

فرا تحلیل (متا-آنالیز) داده‌های بیان ژن بافت پستان آلوده شده با باکتری اشریشیاکلی در گاوهای شیری

سمیه شریفی^۱، عباس پاکدل^{۲*} و اسماعیل ابراهیمی^۳

۱ و ۲. دانشجوی دکتری و دانشیار، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان

۳. دانشیار، گروه بیوتکنولوژی دانشگاه شیراز و عضو گروه تحقیقاتی دانشگاه آدلاید- استرالیا

(تاریخ دریافت: ۱۳۹۶/۲/۷ - تاریخ پذیرش: ۱۳۹۶/۷/۲۸)

چکیده

تشخیص ژن‌های درگیر در صفات پیچیده مانند حساسیت به بیماری نه تنها می‌تواند موجب بهبود تشخیص و پیشگیری از بیماری مورد نظر می‌شود، بلکه در انتخاب راه‌های درمانی کارآمد و همچنین در انتخاب دام‌های مقاوم به اصلاح‌گران کمک خواهد کرد. در این تحقیق با هدف بالا بردن توان تجزیه آماری در شناسایی ژن‌ها و مسیرهای زیستی (بیولوژیکی) درگیر در بیماری ورم پستان، از فراتحلیل به روش Fisher برای یکی کردن *p-value*‌های به دست آمده از تجزیه انفرادی داده‌های شش بررسی ریزآرایه که بیان ژن بافت پستان هنگام درگیری با باکتری اشریشیاکلی (*E. coli*) در گاوهای شیری را بررسی کرده بودند، استفاده شد. شناسایی ژن‌هایی که در هیچ یک از بررسی‌های انفرادی معنی دار نشده بودند می‌تواند تأییدکننده هدف بیان شده باشد که منجر به ارائه مجموعه کامل‌تری از مسیرهای زیستی مرتبط با سامانه ایمنی، التهاب، تجزیه پروتئین (پروتئولیز) و مسیرهای مرتبط با رشد و افزودن و مرگ یاخته‌ای شد. "کنترل مثبت بیان پروموتور RNA پلیمراز II" مسیر جدیدی در رابطه با این بیماری است که با وجود در بر گرفتن بیشترین شمار ژن در این بررسی، در بررسی‌های گذشته مرتبط با ورم پستان گزارش نشده است.

واژه‌های کلیدی: آنتولوژی ژن، اشریشیاکلی، تفاوت بیان ژن، گاو شیری، ورم پستان.

Meta-analysis of transcriptomic data of mammary gland infected by *Escherichia coli* Bacteria in dairy cows

Somayeh Sharifi¹, Abbas Pakdel^{2*} and Esmaeil Ebrahimi³

1, 2. Ph. D. Candidate and Associate Professor, Faculty of Agriculture, Isfahan University of Technology, Isfahan, Iran

3. Associate Professor, Institute of Biotechnology, Shiraz University, Shiraz, Iran, and School of Biological Science, Faculty of Science and Engineering, Flinders University of Adelaide, Australia

(Received: Apr. 27, 2017 - Accepted: Oct. 20, 2017)

ABSTRACT

Identification of disease-causing genes that underlie complex traits such as susceptibility to disease not only can improve diagnosis and the prevention of illness, but also help breeder to select resistance animals against diseases. In the current study to aim the higher power of statistical analysis to identification of genes and biological pathways related to mastitis disease, we used Fisher meta-analysis to combine *p-values* obtained from individual analysis of datasets extracted from 6 microarray-based studies which investigate transcriptomic data of mammary gland tissue infected by *Escherichia coli* (*E. coli*) in dairy cows. Identification of genes that did not show a significant *p-value* in any of the independent studies may confirm the aim and lead to introduce a more complete set of biological pathways involved in this disease such as the pathways related to immune response, inflammation, proteolysis, growth, and death of cell. *Positive regulation of transcription from RNA polymerase II promoter*, is new pathways related to this disease which despite of the enrichment by maximum number of up-regulated genes in this study, have not been reported in previous mastitis studies.

Keywords: Dairy cows, Differential express gene, *Escherichia coli*, gene ontology, mastitis.

* Corresponding author E-mail: pakdel@cc.iut.ac.ir

مقدمه

باکتری‌های گرم منفی مانند *اشریشیاکلی (E. coli)* یکی از بیماری‌گر (پاتوژن)های اصلی مسبب ورم پستان در گاو شیری هستند که زیان اقتصادی شایان توجهی به دامداری‌های گاو شیری تحمیل می‌کند (Halasa et al., 2007). اگرچه ممکن است این بیماری به‌خودی‌خود بهبود یابد (Bannerman et al., 2004) ولی گاهی کشنده بوده و یا به‌وفور منجر به حذف دام می‌شود (Hagiwara et al., 2016). تحقیقات صورت گرفته در این زمینه نشان داده است، آلودگی با *E. coli* پاسخ ایمنی موضعی^۱ و ایمنی عمومی^۲ را در دام تحریک می‌کند (Jiang et al., 2008; Mitterhuemer et al., 2010; Jorgensen et al., 2012). تشخیص ژن‌های درگیر در صفات پیچیده‌ای مانند حساسیت به بیماری‌ها نه تنها می‌تواند منجر به بهبود در تشخیص و پیشگیری از بروز بیماری‌ها شود (Clark & Pazdernik, 2012)، بلکه به اصلاح‌کنندگان در انتخاب دام‌های مقاوم به بیماری نیز کمک خواهد کرد. با تجزیه و تحلیل مجموعه داده‌های^۳ بیان ژن، می‌توان ژن‌هایی که در بافت بیمار نسبت به بافت سالم تفاوت بیانی^۴ نشان داده‌اند را مشخص کرد. در بررسی‌های چندی از داده‌های ریزآرایه^۵، برای بررسی پاسخ بافت پستان پس از آلودگی با یک بیمارگر خاص استفاده شده است، لیکن مقایسه مستقیم نتایج تجزیه و تحلیل این بررسی‌ها به دلیل متفاوت بودن طرح آزمایش و پلتفرم‌های ریزآرایه استفاده شده ممکن نیست. از سویی یکی از کاستی‌های داده‌های بیان ژن، محدود بودن شمار نمونه‌ها نسبت به شمار بالای ژن‌هایی است که به‌طور همزمان در هر پلتفرم اندازه‌گیری می‌شوند. این امر موجب شده است نتایج و یافته‌های هر بررسی تکرارپذیری کمی داشته باشد (Ramamany et al., 2008). امروزه حجم بالایی از داده ریزآرایه در منابع عمومی مانند GEO در سایت NCBI^۶ و Array

Express در سایت EMBL_EBI^۷ برای پژوهشگران قابل دسترس است که امکان بازیابی، یکی کردن و تجزیه دوباره این داده‌ها را برای استخراج داده‌های بیشتر فراهم کرده است (Moreau et al., 2003). فراتحلیل^۸، یک ابزار آماری مناسب برای ترکیب کردن چندین بررسی با فرضیه‌های همسان است که امکان ترکیب داده‌های ریزآرایه از بررسی‌های مختلف را در اختیار ما قرار می‌دهد. در بررسی‌های فراتحلیل صورت گرفته توسط Genini et al. (2011) و Looor et al. (2011) مجموعه‌ای از بررسی‌های مرتبط با ورم پستان برای انجام فراتحلیل استفاده شده است که داده‌های بیان ژن را با استفاده از پلتفرم‌های مختلف ریزآرایه به دست آورده‌اند. لیکن چون هر پلتفرم مجموعه شناساگرهای^۹ مختلفی دارد، یکی کردن آن‌ها موجب از دست رفتن میزان شایان توجهی از داده‌ها شده است، به‌طوری‌که در نهایت فراتحلیل روی ۱۳۱۶۲ شناساگر انجام شد. از سوی دیگر در بررسی این محققان پاسخ بیان ژن یا بیان ژنی مشترک^{۱۰} در بافت پستان چندین گونه (بز، گاو و گوسفند) که توسط بیمارگرهای مختلفی همچون استرپتوکوکوس اوریس^{۱۱}، استرپتوکوکوس اورئوس^{۱۲} و *E. coli* آلوده شده بودند به‌طور همزمان بررسی شد. نتایج تحقیقات نشان داده است، ارتباط بیان ژن در بین نژادهای مختلف چالش‌هایی دارد و ارتباط راست نسخه‌ای (ارتولوژی) بین گونه‌های مختلف امر ساده‌ای نیست. لذا به‌خوبی مشخص شده است که بیان ژن حتی در گونه‌هایی که بسیار به هم نزدیک هستند و نیز در بسیاری از راست نسخه‌ها یا همسان (همولوگ)ها یک ارتباط ساده تک‌به‌تک نیست (Fierro et al., 2008). همچنین در بررسی‌های چندی نشان داده شده، باوجود اینکه گونه‌های متفاوت باکتریایی توانایی ایجاد عفونت در غده پستان گاو را دارند و پاسخ میزبان به این باکتری‌ها را به‌عنوان ورم پستان می‌شناسیم، ولی پاسخ ایمنی

7. <https://www.ebi.ac.uk/arrayexpress/>

8. Meta-analysis

9. Probes

10. Common transcriptional response

11. Streptococcus Uberis

12. Streptococcus Aureus

1. Local immune response

2. Systemic immune response

3. Datasets

4. Differentially expressed

5. Microarray

6. <https://www.ncbi.nlm.nih.gov/gds/>

پس از آلودگی گرفته شده بود. لذا در این تحقیق هر زمان به عنوان بررسی مستقل در نظر گرفته شد (در مجموع ۱۵ مجموعه داده).

پیش پردازش داده‌های ریزآرایه

کیفیت مجموعه داده‌های ریزآرایه پیش و پس از عادی‌سازی توسط PCA^۶ و Box Plot بررسی شد. عادی‌سازی چارکی^۷ و خلاصه‌سازی با کمک الگوریتم RMA با مقیاس لگاریتمی^۸ با کمک بسته نرم‌افزاری LIMMA (Linear Models for Microarray Data,) (Version 3.3.1) در نرم‌افزار R روی تک تک مجموعه داده‌ها اعمال شد (Irizarry et al., 2003). ریزآرایه‌ها به طور معمول مجموعه شناساگری دارند که یک ژن را معرفی می‌کنند. در بسیاری از بررسی‌ها از گرفتن میانگین اعداد بیان ژن و یا از بزرگ‌ترین عدد بیان به عنوان شاخص بیان آن مجموعه شناساگر استفاده می‌شود که معیارهای دقیقی برای انتخاب عدد بیان نیستند. در این بررسی از بین شناساگرهای مختلف معرفی‌کننده یک شناسه ژن، شناساگری با بزرگ‌ترین عدد در محدوده درون چارکی (IQR)^۹ به عنوان عدد بیان آن ژن انتخاب شد (Gentleman et al., 2005). به منظور کاهش نرخ شناسایی خطای نوع اول^{۱۰} تجزیه داده‌های ریزآرایه، ۱۰ درصد از ژن‌های بدون بیان^{۱۱} یا به عبارتی ژن‌هایی که میانگین بیانی پایینی در بیشتر مجموعه داده‌های مورد بررسی داشتند و همچنین ۱۰ درصد از ژن‌های بدون داده‌های مفید^{۱۲}، که شامل ژن‌هایی است که تغییر بسیار جزئی در بیشتر مجموعه داده‌ها داشتند حذف شدند. همه مرحله‌های تطابق و پالایش کردن با استفاده از بسته نرم‌افزاری MetaDe (Huang et al., 2012) در نرم‌افزار R انجام گرفت (Wang et al., 2009b; et al., 2012).

میزبان به طور معنی‌داری بر حسب گونه باکتریایی مورد هجوم متفاوت است (Riollet et al., 2000; Gunther et al., 2011; Schukken et al., 2011). بر همین پایه در بررسی‌های فراتحلیل انجام شده اگرچه مسیرهای زیستی (بیولوژیکی) جدیدی معرفی شده است، ولی نتایج کاملی ارائه نشده است. از سویی امکان شناسایی و جداسازی ژن‌هایی که افزایش یا کاهش بیان داشتند نیز در بررسی‌های این محققان فراهم نشده بود. در این تحقیق برای نخستین بار فراتحلیل داده‌های بیان ژنی پاسخ بافت پستان در گاو شیری در رویارویی با آلودگی با باکتری *E. Coli* به صورت اختصاصی بررسی شده است.

مواد و روش‌ها

داده‌های ریزآرایه

با جستجو در سایت PubMed^۱ و Google Scholar^۲ با کلید واژه‌های *Bos Taurus*، Mastitis و *E-Coli* بررسی‌های مرتبط با ورم پستان در گاو گردآوری شد. دوازده بررسی داده‌های ریزآرایه مرتبط با این تحقیق داشت که شش مورد از آنها برای انجام فراتحلیل انتخاب شدند. داده‌هایی که حذف شدند یا پلتفرم‌های نامعتبر یا غیرتجاری داشتند، یا مستندسازی^۳ کاملی در اختیار ما قرار نمی‌دادند، یا مقاله ثبت‌شده معتبری نداشتند و یا داده‌های کاملی در مورد چگونگی انجام کار ارائه ندادند. در شش بررسی استفاده شده در این تحقیق از پلتفرم افیمتریکس^۴ برای استخراج مجموعه رونوشت‌ها (ترنسکرپتوم) استفاده شده بود. در داده‌های مورد استفاده تنها از نمونه‌هایی که با *E. coli* آلوده شده بودند همراه با نمونه‌های مهار آن‌ها بدون هیچ‌گونه تیمار یا مصرف پادزی (آنتی‌بیوتیک) برای انجام فراتحلیل استفاده شد (جدول ۱). در مجموع ۱۳۰ نمونه برای انجام فراتحلیل استفاده شد (۵۷ نمونه سالم و ۷۳ نمونه آلوده). در بررسی‌های یادشده، نمونه‌گیری‌ها در دوره‌های زمانی^۵ مختلفی

6. Principal Component Analysis

7. Quantile Normalization

8. log scale Robust Multi-array Average

9. InterQuartile Range

10. False Discovery Rate

11. Non-Expressed Genes

12. Non-informative Genes

1. <https://www.ncbi.nlm.nih.gov/pubmed/>

2. <https://scholar.google.com/>

3. Annotation

4. <http://www.affymetrix.com/index.affx>

5. Time series

جدول ۱. خلاصه داده‌های استفاده شده در فراتحلیل از مجموعه داده‌های ریزآرایه
Table 1. Summary of the microarray datasets employed in the meta-analysis

Accession number	Publication	Techniques	Strains of <i>E. coli</i>	Samples (ctr:tr)	Samples used in this study (ctr:tr)*	Time (hours)
GSE15025	(Mitterhuemer <i>et al.</i> , 2010)	In-vivo	1303	15:15	10:10	6,24
GSE24217	(Buitenhuis <i>et al.</i> , 2011)	In-vivo	K2BH2	23:26	23:26	24,192
GSE24560	(Brand <i>et al.</i> , 2011)	In-vitro	-	27:61	10:14	1,6,24
GSE25413	(Gunther <i>et al.</i> , 2011)	In-vitro	1303	6:24	3:12	1,3,6,24
GSE32186	(Gunther <i>et al.</i> , 2012)	In-vitro	1303	6:18	6:6	6,6***
GSE50685	(Sipka <i>et al.</i> , 2014)	In-vivo	ECC-Z	5:15	5:5	24,48

* تعداد نمونه‌های سالم: تعداد نمونه‌های تیمار

** زمان نمونه‌گیری بعد از آلودگی

*** سلول‌های کشت شده به مدت ۳۰ ساعت (آزمایش انتظار کوتاه) یا ۶۰ ساعت (آزمایش انتظار طولانی) بعد از شروع آزمایش

* Number of healthy samples: number of treatment samples.

** Time of sampling after infection.

*** Cells were harvested either 30h (short waiting experiment) or 60 h (long waiting experiment) after the start of the trial.

فراتحلیل

تنها ژن‌هایی که تفاوت بیانی نشان داده‌اند را مشخص می‌کند. به همین دلیل در این تحقیق برای به دست آوردن ژن‌هایی که افزایش بیان و یا کاهش بیان نشان داده‌اند از p -value یک‌سویه استفاده شد و به صورت جداگانه فراتحلیل روی آن‌ها اعمال شد، به طوری که p -value سمت راست ژن‌هایی که افزایش بیان داشته‌اند و p -value سمت چپ ژن‌هایی که کاهش بیان داشته‌اند را به ما نشان می‌دهد. تجزیه بررسی‌های انفرادی و همچنین فراتحلیل آن‌ها با استفاده از بسته نرم‌افزاری MetaDe (Version 1.0.5) در برنامه R (نسخه ۳.۳.۱) انجام گرفت.

درون‌شناسی ژن

ژن‌هایی که پس از فراتحلیل به صورت معنی‌داری دارای بیان متفاوتی بودند، (متا-ژن‌ها) وارد پایگاه داده‌هایی آنلاین DAVID^۳ نسخه ۶.۸ شدند تا با طبقه‌بندی درون گروه‌های مستندسازی شده درون‌شناسی (آنولوژی) ژن بتوان دید جامع‌تری از ژنگان یا ژن نام‌شناسی (ژنومیکس) عملکردی^۴ آن‌ها به دست آورد. درون‌شناسی ژن را می‌توان از سه دیدگاه فرآیند زیستی^۵، اجزای یاخته‌ای^۶ و عملکرد مولکولی^۷ بررسی کرد. در این تحقیق تنها GOهایی که توسط Modified Fisher Exact test غنی شده^۸ بودند و پس از تصحیح توسط روش Benjamini & Hochberg (2009a) با آستانه $q < 0.01$ معنی‌دار اعلام شدند، انتخاب شدند.

در مجموعه داده‌هایی که در این تحقیق برای انجام فراتحلیل استفاده شدند، به طور معمول نمونه‌گیری‌ها هم در آغاز و هم در اوج بیماری و در ساعت‌های مختلفی پس از تلقیح بیمارگر تهیه شده بودند. بسته به هدف بررسی می‌توان دامنه‌های خاصی تعریف کرد و فراتحلیل را در دامنه‌های زمانی مشخص شده به صورت جداگانه انجام داد (Borenstein *et al.*, 2009). لیکن در این بررسی چون هدف یافتن ژن‌ها و مسیرهای زیستی مهم درگیر در بیماری ورم پستان می‌بود و مقایسه بیان ژن‌ها در ساعت‌های اولیه ابتلا به بیماری با حالت حاد یا اوج بیماری مدنظر نبود، هر نوبت زمانی به عنوان بررسی جداگانه در نظر گرفته شد تا با افزایش شمار نمونه‌ها، بتوان ژن‌هایی که تفاوت بیانی بین نمونه‌های سالم و آلوده را نشان می‌دهند، با توان آماری بالاتری شناسایی کرد. بنابراین در مجموع پانزده مجموعه داده برای انجام فراتحلیل استفاده شد.

در آغاز آماره t تغییر یافته^۱ برای مقایسه سطح بیان هر ژن در نمونه‌های بیمار در مقایسه با نمونه‌های سالم در هر بررسی اعمال شد. برای رسیدن به این هدف از عملگر modt در بسته نرم‌افزاری MetaDe استفاده شد. P -value های به دست آمده برای هر ژن در هر بررسی با استفاده از روش Fisher (1950) و با استفاده از بسته (پکیج) MetaDe یکی شدند. برای تصحیح خطای نوع اول برای آزمون‌های چند متغیره^۲ از روش توصیه شده توسط Hochberg & Benjamini (1990) استفاده شد. به طور معمول فراتحلیل به روش یکی کردن p -value

1. Modified T-test
2. Multivariate test

3. <http://david.abcc.ncifcrf.gov/home.jsp>

4. Functional genomics

5. Biological process

6. Cellular component

7. Molecular function

8. Enriched

جدول ۲. اصطلاح‌های درون‌شناختی ژن مرتبط با فرآیند زیستی با استفاده از سایت DAVID و به‌دست‌آمده به کمک modified Fisher Exact test به دنبال تصحیح Benjamini & Hochberg ($q < 0.01$) روی ژن‌های افزایش بیان یافته‌ی به‌دست‌آمده از فراتحلیل

Table 2. Gene ontology related to biological process terms using DAVID online program generated by modified Fisher Exact test followed by the Benjamini & Hochberg correction ($q < 0.01$) on selected up-regulated gene achieved by meta-analysis

عبارات آنتولوژی ژن	تعداد	%	P-value	تصحیح Benjamini & Hochberg
Gene ontology Terms	Count	%	P-value	Benjamini & Hochberg correction
GO:0045944~positive regulation of transcription from RNA polymerase II promoter	70	6.18	0.0000	0.0003
GO:0006954~inflammatory response	56	4.95	0.0000	0.0000
GO:0045087~innate immune response	51	4.51	0.0000	0.0000
GO:0006955~immune response	43	3.80	0.0000	0.0000
GO:0043066~negative regulation of apoptotic process	35	3.09	0.0000	0.0039
GO:0006915~apoptotic process	32	2.83	0.0000	0.0017
GO:0042127~regulation of cell proliferation	27	2.39	0.0000	0.0006
GO:0051607~defense response to virus	26	2.30	0.0000	0.0002
GO:0043123~positive regulation of I-kappaB kinase/NF-kappaB signaling	26	2.30	0.0000	0.0002
GO:0051092~positive regulation of NF-kappaB transcription factor activity	24	2.12	0.0000	0.0000
GO:0032496~response to lipopolysaccharide	23	2.03	0.0000	0.0000
GO:0008360~regulation of cell shape	21	1.86	0.0000	0.0021
GO:0071222~cellular response to lipopolysaccharide	19	1.68	0.0000	0.0001
GO:0001666~response to hypoxia	17	1.50	0.0000	0.0023
GO:0032755~positive regulation of interleukin-6 production	16	1.41	0.0000	0.0000
GO:0030593~neutrophil chemotaxis	16	1.41	0.0000	0.0001
GO:0071356~cellular response to tumor necrosis factor	16	1.41	0.0000	0.0006
GO:0031663~lipopolysaccharide-mediated signaling pathway	15	1.33	0.0000	0.0000
GO:0070098~chemokine-mediated signaling pathway	15	1.33	0.0000	0.0001
GO:0060326~cell chemotaxis	15	1.33	0.0000	0.0003
GO:0071347~cellular response to interleukin-1	15	1.33	0.0000	0.0003
GO:0050728~negative regulation of inflammatory response	15	1.33	0.0000	0.0015
GO:0007229~integrin-mediated signaling pathway	15	1.33	0.0001	0.0071
GO:0050729~positive regulation of inflammatory response	14	1.24	0.0000	0.0004
GO:0006919~activation of cysteine-type endopeptidase activity involved in apoptotic process	13	1.15	0.0000	0.0030
GO:0045071~negative regulation of viral genome replication	12	1.06	0.0000	0.0000
GO:0050727~regulation of inflammatory response	12	1.06	0.0000	0.0034
GO:0051603~proteolysis involved in cellular protein catabolic process	11	0.97	0.0001	0.0062
GO:0071260~cellular response to mechanical stimulus	11	0.97	0.0001	0.0076
GO:0071346~cellular response to interferon-gamma	11	0.97	0.0001	0.0076
GO:0048661~positive regulation of smooth muscle cell proliferation	10	0.88	0.0000	0.0006
GO:0034097~response to cytokine	10	0.88	0.0000	0.0018
GO:0007249~I-kappaB kinase/NF-kappaB signaling	10	0.88	0.0000	0.0030
GO:0033209~tumor necrosis factor-mediated signaling pathway	9	0.80	0.0000	0.0027
GO:2001238~positive regulation of extrinsic apoptotic signaling pathway	9	0.80	0.0000	0.0036
GO:0032722~positive regulation of chemokine production	8	0.71	0.0000	0.0011
GO:0002474~antigen processing and presentation of peptide antigen via MHC class I	8	0.71	0.0000	0.0011

شده است. دو روش دیگر شامل شمارش آراء^۴ (Hedges & Olkin 1980) و یکی کردن مستقیم^۵ (Johnson *et al.*, 2007; Xia *et al.*, 2013) است. روش پایه‌ای فراتحلیل، بر پایه‌ی اندازه‌ اثرگذاری است. لیکن هنگامی از داده‌هایی که از جمعیت‌های متفاوت و یا با کیفیت‌های مختلف به دست آمده‌اند استفاده می‌شود، روش‌هایی مانند روش *p-value* و یا روش رتبه نتایج بهتری خواهند داشت. از سویی در شرایطی که متغیر همبسته^۶ وجود داشته باشد یکی کردن *p-value*ها بهترین کاربرد را خواهد داشت. در شرایطی که تصمیم داریم داده‌های ریزآرایه را به‌طور مستقیم با

نتایج و بحث

فراتحلیل، روشی آماری برای یکی کردن داده‌های بررسی‌های مختلف با فرضیه‌های مرتبط است که روشی ارزشمند برای افزایش توان تجزیه آماری و نیز بالا بردن تکرارپذیری نتایج نسبت به تجزیه بررسی‌های انفرادی است (Ramasamy *et al.*, 2008). پنج روش اصلی برای انجام فراتحلیل وجود دارد که شامل یکی کردن اندازه‌ اثرگذاری^۱ (Borenstein *et al.*, 2009; Yin *et al.*, 2015) *P-value* (Rhodes *et al.*, 2002) و رتبه^۲ (Hong *et al.*, 2006) است؛ که در این سه روش از داده‌های سطح خلاصه‌شده^۳ استفاده

4. Vote counting
5. Direct merging
6. Covariate

1. Effect size
2. Rank
3. Summary level data

شناسایی ژن‌هایی است که تفاوت بیانی نشان می‌دهند. با بررسی فرآیندهای زیستی و شبکه‌های کارکردی توسط تحلیل درون‌شناسی (آنالیز آنتولوژی) ژن، متا-ژن‌هایی که کاهش بیان نشان داده بودند هیچ گروه عملکردی را معنی‌دار نکردند. همان‌طور که در جدول ۲ نمایش داده شده است متا-ژن‌های افزایش بیان یافته در ۳۷ گروه عملکردی زیستی با $q < 0.01$ معنی‌دار شده‌اند. بیشتر گروه‌های زیستی معنی‌دار شده در این بررسی با سامانه ایمنی، التهاب^۶ و پاسخ مرحله^۷ حاد^۸ مرتبط هستند که در جدول ۲ فهرست شده‌اند که از جمله^۹ این‌ها می‌توان به پاسخ به التهاب^{۱۰}، پاسخ ایمنی ذاتی^{۱۱}، پاسخ ایمنی، پاسخ دفاعی به ویروس^{۱۲} و مانند آن اشاره کرد. نشان داده شده است که باکتری‌های گرم منفی مانند *E. coli* موجب پاسخ سریع و بیان میزان بالایی از سایتوکین‌ها و میانجیگری‌های التهابی در شیر می‌شوند (Bannerman, 2009). پاسخ ایمنی ذاتی نخستین مرحله از واکنش بدن به ورود بیمارگرها است و عملکرد آن شناسایی بیمارگرهایی است که بدن با آن‌ها روبه‌رو شده‌اند (Uthaisangsook et al., 2002). پروتئین‌های مرحله^{۱۳} حاد، مولکول‌های سرم هستند که توسط بسیاری از یاخته‌ها به‌ویژه هپاتوسایت‌ها^{۱۱} ترشح می‌شوند که نخستین بار توسط Abernethy & Avery (1941) معرفی شد و در واقع پاسخ اندام (ارگان‌سیم) به آسیب، آلودگی و زخم بافتی و بیماری‌هایی که سامانه ایمنی را درگیر می‌کنند است. پاسخ مرحله^{۱۳} حاد از راه درشت‌خوار (ماکروفاژ)‌های بافت مورد هدف و یا مونوسایت‌های خون که میزان گسترده‌ای از میانجی‌گری‌ها از جمله سایتوکین‌ها را ترشح می‌کنند آغاز می‌شود و شامل یکسری واکنش‌های آبشار مانند^{۱۲} برای محافظت بافت از آسیب‌های بیشتر، حذف عامل آلودگی و افزایش جریان‌های التیام‌بخش^{۱۳}

هم یکی کنیم، باید برای اثرگذاری دسته^۱ تصحیح انجام گیرد (Johnson et al., 2007). در این تحقیق به دلیل وجود گونه‌های مختلف بیمارگر *E. coli* و همچنین نوع تهیه آن‌ها از جمله استفاده از باکتری‌های زنده یا کشته‌شده، متفاوت بودن زمان نمونه‌گیری پس از آلودگی، تفاوت در چگونگی به دست آوردن نقشه بیان ژن^۲ (به‌طوری‌که نیمی از بررسی‌های استفاده‌شده بیان ژن را پس از تزریق باکتری به دام زنده بررسی کرده بودند و نیم دیگر با تزریق باکتری به یاخته‌های بافت پستانی نقشه بیان ژن را به دست آورده بودند)، لذا از فراتحلیل به روش یکی کردن *p-value*ها برای بررسی کارآمدتر ژن‌هایی که تفاوت بیانی در حالت بیمار نسبت به کنترل را نشان داده‌اند استفاده شد. آرایه ژنی گاوی افیمتریکس^۳ ۲۴۱۲۸ مجموعه شناساگر^۴ دارد که شمار ۱۹۱۹۲ از این مجموعه بر پایه آخرین مستندسازی موجود (Annotation from March, 2016) مشخصات Entrez Gene ID داشتند. لذا تنها این شمار برای تجزیه استفاده شدند و دیگر مجموعه شناساگرها حذف شدند. مستندسازی آرایه گاوی افیمتریکس در سایت NetAffx Analysis Center^۵ قابل دسترس است. پس از انطباق ژن‌ها، از بین ۱۹۱۹۲ شناساگر مستندسازی‌شده موجود در آرایه گاوی افیمتریکس، ۱۲۸۸۸ شناسه^۶ منحصربه‌فرد به دست آمد. پس از انجام مراحل مختلف پیش‌پردازش داده‌ها، ۱۰۴۳۹ ژن باقی ماند و فراتحلیل روی این شمار ژن در ۱۳۰ نمونه اجرا شد. پس از انجام فراتحلیل در مجموع ۱۵۹۰ ژن تفاوت بیانی معنی‌دار نشان دادند. از بین این شمار ۴۴۵ ژن کاهش بیان و ۱۱۴۵ ژن افزایش بیان ($q < 0.05$ یک‌سویه) داشتند. شمار ۵۴۵ متا-ژن (۳۴٪) نیز در هیچ‌کدام از بررسی‌های منفرد معنی‌دار نشده بودند و این نتیجه گواهی محکمی بر بالاتر بودن توان فراتحلیل در

6. Inflammation

7. Acute phase response

8. Inflammatory response

9. Innate immune response

10. Defense response to virus

11. Hepatocytes

12. Cascade

13. Healing process

1. Batch effect

2. Transcriptome profile

3. Affymetrix Bovine Genome Array

4. Probe sets

5. [Http://www.affymetrix.com/support/technical/annotationfilesmain.affx](http://www.affymetrix.com/support/technical/annotationfilesmain.affx)

بخش اعظم ژن‌های افزایش بیان یافته را در بر می‌گیرد موجب التهاب و نشانه‌های حاد در کارتیۀ مبتلا شده می‌شود (Mitterhuemer *et al.*, 2010).

از دیگر مسیرهای زیستی مهم و معنی‌دار شده، پاسخ به هیپوکسی^{۱۰} است که نشان داده شده است عامل‌های القاکننده این مسیر، عامل‌های شناخته شده‌ای هستند که در کنترل ایمنی ذاتی و بیان ژن‌های پیش‌التهاب نقش دارند (Imtiyaz & Simon, 2010). افزون بر مسیرهای اصلی و کلیدی درگیر در ورم که در بخش بالا توضیح داده شد، مسیر زیستی مهمی موسوم به "مهار مثبت بیان راه‌انداز (پروموتور) RNA پلیمراز II^{۱۱}" در این بررسی معنی‌دار شد که با وجود پوشش دادن بیشترین شمار ژن در خود (۷۰ ژن) تا به حال در بررسی‌های مرتبط با این بیماری گزارش نشده است. در بررسی‌های پیشین نشان داده شده است که داده‌های ژنتیکی موجود در مولکول DNA یوکاریوت‌ها توسط آنزیم‌های RNA پلیمراز به RNA رونویسی می‌شود که RNA پلیمراز II یکی از آن‌هاست. مهار فعالیت این مولکول به میزان بالایی تحت کنترل ژن‌های انفرادی است که این تنظیم‌های اختصاصی هم برای هموستازی یاخته و هم برای گسترش برنامه‌ریزی شده موجودات چندیاخته‌ای ضروری است (Fuda *et al.*, 2009). نشان داده شده است که ژن‌های پاسخ‌دهنده به تنش مانند موارد پاسخ به آسیب‌های DNA^{۱۲}، پروتئین‌های باز^{۱۳} و مسیرهای سامانه ایمنی موجب متوقف شدن فعالیت RNA پلیمراز II می‌شوند (Guenther *et al.*, 2007). بنابراین به‌طور اصولی سازوکارهایی که تأثیر مثبت بر افزایش بیان این مولکول دارند در حین بروز پاسخ‌های ایمنی ضروری است.

نتیجه‌گیری کلی

در این بررسی با انجام فراتحلیل اختصاصی روی داده‌های بیان ژن گونه گاو آلوده شده به باکتری

به‌منظور برگرداندن خون بندآوری (هموستازی) است که پاسخ موضعی و همچنین عمومی را تحریک می‌کنند (Thulasiraman *et al.*, 2013). دیگر مسیرهای معنی‌دار شده در رابطه با افزونش^۱، فرآیند مرگ تنظیم‌شده یاخته‌ای^۲ شامل: تنظیم منفی روند مرگ تنظیم‌شده یاخته‌ای^۳، تنظیم افزونش یاخته‌ای^۴، تنظیم منفی مسیر پیام‌دهی مرگ تنظیم‌شده برون یاخته‌ای^۵ و پروتئین شکن درگیر در روند زیست‌سوزی (کاتابولیکی) پروتئین یاخته‌ای^۶ هستند. میانجیگری عامل‌های مرگ تنظیم‌شده که موجب محدودیت حیات در یاخته‌های روبه‌رو شده با عفونت‌های باکتریایی و ویروسی می‌شود در بررسی‌های پیشین به‌خوبی مورد تأیید قرار گرفته است. مرگ تنظیم‌شده یاخته‌ای یک عملکرد مهم سامانه ایمنی میزبان در رویارویی با بیمارگرهای وارد شده به یاخته است که به هدف‌های مختلفی صورت می‌گیرد که مهم‌ترین آن‌ها: مرگ یا کاهش پایداری بیمارگرهای وارد شده به یاخته و جلوگیری از انتشار و افزونش میکروب‌ها است (Elmore 2007). بررسی‌های پیشین نشان داده‌اند که ورم پستان ایجاد شده توسط باکتری *E. coli* هم پاسخ ایمنی موضعی و هم ایمنی سامانه‌ای یا همگانی را تحریک می‌کند، به‌طوری‌که پاسخ ایمنی همگانی موجب ایجاد تب و افزایش *TNF-alpha* در خون (Blum *et al.*, 2000) و همچنین موجب افزایش بیان بعضی از ژن‌ها از جمله *Haptoglobin* در کارتیۀ مجاور کارتیۀ آلوده شده (Hiss *et al.*, 2004) می‌شود. از سوی دیگر آلودگی با *E. coli* مسیرهای زیستی وابسته به پردازش و ترشح آنتی‌ژن^۷، تخریب پروتئین^۸، مرگ تنظیم‌شده^۹ یاخته‌ای و ژن‌های سایتوکین‌ها را نیز فعال می‌کند (Mysorekar *et al.*, 2002; Genini *et al.*, 2011). پاسخ ایمنی موضعی که به‌طور معمول

1. Proliferation
2. Apoptotic process
3. Negative regulation of apoptotic process
4. Regulation of cell proliferation
5. Positive regulation of Extrinsic apoptotic signaling pathway
6. Proteolysis involved in cellular protein catabolic process
7. Antigen processing and presentation
8. Protein degradation
9. Apoptosis

10. Response to hypoxia
11. Positive regulation of transcription from RNA polymerase II promoter
12. DNA-damage
13. Unfolded-protein

زیستی که در بررسی‌های گذشته شناسایی شده بود شامل: ارائه مجموعه کاملی از مسیرهای مرتبط با سامانه‌های ایمنی ذاتی و اکتسابی رشد و مرگ یاخته‌های التهاب و تجزیه پروتئین، منجر به کشف و معرفی مسیرهای جدیدی شد که در گزارش‌های انفرادی کمتر گزارش شده یا به کلی گزارش نشده است.

E. coli توانستیم توان تجزیه را با یکی کردن نتایج شش بررسی مرتبط بالا ببریم، به طوری که حدود ۳۴ درصد از ژن‌هایی که در این بررسی تفاوت بیانی معنی‌دار نشان دادند (۵۴۵ ژن از ۱۵۹۰ ژن) در هیچ‌کدام از بررسی‌های انفرادی معنی‌دار نشده بودند. افزایش توان تجزیه و افزایش شمار ژن‌هایی که تفاوت بیان معنی‌دار نشان دادند، ضمن تأیید مسیرهای

REFERENCES

1. Abernethy, T. J. & Avery, O. T. (1941). The Occurrence during Acute Infections of a Protein Not Normally Present in the Blood: I. Distribution of the Reactive Protein in Patients' Sera and the Effect of Calcium on the Flocculation Reaction with C Polysaccharide of Pneumococcus. *Journal of Experimental Medicine*, 73, 173-82.
2. Bannerman, D. D. (2009). Pathogen-dependent induction of cytokines and other soluble inflammatory mediators during intramammary infection of dairy cows. *Journal of Animal Science*, 87, 10-25.
3. Bannerman, D. D., Paape, M. J., Hare, W. R. & Hope, J. C. (2004). Characterization of the bovine innate immune response to intramammary infection with *Klebsiella pneumoniae*. *Journal of Dairy Science*, 87, 2420-32.
4. Blum, J. W., Dosogne, H., Hoeben, D., Vangroenweghe, F., Hammon, H. M., Bruckmaier, R. M. & Burvenich, C. (2000). Tumor necrosis factor-alpha and nitrite/nitrate responses during acute mastitis induced by *Escherichia coli* infection and endotoxin in dairy cows. *Domestic Animal Endocrinology*, 19, 223-35.
5. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom.
6. Brand, B., Hartmann, A., Reipsilber, D., Griesbeck-Zilch, B., Wellnitz, O., Kuhn, C., Ponsuksili, S., Meyer, H. H. & Schwerin, M. (2011). Comparative expression profiling of *E. coli* and *S. aureus* inoculated primary mammary gland cells sampled from cows with different genetic predispositions for somatic cell score. *Genetics Selection Evolution*, 43, 24.
7. Buitenhuis, B., Rontved, C. M., Edwards, S. M., Ingvarsen, K. L. & Sorensen, P. (2011). In depth analysis of genes and pathways of the mammary gland involved in the pathogenesis of bovine *Escherichia coli*-mastitis. *BMC Genomics*, 12, 130.
8. Clark, D. P. & Pazdernik, N. J. (2012). *Molecular Biology*. Elsevier.
9. Elmore, S. (2007). Apoptosis: A Review of Programmed Cell Death. *Toxicologic Pathology*, 35, 495-516.
10. Fierro, A. C., Vandenbussche, F., Engelen, K., de Peer, Y. V. & Marchal, K. (2008). Meta analysis of gene expression data within and across species. *Current genomics*, 9(8), 525-534
11. Fisher, R. A. (1950). *Statistical methods for research workers*. Edinburgh: Genesis Publishing, Oliver and Boyd.
12. Fuda, N. J., Ardehali, M. B. & Lis, J. T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature*, 461, 186-92.
13. Genini, S., Badaoui, B., Sclep, G., Bishop, S. C., Waddington, D., Pinard van der Laan, M. H., Klopp, C., Cabau, C., Seyfert, H. M., Petzl, W., Jensen, K., Glass, E. J., de Greeff, A., Smith, H. E., Smits, M. A., Olsaker, I., Boman, G. M., Pisoni, G., Moroni, P., Castiglioni, B., Cremonesi, P., Del Corvo, M., Foulon, E., Foucras, G., Rupp, R. & Giuffra, E. (2011). Strengthening insights into host responses to mastitis infection in ruminants by combining heterogeneous microarray data sources. *BMC Genomics*, 12, 225.
14. Gentleman, R., Carey, V., Huber, W., Irizarry, R. & Dudoit, S. (2005). *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer.
15. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, 130, 77-88.
16. Gunther, J., Esch, K., Poschadel, N., Petzl, W., Zerbe, H., Mitterhuemer, S., Blum, H. & Seyfert, H. M. (2011). Comparative kinetics of *Escherichia coli*- and *Staphylococcus aureus*-specific activation of key immune pathways in mammary epithelial cells demonstrates that *S. aureus* elicits a delayed response dominated by interleukin-6 (IL-6) but not by IL-1A or tumor necrosis factor alpha. *Infect Immun*, 79, 695-707.

17. Gunther, J., Petzl, W., Zerbe, H., Schubert, H. J., Koczan, D., Goetze, L. & Seyfert, H. M. (2012). Lipopolysaccharide priming enhances expression of effectors of immune defence while decreasing expression of pro-inflammatory cytokines in mammary epithelia cells from cows. *BMC Genomics* 13, 17.
18. Hagiwara, S., Mori, K. & Nagahata, H. (2016). Predictors of fatal outcomes resulting from acute *Escherichia coli* mastitis in dairy cows. *Journal of Veterinary Medical Science*, 78(5), 905-908.
19. Halasa, T., Huijps, K., Osteras, O. & Hogeveen, H. (2007). Economic effects of bovine mastitis and mastitis management. *Veterinary Quarterly*, 29, 18-31.
20. Hedges L. & Olkin I. (1980) Vote-counting methods in research synthesis. *Psychol Bull*, 88, 359-69.
21. Hiss, S., Mielenz, M., Bruckmaier, R. M. & Sauerwein, H. (2004). Haptoglobin concentrations in blood and milk after endotoxin challenge and quantification of mammary Hp mRNA expression. *Journal of Dairy Science*, 87(11), 3778-84.
22. Hochberg, Y. & Benjamini, Y. (1990). More powerful procedures for multiple significance testing. *Statistics in medicine*, 9(7), 811-818.
23. Hong, F., Breitling, R., McEntee, C. W., Wittner, B. S., Nemhauser, J. L. & Chory, J. (2006). RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics*, 22, 2825-7.
24. Huang, D. W., Sherman, B. T. & Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37, 1-13.
25. Huang, D. W., Sherman, B. T. & Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4, 44-57.
26. Imtiyaz, H. Z. & Simon, M. C. (2010). Hypoxia-inducible factors as essential regulators of inflammation. *Curr Top Microbiol Immunol*, 345, 105-20.
27. Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B. & Speed, T. P. (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research*, 31, e15.
28. Jiang, L., Sorensen, P., Rontved, C., Vels, L. & Ingvarsen, K. L. (2008). Gene expression profiling of liver from dairy cows treated intra-mammary with lipopolysaccharide. *BMC Genomics*, 9, 443.
29. Johnson, W. E., Li, C. & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, 8, 118-27.
30. Jorgensen, H. B., Buitenhuis, B., Rontved, C. M., Jiang, L., Ingvarsen, K. L. & Sorensen, P. (2012). Transcriptional profiling of the bovine hepatic response to experimentally induced *E. coli* mastitis. *Physiol Genomics*, 44, 595-606.
31. Mitterhuemer, S., Petzl, W., Krebs, S., Mehne, D., Klanner, A., Wolf, E., Zerbe, H. & Blum, H. (2010). *Escherichia coli* infection induces distinct local and systemic transcriptome responses in the mammary gland. *BMC Genomics*, 11, 138.
32. Moreau, Y., Aerts, S., De Moor, B., De Strooper, B. & Dabrowski, M. (2003). Comparison and meta-analysis of microarray data: from the bench to the computer desk. *Trends Genetics*, 19, 570-7.
33. Mysorekar, I. U., Mulvey, M. A., Hultgren, S. J. & Gordon, J. I. (2002). Molecular regulation of urothelial renewal and host defenses during infection with uropathogenic *Escherichia coli*. *Journal of Biological Chemistry*, 277, 7412-9.
34. Ramasamy, A., Mondry, A., Holmes, C. C. & Altman, D. G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Medicine*, 5, e184.
35. Rhodes, D. R., Barrette, T. R., Rubin, M. A., Ghosh, D. & Chinnaiyan, A. M. (2002). Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Research*, 62, 4427-33.
36. Riollet, C., Rainard, P. & Poutrel, B. (2000). Differential induction of complement fragment C5a and inflammatory cytokines during intramammary infections with *Escherichia coli* and *Staphylococcus aureus*. *Clin Diagn Lab Immunol*, 7, 161-7.
37. Schukken, Y. H., Gunther, J., Fitzpatrick, J., Fontaine, M. C., Goetze, L., Holst, O., Leigh, J., Petzl, W., Schubert, H. J., Sipka, A., Smith, D. G., Quesnell, R., Watts, J., Yancey, R., Zerbe, H., Gurjar, A., Zadoks, R. N., Seyfert, H. M. & Members of the Pfizer mastitis research c. (2011). Host-response patterns of intramammary infections in dairy cows. *Vet Immunol Immunopathol.* 144, 270-89.
38. Sipka, A., Klaessig, S., Duhamel, G. E., Swinkels, J., Rainard, P. & Schukken, Y. (2014)/ Impact of intramammary treatment on gene expression profiles in bovine *Escherichia coli* mastitis. *PLoS ONE*, 9, e85579.
39. Thulasiraman, P., Mukherjee, J., Banerjee, D., Pandiyan, G. D. V., Das, K., Ghosh, P. R. & Das, P. K. (2013). Acute Phase Proteins - a Potent Biomarker for Mastitis. *Vetscan* 7, 6-14.
40. Uthaisangsook, S., Day, N. K., Bahna, S. L., Good, R. A. & Haraguchi, S. (2002). Innate immunity and its role against infections. *Ann Allergy Asthma Immunol*, 88, 253-64; quiz 65-6, 318.
41. Wang, X., Kang, D. D., Shen, K., Song, C., Lu, S., Chang, L. C., Liao, S. G., Huo, Z., Tang, S., Ding, Y., Kaminski, N., Sibille, E., Lin, Y., Li, J. & Tseng, G. C. (2012). An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. *Bioinformatics*, 28, 2534-6.

42. Xia, J., Fjell, C. D., Mayer, M. L., Pena, O. M., Wishart, D. S. & Hancock, R. E. W. (2013). INMEX-a web-based tool for integrative meta-analysis of expression data. *Nucleic Acids Res.* 41, web server issue w63-w70.
43. Yin, L., Coelho, S. G., Valencia, J. C., Ebsen, D., Mahns, A., Smuda, C., Miller, S. A., Beer, J. Z., Kolbe, L. & Hearing, V. J. (2015). Identification of Genes Expressed in Hyperpigmented Skin Using Meta-Analysis of Microarray Data Sets. *Journal of Investigative Dermatology*, 135, 2455-63.