

# Regulatory Network of miRNA, lncRNA, Transcription Factor and Target Immune Response Genes in Bovine Mastitis

Olanrewaju Morenikeji (✉ [obm3@pitt.edu](mailto:obm3@pitt.edu))

University of Pittsburgh at Bradford

Ashley Tucker

Rochester Institute of Technology

Nicole Salazar

Rochester Institute of Technology

Adeola Ayoola

Chinese Academy of Sciences

Erdogan Memili

Mississippi State University

Bolaji Thomas

Rochester Institute of Technology

---

## Research Article

**Keywords:** lncRNA, miRNA, transcription factor, prediction, markers, regulation, bovine mastitis

**Posted Date:** June 14th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-590138/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Pre- and post-transcriptional modifications of gene expression are emerging as foci of disease studies, with some studies revealing the importance of non-coding transcripts, like long non-coding RNAs (lncRNAs) and microRNAs (miRNAs). We hypothesize that TFs, lncRNAs and miRNAs can modulate the immune response in bovine mastitis and potentially serve as disease biomarkers and/or drug targets. With computational analyses, we identified candidate genes potentially regulated by miRNAs and lncRNAs base pair complementation and thermodynamic stability of binding regions. Remarkably, we found six miRNAs, two being bta-miR-223 and bta-miR-24-3p, to bind to several targets. NONBTAT027932.1 and XR\_003029725.1, were identified to target several genes. Functional and pathway analyses revealed lipopolysaccharide-mediated signaling pathway, regulation of chemokine (C-X-C motif) ligand 2 production and regulation of IL-23 production among others. The overarching interactome deserves further *in vitro/in vivo* explication for specific molecular regulatory mechanisms during bovine mastitis immune response and could lay the foundation for development of disease markers and therapeutic intervention.

## Introduction

Bovine mastitis, caused by a variety of organisms, including Staphylococci, Streptococci, and Enterobacteria, is a mammary gland infection posing a great threat to the dairy industry, and is the costliest disease affecting cattle, due to its impact on the quality and quantity of milk production<sup>1,2</sup>. The immune response to pathogens in the mammary gland is highly complex, having to first identify the pathogen, recruit the appropriate immune cells, remove the pathogen, and then terminate the inflammatory response<sup>3</sup>. Recent studies have shown that mammary epithelial cells (MEC) initiate this process by changing the expression of various genes, responsible for a wide range of immunological and inflammatory responses<sup>4</sup>. The regulatory elements involved in the over- or under-expression of these genes during bovine mastitis are yet to be fully understood and may be a potential area for initiating prevention and control of a global, economically devastating disease in the dairy industry.

Of recent, attention has shifted to regulatory elements, such as microRNAs (miRNAs), long non-coding RNAs (lncRNAs), and transcription factors (TFs), possibly influencing the expression of host immune genes<sup>5,6</sup>. MiRNAs are short, non-coding RNA that bind to the 3' untranslated region (UTR), 5'UTR and coding regions of an mRNA to post-transcriptionally modify genes, repressing their translation into their final protein product<sup>5,7</sup>. Likewise, lncRNAs are long non-coding RNA transcripts, harboring a 5' cap and 3' poly (A) tail but no functional open reading frame, and as such cannot encode proteins<sup>8,9</sup>. This heterogeneous class of RNAs can modulate gene expression and protein synthesis at the transcriptional and post-transcriptional levels via complementary base pairing<sup>10</sup>. lncRNAs play key roles in regulating several cellular and developmental processes, including genomic imprinting, DNA methylation, splicing and chromatin modification<sup>11,12</sup>. With an abundance of binding sites for miRNA and mRNA, lncRNA can act as ceRNA (competing endogenous RNA) and are significant regulatory elements in post-

transcriptional gene expression<sup>13,14</sup>. For example, lncRNA adipocyte differentiation-associated long noncoding RNA (ADNCR) targets miR-204 significantly regulating Sirtuin 1 (SIRT1) gene expression at the mRNA and protein level, inhibiting adipogenesis in the process<sup>15</sup>. Additionally, Li and colleagues showed that lncRNA H19 targets miRNAs let-7b and miR-200b, competes endogenously with transcription factor Sp1 (SP1), thereby regulating the expression of TGF- $\beta$ -receptor 2 (TGFB2) gene in cancers<sup>9</sup>. On the other hand, TFs are proteins capable of altering or activating gene-expression level<sup>16,17</sup>.

Here, we propose the need to understand the crosstalk between miRNA, lncRNA and TFs, and the role they play in regulating genes associated with bovine mastitis, including the possibility of translation to drug treatment or disease prevention therapies. To do this, we identified candidate miRNAs, lncRNAs and TFs, and their potential role in bovine mastitis via *in silico* analyses. Current algorithms, relying on base pair complementation, evolutionary conservation, and thermodynamic stability of binding regions, have been shown to be useful in predicting miRNA, lncRNA and TF binding sites on target genes<sup>18,19</sup>. Multiple miRNA databases such as miRWalk<sup>20</sup>, miRNet<sup>21</sup>, and TargetScan<sup>22</sup> compute potential miRNA-mRNA interactions, while the role of individual miRNA can be inferred through functional analysis with Gene Ontology (GO)<sup>23</sup>. Similarly, lncRNA prediction binding software including NONCODE<sup>24</sup>, lncRNA2Target<sup>25</sup>, and lncTar<sup>19</sup> have been useful in guiding bench experiments thereby saving time and resources. Knowing the interconnectivity of key regulatory elements involved in bovine mastitis is potentially useful in disease detection, especially in subclinical cases. If a regulatory element plays a key role in disease pathogenesis or in the immune response, this may be an ideal drug target for treatment.

## Results

### Dataset of candidate genes associated with immune response to bovine mastitis

Our initial search on PubMed using the keywords “bovine mastitis” and “gene expression” generated 276 articles. Following filtering, a total of 109 articles were selected and added to the pipeline for candidate genes identification. These articles, including experimental, meta-analysis, and reviews mentioned 919 individual genes associated with host immune response to bovine mastitis. Of importance to our analysis are genes with multiple mentions, including CXCL8 and IL-6 referenced in 41 and 29 articles, respectively (Fig 1b). Comparing the 919 genes found during our search with the 20 genes generated from Genomatix, yielded 16 genes that were common to both search methods (Fig S1). These genes are MYD88, CD4, IL-10, IFN $\gamma$ , IL-4, ICAM1, CXCL8, TLR4, TNF $\alpha$ , IL-18, TLR2, CD86, CCL2, IL-6, CSF2, and CD14. The interactions between these candidate genes reveal TLR4 and TLR2 are considerably more linked to other genes (Fig 2), while CD4 on the other hand had the least interaction.

### miRNAs bind to the candidate genes in bovine mastitis

Our analyses explored the complementary binding of miRNAs to the 3'UTR, CDS, and 5'UTR regions of the candidate genes. We report a total of 768 miRNAs binding either the 3' UTR, CDS, or 5' UTR of the 16 target genes (Tables S2). With Bioinformatics and Evolutionary Genomics (BEG), five miRNAs were

uncovered to bind all three regions within the same gene; bta-miR-2338 and bta-miR-2433 to CD14, bta-miR-2373-3p to ICAM1, bta-miR-2328-3p and bta-miR-2356 to TLR4 (Fig 3). In addition, a substantial number of miRNAs bound to two or more regions within the same gene (Fig 3). Displaying all the miRNA predicted to bind to three or more target genes is Fig 4, which also depicts whether the miRNA binds one (green), two (blue), or three (red) of the possible regions (3'UTR, 5'UTR, and CDS). The complete list of miRNAs binding three or more target genes is depicted in Table S3. Likewise, we found six miRNAs; bta-miR-24-3p, bta-miR-149-5p, bta-miR-223, bta-miR-185, bta-miR-874, and bta-miR-328, commonly predicted by the three software as shown (Fig S2; Table S2).

### **lncRNAs interact with mRNA target inferring their functionalities**

NONCODE database search revealed 22 bovine lncRNAs that matched our prediction criteria. Genomic location and strand sense for each lncRNAs are shown in Table S4. From the analysis, we identified 20 lncRNAs that bind to our target mRNAs with a ndG < -0.05 (Fig 5; Table S4), with 8 of them targeting more than 13 of disease associated genes. Of the 8 lncRNA, six (NONBTAT001181.2, XR\_003029725.1, NONBTAT010129.2, XR\_003030515.1, XR\_003033296.1, NONBTAT027932.1) targeted all of the 16 genes, while NONBTAT027932.1, XR\_003033296.1, and XR\_003029725.1 could bind at least one target gene with an average ndG greater than -0.08. NONBTAT001181.2-IL-18 and XR\_003029725.1 showed a high affinity for IL-18 with thresholds less than -0.1. Additionally, NONBTAT027932.1 could bind MYD88, CD4, ICAM1, CD14, TLR2, and TNF $\alpha$  with binding free energy less than -0.1 (Table S4). Eight lncRNAs that bind 13 or more candidate genes with lower ndGs were selected for network and conservational analysis (Table 3). Most of the identified lncRNAs are not yet annotated, therefore, we inferred their gene ontology and pathways based on their potential mRNA targets/genes. Identified ontologies and associated pathways are shown in Table 4 and Figures 6a and b.

### **Biological, molecular and cellular function of target genes and pathway analysis**

The pathway analysis of 16 target genes identified 25 biological processes (Fig 6a). Myeloid leukocyte activation with a p-value of 4.92E-13, vascular endothelial growth factor (VEGF) production, regulation of chemokine (C-X-C motif) ligand 2 production and lipopolysaccharide (LPS)-mediated signaling pathway (p-value; 2.58E-13) (Table 4) are identified. Of importance are the three significant biological processes concomitantly identified from the three databases including LPS-mediated signaling pathway, regulation of IL-23 production, and positive regulation of chemokine production. DAVID analysis identified eight biological processes, three molecular functions and one cellular component, while 11 disease pathways were identified including malaria and rheumatoid arthritis that featured most candidate genes during bovine mastitis with p-values of 9.20E-19 and 4.60E-16, respectively (Fig 6b; Table S5).

### **Catalog of transcription factors, biological processes and molecular function**

Seventeen TFs were identified as significant by at least two software (Table 5; Table S6). All 17 TFs appear to have the potential to regulate the transcription of the 16 genes including CD4, MYD88, ICAM1, TLR4, TLR2, IL-18, and CD86. The genomic location and other details of the 17 TFs are presented in Table

5. In addition, functional analysis identified 3 biological processes, 2 molecular functions, and 1 cellular component involving two to three TFs, with a p-value less than 0.0005 (Table 6).

### **Sequence conservation of bovine miRNA and lncRNAs in other species**

Sequence analysis was performed to assess the evolutionary conservation of bovine lncRNAs and miRNAs across various species (Table S1). We examined the significant miRNA (n=6) and lncRNA (n=8) identified previously with their homologous sequence from 14 other species. Our analysis of the six miRNAs reveal bta-miR-223 to be highly conserved in all 15 species, without a single deviant polymorphism (Fig S3). The evolutionary closeness of cow and wild yak is also seen in bta-miR-149-5p, bta-miR-328, and bta-miR-185, but are distantly related in bta-miR-24-3p and bta-miR-874 (Fig S4a-c). The phylogenetic analysis of bta-miR-223 defines cow having the closest identity with wild yak and the remaining ruminants (goat and sheep) as orthologs (Fig S4b). The alignment of bta-miR-874 showed high conservation between cow, mouse, and megabat, with phylogenetic analysis having these three sharing a common node. No other tree had these three species this close in proximity.

The length of lncRNA allow for greater variation within the gene throughout evolution. Analysis of the evolutionary alignment reveal that each lncRNA was not completely conserved in the 15 species; however, the majority have similar base pair sequences in conserved regions. Out of the 8 lncRNA, 7 have bovine lncRNA to be most closely related to wild yak, followed by other ruminants (Fig S5a-d). The tree of NONBTAT027932.1 show wild yak to be relatively distant from cow, sheep, and goat compared to the species relation in the other lncRNA analysis. Five out of eight lncRNA have human, gorilla, and chimpanzee clustered with the same origin.

### **Dysregulation of lncRNA-miRNA interaction on immune response in bovine mastitis**

Our analysis showed several miRNA potential binding sites on lncRNA, based on the complementary base pairing and normalized binding free energies (ndG) values. Of interest, we found 8 out of 22 lncRNAs possessing binding sites for the entire length of these miRNAs, with the exception of 4 out of the 48 pairings (Fig 7). The ndG values range from -0.1774 to -2.875. lncRNA XR\_003033296.1 and bta-miR-223 had the greatest ndG while lncRNA XR\_234647.4 and bta-miR-328 had the lowest ndG. The lower the ndG, the greater the possibility that the two RNA will bind. The vast majority of miRNA bind their entire length to the lncRNA except four, the shortest binding length is 10 (out of 22 base pairs) between bta-miR-223 and XR\_234647.4. Additional results on lncRNA-miRNA binding are presented in Table S7.

### **Combined network and molecular interactome of lncRNAs, miRNAs, TFs and immune genes**

Figures 8 through 11 show molecular interactions between miRNAs, lncRNAs, TFs, target genes and pathways. Fig 8 in particular show the connections between the bovine mastitis target genes, miRNA, and the pathways/biological processes of the genes. CD4 (36) and TLR4 (36) have the most miRNAs targeting them, as opposed to IL-4, CSF2, and IL-18 (one, two and two miRNAs respectively). Some miRNAs are seen targeting more than one gene such as bta-miR-2294 (targeting CD4 and CD86) and bta-

miR-2374 (targeting TLR4 and CD4). Interestingly, we found some miRNAs that bind to multiple genes involved in the same pathway, such as bta-miR-671 targeting TLR4 and IFN $\gamma$ , both of which are involved in macrophage activation; bta-miR-2413 targeting IFN $\gamma$ , IL-10, and TLR4, all of which are involved in the regulation of nitric oxide biosynthetic process.

The functional annotation of each lncRNA was inferred from the target genes using GO Enrichment Analysis and the DAVID KEGG algorithm program. Several of these lncRNAs are involved in multiple ontologies (Fig 9). Specifically, NONBTAT001181.2 could bind to IL-6 to regulate its expression, thereby perturbing functional pathways like vascular endothelial growth factor, regulation of cytokine production, which are involved in inflammatory response, and regulation of tyrosine phosphorylation of STAT protein pathways. LncRNA XR\_003030515.1 targets TLR4 and is connected to cellular response to lipoteichoic acid, macrophage activation, NAD $^{+}$  nucleosidase activity, and regulation of interleukin-23 production. NONBTAT027932.1 lncRNA targets CD14 and is involved in positive regulation of NIK/NF- $\kappa$ B signaling, regulation of interleukin-8 production, regulation of cytokine secretion, LPS-mediated signaling pathway, lipopeptide binding, and LPS receptor complex (Fig 9). Figure 10 shows the interconnections between the 17 TFs, 16 target genes, and the pathways they are involved in. Both CREB1 and JUN have numerous connections directed toward the target genes and pathways.

The final immunoregulatory integrated network in our analysis contains miRNAs, lncRNAs, TFs, target genes, and biological pathways and their extensive interconnections (Fig 11). We show an important connectivity between lncRNA, miRNA, bovine mastitis immune genes and that their biological functions have all been mapped on the network. The outer circle contains mostly miRNAs, with the exception of 3 TFs, that bind only to a single target gene. The inner circle consists of miRNAs binding to more than one target gene, lncRNAs, TFs, target genes, and biological pathways. LncRNAs are found densely in the top inner circle and can be sporadically located in the outer circle. Our results indicate that the majority of lncRNAs, miRNAs, and TFs interactions overlapped and are involved in several pathways, including biological processes, cellular components, and molecular functions.

## Discussion

Bovine mastitis is detrimental to the dairy industry with significant impact on farms in developing countries<sup>26,27</sup>. Mastitis can become chronic and highly contagious, and the effective treatment options available are suboptimal, one reason being the cost associated with antibiotic residue in the milk<sup>28</sup>. Regulatory factors have been associated with other diseases, including atherosclerosis, identifying a lncRNA that may be an intuitive preventive and therapeutic treatment option<sup>29</sup> but not in bovine mastitis. Novel key lncRNA and TFs were identified in lung adenocarcinoma regulatory networks, offering new drug targets for therapeutic treatment<sup>30</sup>. In another study, miR-99b was found to repress the expression of IGF1R and mTOR, ultimately acting as an onco-suppressor in clear cell renal cell carcinoma<sup>31</sup>. To our knowledge, this is the first report examining all three regulatory elements (miRNA, lncRNA and TFs) in the

context of bovine mastitis, with the aim to identify elements influencing expression of candidate genes regulating immune response.

Through our meta-analysis and computational mining, we generated a list of genes playing critical roles in immune response during bovine mastitis. TLR2 and TLR4 were substantially connected to other candidate genes through our network, which is expected given their involvement in pro-inflammatory and innate immune response<sup>32</sup>. Specifically, TLR4 recognizes lipopolysaccharides found in the outer membrane of gram-negative bacteria and TLR2 recognizes lipoteichoic acid found in the cell wall of gram-positive bacteria<sup>33</sup>. Pathogen recognition is imperative to the initiation of the immune response, and without TLR4 and TLR2, identification of bacteria in bovine mammary epithelial cells would be impaired. IL-8, IL-6 and TNF $\alpha$ , the three most common genes in our meta-analysis, are all inflammatory cytokines or chemokines, and have been shown to be upregulated in bovine mammary epithelial cells stimulated with LPS<sup>34</sup>.

Of importance, our results identified six miRNAs binding to several target genes with the potential to regulate their expression during infection. Interestingly, previous studies have shown bta-miR-24-3p, bta-miR-328, bta-miR-223, and bta-miR-185 to be associated with bovine immune response. Bta-miR-24-3p has been associated with serum antibody response to *Mycoplasma bovis* in beef cattle<sup>35</sup>. Bta-miR-185 had increased expression in milk from a *S. aureus* infected cow compared to a healthy cow. Ma et al.<sup>36</sup> identified bta-miR-185 targeting five genes including MLLT3, DYRK1B and NPR2. Other studies have demonstrated the involvement of bta-miR-223 in bovine mastitis immune response<sup>32,37</sup>, but also in other immune responses to lung infection, bacterial peritonitis, and cutaneous leishmaniasis<sup>38-41</sup>. The other two miRNAs identified in our study, bta-miR-149-5p and bta-miR-874, have been associated with fat levels in beef cattle<sup>42,43</sup>. Our findings, along with previous publications, collectively demonstrate the involvement of these miRNAs in immune response and should be strongly considered for further laboratory studies. In addition, our study demonstrated that miRNAs that could bind to all three regions on a single gene had a greater probability of binding the target gene, and ultimately perturb the disease pathways the target genes are involved in.

Furthermore, we identified important interactions between certain lncRNAs and mRNAs that could mediate host immune response at the post-transcriptional level. lncRNAs are becoming known to be involved in several biogenetic processes that may regulate disease pathogenesis<sup>14,44</sup>. We hypothesized that lncRNAs that bind to multiple mRNA target sequences could serve as drug targets and have the potential to modify or regulate immune signaling pathways, ameliorating the over- or under-production of cytokines or chemokine during infection. Importantly, lncRNAs could directly bind to the mRNA thereby getting it degraded and prevent translation to the expected protein. Of interest, we also found several lncRNAs with binding sites for miRNAs, for example XR\_003033296.1-bta-miR-223 and lncRNA XR\_234647.4-bta-miR-328 interactions thereby providing a way to prevent the negative regulatory effect of miRNA on the mRNA targets during disease perturbation. Gene ontology analysis reveal several of these miRNAs and lncRNAs to be associated with myeloid leukocyte activation, LPS-mediated signaling

pathway, and positive regulation of IL-8 production. These miRNAs and lncRNAs interactions may be an avenue to manipulate the regulatory elements involved in immune response to bovine mastitis. The knowledge is potentially useful in disease detection/diagnosis, especially in subclinical cases and may be an ideal target for treatment<sup>15,45</sup>.

Additionally, our pathway analysis revealed seven genes, IL-4, IFN $\gamma$ , TLR4, CXCL8, IL-18, CSF2 and TNF $\alpha$ , are involved in the morphological and behavioral changes of macrophages, monocytes, granulocytes, and dendritic cells (DCs) and their recruitment into infection sites. Therefore, the miRNA binding to these seven gene transcripts, degrading their mRNA in the process, would have a deleterious effect on the immune response needed to eliminate the pathogen. Our study indicated bta-miR-328 and bta-miR-223 bind to three out of the seven genes involved in this pathway. A recent study found the overexpression of bta-miR-223 in Mac T cells attenuating the inflammatory response induced by lipoteichoic acid derived from *S. aureus*, by targeting the Cbl proto-oncogeneB (CBLB), positively correlated with the expression of P13K, AKT, and NF- $\kappa$ B p65<sup>37</sup>. Fang and others concluded that bta-miR-223 is a central post-transcriptional regulator in mammary tissue when infected with *S. aureus* by altering the expression of key innate immune response genes, one being CXCL14<sup>32</sup>. Other studies have also identified bta-miR-223 as a key miRNA expressed in mammary tissue during bovine mastitis<sup>46</sup>. Given our findings and previous reports, we surmise that these six miRNAs, especially bta-miR-223 would play significant dysregulatory role on immune response during bovine mastitis and could serve as biomarker or drug target, deserving further *in vivo* and *in vitro* investigation to clarify.

Our study showed seven significant transcription factors including SP1, SF1 and JUN targeting several genes within the biological pathways. SP1, JUN and CREB are commonly found in almost all the significant biological pathways (molecular function, biological and cellular processes), resulting from their function as activators and repressors. SP1 has been identified as one of the first eukaryotic trans activators to regulate numerous mammalian genes, often in association with other transcription factors, playing major roles, including cell cycle regulation, hormonal activation and apoptosis<sup>47,48</sup>. JUN directly transcribe the expression of epidermal growth factor receptor thereby positively regulating the proliferation and differentiation of keratinocytes<sup>49</sup>.

Nuclear transcription factor- $\kappa$  (NF- $\kappa$ B) signaling is a major player in the regulation of innate and adaptive immune responses and is widely studied for the role it plays in mastitis progression and pathogenesis. Many inflammatory genes contain binding sites for NF- $\kappa$ B in their promoter regions, thus active NF- $\kappa$ B regulates expression by directly binding to target genes in the nucleus<sup>4</sup>. Cytokines (IL-6, IL-1 $\beta$ , TNF $\alpha$ ) and chemokine (CXCL8) are associated with increased expression in LPS-stimulated mammary glands<sup>8,50,51</sup>. These pro-inflammatory cells disrupt the blood-milk barrier causing slowed milk production and painful clinical manifestations in cattle with mastitis. Given the immunological and inflammatory importance, lncRNAs targeting mediators (TLR2, TLR4, MYD88, CD14, ICAM-1, IL-6, IL-1 $\beta$ , TNF $\alpha$ , and CXCL8) of the NF- $\kappa$ B signaling pathway may act as biomarkers for disease diagnosis and drug therapeutics by blocking lncRNAs from interacting with the target genes.



A variety of growth factors and inflammatory signals induces CREB and subsequently mediate the transcription of genes that contain a cAMP-responsive element. Interestingly, we report that miRNAs could bind and regulate IL-6, IL-10, and TNF $\alpha$ , which possess the cAMP-responsive element<sup>52</sup>. IRF1, through the regulation of autoimmunity, inflammation, viral infections, and innate and adaptive immune responses, is also responsible for the protection of host cells<sup>53</sup>.

Conservation analysis reveal all six miRNAs to be almost entirely conserved between the 15 species. The bta-miR-223 was highly conserved in all 15 species, without a single base pair difference. The conservation of these miRNAs throughout evolution indicates their importance and benefits to the host, further validating our results. These findings also suggest a drug targeting one of these miRNAs can be used across multiple species. This also suggests the possibility of experimental drug trials on smaller animals that have a higher rate of reproduction, such as a mouse or rat, following *in vivo/in vitro* elucidation.

Network interactome of miRNA, target genes, pathways, and cellular functions give a visualization of the connections between them, allowing for comprehension of the broad picture- regulatory element to target gene to biological pathway. For example, bta-miR-671 is predicted to bind to IFN $\gamma$  and TLR4 in at least two of the three mRNA regions, and both IFN $\gamma$  and TLR4 are involved in the positive regulation of tumor necrosis factor superfamily cytokine production. Therefore, bta-miR-671 may be a strong target gene as it affects a biological pathway through two different genes. One of our identified miRNAs, bta-miR-24-3p binds to CD14 in at least two regions. CD14 is involved in cellular response to lipoteichoic acid, Toll-like receptor signaling pathway, positive regulation of tumor necrosis factor superfamily cytokine production, positive regulation of cytokine secretion, positive regulation of NIK/NF- $\kappa$ B signaling, regulation of interleukin-8 production, and LPS-mediated signaling pathway, all of which play major roles in immune response. Bta-miR-149-3p targets both CXCL8 and TLR4, both of which are involved in myeloid leukocyte activity. This interactome ties our findings together and allows for miRNA to be seen as possible therapeutic targets.

## Conclusions

The present study provides computational evidence that miRNA, lncRNA, and TFs could regulate the host immune response to bovine mastitis. These regulatory element interactomes provide an insight into immune gene dysregulation during bovine mastitis. The knowledge, which is useful in disease diagnosis, especially in subclinical cases, may be ideal targets as an antibiotic for treatment. Meta-analysis indicated 16 immune-related genes as key components involved in host immune response to bovine mastitis. We identified 6 miRNA, 8 lncRNA, and 17 TFs that could potentially regulate expression of our candidate genes and were deemed significant to our study. Of importance, we reported miRNA bta-miR-223 to be highly conserved across species and targeting multiple genes. Likewise, 6 lncRNAs including NONBTAT027932.1 and XR\_003029725.1 bind to all candidate genes while SP1, JUN, and CREB transcription factors were involved in a majority of ontology processes. The crosstalk between these regulatory elements was particularly important to our study, providing a new direction for disease studies

in farm animals. Functional network analyses reveal complex interactions, suggesting that our identified significant elements would be effective drug targets for treatment and prevention through vaccination. To validate these predictions, further work is needed with *in vitro* and *in vivo* analyses that examine the expression or regulatory role of miRNA, lncRNA and TFs on host immune response in bovine mastitis.

## Methods

### Extraction and verification of significantly regulated genes in bovine mastitis

The analytical pipeline employed for this study, starting from literature search to interaction networks (Fig 1a), including some modifications have been described previously<sup>7,54</sup>. Briefly, to obtain experimentally validated genes associated with bovine mastitis, we carried out an extensive search for significantly mentioned genes, utilizing several databases. First, LitInspector (Genomatix version 3.10, Munich, Germany) was used to scan all literature-curated datasets derived from high and low throughput experiments with a Medical Subject Headings (MeSH) on bovine mastitis, as described<sup>7</sup>. We obtained a MeSH-Term ID of C22.196.5819 and a p-value of 48E-31. Identified gene sets were filtered based on the number of reports from Google Scholar and Web of Science articles, pathway enrichment and gene ontology (GO)<sup>55</sup>. Second, we carried out a meta-analysis on PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>) with combination of keywords including bovine mastitis + gene expression, to catalog genes from published literature in the database. We matched the identified genes from both search methods with Bioinformatics and Evolutionary Genomics (BEG) Venn diagram generator (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) to identify common genes. The overlapped genes between the two lists were selected as our candidate genes (total of 16 genes) and were included for further analysis (Table 1). We then constructed a protein-protein interaction network of the candidate genes with Search Tool for the Retrieval of Interacting Genes database (STRING-DB; [string-db.org](http://string-db.org)) to determine molecular interactions and possible mechanisms for co-expression and interconnectivity, from both text mining and experimentally determined datasets<sup>56</sup>.

### Prediction of miRNA for bovine mastitis candidate genes

The locations of the 16 identified genes from our analysis were determined within the *Bos taurus* reference genome (ARS-UCD1.2) and complete transcript sequences were downloaded from GenBank. The transcript information and accession numbers are listed in Table 1. Furthermore, using *Bos taurus* as the reference organism, we searched for miRNAs that could bind complementarily against the entire region of each of our candidate genes, including the 5'-UTR, CDS, and 3'-UTR, using the prediction software miRWalk<sup>20</sup>. MiRNA binding to any of the three regions with a maximum p-value of 1 was used for further selection. To evade false positives in our analysis, the generated list of miRNAs was cross referenced with two other prediction software, miRNet<sup>21</sup> and TargetScan<sup>22</sup> as described<sup>7</sup>. Using Venn generator, the overlap of predicted miRNA from the three software were selected as significant for further analysis (Table 2).

## Identification of lncRNAs associated with immune response during bovine mastitis

In order to identify putative lncRNAs within cattle genome that may be binding to the transcript sequences thereby regulating expression, we performed a genome-wide search based on the publicly available curated data from four databases: NONCODE,<sup>24</sup> a comprehensive knowledge based ncRNA database that specifically focuses on lncRNA; LncTar<sup>19</sup> a lncRNA-RNA prediction software using the minimum binding free energy between RNA sequences; Ensembl (<http://useast.ensembl.org/index.html>), a vertebrate genome browser supporting analysis software, and LncRNA2Target v2.0<sup>25</sup>, a comprehensive database of lncRNA-target relationships, manually curated from published lncRNA-focused articles. Briefly, we searched each database for lncRNAs using cow as reference. Due to the high number of lncRNAs within the genome, we made some assumptions and applied the most stringent filter criteria using: 1. The number of exons, believing more exons imply high quality transcripts; we selected only lncRNAs with exon number greater than two. 2. Longer transcript length gives better quality transcripts; transcript length >200 bp were selected. 3. The prediction score; a lower score means a higher quality transcript; prediction score less than -0.2 was used. 4. Finally, we selected only lncRNAs that meet criteria 1-3 and also found in minimum of three out of the four databases. Briefly, complete sequences of each lncRNA were retrieved from NONCODE and each sequence searched on the bovine genome ARS-UCD1.2. using the Ensembl BLAT/BLAST genomic sequencing tools. NCBI BLAST was used to identify available accession numbers for each lncRNA (Table 3). Using LncTar, each mRNA gene was used as target to determine if bovine lncRNAs binds to the mRNA target. Based on the normalized free binding energy (ndG) value from the nucleotide base pairing, dataset with less than -0.05 were considered significant and used as candidates for further analysis<sup>19</sup>. To avoid false positive lncRNAs, we used LncRNA2Target v2.0 as a second lncRNA-mRNA prediction software with human homology for the 16 target genes to determine bovine lncRNA binding sites. We located all identified human lncRNAs counterparts on the bovine genome ARS-UCD1.2. The lncRNAs found in bovine genome ARS-UCD1.2 and the human genome GRCh38 e were significant to our study.

## Prediction of significant transcription factors

Transcription factors and their corresponding binding sites were predicted for the genes using different prediction software; AnimalTFDB<sup>57</sup> and GeneXplain<sup>58</sup>, so as to increase result accuracy. AnimalTFDB was employed using the default settings with the p-value cut off at the minimum of e-06. For GeneXplain, both Match and P-Match algorithms were adjusted to use the cut off selection “1.0 and 1.0” as matrix similarity cut-off and core similarity cut-off. We included only seventeen TFs that appeared in the results of at least two software for further analysis.

## Functional and pathway analysis of candidate genes associated with immune response to bovine mastitis

Functional enrichment analysis, using geneontology.org, was performed on the 16 candidate genes to determine the molecular functions, biological processes and cellular components the genes are involved

in as described<sup>14,59</sup>, with slight changes. Using the fold enrichment of greater than 100, the false discovery rate (FDR) of  $<10^{-5}$ , and the p-value of  $<10^{-7}$ , we catalogue the biological processes and cellular components. The molecular functions were selected based on the following criteria; fold enrichment  $>100$ , FDR  $<0.001$ , and the p-value was  $<0.01$ . To validate these findings, we carried out functional analysis with Database for Annotation, Visualization and Integrated Discovery (DAVID)<sup>60</sup>, using the same parameters as stated previously. Results common in both programs were used for further analysis. Significant pathways were generated using KEGG program algorithms with fold enrichment of  $>40$ , while the p-value and FDR were  $<10^{-4}$ .

### **Sequence conservation analysis of significant miRNA and lncRNA**

Using miRBase<sup>61</sup>, candidate miRNA sequences were retrieved and ascertained using BLAST against cow, dog, cat, pig, human, gorilla, chicken, goat, rat, megabat, horse, mouse, wild yak, sheep, and chimpanzee for conservation analysis (Table S1). MEGAX<sup>62</sup> and MultiAlin<sup>63</sup> were used to align miRNA sequences from each species to determine the regions of conservation and construct a phylogenetic tree. The tree was exported as a Newick text-file and uploaded to iTOL<sup>64</sup> for further editing. Furthermore, lncRNA sequences extracted from NONCODE were compared using multiple sequence alignment and phylogenetic analyses to detect evolutionary profile across 15 species using MEGA X. The phylogenetic trees were also imported to iTOL for further editing and visualization.

### **Prediction of lncRNA-miRNA interactions associated with bovine mastitis pathogenesis**

We hypothesized that a complementary base pairing or binding of miRNA and lncRNA would be an important mechanism to manipulate the regulatory elements involved in immune response to bovine mastitis. To do this, we predicted the possible miRNA binding sites on the lncRNA sequences and determine their normalized binding free energy (ndG) based on the complementary base pairing, using LncTar<sup>19</sup>. Furthermore, miRNA, lncRNA, their target genes, and the biological processes were connected on a critical network using Cytoscape version 3.7.2, as described previously<sup>54,65</sup>.

### **Combined interactome network of lncRNAs, miRNAs, TFs and immune gene targets**

To provide information on the connectivity as well as to identify the molecular mechanisms for gene regulation between the candidate genes, miRNAs, lncRNAs, and TFs, we created an interactome network with Cytoscape software, version 3.7.2. Network of different nodes were created based on all identified RNAs, TFs and associated pathways, while the network edges determine the relationship between the regulatory elements and pathways. Our interactome modeling took the advantage of modeling information in process pathways of Database for Annotation, Visualization and Integrated Discovery (DAVID), Gene ontology (GO) and KEGG ([www.genome.jp/kegg](http://www.genome.jp/kegg)), creating new global biological associations within the dataset.

## **Declarations**

## Data availability

The datasets analyzed in this study are available from multiple repositories (see accession numbers in Tables 1, 2, and 5). Additional accession numbers are provided in Tables S1 and S4.

## Acknowledgements

Part of this work was presented at the Virtual Annual Conference of the American Association of Immunologists, May 10-15, 2021. OBM was supported by an AAI Minority Scientist Travel Award, while ART and BNT were funded through an AAI Undergraduate Faculty Travel Award. Ongoing support by the College of Health Sciences and Technology, Rochester Institute of Technology is acknowledged.

## Author Contributions

OBM and BNT conceptualized and designed the experiments; OBM, ART, NAS and AA carried out the experiments, analyzed the data and drafted the manuscript; ART, EM, BNT and OBM revised the manuscript, contributed to the discussion and scientific content. All authors read and approved the final version of the manuscript.

## Competing interests

Authors declare no competing interests

## References

1. Lavon Y, Leitner G, Kressel Y, Ezra E, Wolfenson D. Comparing effects of bovine *Streptococcus* and *Escherichia coli* mastitis on impaired reproductive performance. *J Dairy Sci.* **102**, 10587–98 (2019).
2. Rault L, Lévêque P-A, Barbey S, Launay F, Larroque H, Le Loir Y, et al. Bovine teat cistern microbiota composition and richness are associated with the immune and microbial responses during transition to once-daily milking. *Front Microbiol.* **11**, 602404 (2020).
3. McConnel CS, Crisp SA, Biggs TD, Ficklin SP, Parrish LM, Trombetta SC, et al. A fixed cohort field study of gene expression in circulating leukocytes from dairy cows with and without mastitis. *Front Vet Sci.* **7**, 559279 (2020).
4. Khan MZ, Khan A, Xiao J, Ma J, Ma Y, Chen T, et al. Overview of research development on the role of NF- $\kappa$ B signaling in mastitis. *Animals.* **10**, 1625 (2020).
5. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biol.* **20**, 18 (2019).
6. Tong C, Chen Q, Zhao L, Ma J, Ibeagha-Awemu EM, Zhao X. Identification and characterization of long intergenic noncoding RNAs in bovine mammary glands. *BMC Genomics.* **18**, 468 (2017).

7. Morenikeji OB, Hawkes ME, Hudson AO, Thomas BN. Computational network analysis identifies evolutionarily conserved miRNA gene interactions potentially regulating immune response in bovine trypanosomosis. *Front Microbiol.* **10**, 2010 (2019).
8. Wang H, Wang X, Li X, Wang Q, Qing S, Zhang Y, et al. A novel long non-coding RNA regulates the immune response in MAC -T cells and contributes to bovine mastitis. *FEBS J.* **286**, 1780–95 (2019).
9. Li A, Mallik S, Luo H, Jia P, Lee D-F, Zhao Z. H19, a long non-coding RNA, mediates transcription factors and target genes through interference of microRNAs in pan-cancer. *Molecular Therapy - Nucleic Acids.* **21**, 180-191 (2020).
10. Dahariya S, Paddibhatla I, Kumar S, Raghuwanshi S, Palapati A, Gutti RK. Long non-coding RNA: classification, biogenesis and functions in blood cells. *Molecular Immunology.* **112**, 82–92 (2019).
11. Chen J, Wang Y, Wang C, Hu J-F, Li W. LncRNA functions as a new emerging epigenetic factor in determining the fate of stem cells. *Front Genet.* **11**, 277 (2020).
12. Ou-Yang L, Huang J, Zhang X-F, Li Y-R, Sun Y, He S, et al. LncRNA-disease association prediction using two-side sparse self-representation. *Front Genet.* **10**, 476 (2019).
13. Yue B, Li H, Liu M, Wu J, Li M, Lei C, et al. Characterization of lncRNA–miRNA–mRNA network to reveal potential functional ceRNAs in bovine skeletal muscle. *Front Genet.* **10**, 91 (2019).
14. Morenikeji OB, Bernard K, Strutton E, Wallace M, Thomas BN. Evolutionarily conserved long non-coding RNA regulates gene expression in cytokine storm during COVID-19. *Front Bioeng Biotechnol.* **8**, 582953 (2021).
15. Li M, Sun X, Cai H, Sun Y, Plath M, Li C, Lan X, Lei C, Lin F, Bai Y, Chen H. Long non-coding RNA ADNCR suppresses adipogenic differentiation by targeting miR-204. *Biochim Biophys Acta.* **1859**, 871-82 (2016).
16. Jiang M, Cheng Y, Wang D, Lu Y, Gu S, Wang C, et al. Transcriptional network modulated by the prognostic signature transcription factors and their long noncoding RNA partners in primary prostate cancer. *EBioMedicine.* **63**, 103150 (2020).
17. Madsen JGS, Rauch A, Van Hauwaert EL, Schmidt SF, Winnefeld M, Mandrup S. Integrated analysis of motif activity and gene expression changes of transcription factors. *Genome Res.* **28**, 243–55 (2018).
18. Hanif Q, Farooq M, Amin I, Mansoor S, Zhang Y, Khan QM. In silico identification of conserved miRNAs and their selective target gene prediction in indicine (*Bos indicus*) cattle. *PLOS ONE.* **13**, e0206154 (2018).

19. Li J, Ma W, Zeng P, Wang J, Geng B, Yang J, et al. LncTar: a tool for predicting the RNA targets of long noncoding RNAs. *Briefings in Bioinformatics*. **16**, 806–12 (2015).
20. Sticht C, Torre CDL, Parveen A, Gretz N. MiRWalk: an online resource for prediction of microRNA binding sites. *PLOS ONE*. **13**, e0206239 (2018).
21. Chang L, Zhou G, Soufan O, Xia J. MiRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Res*. **48**, W244–51 (2020).
22. Agarwal V, Bell GW, Nam J-W, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, e05005 (2015).
23. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. **47**, D330–8 (2019).
24. Zhao Y, Li H, Fang S, Kang Y, wu W, Hao Y, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res*. **44**, D203–8 (2016).
25. Cheng L, Wang P, Tian R, Wang S, Guo Q, Luo M, et al. LncRNA2Target v2.0: a comprehensive database for target genes of lncRNAs in human and mouse. *Nucleic Acids Res*. **47**, D140–4 (2019).
26. Abebe R, Hatiya H, Abera M, Megersa B, Asmare K. Bovine mastitis: prevalence, risk factors and isolation of *Staphylococcus aureus* in dairy herds at hawassa milk shed, South Ethiopia. *BMC Veterinary Research*. **12**, 270 (2016).
27. Heikkilä AM, Liski E, Pyörälä S, Taponen S. Pathogen-specific production losses in bovine mastitis. *J Dairy Sci*. **101**, 9493–9504 (2018).
28. Skarbye AP, Krogh MA, Sørensen JT. The effect of individual quarter dry-off in management of subclinical mastitis on udder condition and milk production in organic dairy herds: a randomized field trial. *J Dairy Sci*. **101**, 11186–98 (2018).
29. Sun C, Fu Y, Gu X, Xi X, Peng X, Wang C, et al. Macrophage-enriched lncRNA RAPIA: a novel therapeutic target for atherosclerosis. *Arterioscler Thromb Vasc Biol*. **40**, 1464–78 (2020).
30. Li T, Lu H, Wang X, Gao Q, Dai Y, Shang J, et al. Molecular characteristics of *Staphylococcus aureus* causing bovine mastitis between 2014 and 2015. *Front Cell Infect Microbiol*. **7**, 127 (2017).
31. Lin T, Yang Y, Ye X, Yao J, Zhou H. Low expression of miR-99b promotes progression of clear cell renal cell carcinoma by up-regulating IGF1R/Akt/mTOR signaling. *Int J Clin Exp Pathol*. **13**, 3083–91 (2020).
32. Fang L, Hou Y, An J, Li B, Song M, Wang X, et al. genome-wide transcriptional and post-transcriptional regulation of innate immune and defense responses of bovine mammary gland to

*Staphylococcus aureus*. *Front Cell Infect Microbiol*. **6**, 193 (2016).

33. Bulgari O, Dong X, Roca AL, Caroli AM, Loor JJ. Innate immune responses induced by lipopolysaccharide and lipoteichoic acid in primary goat mammary epithelial cells. *J Anim Sci Biotechnol*. **8**, 29 (2017).
34. Wang G, Zheng X, Zheng Y, Cao R, Zhang M, Sun Y, et al. Construction and analysis of the lncRNA-miRNA-mRNA network based on competitive endogenous RNA reveals functional genes in heart failure. *Mol Med Rep*. **19**, 994–1003 (2019).
35. Casas E, Cai G, Kuehn LA, Register KB, McDanel TG, Neill JD. Association of microRNAs with antibody response to *Mycoplasma bovis* in Beef Cattle. *PLOS ONE*. **11**, e0161651 (2016).
36. Ma S, Tong C, Ibeagha-Awemu EM, Zhao X. Identification and characterization of differentially expressed exosomal microRNAs in bovine milk infected with *Staphylococcus aureus*. *BMC Genomics* **20**, 934 (2019).
37. Han S, Li X, Liu J, Zou Z, Luo L, Wu R, et al. Bta-miR-223 targeting CBLB contributes to resistance to *Staphylococcus aureus* mastitis through the PI3K/AKT/NF- $\kappa$ B pathway. *Front Vet Sci*. **7**, 529 (2020).
38. Deny M, Romano M, Denis O, Casimir G, Chamekh M. Progressive control of streptococcus agalactiae-induced innate inflammatory response is associated with time course expression of microRNA-223 by neutrophils. *Infect Immun*. **88**, 2 (2020).
39. Brook AC, Jenkins RH, Clayton A, Kift-Morgan A, Raby A-C, Shephard AP, et al. Neutrophil-derived miR-223 as local biomarker of bacterial peritonitis. *Sci Rep*. **9**, 10136 (2019).
40. Tay HL, Kaiko GE, Plank M, Li J, Maltby S, Essilfie A-T, et al. Antagonism of miR-328 increases the antimicrobial function of macrophages and neutrophils and rapid clearance of non-typeable *Haemophilus influenzae* (NTHi) from infected lung. *PLOS Pathogens*. **11**, :e1004549 (2015).
41. Mendonça LS, Santos JM, Kaneto CM, de Carvalho LD, Lima-Santos J, Augusto DG, Carvalho SM, Soares-Martins JA, Silva-Jardim I. Characterization of serum cytokines and circulating microRNAs that are predicted to regulate inflammasome genes in cutaneous leishmaniasis patients. *Experimental Parasitol*, 210, 107846 (2020).
42. Khan R, Raza SHA, Junjvlieke Z, Xiaoyu W, Wang H, Cheng G, et al. Bta-miR-149-5p inhibits proliferation and differentiation of bovine adipocytes through targeting CRTCs at both transcriptional and posttranscriptional levels. *J Cell Physiol*. 235, (2020).
43. Guo L, Zhao Y, Yang S, Zhang H, Chen F. An integrated analysis of miRNA, lncRNA, and mRNA expression profiles. *Sci Rep*. **7**, 44026 (2017).



44. Zhang S, Kang Z, Cai H, Jiang E, Pan C, Dang R, et al. Identification of novel alternative splicing of bovine lncRNA lncFAM200B and its effects on preadipocyte proliferation. *J Cell Physiol.* **236**, 601–11 (2021).
45. Li N, Wu S, Yu L. The associations of long non-coding RNA taurine upregulated gene 1 and microRNA-223 with general disease severity and mortality risk in sepsis patients. *Medicine.* **99**, e23444 (2020).
46. Li L, Huang J, Zhang X, Ju Z, Qi C, Zhang Y, et al. One SNP in the 3'-UTR of HMGB1 gene affects the binding of target bta-miR-223 and is involved in mastitis in dairy cattle. *Immunogenetics.* **64**, 817–24 (2012).
47. Black AR, Black JD, Azizkhan-Clifford J. Sp1 and krüppel-like factor family of transcription factors in cell growth regulation and cancer. *J Cell Physiol.* **188**, 143–60 (2001).
48. Wierstra I. Sp1: emerging roles—beyond constitutive activation of TATA-less housekeeping genes. *Biochemical and Biophysical Research Communications.* **372**, 1–13 (2008).
49. Zenz R, Scheuch H, Martin P, Frank C, Eferl R, Kenner L, et al. C-Jun regulates eyelid closure and skin tumor development through EGFR signaling. *Developmental Cell.* **4**, 879–89 (2003).
50. Liu T, Zhang L, Joo D, Sun S-C. NF-κB signaling in inflammation. *Signal Transduct Target Ther.* **2**, 17023 (2017).
51. Wu J, Li L, Sun Y, Huang S, Tang J, Yu P, et al. Altered Molecular Expression of the TLR4/NF-κB Signaling Pathway in Mammary Tissue of Chinese Holstein Cattle with Mastitis. *PLOS ONE.* **10**, (2015).
52. Wen AY, Sakamoto KM, Miller LS. The Role of the Transcription Factor CREB in Immune Function. *J Immunol.* **185**, 6413–9 (2010).
53. Burimuah V, Sylverken A, Owusu M, El-Duah P, Yeboah R, Lamptey J, et al. Molecular-based cross-species evaluation of bovine coronavirus infection in cattle, sheep and goats in Ghana. *BMC Veterinary Research.* **16**, 405 (2020).
54. Morenikeji OB, Wallace M, Strutton E, Bernard K, Yip E, Thomas BN. Integrative network analysis of predicted miRNA-targets regulating expression of immune response genes in bovine coronavirus infection. *Front Genet.* **11**, 584392 (2020).
55. Buza N, Xu F, Wu W, Carr RJ, Li P, Hui P. Recurrent chromosomal aberrations in intravenous leiomyomatosis of the uterus: high-resolution array comparative genomic hybridization study. *Hum Pathol.* **45**, 1885-92 (2014).
56. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide

experimental datasets. *Nucleic Acids Res.* **47**, D607–13 (2019).

57. Hu H, Miao Y-R, Jia L-H, Yu Q-Y, Zhang Q, Guo A-Y. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res.* **47**, D33–8 (2019).
58. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E. TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **34(Database issue)**, D108-110 (2006).
59. Medina IR, Lubovac-Pilav Z. Gene co-expression network analysis for identifying modules and functionally enriched pathways in type 1 diabetes. *PLOS ONE.* **11**, e0156006 (2016).
60. Jiao X, Sherman B, Huang da W, Stephens R, Baseler M, Lane H, et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics.* **28**, 1805–6 (2012).
61. Kozomara A, Birgaoanu M, Griffiths-Jones S. MiRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–62 (2019).
62. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol.* **35**, 1547–9 (2018).
63. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881–90 (1988).
64. Letunic I, Bork P. Interactive Tree of Life (iTOL) v4: recent updates and developments. *Nucleic Acids Res.* **47**, W256-W259 (2019).
65. Luo X, Tu T, Zhong Y, Xu S, Chen X, Chen L, et al. CeRNA network analysis shows that lncRNA CRNDE promotes progression of glioblastoma through sponge mir-9-5p. *Front Genet.* **12**, 617350 (2021).

## Tables

**Table 1.** List of candidate bovine mastitis genes and their accession number, genomic location, and strand type.

Candidate Gene	Accession Number	Genomic Location	Strand (+/-)
MYD88	NM_001014382	22:11625515-11629955	+
CD4	NM_001103225	5:103631360-103654878	-
IL-10	NM_174088	16:4550836-4555318	-
IFN $\gamma$	NM_174086	5:45624462-45629336	+
IL-4	NM_173921	7:21696248-21704136	-
ICAM1	NM_174348	7:14787173-14797855	+
CXCL8	NM_173925	6:88810418-88814572	+
TLR4	NM_174198	8:107057826-107068836	+
TNF $\alpha$	NM_173966	23:27716168-27719047	-
IL-18	NM_174091	15:22475988-22502331	-
TLR2	NM_174197	17:3954019-3967242	-
CD86	NM_001038017	1:66543642-66610926	+
CCL2	NM_174006	19:15902777-15905368	-
IL-6	NM_173923	4:31454749-31459131	+
CSF2	NM_174027	7:22398938-22401303	-
CD14	NM_174008	7:51762895-51765768	-

**Table 2.** Significant miRNA predicted by all three software and their mRNA sequence, accession number, genomic location, strand type and length.

miRNA	mRNA Sequence	Accession	Genomic Location	Strand (+/-)	Length (bp)
bta-miR-149-5p	UCUGGCUCCGUGUCUUCACUCCC	MI0021115	3:120921662-120921751	+	23
bta-miR-185	UGGAGAGAAAGGCAGUUCCUGA	MI0009758	17:75173467-75173545	+	22
bta-miR-223	UGUCAGUUUGUCAAAUACCCCA	MI0009782	X:94,562,822-94,562,929	-	22
bta-miR-24-3p	UGGCUCAGUUCAGCAGGAACAG	MIMAT0003840	-	-	22
bta-miR-328	CUGGCCCUCUCUGCCCUUCCGU	MI0009800	18:35062240-35062331	-	22
bta-miR-874	CUGCCCUGGCCCGAGGGACCGA	MI0009900	7: 50886238-50886313	-	22

These miRNAs are from the Venn diagram generated through the overlap of the three softwares-miRWalk, miRNet an TargetScan

**Table 3. Significant lncRNA predicted to target the most candidate genes**

List of lncRNA	Number of target genes	List of target genes
NONBTAT001181.2	16	MYD88, CD4, IFN $\gamma$ , IL-4, ICAM1, IL-18, CD86, CSF2, CCL2, TLR4, CXCL8, IL-10, IL-6, CD14, TLR2, TNF $\alpha$
XR_003029725.1	16	MYD88, CD4, IFN $\gamma$ , IL-4, ICAM1, IL-18, CD86, CSF2, CCL2, TLR4, CXCL8, IL-10, IL-6, CD14, TLR2, TNF $\alpha$
NONBTAT010129.2	16	MYD88, CD4, IFN $\gamma$ , IL-4, ICAM1, IL-18, CD86, CSF2, CCL2, TLR4, CXCL8, IL-10, IL-6, CD14, TLR2, TNF $\alpha$
XR_003030515.1	16	MYD88, CD4, IFN $\gamma$ , IL-4, ICAM1, IL-18, CD86, CSF2, CCL2, TLR4, CXCL8, IL-10, IL-6, CD14, TLR2, TNF $\alpha$
XR_003033296.1	16	MYD88, CD4, IFN $\gamma$ , IL-4, ICAM1, IL-18, CD86, CSF2, CCL2, TLR4, CXCL8, IL-10, IL-6, CD14, TLR2, TNF $\alpha$
NONBTAT027932.1	16	MYD88, CD4, IFN $\gamma$ , IL-4, ICAM1, IL-18, CD86, CSF2, CCL2, TLR4, CXCL8, IL-10, IL-6, CD14, TLR2, TNF $\alpha$
XR_234647.4	13	MYD88, CD4, IFN $\gamma$ , IL-4, ICAM1, IL-18, CSF2, TLR4, CXCL8, IL-10, CD14, TLR2, TNF $\alpha$
NONBTAT013032.2	13	MYD88, CD4, ICAM1, IL-18, CD86, CSF2, CCL2, TLR4, CXCL8, IL-10, CD14, TLR2, TNF $\alpha$

**Table 4.** Gene Ontology biological processes, molecular functions, and cellular components of bovine mastitis target genes

GO biological process	Gene (s)	Raw p-value
LPS-mediated signaling pathway	CCL2, TLR4, MYD88, CD14, IL-18, TNF $\alpha$	2.58E-13
myeloid leukocyte activation	IL-4, IFN $\gamma$ , TLR4, CXCL8, IL-18, CSF2, TNF $\alpha$	4.92E-13
positive regulation of interleukin-8 production	TLR2, TLR4, MYD88, CD14, IL-6, TNF $\alpha$	1.30E-12
regulation of interleukin-8 production	TLR2, TLR4, MYD88, CD14, IL-6, TNF $\alpha$	6.92E-12
positive regulation of interleukin-6 production	TLR2, IFN $\gamma$ , TLR4, MYD88, IL-6, TNF $\alpha$	2.01E-11
regulation of interleukin-23 production	IFN $\gamma$ , TLR4, MYD88, CSF2	1.34E-10
regulation of cytokine secretion	TLR2, IFN $\gamma$ , CD14, IL-10, TNF $\alpha$	2.50E-10
positive regulation of tyrosine phosphorylation of STAT protein	IL-4, IFN $\gamma$ , IL-6, IL-18, CSF2	4.67E-10
positive regulation of NIK/NF-kappaB signaling	TLR2, TLR4, CD14, IL-18, TNF $\alpha$	5.65E-10
regulation of chemokine production	TLR2, TLR4, MYD88, IL-6, TNF $\alpha$	8.12E-10
regulation of tyrosine phosphorylation of STAT protein	IL-4, IFN $\gamma$ , IL-6, IL-18, CSF2	1.14E-09
positive regulation of cytokine production involved in inflammatory response	TLR4, MYD88, IL-6, TNF $\alpha$	4.01E-09
positive regulation of cytokine secretion	IFN $\gamma$ , CD14, IL-10, TNF $\alpha$	4.70E-09
regulation of interleukin-17 production	IFN $\gamma$ , TLR4, MYD88, IL-18	6.36E-09

macrophage activation	IL-4, IFN $\gamma$ , TLR4, TNF $\alpha$	1.40E-08
regulation of nitric oxide biosynthetic process	IFN $\gamma$ , TLR4, IL-10, TNF $\alpha$	1.76E-08
positive regulation of interleukin-23 production	IFN $\gamma$ , MYD88, CSF2	2.30E-08
positive regulation of chemokine production	TLR2, TLR4, IL-6, TNF $\alpha$	2.98E-08
response to lipoteichoic acid	TLR2, TLR4, CD14	3.45E-08
cellular response to lipoteichoic acid	TLR2, TLR4, CD14	3.45E-08
positive regulation of tumor necrosis factor superfamily cytokine production	IFN $\gamma$ , TLR4, MYD88, CD14	3.61E-08
regulation of cytokine production involved in inflammatory response	TLR4, MYD88, IL-6, TNF $\alpha$	3.96E-08
toll-like receptor signaling pathway	TLR2, TLR4, MYD88, CD14	5.16E-08
vascular endothelial growth factor production	IL-6, TNF, TNF $\alpha$	6.78E-08
regulation of chemokine (C-X-C motif) ligand 2 production	TLR4, MYD88, TNF $\alpha$	6.78E-08
<b>GO molecular function</b>	<b>Gene (s)</b>	<b>Raw p-value</b>
NAD(P)+ nucleosidase activity	TLR2, TLR4	2.64E-05
NAD+ nucleotidase, cyclic ADP-ribose generating	TLR2, TLR4	2.64E-05
NAD+ nucleosidase activity	TLR2, TLR4	2.64E-05
Toll-like receptor binding	TLR2, MYD88	3.86E-05
lipopeptide binding	TLR2, CD14	8.93E-05
pattern recognition receptor activity	TLR2, TLR4	9.97E-05

GO cellular component	Gene (s)	Raw p-value
LPS receptor complex	TLR4, CD14	8.80E-06
phagocytic cup	TLR4, TNF $\alpha$	1.47E-04

**Table 5:** List of transcription factors predicted by AnimalTFDB and GeneXplain, their genomic location and strands.

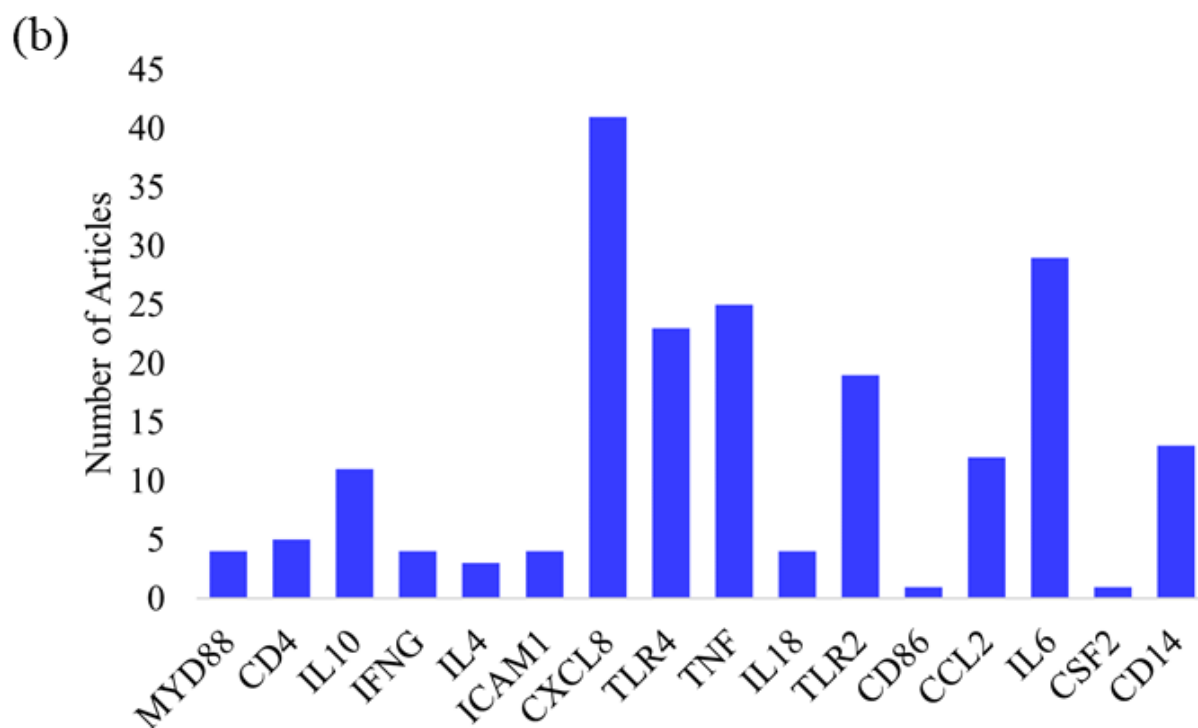
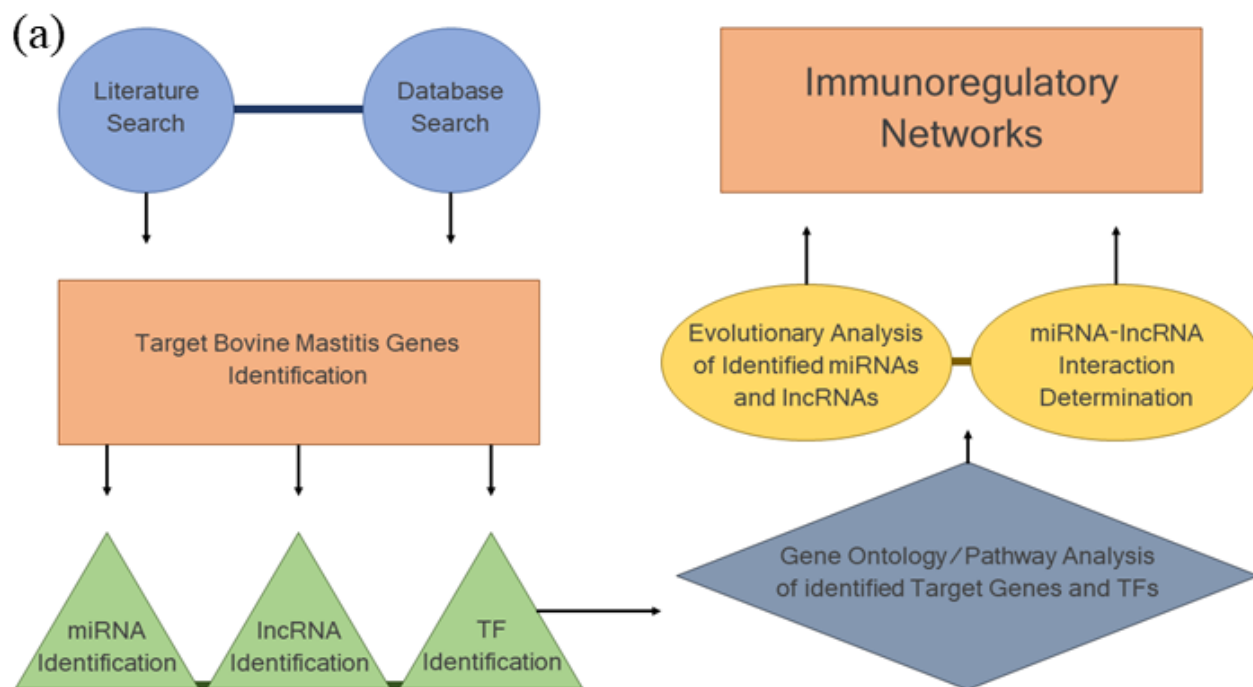


COMMON TFs from the TWO Database	Accession Number	Genomic Location	Strand (+/-)
BCL6	ENSBTAG000000001511	1:79557257-79612240	(-)
CREB1	ENSBTAG000000005474	2:95845037-95898849	(-)
FOXA2	ENSBTAG000000012407	13:41542415-41547044	(-)
FOXM1	ENSBTAG000000015875	5:106886408-106900942	(-)
JUN	ENSBTAG000000004037	3:87265922-87268047	(-)
GATA1	ENSBTAG000000003184	X:86899691-86906632	(-)
HSF1	ENSBTAG000000020751	14:612908-634769	(+)
IRF1	ENSBTAG000000031231	7:21938453-21946840	(+)
SP1	ENSBTAG000000003021	5:26607078-26643492	(+)
SP2	ENSBTAG000000013740	19:38533093-38565741	(+)
SMAD1	ENSBTAG000000002835	17:12739131-12825260	(+)
TBP	ENSBTAG000000007686	9:104096534-104106636	(-)
TCF12	ENSBTAG000000002586	10:53023683-53085010	(-)
ZNF143	ENSG000000166478	11:9458955-9529888	(-)
ZNF274	ENSBTAG000000013353	18:65546272-65569269	(-)
ZNF384	ENSBTAG000000017072	5:103755842-103776263	(-)
ZNF92	ENSG000000146757	7:65373253-65401682	(-)

**Table 6.** Gene Ontology biological processes, molecular functions, and cellular components of predicted TFs.

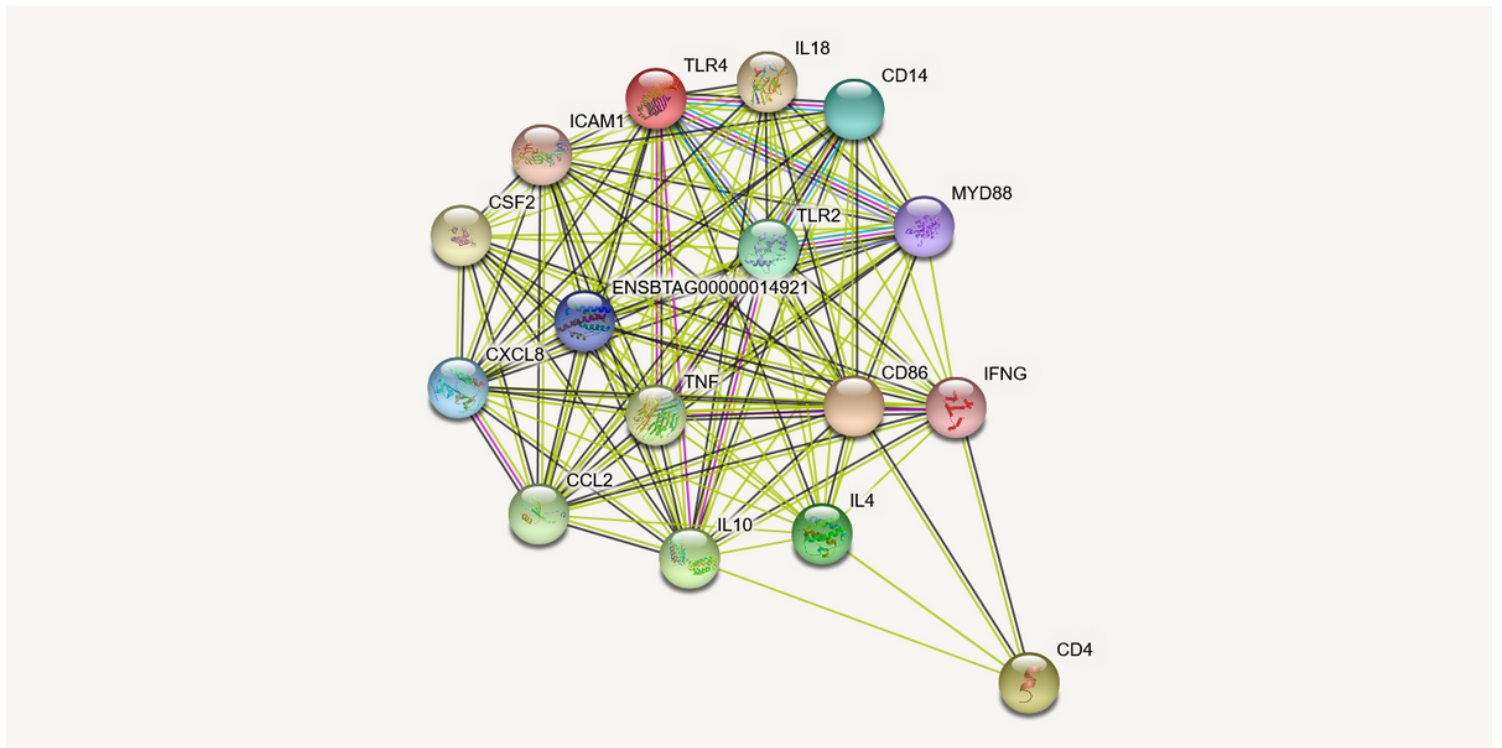
GO biological process	TFs	Raw p-value
Positive regulation by host of viral transcription	JUN, SP1	5.94E-05
Regulation of regulatory T cell differentiation	BCL6, IRF1	1.11E-04
Negative regulation of cell aging	FOXM1, BCL6	1.37E-04
<b>GO molecular function</b>		
Intronic transcription regulatory region sequence-specific DNA binding	BCL6, HSF1	4.31E-05
cAMP response element binding	JUN, CREB1	7.83E-05
<b>GO cellular component</b>		
Euchromatin	JUN, SP1, CREB1	1.11E-06

## Figures



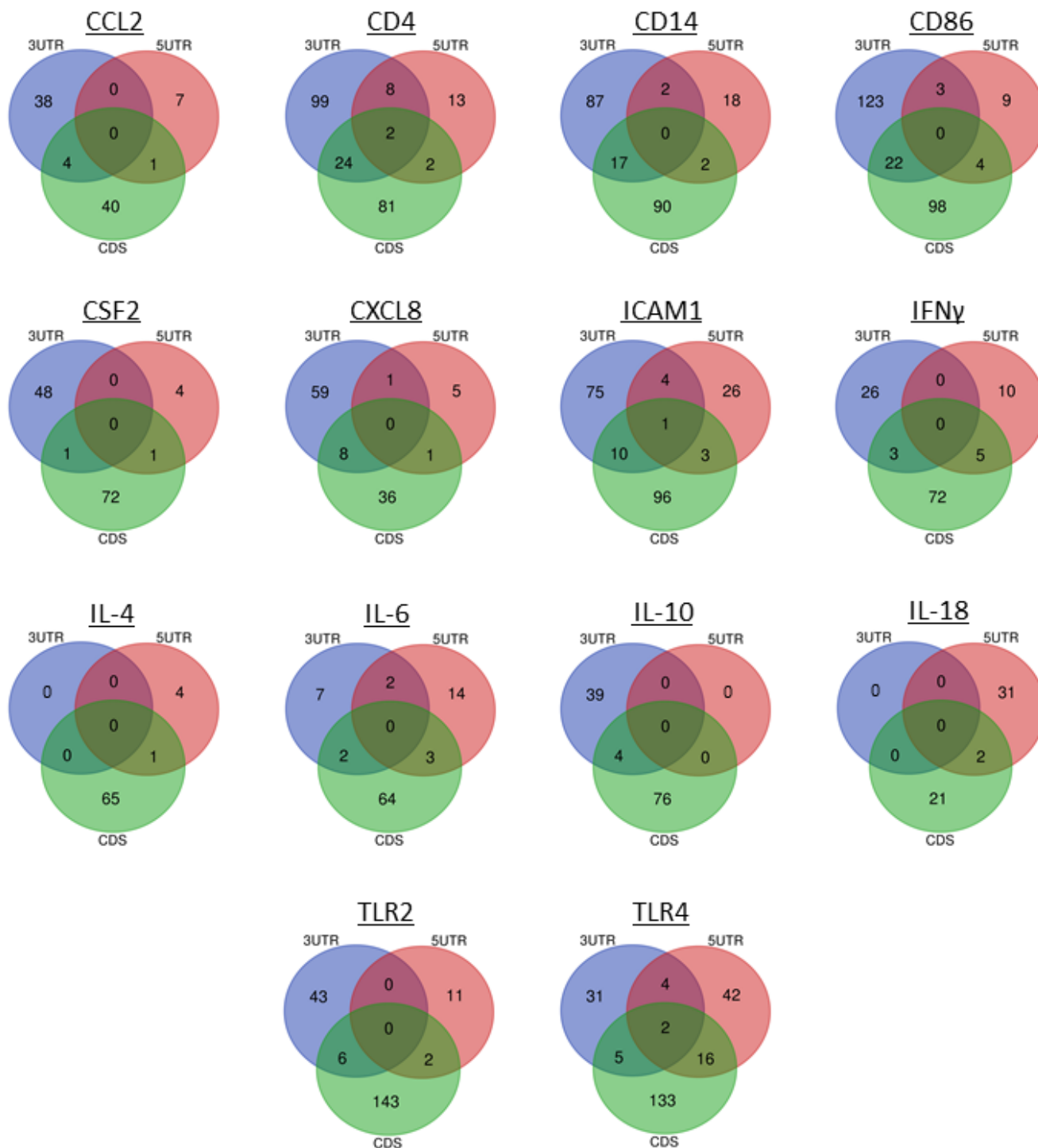
**Figure 1**

Flowchart of project steps and analysis procedures (A) and number of articles mentioning each of the 16 target bovine mastitis genes (B).



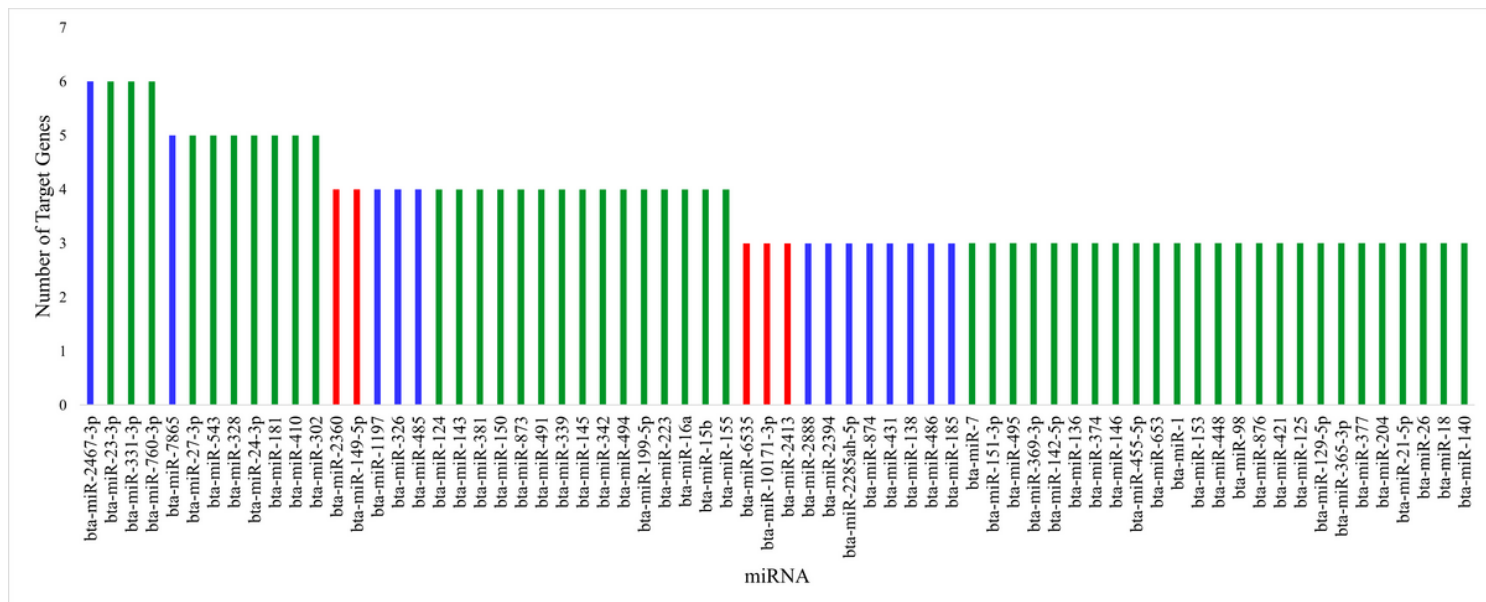
**Figure 2**

Network of 16 target bovine mastitis genes created using STRING. The Ensembl ID in the network refers to IL-6.



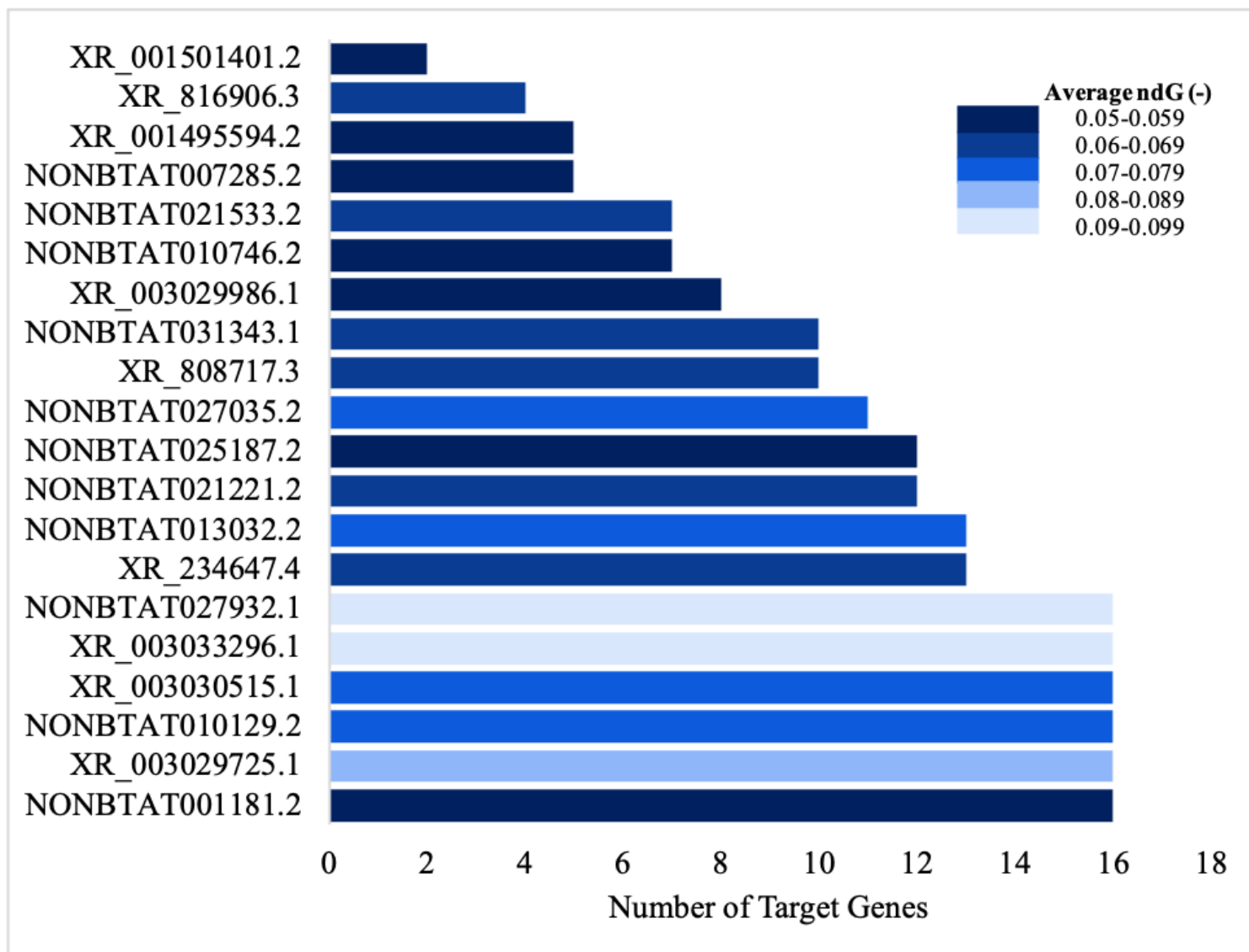
**Figure 3**

Venn Diagrams for 14 out of 16 target bovine mastitis genes showing the number of miRNA binding to the 3' UTR, CDS, and 5' UTR. The middle region of each diagram represents miRNA that bind to all three regions of the target gene.



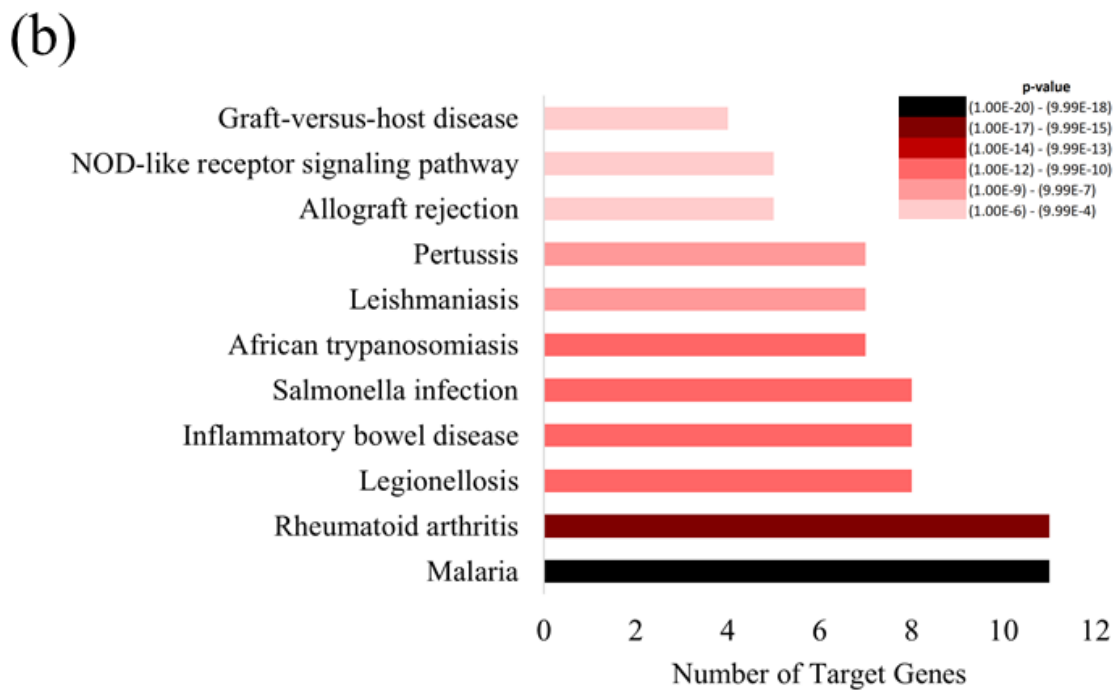
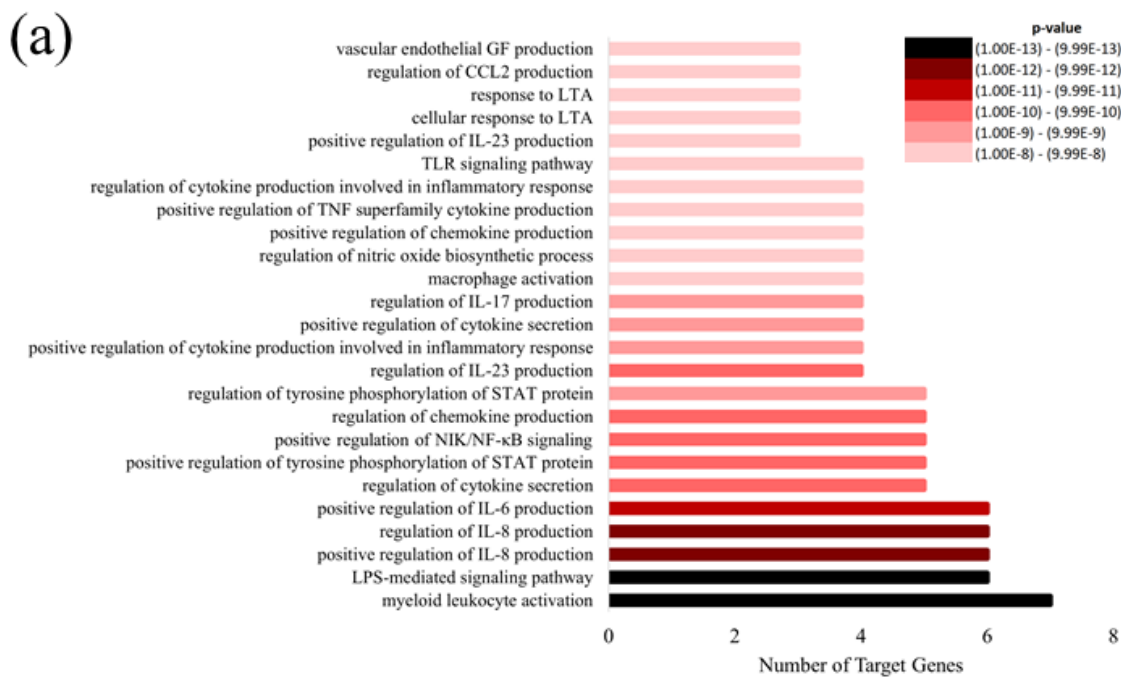
**Figure 4**

miRNA predicted to bind to three or more target genes. The red bars are miRNA predicted to bind to 3' UTR, CDS, and 5' UTR. The blue bars represent miRNA binding to two of the three regions. The green bars are miRNA predicted to bind to one of the three regions.



**Figure 5**

The 20 lncRNA and the number of target genes they are predicted to bind to. The darkest shade of blue corresponds to the lowest (-) ndG and the lightest shade of blue corresponds to the highest (-) ndG.



**Figure 6**

The biological processes and the number of target bovine mastitis genes involved (A) and the pathways and the corresponding number of target genes (B). In both Fig 6a and Fig 6b, the black bars indicate biological processes/pathways predicted with the lowest p-value and the lightest red bar represent predictions with the greatest p-value.





Network interactome of miRNA, their target genes, and the gene ontologies of the target genes. The pink diamonds are miRNA binding to a single target gene while the red diamonds are miRNA binding to two or more target genes. The green circles are the 16 target genes and the blue rectangles are the biological processes, molecular functions, cellular components, and pathways.

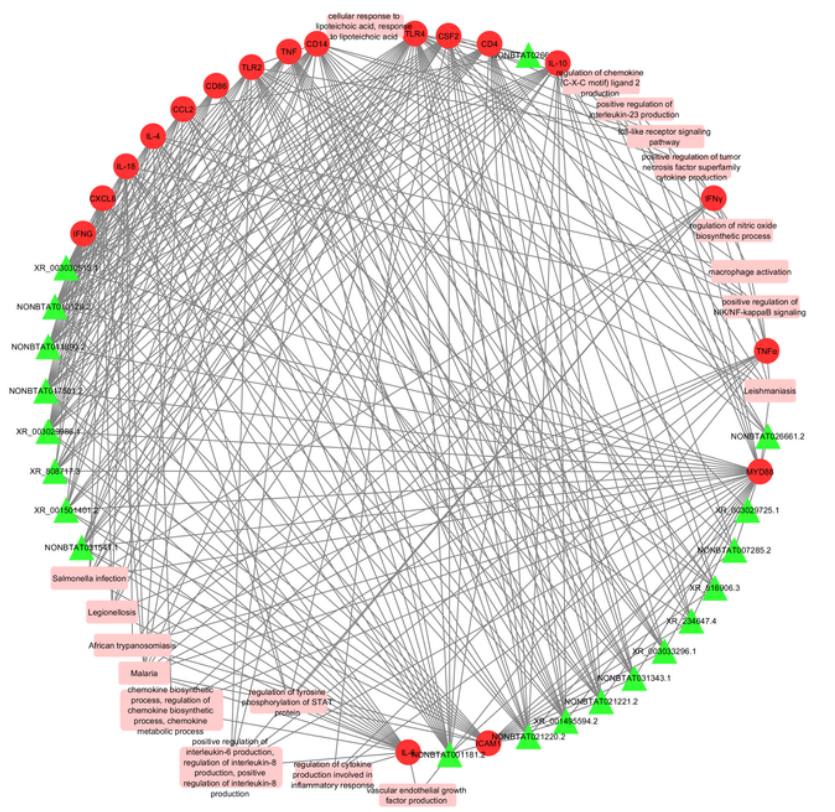
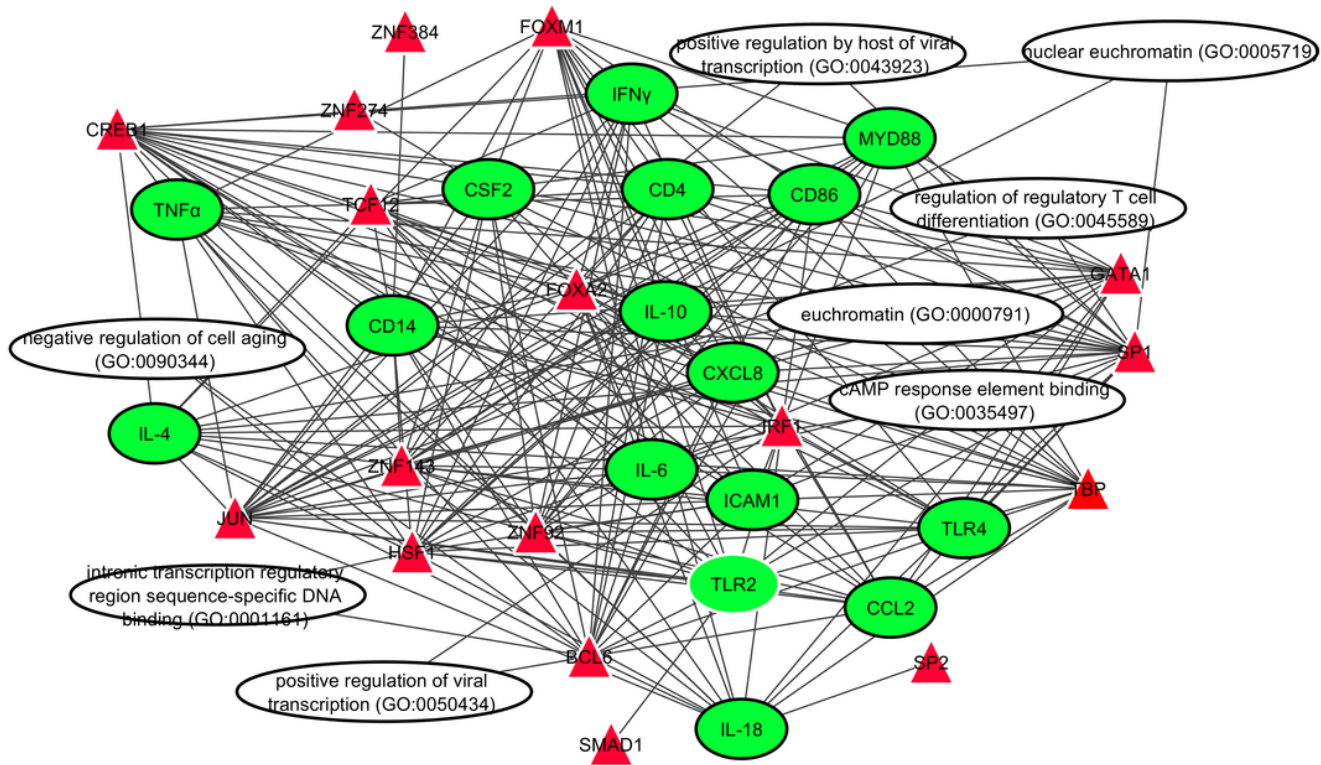


Figure 9

A network incorporating lncRNA, their target genes and corresponding gene ontologies. The green triangles are the lncRNA, red circles represent the target bovine mastitis genes, and the pink rectangles are the biological processes, molecular functions, cellular components, and pathways.



**Figure 10**

Network interactome of TFs, bovine mastitis target genes, and gene ontologies. The red triangles are TFs while the green circles are the target bovine mastitis genes. The while ovals represent the biological processes, molecular functions, cellular components, and pathways.



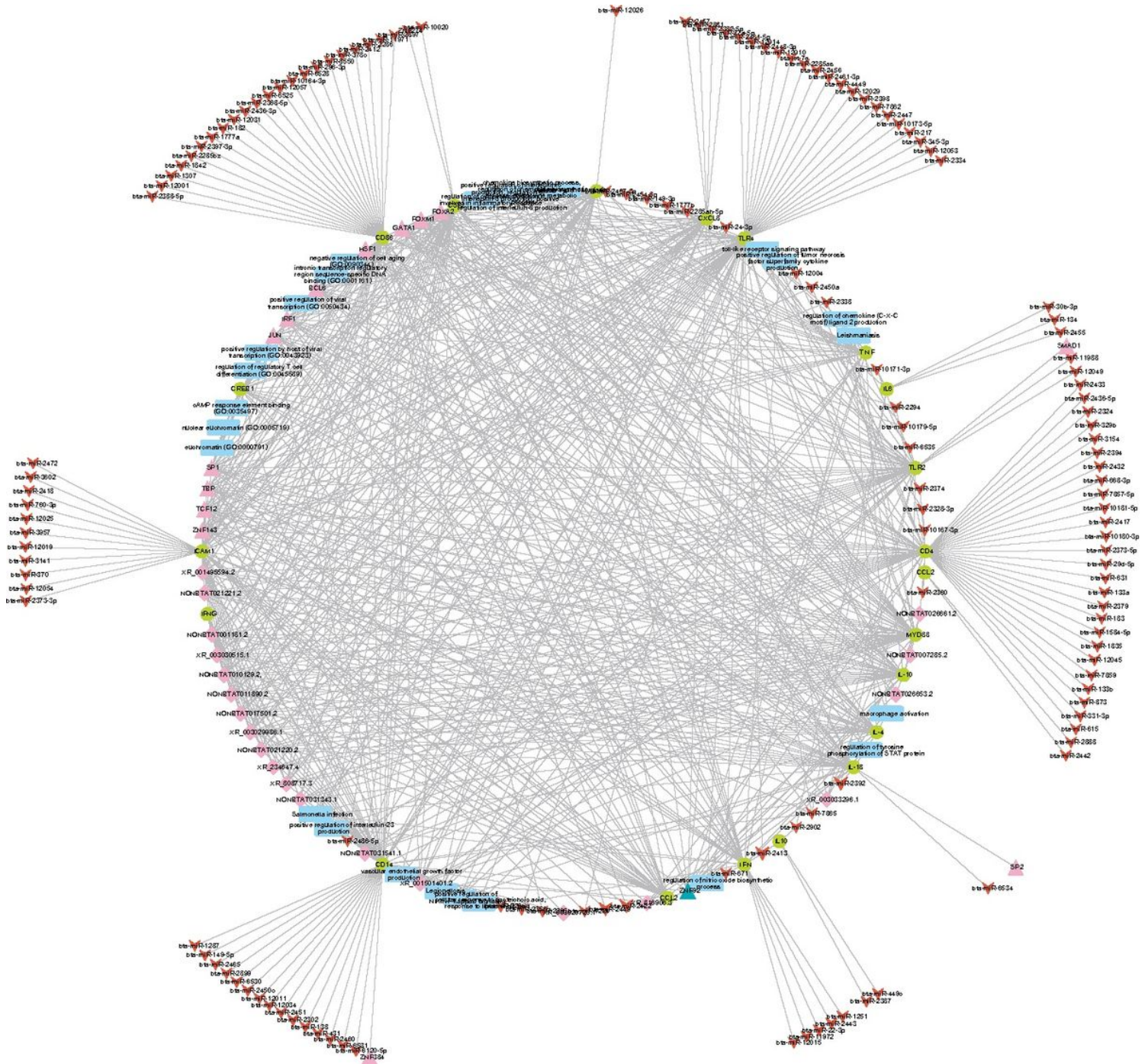


Figure 11

A network incorporating miRNA, lncRNA, TFs, target genes, and gene ontologies. The red arrowheads represent miRNA, the purple diamonds lncRNA, the teal triangles TFs, the green circles target genes, and the blue rectangles biological processes, molecular functions, cellular components, and pathways.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplFigure1.pdf](#)
- [SupplFigure2.pdf](#)
- [SupplFigure3.pdf](#)
- [SupplFigure4a.pdf](#)
- [SupplFigure4b.pdf](#)
- [SupplFigure4c.pdf](#)
- [SupplFigure5a.pdf](#)
- [SupplFigure5b.pdf](#)
- [SupplFigure5c.png](#)
- [SupplFigure5d.pdf](#)
- [SupplementaryTable1.docx](#)
- [SupplementaryTable2.docx](#)
- [SupplementaryTable3.docx](#)
- [SupplementaryTable4.docx](#)
- [SupplementaryTable5.docx](#)
- [SupplementaryTable6.docx](#)
- [SupplementaryTable7.docx](#)