

Identification of key genes in bovine mammary epithelial cells challenged with *Escherichia coli* and *Staphylococcus aureus* by integrated bioinformatics analysis

Zeshi Chen¹ , Zhengyu Bian¹ , Xiuli Chen² , Boling Li³ , Lianbin Li^{1*} 

¹ Hainan University, College of Animal Science and Technology, Key Laboratory of Tropical Animal Breeding and Epidemic Disease Research of Hainan Province, Haikou, China.

² Animal Disease Prevention and Control Center of Qiongzhan, Haikou, China.

³ Hainan Animal Husbandry Technology Promotion Station, Haikou, China.

*Corresponding author:

lianbin@hainanu.edu.cn

Received: April 30, 2022

Accepted: November 24, 2022

How to cite: Chen, Z.; Bian, Z.; Chen, X.; Li, B. and Li, L. 2023. Identification of key genes in bovine mammary epithelial cells challenged with *Escherichia coli* and *Staphylococcus aureus* by integrated bioinformatics analysis. *Revista Brasileira de Zootecnia* 52:e20220064. <https://doi.org/10.37496/rbz5220220064>

Copyright: This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



ABSTRACT - This study integrated four microarray datasets by Robust Rank Aggregation (RRA) method to identify the differentially expressed genes (DEG) in bovine mammary epithelial (BME) cells in response to *Escherichia coli* and *Staphylococcus aureus* infection. Furthermore, the GO function and KEGG pathway enrichment analysis of the integrated DEG were performed. Finally, the protein-protein interaction (PPI) network was constructed. A total of 72 integrated DEG were identified from the four datasets. The most significantly enriched terms within the integrated DEG were mainly involved in the immune response. The PPI network of DEG was constructed with 53 nodes. Seventeen genes, which constitute a significant module, were identified as hub genes. Among them, *CD40*, *CXCL6*, and *NFKBIZ* were further screened as the key genes and have the potential to become biomarkers of *E. coli* and *S. aureus* mastitis, considering the specificity of biomarkers for diseases. The identified key genes and pathways in this study can assist in the search for biomarkers for mastitis diagnosis and disease resistance breeding.

Keywords: biomarkers, differentially expressed genes, *E. coli*, mastitis, *S. aureus*

1. Introduction

Cattle breeding, including dairy and beef cattle breeding, is the dominant livestock industry in many countries. It provides a significant source of nutrition to humans. However, pathogenic microbial infection is an important obstacle hindering the development of cattle breeding, causing significant losses and animal welfare problems. For example, bovine mastitis, an inflammatory disease caused mostly by the dysbiosis of the mammary gland microbiota (Steinberg et al., 2022), is the costliest disease in dairy farming. Bovine mastitis can cause a series of problems to the production system, such as a decrease in milk quality and production, changes in milk composition, and increased costs for treatments, labor and culling. These problems not only lead to huge economic losses but also create concerns for animal welfare and human health (Islam et al., 2020).

Escherichia coli and *Staphylococcus aureus* are the common pathogenic bacteria that lead to mastitis in dairy cows (Lin et al., 2021). A large number of studies have confirmed *E. coli* often leads to an acute and severe form of mastitis, while *S. aureus* leads to chronic but sub-clinical mastitis (Petzl et al., 2018). In addition, mastitis caused by *S. aureus* is frequently infectious (Leuenberger et al., 2019), whereas *E. coli* is usually environmental (Zaatout, 2022). However, despite many efforts dedicated to understanding the pathogenesis of *E. coli* and *S. aureus* mastitis, the pathogenic mechanism is still poorly understood. In addition, the current treatment of mastitis caused by pathogens is based on

antibiotic therapy, which has disadvantages including the development of resistance, causing drug residues in milk (Younis et al., 2016) and increasing milk production costs. Hence, new alternative options are urgently needed for mastitis treatment and prevention. Better knowledge on the pathogenic mechanism of mastitis is required to allow the development of new effective curative strategies. Moreover, the comprehensive molecular investigation on host-pathogen interactions may aid in the identification of new biomarkers for the rapid detection of mastitis caused by pathogens.

The expression profile of genes associated with *E. coli* and *S. aureus* mastitis infection in cows have provided very useful information for the understanding of the infectious process and screening differentially expressed genes (DEG) as rapid detection biomarkers. Many studies on data analysis of gene expression profile on mastitis infection of cows have been carried out, and hundreds of DEG have been identified (Han, 2019). For instance, a study showed that 447 and 520 DEG are involved in bovine mammary epithelial (BME) cells line developed by the authors after lipopolysaccharide and *S. aureus* stimulation, respectively (Islam et al., 2020). Another study identified 92 and 81 DEG in the primary BME cells challenged with *E. coli* or *S. aureus* in all periods (Han, 2019). It can be seen that in different studies, the DEG identified are not completely consistent. The reason is that different independent experiments are always limited by the heterogeneity of experimental samples with the environment, breed, population, and specific animal differences. Therefore, it is difficult to fully elucidate the molecular mechanism of this disease by using an expression profile obtained in a single independent study.

The Robust Rank Aggregation (RRA) method is a suitable and effective solution to obtain integrated DEG from different datasets performed by different technological platforms (Vösa et al., 2014). This method checks the ranking of each gene in each dataset and is based on the assumption that each gene identified in each experiment is randomly arranged. Therefore, the RRA method compares the ranking of the randomly sorted list with the baseline case, and a higher gene ranking is associated with a lower P-value (Gao et al., 2018). The RRA method has been used in many studies to integrate different genes under different conditions (Gao et al., 2018; Subramanian and Natarajan, 2021). However, this method has not been used to integrate and analyze the key genes in the state of mastitis. Hence, this study used the RRA method to integrate and analyze the results of multiple gene expression datasets to better understand the molecular mechanism of mastitis.

In this study, two Gene Expression Series, namely GSE24560 and GSE25413, were selected from the Gene Expression Omnibus (GEO) databases. Through the analysis of the database, this study identified the key genes and pathways in *E. coli* and *S. aureus* mastitis of dairy cows, to further clarify the immune response mechanism of dairy cows and explore candidate genes for dairy cow mastitis biomarkers. Therefore, this research focuses on finding the common points of the two bacterial infections, such as finding the common triggering mechanism and the common key genes in the body's response to the two bacterial infections.

2. Material and Methods

2.1. Dataset collection and analysis

The GSE24560 and GSE25413 were selected from the GEO databases (<https://www.ncbi.nlm.nih.gov/geo/>). Each dataset contains the expression profile of normal and challenged BME cells with *E. coli* or *S. aureus*, respectively. Each dataset is divided into two subsets according to whether it is processed by *E. coli* or *S. aureus* and renamed as GSE24560E, GSE24560S, GSE25413E, and GSE25413S. These studies were based on the platform of the GPL2112 (Affymetrix Bovine Genome Array). In total, 24 normal data, 26 challenged with *E. coli* and 26 challenged with *S. aureus*, were obtained (Table 1). These studies were all generated using primary bovine mammary epithelial cells. Datasets GSE24560E and GSE24560S were obtained after cells (3×10^5 cells/well) were treated with heat-inactivated *E. coli* isolates and *S. aureus* M60 (100 μ L of bacterial-solution) derived from bovine milk samples of mastitis for 6 and 24 h, respectively. Datasets GSE25413E and GSE25413S

Table 1 - Summary of the microarray datasets included in the analysis

| Accession number | Dataset name | Pathogen | Treatment time (h) | Samples ¹ (Con:Tr) | Reference |
|------------------|--------------|---|--------------------|-------------------------------|-----------------------|
| GSE24560 | GSE24560E | Heat-inactivated <i>E. coli</i> isolates | 6 and 24 | 18:20 | Brand et al. (2011) |
| | GSE24560S | Heat-inactivated <i>S. aureus</i> M60 | 6 and 24 | 18:20 | |
| GSE25413 | GSE25413E | Heat-inactivated <i>S. aureus</i> strain 1027 | 6 and 24 | 6:6 | Gunther et al. (2011) |
| | GSE25413S | Heat-inactivated <i>E. coli</i> strain 1303 | 6 and 24 | 6:6 | |

¹ Data were all generated using primary bovine mammary epithelial cells; the pathogens were derived from samples of mastitis; data of *E. coli* and *S. aureus* treatment (6 and 24 h) group and normal control group in each dataset were selected; Con:Tr - number of normal control samples: number of treatment samples.

were obtained after cells were treated with heat-inactivated particles of the mastitis-causing pathogens *E. coli* strain 1303 and *S. aureus* strain 1027 (10^7 particles/mL) for 6 and 24 h, respectively (Brand et al., 2011; Gunther et al., 2011).

Then, the GEO2R online analysis tool (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>) was used to normalize the data and compare the DEG in the samples of the normal control group and the treatment group. The significant DEG were selected by a threshold of P-value <0.05 and $|\log_{2}FC| > 0.585$. The upregulated and downregulated DEG lists were saved as TXT files independently for subsequent integration analysis.

2.2. Integration of microarray data

The TXT files of all DEG lists sorted by $|\log_{2}FC|$ were integrated using the RRA R package (version 1.1), and the integrated DEG were saved for subsequent analysis. This method aggregates the provided gene rankings into a list, and each gene is assigned a P-value. P<0.05 was chosen as a threshold for defining significant genes which were identified as the integrated DEG.

2.3. Enrichment analyses of DEG

The Gene Ontology (GO) annotation and notably relevant Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of the integrated DEG were performed using DAVID 6.8 database (<https://david.ncifcrf.gov/>). Gene counts >5 and P<0.01 were chosen as the threshold for defining statistical significance. R software (version 4.1.0) and ggplot2 package (version 3.3.5) were used for data visualization.

2.4. Protein-protein interaction network construction and analysis of the module

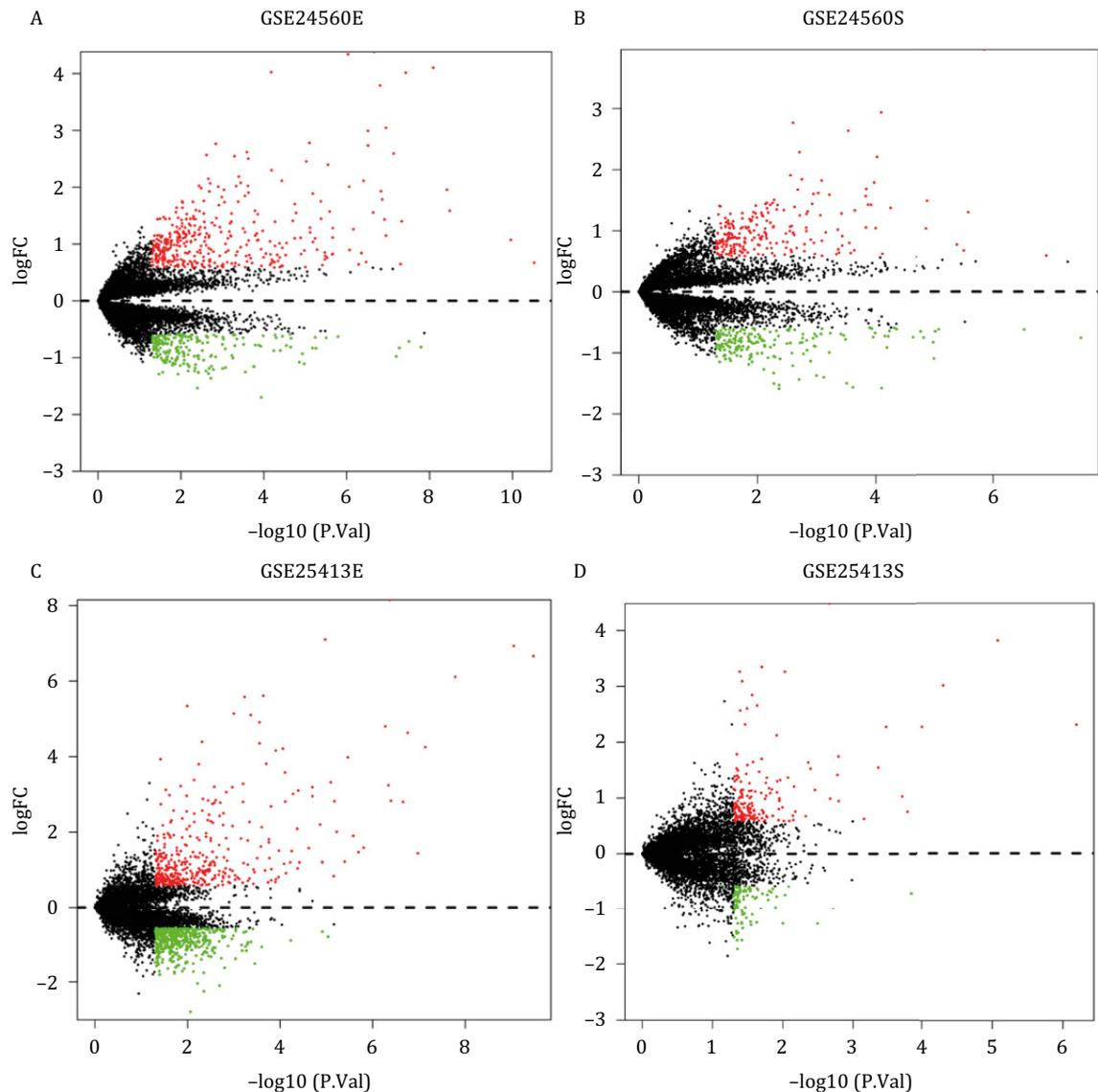
The STRING database v11.5 (<https://string-db.org/>) was used to analyze the interaction relationships between integrated DEG. Based on this information, a protein-protein interaction (PPI) network was visualized by Cytoscape, and Centiscape, a plugin in Cytoscape, was used to calculate the degree of each node. Then, the PPI network was analyzed by another Cytoscape plugin-MCODE to search the module of highly interconnected nodes.

3. Results

3.1. Identification of primitive DEG

By employing the GEO2R tool, the expression datasets were first normalized. Then, the primitive DEG were identified in the *E. coli*-infected (26 samples) and the *S. aureus*-infected (26 samples) group, respectively. Compared with the normal control group, the infected group in GSE24560E contained

390 upregulated genes, 209 downregulated genes, and a total of 599 DEG; the infected group in GSE24560S contained 227 upregulated genes, 195 downregulated genes, and a total of 422 DEG; the infected group in GSE25413E contained 387 upregulated genes, 497 downregulated genes, and a total of 884 DEG; the infected group in GSE25413S contained 179 upregulated genes, 90 downregulated genes, and a total of 269 DEG (Figure 1).



A: GSE24560E (heat-inactivated *E. coli* isolates); B: GSE24560S (heat-inactivated *S. aureus* M60); C: GSE25413E (heat-inactivated *E. coli* strain 1303); D: GSE25413S (heat-inactivated *S. aureus* strain 1027). Red dots represent the upregulated genes, green dots represent the downregulated genes, and black spots represent stable genes.

Figure 1 - Volcano plot of the primitive differentially expressed genes between the two groups of samples in each dataset.

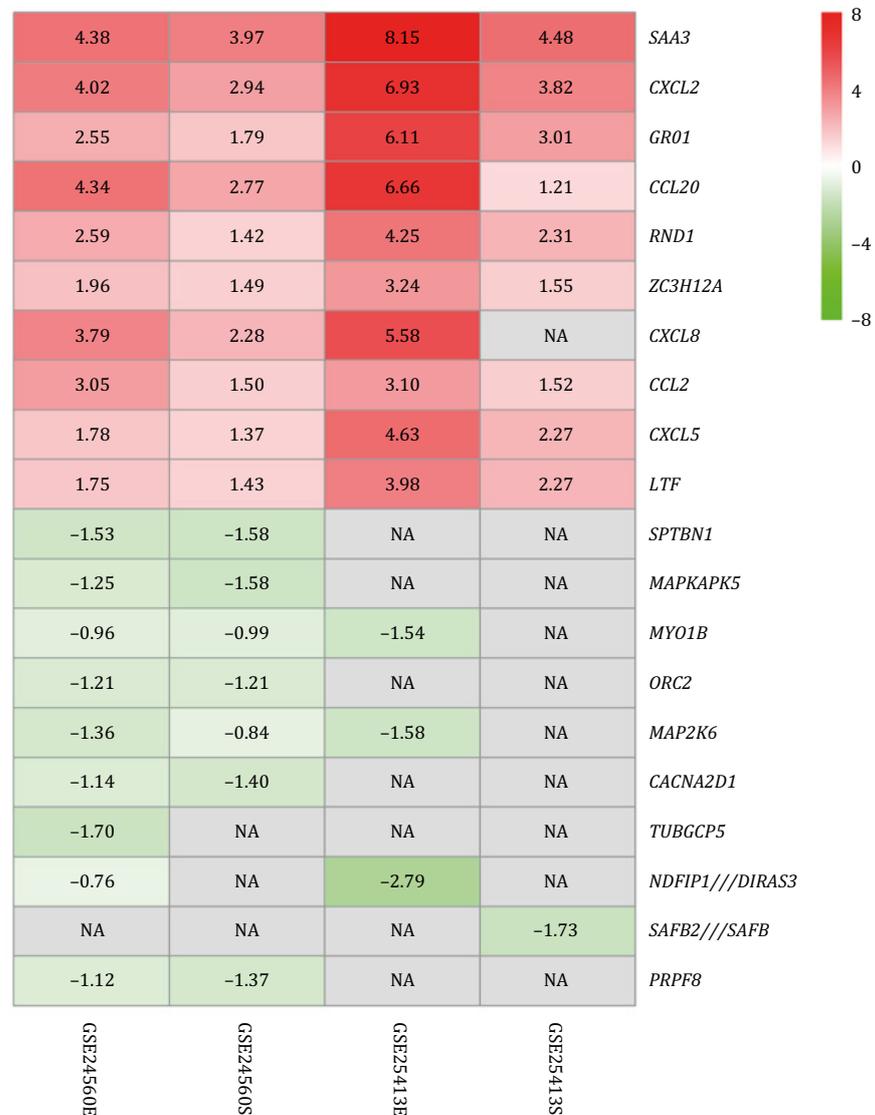
3.2. Identification of integrated DEG

The primitive DEG of the four datasets were saved in a TXT file according to the upregulated and downregulated genes separately, and the genes were sorted by $|\logFC|$ order. These primitive DEG were analyzed by the RRA package in R software. If the gene ranks high in as many datasets as possible, its associated P-value will be lower. In this study $P < 0.05$ was chosen as a threshold. Through

the RRA analysis, 59 upregulated genes, 13 downregulated genes, and a total of 72 integrated DEG were identified (Table 2). The top 10 upregulated genes and the top 10 downregulated genes were mapped on a heat map (Figure 2).

Table 2 - Seventy-two integrated differentially expressed genes (DEG), consisting of 59 upregulated genes and 13 downregulated genes, were identified by the Robust Rank Aggregation method in the *E. coli* or *S. aureus* treatment samples, compared with healthy samples

| Integrated DEG | Gene symbol |
|----------------|---|
| Upregulated | <i>SAA3, CXCL2, GRO1, CCL20, RND1, ZC3H12A, CXCL8, CCL2, CXCL5, LTF, CCL5, NFKBIZ, S100A9, MX2, HP, IL1A, IL6, IL1B, IFI44, CD40, BCL2A1, ISG15, LOC100298356, S100A8, NFKB2, LGALS9, CFB, RTP4, MX1, TNF, CYP1A1, LTB, RSAD2, SOD2, NFKBIA, TMEM140, TNIP1, IKKBE, PLPP3, TLR2, LPO, SLC6A14, BIRC3, C2, MBP, CHI3L1, SLC46A2, PPP1R3C, ARRDC2, NOS2, USP18, LOC100847724, DHX58, DAPP1, NFKBIE, SCML4, PLAC8, CDC42EP2, TAP</i> |
| Downregulated | <i>SPTBN1, MAPKAPK5, MYO1B, ORC2, MAP2K6, CACNA2D1, TUBGCP5, NDFIP1///DIRAS3, SAFB2///SAFB, PRPF8, APP, FRMD3, ZEB1</i> |

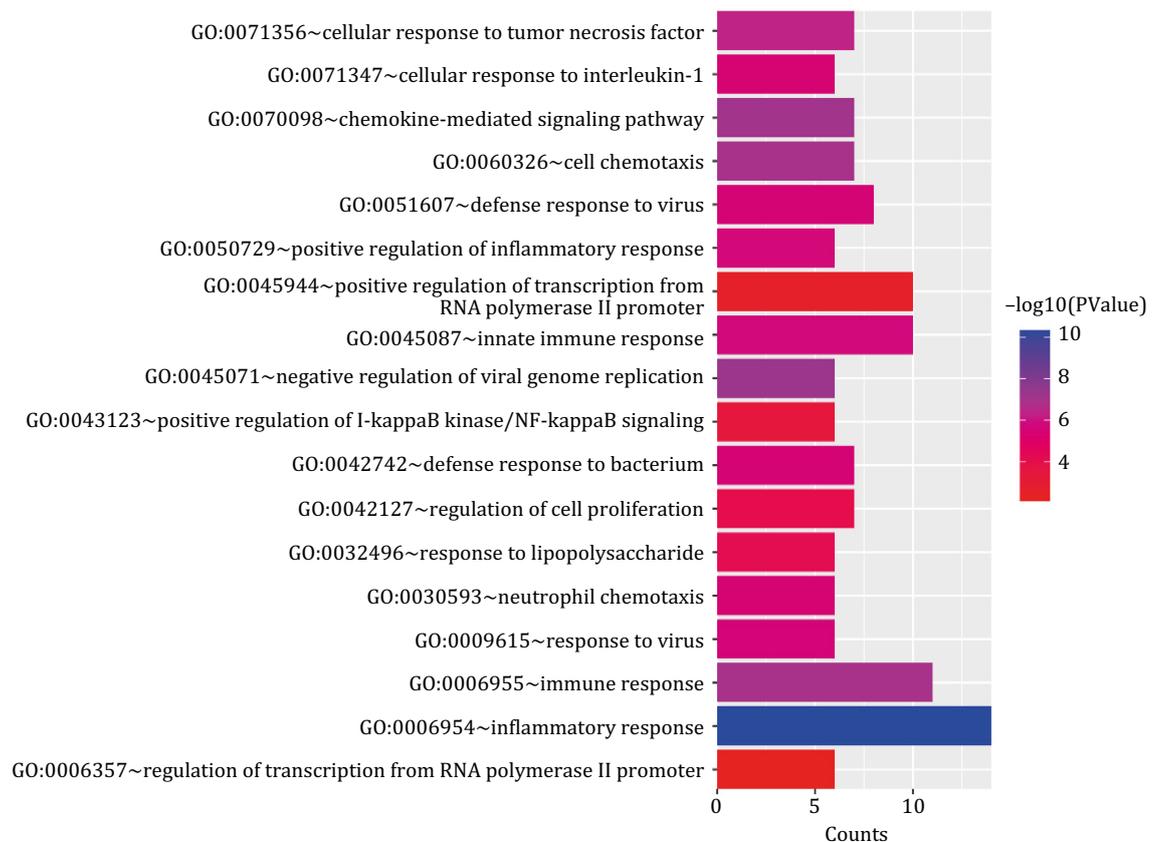


The ordinate represents the gene symbol, the abscissa represents the datasets, red represents upregulated genes, green represents downregulated genes; NA - not identified as DEG in this dataset; the number in the box represents the Log FC value.

Figure 2 - Log FC heatmap of top 10 upregulated and downregulated genes of the integrated differentially expressed genes (DEG) in each expression microarray.

3.3. Gene ontology enrichment analysis

The GO biological process (BP) terms and KEGG pathways were enriched for the 72 integrated DEG using the DAVID database. In the infected group, 18 BP terms and 24 KEGG pathways were enriched with $P < 0.01$ and gene counts > 5 as the cutoff points (Figure 3). From this, the integrated DEG were mainly involved in the immune process such as inflammatory response, chemokine-mediated signaling pathway, immune response, and innate immune response. The top 20 KEGG pathways were analyzed according to the P-value of the integrated DEG (Figure 4). The integrated DEG were mainly enriched in five pathways (counts > 10), namely, TNF signaling pathway, Toll-like receptor signaling pathway, Rheumatoid arthritis, Influenza A, and Cytokine-cytokine receptor interaction.

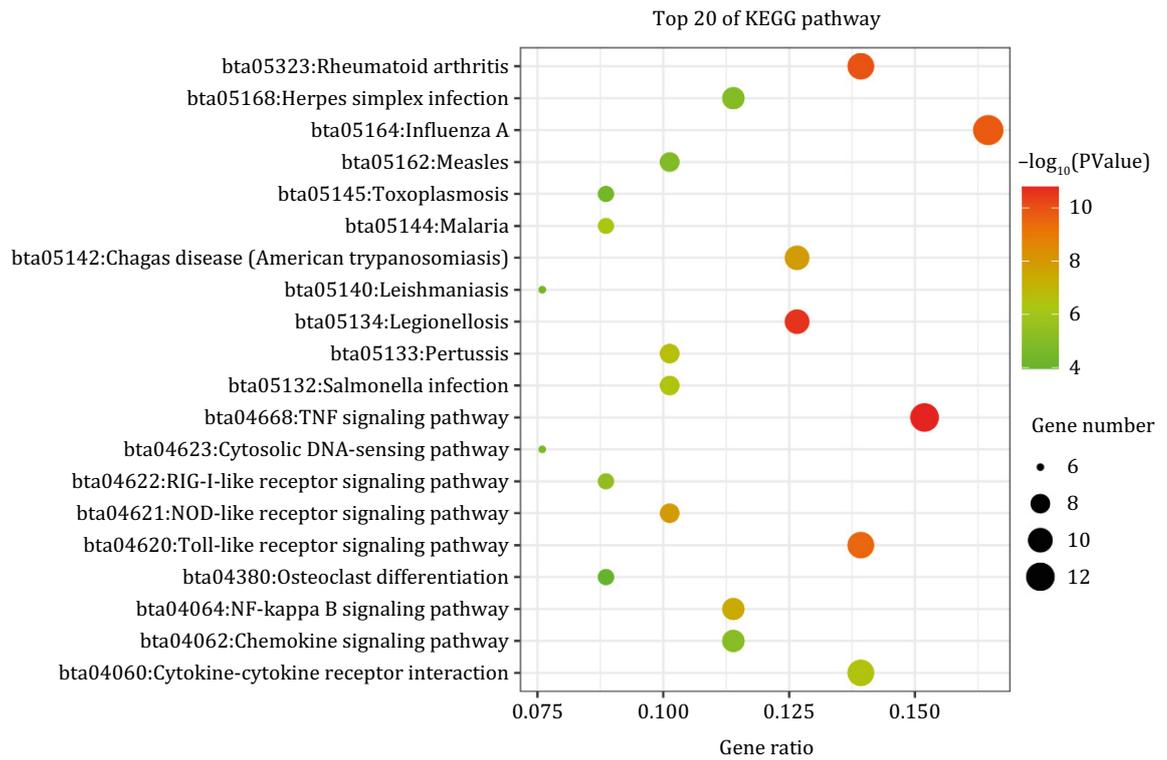


In the infected group, 18 BP terms were enriched terms with $P < 0.01$ and gene counts > 5 as the cutoff points. The integrated DEG mainly involved in the immune process.

Figure 3 - Enriched Gene Ontology (GO) terms of the integrated differentially expressed genes (DEG).

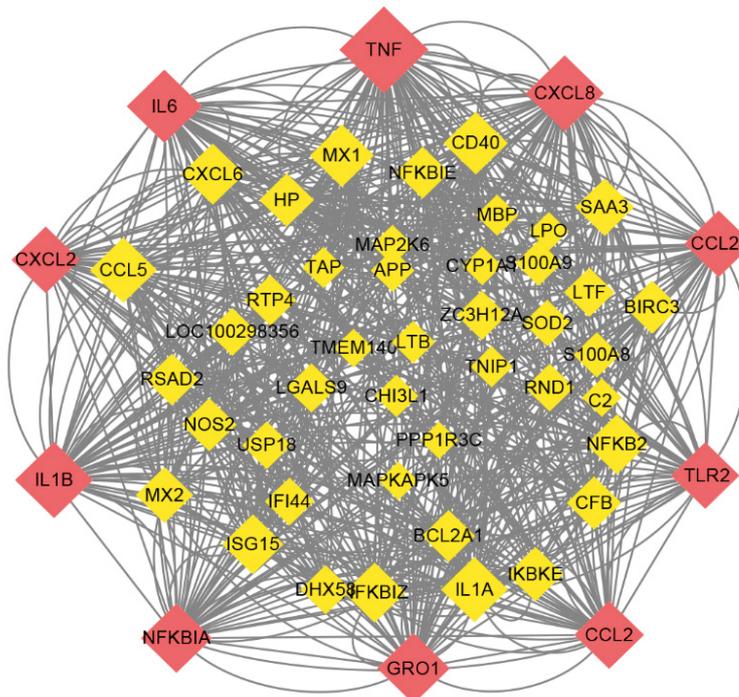
3.4. Protein-protein interaction network construction and hub gene identification

Protein interactions among integrated DEG were predicted with STRING and Cytoscape stools. A total of 53 nodes and 710 edges were involved in the PPI network (Figure 5). The top 10 genes with high degrees identified by the Cytoscape were marked with red (Figure 5). In addition, the most significant module, which has the highest score through MCODE analysis, was screened from the PPI network (Figure 6). There are 17 nodes in this module, which are mainly related to inflammatory response, immune response, TNF signaling pathway, and Legionellosis, etc. (Table 3). In addition, all the 17 genes in this module belong to high degree genes and include the top 10 genes with degree value. These 17 genes were identified as hub genes (Table 4).



KEGG - Kyoto Encyclopedia of Genes and Genomes. The integrated DEG were mainly enriched in five pathways (counts>10), namely, TNF signaling pathway, Toll-like receptor signaling pathway, Rheumatoid arthritis, Influenza A, and Cytokine-cytokine receptor interaction.

Figure 4 - Top 20 KEGG pathways analyzed according to the P-value of the integrated differentially expressed genes (DEG).



The top 10 genes with high degree identified by the Cytoscape are marked with red. A total of 53 nodes and 710 edges were involved in the PPI network.

Figure 5 - Protein-protein interaction (PPI) network of the integrated differentially expressed genes.

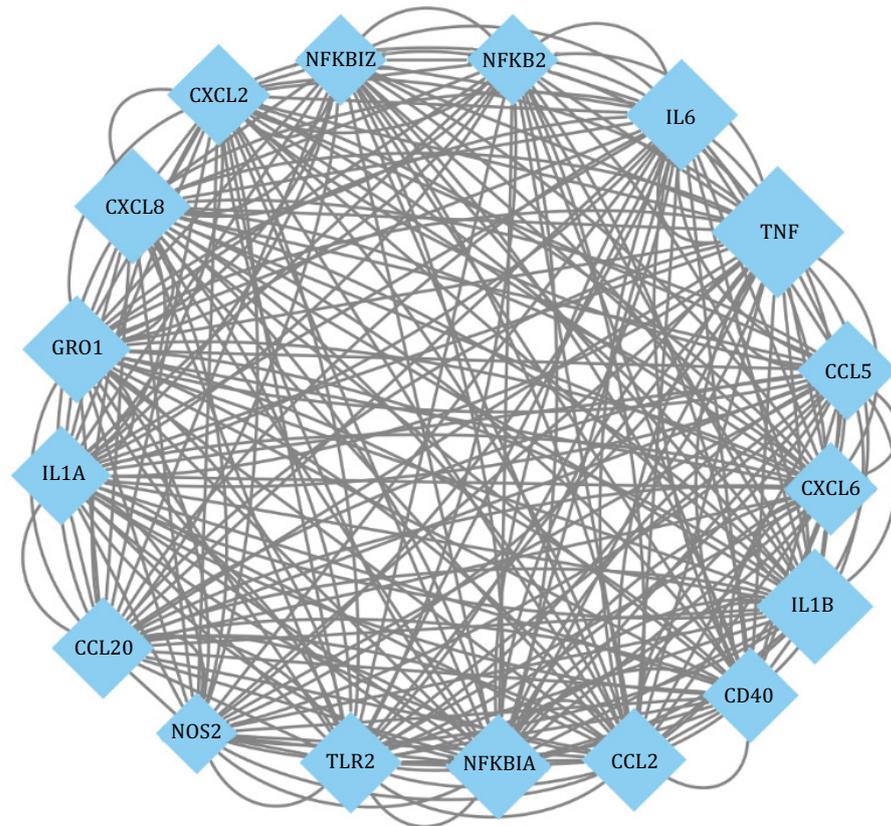


Figure 6 - Protein-protein interaction (PPI) network of the module (= 17 nodes) that has the highest score through MCODE analysis.

Table 3 - Gene ontology (GO; BP) and KEGG pathway analysis of genes in selected module

| Category | Term | P-value | Gene |
|--------------|--|----------|---|
| BP | GO:0006954~inflammatory response | 1.13E-16 | <i>CD40, CXCL8, NOS2, CCL20, CCL5, NFKBIZ, CCL2, GRO1, CXCL2, NFKB2, TLR2</i> |
| | GO:0006955~immune response | 6.48E-11 | <i>IL1A, CD40, CXCL8, IL1B, CCL2, GRO1, CXCL2, TNF</i> |
| | GO:0070098~chemokine-mediated signaling pathway | 1.55E-09 | <i>CXCL8, CCL20, CCL5, CCL2, CXCL2</i> |
| | GO:0060326~cell chemotaxis | 2.70E-07 | <i>CCL20, CCL2, GRO1, CXCL2</i> |
| | GO:0071347~cellular response to interleukin-1 | 2.70E-07 | <i>IL6, CXCL8, CCL20, CCL5, CCL2</i> |
| KEGG pathway | bta05134:Legionellosis | 9.51E-16 | <i>NFKBIA, IL6, CXCL8, IL1B, GRO1, CXCL2, TNF, NFKB2, TLR2</i> |
| | bta04668:TNF signaling pathway | 3.96E-13 | <i>NFKBIA, IL6, CCL20, CCL5, IL1B, CCL2, GRO1, CXCL2, TNF</i> |
| | bta05142:Chagas disease (American trypanosomiasis) | 6.52E-13 | <i>NFKBIA, IL6, CXCL8, NOS2, CCL5, IL1B, CCL2, TNF, TLR2</i> |
| | bta05323:Rheumatoid arthritis | 1.03E-11 | <i>IL1A, IL6, CXCL8, CCL20, CCL5, IL1B, CCL2, TNF, TLR2</i> |
| | bta04620:Toll-like receptor signaling pathway | 2.34E-11 | <i>NFKBIA, CD40, IL6, CXCL8, CCL5, IL1B, TNF, TLR2</i> |

KEGG - Kyoto Encyclopedia of Genes and Genomes.

Table 4 - Hub genes in selected module with degree values and mean LogFC

| Gene symbol | Gene description | Mean LogFC | Degree |
|---------------|--|------------|--------|
| <i>TNF</i> | Tumor necrosis factor | 1.8 | 76 |
| <i>IL1B</i> | Interleukin-1 beta | 2.8 | 60 |
| <i>CXCL8</i> | Interleukin-8 | 3.9 | 60 |
| <i>IL6</i> | Interleukin-6 | 3.0 | 56 |
| <i>GRO1</i> | Growth-regulated protein homolog beta | 3.4 | 52 |
| <i>NFKBIA</i> | <i>Bos taurus</i> nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor | 1.7 | 50 |
| <i>CCL2</i> | C-C motif chemokine 20 | 2.3 | 48 |
| <i>CXCL2</i> | Growth-regulated protein homolog gamma | 4.4 | 46 |
| <i>CCL20</i> | C-C motif chemokine 20 | 3.7 | 46 |
| <i>TLR2</i> | Cooperates with LY96 to mediate the innate immune response to bacterial lipoproteins and other microbial cell wall components. | 2.0 | 46 |
| <i>IL1A</i> | Interleukin-1 alpha | 2.7 | 44 |
| <i>CCL5</i> | C-C motif chemokine 5 | 4.6 | 44 |
| <i>CD40</i> | Receptor for TNFSF5/CD40LG | 1.5 | 40 |
| <i>CXCL6</i> | C-X-C motif chemokine 6 | 2.5 | 38 |
| <i>NFKB2</i> | Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 | 1.2 | 36 |
| <i>NFKBIZ</i> | NF-kappa-B inhibitor zeta | 1.8 | 36 |
| <i>NOS2</i> | Produces nitric oxide (NO) | 3.6 | 26 |

4. Discussion

Despite many efforts dedicated to understanding the pathogenesis of mastitis, the infectious process is still poorly understood (Tartaglia et al., 2018), and the detrimental economic losses caused by cow mastitis remain high (Petzl et al., 2018). The detailed investigation of gene expression response of BME cells to pathogen infections can be of great importance for understanding the innate immune defense mechanisms of dairy cows, and for exploring the candidate genes as biomarkers of mastitis.

Numerous studies on gene expression analysis of cow mastitis have been carried out, which often focus on studying the gene expression response of single bacterial mastitis alone (Sharifi et al., 2018; Li et al., 2019), or the difference between *S. aureus* mastitis and *E. coli* mastitis (Gunther et al., 2017; Han, 2019; Islam et al., 2020). To the best of our knowledge, there is no research on gene expression after co-stimulation of mixed bacteria in BME cells. However, in reality, a large part of dairy cow mastitis may be caused by mixed bacteria. The latest research found that mastitis in dairy cows is a dysbiosis of the mammary gland microbiota (Steinberg et al., 2022). The results of studying the difference between the two types of bacterial mastitis may have a limited effect on the treatment of infections caused by mixed bacteria. To obtain unique genes and pathways associated with the two types of bacterial mastitis, for the first time, we used the RRA method to integrate the results of multiple gene expression datasets about mastitis.

Our results showed that the integrated DEG were mainly enriched to the GO terms of the immune response, inflammation, and TNF signaling pathway. This result is similar to the biological process caused by *E. coli* stimulation (Li et al., 2019). Moreover, these integrated DEG associated with *S. aureus* and *E. coli* mastitis were involved in signaling pathways which were also similar to those caused by *E. coli* mastitis (Li et al., 2019). In addition to this consistency, more importantly, we found that these integrated DEG were enriched in the RIG-I-like receptor pathway and ranked high, but in the study of *E. coli* mastitis, this signaling pathway did not appear in the top 20 (Li et al., 2019). RIG-I

like receptors are a type of pattern recognition receptor that can induce the production of interferon and proinflammatory cytokines, and are critical for the development of anti-pathogen immunity (Kato et al., 2011). The result of this study indicates RIG-I-like receptor pathway may be closely associated with the response of BME cells to these two pathogens.

In addition to the pathway, this study also identified 17 hub genes associated with *S. aureus* and *E. coli* mastitis namely *TNF*, *IL1B*, *CXCL8*, *IL6*, *GRO1*, *NFK-BIA*, *CCL2*, *CXCL2*, *CCL20*, *TLR2*, *IL1A*, *CCL5*, *CD40*, *CXCL6*, *NFKB2*, *NFKBIZ*, and *NOS2*. Some of these genes have been identified as key genes for mastitis in other studies. For example, Sharifi et al. (2018) found *CXCL8*, *CXCL2*, and *GRO1* were the top three genes related to *E. coli* mastitis, and the study of Han (2019) confirmed that *CXCL2* and *GRO1* were key genes in both the *E. coli* and *S. aureus*-mastitis. These similar results further verify the correctness of our integrated analysis.

Among these 17 hub genes, *CD40*, *CXCL6*, and *NFKBIZ*, through our analysis, were identified as key genes and had the potential to be biomarkers of *E. coli* and *S. aureus* mastitis. An important reason for choosing these three genes is that KEGG pathways analysis proved that other genes were involved in signaling pathways triggered by other diseases (Table 3). For example, *NFKBIA*, *IL6*, *CXCL8*, *TNF*, and *TLR2* were involved in Chagas disease, legionellosis, and so on. Cattle can also get these diseases in some cases; for example, *Legionella pneumophila* can cause cattle natural pneumonia (Mondino et al., 2020). This result indicates that those genes may not be suitable as a molecular biomarker of mastitis because biomarker should be specific for disease and should remain unchanged by unrelated disorders (Duarte et al., 2015). Therefore, after removing genes that cross with other diseases, *CD40*, *CXCL6*, and *NFKBIZ* have the potential to be biomarkers of *E. coli* and *S. aureus* mastitis in this study.

CD40 belongs to the tumor necrosis factor receptor family and can promote CD8+ T cell-mediated immunity against tumor cells and some intracellular pathogens (Wallemacq et al., 2012). A recent study showed that *S. aureus* superantigens can stimulate epithelial cells through *CD40* to produce chemokines (Schlievert et al., 2019), which play a crucial role in regulating a variety of inflammatory reactions. Wallemacq et al. (2012) successfully reduced the severity of *S. aureus* mastitis by blocking *CD40* function (Wallemacq et al., 2012). Consistent with the above research, in this study, *CD40* was identified as a key gene. But how *CD40* plays an important role in *E. coli* mastitis is worthy of further study.

CXCL6 is a member of the C-X-C motif chemokine family and has chemotactic granulocytes and immune functions. There was evidence that indicated the *CXCL6* level is increased during *E. coli* mastitis (Li et al., 2019) and *S. aureus* mastitis (Kiku et al., 2016). *CXCL6* is secreted through a signaling cascade of ligand/receptor reactions and serves as important mediator of the immune reaction in the innate immune system response. Combined with the analysis results of this experiment, it is further confirmed that *CXCL6* plays critical roles in the response to pathogen infection in BME cells.

I κ B ζ , encoded by the *NFKBIZ* gene, is a member of the nuclear I κ B protein family and plays a role in inflammatory responses through interaction with NF κ B proteins (Capoferri et al., 2021). Several previous studies revealed a critical role of I κ B ζ signaling in the regulation of immune responses in cow mastitis. For example, I κ B ζ is crucial for induced IL-6 expression but is also known to block TNF-synthesis in BME cells challenged with *S. aureus* (Gunther et al., 2011). Moreover, a study identified *NFKBIZ* as a highly connected gene with *Streptococcus uberis* mastitis and introduced this gene as a potential marker of mastitis resistance in dairy heifers (Bakhtiarizadeh et al., 2020). In this study, *NFKBIZ* was also identified as a hub gene with a continuous increase in expression in the BME cells to cope with *E. coli* and *S. aureus* infection. These results suggest *NFKBIZ* is critical for mastitis pathogenesis and has the potential to be a biomarker of *E. coli* and *S. aureus* mastitis.

Due to the significant expenses of mastitis, early detection is critical. Most existing detection technologies, however, do not match the high accuracy necessary for clinical diagnosis of mastitis (Jensen et al., 2016). As a result, it is critical to build a new diagnosis system that can be converted to quick, on-farm diagnostic methods. The development of novel biomarkers for mastitis diagnosis

has been identified as a mastitis detection trend (Duarte et al., 2015). A biomarker is a characteristic that can be measured and evaluated as an indicator of pathological processes, or pharmacological responses to therapeutic interventions. Genes coding for proteins such as Haptoglobin and Lingual antimicrobial peptide have been identified as potential biomarkers for mastitis detection (Sharifi et al., 2018). The genes identified in this experiment have the potential to build test strips that comprise many biomarkers on a single test strip, thereby improving diagnostic efficiency.

5. Conclusions

Our research identified key candidate genes (*CD40*, *CXCL6*, and *NFKBIZ*) and pathway (RIG-I-like receptor pathway) that have critical roles in both the *E. coli* and *S. aureus* mastitis. This study lays the foundation for further mechanistic studies about this disease and will assist in the search for biomarkers for this disease diagnosis and disease resistance breeding.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization: L. Li. Data curation: Z. Chen, Z. Bian and L. Li. Formal analysis: Z. Chen, Z. Bian and L. Li. Funding acquisition: L. Li. Visualization: X. Chen and B. Li. Writing – original draft: L. Li.

Acknowledgments

This research was funded by the Hainan Provincial Basic and Applied Basic Research Program (Natural Science Foundation) for High-level Talents in 2019 (2019RC090) and the Scientific Research Foundation of Hainan University, grant number KYQD(ZR) 1940.

References

- Bakhtiarzadeh, M. R.; Mirzaei, S.; Norouzi, M.; Sheybani, N. and Vafaei Sadi, M. S. 2020. Identification of gene modules and hub genes involved in mastitis development using a systems biology approach. *Frontiers in Genetics* 11:722. <https://doi.org/10.3389/fgene.2020.00722>
- Brand, B.; Hartmann, A.; Reipsilber, D.; Griesbeck-Zilch, B.; Wellnitz, O.; Kuhn, C.; Ponsuksili, S.; Meyer, H. H. and 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. <https://doi.org/10.1186/1297-9686-43-24>
- Capoferri, R.; Cremonesi, P.; Castiglioni, B.; Pisoni, G.; Roccabianca, P.; Riva, F.; Filipe, J.; del Corvo, M.; Stella, A.; Williams, J. L.; Rupp, R. and Moroni, P. 2021. Comparison of the response of mammary gland tissue from two divergent lines of goat with high and low milk somatic cell scores to an experimental *Staphylococcus aureus* infection. *Veterinary Immunology and Immunopathology* 234:110208. <https://doi.org/10.1016/j.vetimm.2021.110208>
- Duarte, C. M.; Freitas, P. P. and Bexiga, R. 2015. Technological advances in bovine mastitis diagnosis: an overview. *Journal of Veterinary Diagnostic Investigation* 27:665-672. <https://doi.org/10.1177/1040638715603087>
- Gao, X.; Chen, Y.; Chen, M.; Wang, S.; Wen, X. and Zhang, S. 2018. Identification of key candidate genes and biological pathways in bladder cancer. *PeerJ* 6:e6036. <https://doi.org/10.7717/peerj.6036>
- Gunther, J.; Esch, K.; Poschadel, N.; Petzl, W.; Zerbe, H.; Mitterhuemer, S.; Blum, H. and 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. *Infection and Immunity* 79:695-707. <https://doi.org/10.1128/IAI.01071-10>
- Gunther, J.; Petzl, W.; Bauer, I.; Ponsuksili, S.; Zerbe, H.; Schuberth, H. J.; Brunner, R. M. and Seyfert, H. M. 2017. Differentiating *Staphylococcus aureus* from *Escherichia coli* mastitis: *S. aureus* triggers unbalanced immune-dampening and host cell invasion immediately after udder infection. *Scientific Reports* 7:4811. <https://doi.org/10.1038/s41598-017-05107-4>

- Han, H. 2019. Identification of several key genes by microarray data analysis of bovine mammary gland epithelial cells challenged with *Escherichia coli* and *Staphylococcus aureus*. *Gene* 683:123-132. <https://doi.org/10.1016/j.gene.2018.10.004>
- Islam, M. A.; Takagi, M.; Fukuyama, K.; Komatsu, R.; Albarracin, L.; Nochi, T.; Suda, Y.; Ikeda-Ohtsubo, W.; Rutten, V.; van Eden, W.; Villena, J.; Aso, H. and Kitazawa, H. 2020. Transcriptome analysis of the inflammatory responses of bovine mammary epithelial cells: exploring immunomodulatory target genes for bovine mastitis. *Pathogens* 9:200. <https://doi.org/10.3390/pathogens9030200>
- Jensen, D. B.; Hogeveen, H. and De Vries, A. 2016. Bayesian integration of sensor information and a multivariate dynamic linear model for prediction of dairy cow mastitis. *Journal of Dairy Science* 99:7344-7361. <https://doi.org/10.3168/jds.2015-10060>
- Kato, H.; Takahashi, K. and Fujita, T. 2011. Rig-I-like receptors: cytoplasmic sensors for non-self RNA. *Immunological Reviews* 243:91-98. <https://doi.org/10.1111/j.1600-065X.2011.01052.x>
- Kiku, Y.; Nagasawa, Y.; Tanabe, F.; Sugawara, K.; Watanabe, A.; Hata, E.; Ozawa, T.; Nakajima, K. I.; Arai, T. and Hayashi, T. 2016. The cell wall component lipoteichoic acid of *Staphylococcus aureus* induces chemokine gene expression in bovine mammary epithelial cells. *Journal of Veterinary Medical Science* 78:1505-1510. <https://doi.org/10.1292/jvms.15-0706>
- Leuenberger, A.; Sartori, C.; Boss, R.; Resch, G.; Oechslin, F.; Steiner, A.; Moreillon, P. and Graber, H. 2019. Genotypes of *Staphylococcus aureus*: On-farm epidemiology and the consequences for prevention of intramammary infections. *Journal of Dairy Science* 102:3295-3309. <https://doi.org/10.3168/jds.2018-15181>
- Li, L.; Chen, X. and Chen, Z. 2019. Identification of key candidate genes in dairy cow in response to *Escherichia coli* mastitis by bioinformatical analysis. *Frontiers in Genetics* 10:1251. <https://doi.org/10.3389/fgene.2019.01251>
- Lin, C.; Zhu, Y.; Hao, Z.; Xu, H.; Li, T.; Yang, J.; Chen, X.; Chen, Y.; Guo, A. and Hu, C. 2021. Genome-wide analysis of lncRNA in bovine mammary epithelial cell injuries induced by *Escherichia coli* and *Staphylococcus aureus*. *International Journal of Molecular Sciences*. 22:9719. <https://doi.org/10.3390/ijms22189719>
- Mondino, S.; Schmidt, S.; Rolando, M.; Escoll, P.; Gomez-Valero, L. and Buchrieser, C. 2020. Legionnaires' disease: State of the art knowledge of pathogenesis mechanisms of *Legionella*. *Annual Review of Pathology: Mechanisms of Disease* 15:439-466. <https://doi.org/10.1146/annurev-pathmechdis-012419-032742>
- Petzl, W.; Zerbe, H.; Günther, J.; Seyfert, H.-M.; Hussien, J. and Schuberth, H.-J. 2018. Pathogen-specific responses in the bovine udder. Models and immunoprophylactic concepts. *Research in Veterinary Science* 116:55-61. <https://doi.org/10.1016/j.rvsc.2017.12.012>
- Schlievert, P. M.; Cahill, M. P.; Hostager, B. S.; Brosnahan, A. J.; Klingelutz, A. J.; Gourronc, F. A.; Bishop, G. A. and Leung, D. Y. M. 2019. Staphylococcal superantigens stimulate epithelial cell through CD40 to produce chemokines. *mBio* 10:e00214-19. <https://doi.org/10.1128/mBio.00214-19>
- Sharifi, S.; Pakdel, A.; Ebrahimi, M.; Reecy, J. M.; Fazeli Farsani, S. and Ebrahimie, E. 2018. Integration of machine learning and meta-analysis identifies the transcriptomic bio-signature of mastitis disease in cattle. *PLoS One* 13:e0191227. <https://doi.org/10.1371/journal.pone.0191227>
- Steinberg, R. S.; Silva e Silva, L. C.; Souza, M. R.; Reis, R. B.; Silva, P. C. L.; Lacorte, G. A.; Nicoli, J. R.; Neumann, E. and Nunes, A. C. 2022. Changes in bovine milk bacterial microbiome from healthy and subclinical mastitis affected animals of the Girolando, Gyr, Guzera, and Holstein breeds. *International Microbiology* 25:803-815. <https://doi.org/10.1007/s10123-022-00267-4>
- Subramanian, D. and Natarajan, J. 2021. Integrated meta-analysis and machine learning approach identifies acyl-CoA thioesterase with other novel genes responsible for biofilm development in *Staphylococcus aureus*. *Infection, Genetics and Evolution* 88:104702. <https://doi.org/10.1016/j.meegid.2020.104702>
- Tartaglia, N. R.; Breyne, K.; Meyer, E.; Cauty, C.; Jardin, J.; Chrétien, D.; Dupont, A.; Demeyere, K.; Berkova, N.; Azevedo, V.; Guédon, E. and Le Loir, Y. 2018. *Staphylococcus aureus* extracellular vesicles elicit an immunostimulatory response *in vivo* on the murine mammary gland. *Frontiers in Cellular and Infection Microbiology* 8:277. <https://doi.org/10.3389/fcimb.2018.00277>
- Võsa, U.; Kolde, R.; Vilo, J.; Metspalu, A. and Annilo, T. 2014. Comprehensive meta-analysis of microRNA expression using a robust rank aggregation approach. p.361-373. In: *RNA mapping. Methods in Molecular Biology*, vol. 1182. Alvarez, M. and Nourbakhsh, M., eds. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-1062-5_28
- Wallemacq, H.; Bedoret, D.; Pujol, J.; Desmet, C.; Drion, P. V.; Farnir, F.; Mainil, J.; Lekeux, P.; Bureau, F. and Fiévez, L. 2012. CD40 triggering induces strong cytotoxic T lymphocyte responses to heat-killed *Staphylococcus aureus* immunization in mice: A new vaccine strategy for staphylococcal mastitis. *Vaccine* 30:2116-2124. <https://doi.org/10.1016/j.vaccine.2012.01.039>
- Younis, S.; Javed, Q. and Blumenberg, M. 2016. Meta-analysis of transcriptional responses to mastitis-causing *Escherichia coli*. *PLoS One* 11:e0148562. <https://doi.org/10.1371/journal.pone.0148562>
- Zaatout, N. 2022. An overview on mastitis-associated *Escherichia coli*: Pathogenicity, host immunity and the use of alternative therapies. *Microbiological Research* 256:126960. <https://doi.org/10.1016/j.micres.2021.126960>