زیادی در بیولوژیک ملکولی داشته باشند (Gupta و همکاران، ۲۰۰۷). ازجمله تکنیک‌هایی که برای مدل‌سازی شبکه‌های تنظیمی ژنی استفاده می‌شود خوشه‌بندی داده‌ها می‌باشد که برای خوشه‌بندی داده‌ها از اطلاعات بیولوژیکی موجود در ژن‌ها استفاده می‌شود (Shih و همکاران، ۲۰۰۴). خانواده ژنی به گروهی از ژن‌ها متعلق می‌شود که از نظر بیولوژیکی و عملکرد سلولی به‌هم شبیه هستند و برای یافتن اختلاف بیان ژن‌ها و بیان فنوتیپ‌ها از ترکیب دانش بیولوژیکی و تحلیل‌های آماری استفاده می‌کنند. اطلاعات مربوط به این خانواده‌ها، در پایگاه داده‌ای مانند Gene Ontology (GO), Kyoto Encyclopedia of Genomes Genemap (KEGG) در دسترس است (Khatri و همکاران، ۲۰۰۵). آنتولوژی ژن یکی از مهم‌ترین ابزارهای کمکی مربوط به ارائه و پردازش اطلاعات ژن‌ها، محصولات ژن و عملکردشان می‌باشد. از مزایای آنتولوژی ژن، اولاً از طریق آنالیز آماری ژن‌های مهم را شناسایی می‌کند و سپس این‌که محصولات ژن‌ها کجا هستند و چه کارایی انجام می‌دهند و در چه فرآیندهای بیولوژیکی شرکت می‌کنند را مشخص می‌کند. ثانیاً می‌توان مطابق با عملکرد ژن‌ها یا فرآیندهای بیولوژیکی داده‌ها را گروه‌بندی نمود، این مزیت بسیار مهم است زیرا اطلاعات به شیوه کاملاً معنی‌داری سازماندهی می‌شوند، بنابراین می‌توان درک بهتری از داده‌ها به‌دست آورد (Li و همکاران، ۲۰۰۴). تحلیل خانواده‌های ژنی در مقایسه با تحلیل ژن‌های منفرد هماهنگی بیش‌تری دارد زیرا در تحلیل خانواده‌های ژنی، فعالیت گروهی ژن‌ها درنظر گرفته می‌شود و از نظر بیولوژیکی تفسیر حاصل از خانواده‌ها ژنی ساده‌تر است، زیرا قرار گرفتن ژن‌ها در یک خانواده براساس نقش آن‌ها در سلول است، به این ترتیب درک شبکه پیچیده فرآیندهای سلولی نیز ساده‌تر می‌گردد (Dopazo، ۲۰۰۶). شیر گاو مایعی بسیار مغذی است که در بدن گاو تشکیل می‌شود و خواص و فواید زیادی دارد، از شیر فرآورده‌های غذایی مختلفی تهیه می‌شود. این فراورده‌ها را محصولات لبنی می‌گویند که بخش مهمی از رژیم غذایی یک فرد می‌باشدو می‌توان گفت تمامی مواد مورد نیاز بدن انسان را در خود دارد (Kharrati و همکاران، ۲۰۱۲). استراتژی‌هایی که از طریق آن می‌توان تولید شیر را تحت تأثیر قرار داد شناسایی ژن‌هایی است که روی تولید و ترکیب شیر اثر می‌گذارند و استراتژی دیگرشناسایی ریزRNAهایی است که ژن‌های هدف را تحت تأثیر قرار می‌دهند (Cole و همکاران، ۲۰۱۳). تولید شیر فرایندی فیزیولوژیکی می‌باشد که تحت تأثیر ریبونوکلئیک اسیدهای مداخله‌گر مانند ریزRNAها قرار می‌گیرد این ملکول‌ها با تأثیر بر بیان ژن فرایندهای مختلف مانند تکثیر سلول‌های آلوئل، توسعه بافت پستانی، متابولیسم اسیدهای آمینه، تکثیر، تمایز، رشد، توسعه سلولی و سنتز ترکیبات شیر را تحت تأثیر قرار می‌دهد (Friedman و همکاران، ۲۰۰۹). هدف از مطالعه حاضر تجزیه و تحلیل داده‌های بیان ریزRNA مربوط به بافت پستانی گاو به‌منظور خوشه‌بندی ژن‌های هدف ریزRNAهای مؤثر بر تولید و ترکیب شیر با استفاده از دانش بیولوژیکی فراهم‌شده از ساختار درختی آنتولوژی ژن و بررسی ژن‌ها در بخش‌های بیولوژیکی، ساختاری و سلولی هم‌چنین شناسایی فرآیندهای مؤثر می‌باشد.

مواد و روش‌ها

اغلب محققان برای دست‌یابی به ریزRNAها و تحلیل داده‌ها از داده‌های بانک‌های اطلاعاتی اینترنتی نظیر ژن بانک استفاده کنند. در این مطالعه ابتدا داده‌های ریزRNAهای مربوط به غده پستانی گاو با کد شناسایی E-GEOD-61227 از شبکه جهانی اینترنت (پایگاه GEO) بانک اطلاعاتی NCBI دانلود شد. بخش GEO یک پایگاه داده‌ای است که داده‌های مرتبط به مطالعات بیان ریزRNA با روش‌های مختلف ذخیره کرده و دارای ابزاری جهت تجزیه و تحلیل آن‌ها به‌منظور تشخیص ریزRNAها با بیان متفاوت است (Barrett و همکاران، ۲۰۱۳). بعد از مشخص شدن میزان بیان ریزRNAها، تعیین ژن‌های هدف به کمک پایگاه داده‌ای miRwalk انجام گردید. این پایگاه داده‌ای قادر به شناسایی ژن‌های هدف از طریق پیش‌بینی و مشاهدات آزمایشگاهی می‌باشد. این پایگاه داده‌ای پیش‌بینی ژن‌های هدف را می‌تواند با استفاده از

(DOI): 10.22034/aej.2020.118071

الگوریتم‌های موجود در ده پایگاه اطلاعاتی دیگر (Mircrit4، PITA، miRanda، mirbridge، miRDB، miRNAMap، pictar2، RNA22 و Targetscan) انجام دهد (Sticht و همکاران، ۲۰۱۸). در این مرحله ریزRNAهایی مورد توجه قرار گرفت که ژن‌های هدف آن‌ها هم از طریق miRwalk تأیید شدند هم از طریق هر ده پایگاه داده‌ای که توسط miRwalk قابل بررسی بودند، مورد تأیید قرار گرفت. پس از شناسایی و مشخص شدن ژن‌های هدف فرایند خوشه‌بندی ژن‌ها بر مبنای دانش بیولوژیکی با استفاده از نرم‌افزار آنلاین AgriGO (http://bioinfo. Cau. Edu.cn/agriGO) انجام شد. روش خوشه‌بندی براساس دانش بیولوژیکی به ما امکان شناسایی فرآیند سیکل سلولی را می‌دهد. لذا در این روش از اطلاعات بیولوژیکی موجود (آنتولوژی ژن) برای خوشه‌بندی ژن‌ها استفاده می‌شود. بنابراین هر یک از ژن‌ها می‌توانند چند ترم GO (فعالیت سلولی) مختلف داشته باشند که این می‌تواند نشان‌دهنده شرکت آن‌ها در چند فرآیند بیولوژیکی متفاوت باشد. هر خوشه، متناظر با یک ترم GO خاص می‌باشد و ژن‌هایی در یک خوشه قرار می‌گیرند که این GO را در مجموعه GOهای خود داشته باشند، لذا تعداد ژن‌های موجود در خوشه‌های حاصل متفاوت است و خوشه‌ها می‌توانند با هم هم‌پوشانی داشته باشند، زیرا برخی از ژن‌ها بیش از یک ترم GO دارند (Smith و همکاران، ۲۰۰۳).

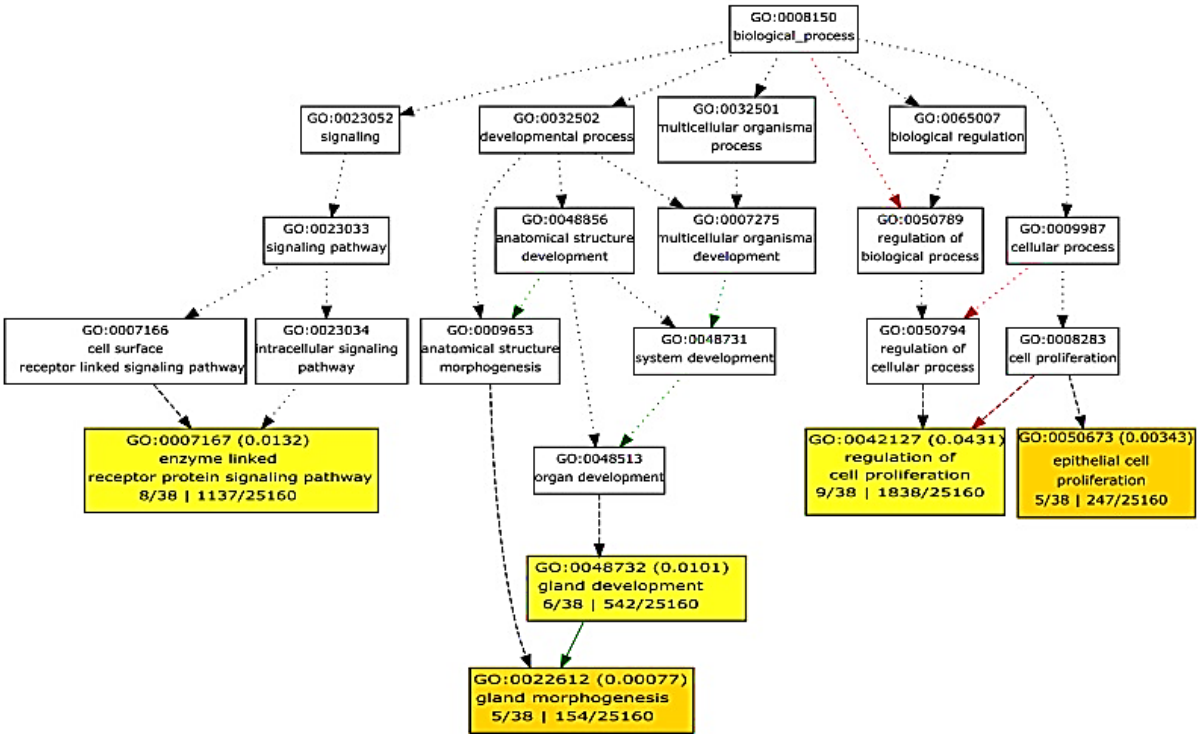
نتایج

این مطالعه جهت تجزیه و تحلیل عوامل ملکولی مؤثر بر تولید شیر و فعالیت طبیعی غدد پستانی با استفاده از دادهای ریزRNAها انجام شد. در مرحله اول به بررسی بیان ریزRNAها پرداخته شد. ریزRNAهایی که (Log FC< - 1) بیان پایینی دارند و کم‌تر از بیان ژن‌های هدف جلوگیری می‌کنند که تعداد آن‌ها ۵ عدد است و ریزRNAهایی که (Log FC>1) بیان بالایی دارند و ژن‌های هدف را بیش‌تر تحت تأثیر قرار می‌دهند و ژن‌ها کم‌تر بیان می‌شوند که تعداد آن‌ها ۱۸ عدد است. بعد از محاسبه میزان بیان ریزRNAها، ژن‌های هدف آن‌ها شناسایی شدند. با توجه به این‌که این ژن‌ها در طول مسیر تکامل تغییر نکرده‌اند نقش مهم‌تری می‌توان برای آن‌ها در بافت مربوطه درخصوص فرآیندهای فیزیولوژیکی آن بافت درنظر گرفت. ۹۳۶ ژن تحت کنترل این ریزRNAها بودند که از این تعداد ۸۶ ژن بیش‌ترین تأثیر بر روی غدد پستان و تولید شیر داشتند که وارد نرم‌افزار AgriGO شدند. در این مطالعه به بررسی ژن‌ها در بخش‌های بیولوژیکی، ساختاری و سلولی با استفاده از نرم‌افزار Agri Go پرداخته شد. بررسی آنتولوژی تک تک ژن‌ها و گروه‌بندی نمودن سر دسته‌های اصلی جهت پی بردن به مهم‌ترین مسیرهای بیان و عملکرد مجموع این ژن‌ها می‌باشد. ساختار درختی آنتولوژی ژن به‌صورت یک گراف قابل نمایش است که در شکل‌های ۱، ۲، ۳ مشاهده می‌شود. خوشه‌هایی که با رنگ زرد نشان داده شده‌اند از نظر بیولوژیکی معنی‌دار می‌باشند. نتایج کلی آنتولوژی ژن‌های هدف مؤثربر تولید شیر به‌ترتیب GO-ID،Level،P-value در جدول ۱ آورده شده‌اند.

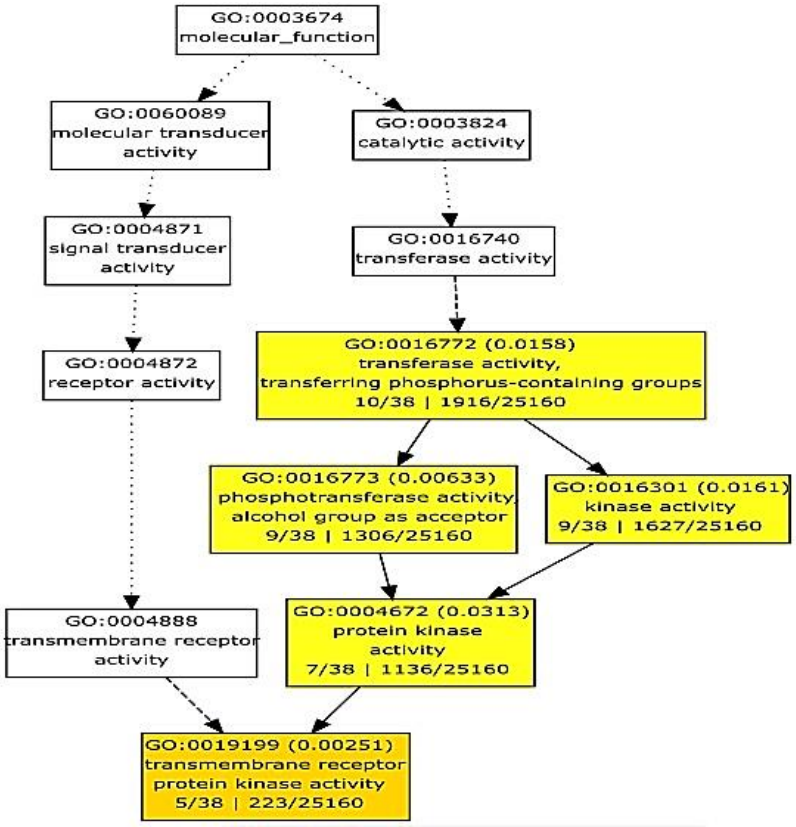
جدول۱: آنتولوژی ژن‌های هدف مؤثر بر تولید شیر با GO level≥3

Ontology	GO ID	GO term	Level	P Value
Biological-Process	۰/۰۰۱۴	۹	Regulation of cell proliferation	GO:۰۰۴۲۱۲۷
Biological-Process	۰/۰۰۰۲۶	۸	Enzyme linked	GO:۰۰۰۷۱۶۷
Biological-Process	۰/۰۰۱۶	۶	Gland development	GO:۰۰۴۸۷۳۲
Biological-Process	۰/۰۰۰۰۰۴	۵	Gland morphogenesis	GO:۰۰۲۲۶۱۲
Biological-Process	۰/۰۰۰۰۰۰۳۷	۵	Epithelial cell proliferation	GO:۰۰۵۰۶۷۳
Mulecular-function	۰/۰۰۱۴	۷	Protein Kinase activity	GO:۰۰۴۶۷۲
Mulecular-function	۰/۰۰۰۱۲	۵	Phosphotransferase activity	GO:۰۰۱۶۷۷۳
Mulecular-function	۰/۰۰۰۵۹	۵	Kinase activity	GO:۰۰۱۶۳۰۱
Mulecular-function	۰/۰۰۰۰۰۲۳	۵	Transmembransereseotore Protein Kinase activity	GO:۰۰۱۹۱۹۹
Mulecular-function	۰/۰۰۴۳	۴	Transferase activitiy	GO:۰۰۱۶۷۷۲
Cellular-component	۰/۰۰۰۰۰۰۴۹	۴	Plasma membrance	GO:۰۰۰۵۸۸۶
Cellular-component	۰/۰۰۰۵۵۵	۴	Intracellular membrance bunded organelle	GO:۰۰۴۳۲۳۱
Cellular-component	۰/۰۰۰۵۵۵	۳	membrance bunded organelle	GO:۰۰۴۳۲۲۷

(DOI): 10.22034/aej.2020.118071

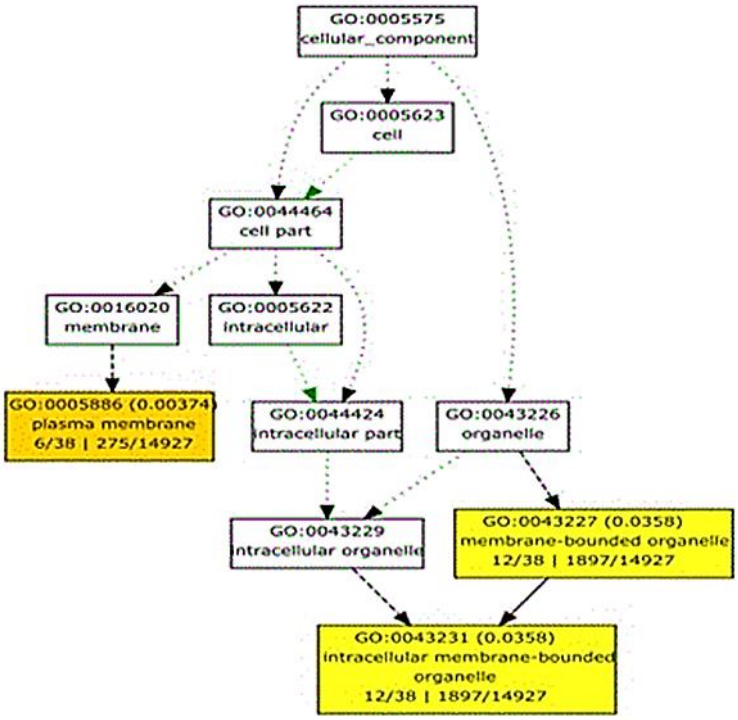


شکل۱: نتایج هستی‌شناسی براساس فرآیندهای بیولوژیکی به‌دست آمده توسط نرم‌افزار AgriGO



شکل۲: نتایج هستی شناسی براساس فعالیت مولکولی به‌دست آمده توسط نرم‌افزار AgriG

(DOI): 10.22034/aej.2020.118071



شکل۳: نتایج هستی شناسی براساس اجزا سلولی به‌دست آمده توسط نرم‌افزار AgriGO

بحث

آنتولوژی‌ژن(GO) یک لغت‌نامه برای توصیف محصولات ژن است و تلاشی برای هماهنگ کردن توصیف محصولات ژن‌ها در پایگاه‌های داده‌ای مختلف می‌باشد. GO شامل سه زیر آنتولوژی است که شامل فرآیند بیولوژیکی (BP)، عملکرد ملکولی (MF) و جزء سلولی (CC) می‌باشد. فرآیند بیولوژیکی نشان‌دهنده فعالیت بیولوژیکی است که ژن در آن شرکت می‌کند مانند سیکل سلولی عملکرد مولکولی بیان‌کننده فعالیت‌هایی است که در سطح سلول رخ می‌دهند و جزء سلولی تعیین می‌کند که محصول ژن در کدام قسمت سلول یافت می‌شود (مثلاً نوکلئوزوم و ریبوزوم)(Yavari و همکاران، ۲۰۰۸). همان‌طورکه در شکل ۱ مشاهده می‌شود فرایند بیولوژیکی نهایی و در سطح بسیار معنی‌دار مورفوژنز بافت پستانی (۵ ژن) است و فعالیت بیولوژیکی دیگری که بسیار معنی‌دار است،تکثیر سلول‌های اپیتلیال (۵ژن) است که این سلول‌های اپیتلیال غدد پستانی در تکامل و توسعه پستان نقش حائز اهمیتی دارند که خود این مسیر تکامل و توسعه بافت هم در بالا دست مسیر مورفوژنز بافت پستانی قرار دارد. در نتیجه ژن‌های هدف شرکت کننده در این فعالیت‌های بیولوژیکی از طریق فعال‌سازی آن‌ها بر توسعه بافت پستانی و مورفولوژی تولیدکننده شیر اثرگذار هستند.ژن‌هایی که به‌عنوان فاکتور رونویسی یا کوفاکتورفعالیت داشته باشند می‌توانند نقش تنظیم‌کنندگی داشته باشند از این‌رو برای تأیید نقش تنظیم‌کنندگی ۵ ژن هدف شناسایی شده در فرایند مورفوژنز بافت پستانی از پایگاه Animal transcription factor database در گونه گاو استفاده شد و نتایج نشان داد که دو ژن FGFR2 و EGFR به‌عنوان فاکتور رونویسی می‌باشندو این ژن‌ها می‌توانندبیان‌سایر ژن‌ها را تحت تأثیر قرار دهند.برای مطالعه‌بیش‌تر این ژن‌ها دمین عملکردی آن‌ها موردبررسی قرارگرفت. نتایج بررسی نشان داد که دمین حفاظت شده برای این ژن‌ها دمین پروتئین کیناز می‌باشد که دارای نقش کاتالیزوری است. پروتئین کینازها گروهی از آنزیم‌ها هستند، که گروه فسفات را به‌پروتئین‌ها طی‌عملی به‌نام فسفوریلاسیون‌منتقل می‌کنند. این‌عمل کلید روشن و خاموش کردن بسیاری از فرایندهای بیولوژیکی هم‌چون متابولیسم، رونویسی، پیشرفت چرخه سلولی و مرگ سلولی و تمایز است (wang و همکاران، ۲۰۱۱). دمین دیگری تیروزین کیناز است‌که این دمین زیرگروهی از خانواده‌پروتیئن‌کیناز است و در انتقال فسفات از ATP به پروتئین نقش دارد (Dengjel و همکاران، ۲۰۰۹). مهم‌ترین فعالیت‌های مولکولی فعال شده توسط ژن‌های هدف که توسط پایگاه اطلاعاتی AgriGo شناسایی گردید درشکل ۲ آورده شده است و فعالیت‌های معنی‌دار که به رنگ زرد نمایش داده‌اند می‌توانند

(DOI): 10.22034/aej.2020.118071

به‌عنوان فعالیت‌های مولکولی فعال در بافت پستان در شرایط طبیعی و درحال تولیدشیر باشند. نتایج آنتولوژی ژن براساس فعالیت مولکولی نشان داد که سیگنال‌دهی با استفاده از پروتئین‌های گیرنده کینازی بیش‌ترین سطح معنی‌داری به‌خود اختصاص داده است. در واقع این گیرنده کینازها مسئول اضافه شدن گروه فسفات در فرآیند فسفرزایی به پروتئین‌ها بوده که این فرآیند باعث فعال شدن آبشارهای سیگنال‌دهی است(Yang و همکاران، ۲۰۰۸). نتایج بخش فعالیت مولکولی در پژوهش حاضر در این راستا است زیرا معنی‌دارترین فعالیت‌ها مرتبط با فعالیت‌های کینازی و انتقال گروه فسفر است و ژن‌های هدف شرکت کننده در این فعالیت‌ها عبارتند از EGFR، FGFR2، BMPR1A و CSNK1D که نتایج بررسی نقش کوفاکتوری این ژن‌ها با استفاده از دیتابیس مربوط به فاکتورهای رونویسی حیوانات نشان داد که فقط ژن‌های EGFR، FGFR2 نقش کوفاکتوری دارند که در بخش تفسیر نتایج فرآیندهای بیولوژی به آن اشاره شد. پروتئین کینازها آنزیم‌های مهمی در تولید شیر هستند و بیش از پانصد ژن این پروتئین را کد می‌کند و عمل این آنزیم‌ها در واقع کنترل‌کننده فعالیت‌های سلولی هم‌چون متابولیسم، رونویسی، پیشرفت چرخه سلولی و مرگ سلولی و تمایز است (wang و همکاران، ۲۰۱۱). نتایج آنالیز آنتولوژی ژن بر اساس اجزا سلولی در شکل ۳ نشان داد که بیش‌ترین شمار محصولات ژن‌های هدف با سطح بسیار معنی‌داری در بخش داخلی غشا سلولی ارگانل و غشا پلاسمایی (۱۲ ژن) قرار دارند. نتایج بررسی نقش کوفاکتوری ژن‌های معنی‌دار در این خوشه نشان داد که ژن‌های EGFR، FGFR2، CREB، VDR، SP1، MSX1 و HIF1A دارای نقش کوفاکتوری‌بوده‌وبه‌ترتیب دمین‌های عملکردی آن‌ها Protein Kinase، Protein Tyrosine Kinase، ATP Binding، b_ZIP، zf-C4، Hormone_recep، zf-H2C2_2 و Homeobox می‌باشدو این ژن‌ها تنظیم‌کننده فرایند بیولوژیکی تکثیر سلولی که در خوشه فعالیت‌های بیولوژیکی به‌صورت معنی‌دار نمایش داده شده است می‌باشند، بنابراین این ژن‌ها با نقش کوفاکتوری که دارند می‌توانند در فرایند تولید شیر نقش به‌سزایی داشته‌باشند. به‌طور خلاصه نتایج این مطالعه نشان داد که اکثر ژن‌های موجود در هر خوشه براساس فعالیت‌های بیولوژیکی در رشد و توسعه غده پستانی فعال می‌باشند و براساس فعالیت ملکولی پروتئین‌های گیرنده کینازی بیش‌ترین نقش دارند و براساس اجزای سلولی بیش‌تر ژن‌های هدف ریزRNAها مؤثر بر تولید شیر در بخش داخلی‌غشای‌سلولی ارگان‌ها و غشای پلاسمایی قرار دارندو فاکتورهای رونویسی تنظیم کننده آن‌ها شامل ژن‌های EGFR، FGFR2،CREB، VDR،SP1، MSX1و HIF1A می‌باشد. باتوجه به این‌که اخیراً انفجاری وسیع‌در پیشرفت تکنیک‌های بیوتکنولوژی‌جهت‌دستیابی‌به‌جنبه‌های مختلف ژن صورت گرفته است، می‌توان با نگاه سیستمی از این فاکتورهای رونویسی به‌عنوان عوامل تنظیم کننده فرآیند تولید شیر در بافت غده پستانی برای درک بهتر فرآیند تولید شیر و بهبود آن بهره برد و توانایی در پیش‌بینی صحیح عملکرد ژن‌ها می‌تواند اقدام مفیدی در ارائه استراتژی پرورش و اصلاح نژاد گاو شیری داشته باشد.

منابع

۱. **Bar-Joseph, Z., 2004.** Analyzing time series gene expression data. Bioinformatics. Vol. 20, No. 16, pp: 2493-2503.

۲. **Barrett, T.; Wilhite, S.; Ledoux, P.; Evangelista, C. and Kim, I., 2013.** Ncbi Geo Archive for functional genomics data sets-update. Nucleic acids research. Vol. 41, No. 1, pp: D991-D995.

۳. **Cole, J.B.; Lewis, R.M.; Maltecca, C.; Newman, S.; Olson, K. and Tait, J.R., 2013.** Breeding and genetics sympuosium Systems biology in animal breeding identifying relationships among markers, genes and phenotypes. Journal of Animals Science. Vol. 91, No. 2, pp: 521-522.

۴. **Dengjel, J.; Kratchmarova, I. and Blagoev, B., 2009.** Receptor tyrosine kinase signaling: a view from quantitative proteomics. Molecular biosystems. Vol. 10, No. 5, pp: 12-21.

۵. **Dopazo, J., 2006.** Functional interpretation of microarray experimentsOmics: a journal of integrative biology. Vol. 10, No. 3, pp: 398-410.

۶. **Finucane, K.A.; McFadden, T.B.; Bond, J.P.; Kennelly, J.J. and Zhao, F.Q., 2008.** Onseft of lactation in the bovine mammary gland: gene expression profiling indicates a strong inhibition of Gene expression in cell proliferation. Functional and integrative genomics. Vol. 8, No. 3, pp: 251-264.

۷. **Friedman, R.C.; Farh, K.K.H.; Burge, C.B. and Bartel, D.P. 2009.** Most mammalian mRNAs are conserved targets of microRNAs. Genomics research. Vol. 19, No. 1, pp: 92-105.

۸. **Goujon, M.; Mcwilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J. and Lopez, R., 2010.** A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic acids research. Vol. 38, No. 2, pp: 695-699.

۹. **Gupta, R. and Achenie, L.E., 2007.** A network model for gene regulation. Computers and chemical engineering. Vol. 31, No. 8, pp: 950-961.

۱۰. **Kharrati, K.H.; Mohammadabadi, M.R.; Ansari, M.S.; Esmailizadeh, A.K.; Tarang, A. and Nikbakhti, M., 2012.** Effect of DGATI variants on milk composition traits in Iranian Holstein cattle population Iran. Journal of animal Science. Vol. 3, No. 3, pp: 185-192.

۱۱. **Khatri, P. and Draghici, S., 2005.** Ontological analysis of gene expression data: current tools, limitations, and open problems. Bioinformatics. Vol. 21, No.1, pp: 3587-3595.

۱۲. **Li, S.; Becich, M.J. and Gilbertson, J., 2004.** Microarray data mining using gene ontology. In MEDINFO. pp: 778-782.

۱۳. **Niemi, J., 2007.** Accuracy of the Bayesian Network Algorithms for inferring gene regulatory network in engineering physics and mathematics systems analysis laboratory: Helsinki university of technology. Vol. 2, No. 108, pp: 1-21.

۱۴. **Rawool, B. and Venkatesh, V., 2007.** Steady state approach to model gene regulatory networks simulation of microarray experiments. Biosystems. Vol. l90, No. 3, pp: 636-655.

۱۵. **Shin, K.C.; Chen, R.M.; Hu, R.M.; Liu, F.M.; Chen, H.K. and Tsai, J.J., 2004.** Prediction of gene regulatory networks using differential expression of cDNA microarray data. In IEEE Sixth international symposium on multimedia software engineering. pp: 378-385.

۱۶. **Smith, B.; Williams, J. and Schulze, K., 2003.** The Ontology of the Gene Ontology. In American medical informatics association Annuesymp proc. pp: 609-613.

۱۷. **Sticht, C.; De, C.; Parveen, A. and Grtez, N., 2018.** miRwalk; an online resource for prediction of microRNA binding sites. Plos One. Vol. 13, No. 10, pp: e0206239.

۱۸. **Wang, C.; Jing, R.; Mao, X.; Chang, X. and Li, A., 2011.** TaABC1, a member of the activity of bc1 complex protein kinase family from common wheat, confers enhanced tolerance to abiotic stresses in Arabidopsis. Journal of Experimental Botany. Vol. 62, No. 1, pp: 1299-1311.

۱۹. **Yang, J. and Weinberg, R.A., 2008.** Epithelial mesen chymal transition at the crossroads of development and tumor metastasis. Developmental Cell. Vol. 14, No. 6, pp: 818-829.

۲۰. **Yavari, F.; Towhidkhad, F. andGharibzadeh, S. 2008.** Modelling large-sccale gene regulatory networks using gene ontology-based clustering and dynamic Bayesian network. In the 2nd international conference on bioinformatics and biomedical engineering. pp: 297-308.

(DOI): 10.22034/aej.2020.118071

Scientific Research Journal of Animal Environment Vol. 12, No. 3, Autumn 2020

Clustering based on the ontology of microRNAs target genes affecting milk production

- **Saeideh Eskandarinasab:** Department of Animal Science, Yasouj University, Yasouj, Iran
- **Zahra Roudbari*:** Department of Animal Science, Faculty of Agriculture, University of Jiroft, Jiroft, Iran
- **Mohammad Reza Bahreini Behzadi:** Department of Animal Science, Yasouj University, Yasouj, Iran

Received: December 2019 Accepted: March 2020

Keyword: Ontology Gene, MicroRNA, Gene Cluster, Milk production

Abstract

Cow's milk is the most nutritious drinks and it's an important part of a person's diet, and it has all the materials the human body needs. Identification of genes related to milk production and composition is a strategy which through can be effectively influenced milk production. Another strategy is to identify microRNAs that affect the expression of effective genes. In this study, we were downloaded the microRNAs data related to the mammary gland of dairy cattle with E-GEOD-61227 accession number from the GEO database. After estimating the expression level of microRNAs, target genes for microRNAs were identified using the miRwalk database. In the next step, AgriGO software was used for clustering of genes based on the tree structure of the gene ontology, which includes three sub-ontology of biological process, molecular function, and cellular component. In the present study, based on the clustering of target genes, the results showed that the biological process of gland morphogenesis and epithelial cell proliferation were highly significant. These processes play an important role in the development of the mammary glands and the results based on molecular activity showed that the signaling using transmembrane receptor protein kinase had the highest significant level. Also, the results of the cellular compartment showed that most of the target genes are located in the intracellular membrane-bounded organelle and the plasma membrane. Therefore, the target genes that regulate these significant processes have the potential to play a significant role in finding a targeted solution to improve milk production.

* Corresponding Author's email: roudbari.zahra@ujiroft.ac.ir