Article

In silico identification and characterization of bovine mastitis related inflammatory genes and their possibilities to control mastitis

Shraddha Vishwakarma, **Neelam Krishna**, **Pramod Katara**
Computational Omics Lab, Centre of Bioinformatics, IIDS, University of Allahabad, Prayagraj (UP), India
E-mail: pkatara@allduniv.ac.in

Received 10 September 2023; *Accepted 5 October 2023*; *Published online 10 November 2023*; *Published 1 December 2023*

Abstract
Bovine mastitis, a bout of inflammation of the mammary gland in high-yield dairy cows, frequently brings on by bacterial infections. One of the biggest illnesses that harm the dairy sector financially is mastitis, which lowers milk production and quality. Taking into consideration the impact of mastitis on milk production and dairy industry, research studies are already underway, but till now no concrete solution is there. It is assuming that molecular level understanding behind the mastitis can boost the process to control the mastitis at some extent. Thus, the goal of this scientific study was to identify immunogenic genes and their functional characterization. Along with that study also attempts to explore the possibilities for the use of immunogenic proteins as therapeutic targets to control mastitis. For the purpose cDNA microarray data of bovine mammary epithelial cells after the in vitro stimulation with *Escherichia coli* was considered. This investigation proceeds to study immunological responses on mammary epithelial (BME) cells. Using network based approach the investigation pinpointed 25 essential key genes for *E. coli* mastitis and their contribution related to the immune response. Through molecular docking, study also reported 33 herbal compounds with considerable affinity against toll-like receptor (TLR) and bovine granulocyte-colony stimulating factor proteins which are critical for mastitis.

Keywords bovine mastitis; immunogenic genes; toll-like receptors; inflammation; mammary gland; *E. coli*; Neem; Aloe; Basil.

Network Biology
ISSN 2220-8879
URL: http://www.iaees.org/publications/journals/nb/online-version.asp
RSS: http://www.iaees.org/publications/journals/nb/rss.xml
E-mail: networkbiology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

All over the world, there has been increasing demand from consumers for healthy dairy products. Imperatively to say, there is a need to pay attention to animal health and it has an expensive problem in global milk production (Palii et al., 2020; Krishna et al., 2022). Inflammatory disease, i.e., bovine mastitis, is one of the common diseases that caused significant economic losses for the global dairy industry (Barkema, 2015). There are two sub-type of mastitis in the dairy world, i.e., Clinical and subclinical, which manifest the degree of inflammation, and their causative factors primarily depend on bacterial pathogens (Andrews et al., 2003). Predisposing factors

and pathogen types are crucial and have a critical role in severity and outbreak of complex and various etiological nature of diseases whereas environment, host, pathogen combination and disease severity makes it more complex (De Vliegher, 2012). Mastitis is caused by teat infection through bacterial pathogens, other factors such as unhygienic milking processes, biological agents and management techniques of the herd also contributes in infection. All these situations are responsible for the reduction of milk production (Hortet et al., 1998). Recommendations are made to maintain the hygiene in crowds and vaccinate bovine with antibiotic injection, although milk production is affected by the use of injection or antibiotics (Schepers and Dijkhuizen 1991; Porter et al., 2016). Bovine mastitis is a frequently received inflammatory condition involving the mammary gland in dairy cows. It may compromise milk supply and quality while also costing more to cure (Bhattarai et al., 2018). Research regarding the molecules, i.e., genes or proteins involve in mastitis related inflammatory process might can boost the understanding of mastitis and help to handle it.

Pathogen-associated molecular patterns (PAMPs) are flagged by the family of proteins known as toll-like receptors (TLRs), and they subsequently known to trigger an immune response (Maurić et al., 2023). TLRs have been recognized to the pathophysiology of bovine mastitis. TLRs members that have been most extensively looked over in bovine mastitis are TLR2 and TLR4. While TLR4 acknowledges lipopolysaccharides (LPS) from gram-negative bacteria, TLR2 recognizes lipoproteins and peptidoglycans from gram-positive bacteria (de Mesquita et al., 2012). Activation of TLR2 and TLR4 promotes the synthesis of chemokines, which attract immune cells to the site of infection, pro-inflammatory cytokines, i.e., TNF- alpha, IL-1beta, and IL-6, are also known to get trigger by TLR2 and TLR4. However, excessive TLR activation might end up in tissue damage and excessive inflammation, which may aggravate the severity of mastitis (Akhtar et al., 2020). Involvement of TLR in mastitis make it one of the potential target, thus, targeting TLRs, for therapeutic purpose, might constitute a viable method for the management of bovine mastitis. As an instance, some research investigated the use of TLR agonists or antagonists as immunomodulators to boost or decrease the immunological response in cows with mastitis (Li et al., 2021; Orsatti et al., 2010).

Advancement of bioinformatics resources, mainly database, offering range of information gathered from different experiments and high throughput technologies (Katara, 2014; Yadav et al., 2020). Availability of cDNA microarray data on bovine mastitis offers genome wide gene expression studies which can be utilized to understand and deduce role of genes in mastitis. *In silico* approaches, mainly, gene enrichment and molecular network can further extend the possibilities to get more inside of molecular process behind the disease (Verma et al., 2020).

Considering the scenario current *in silico* study aims to categorize genes involved in inflammatory process, and identifies immune-related therapeutic targets to control the disease. Along with these, study also attempts to identify potential therapeutic herbal compounds, i.e., Aloe (*Aloe barbadensis miller*), Basil (*Ocimum basilicum*) and Neem (*Azadirachta indica*) against the predicted targets to control the mastitis (Mainardi et al., 2008). To fulfill the objectives, study considered cDNA microarray data and utilized bioinformatics tools for data analysis, network construction, and molecular docking purpose.

2 Materials and Methodology

2.1 Datasets and resources

cDNA microarray based gene expression profile data for bovine mastitis infected with *E. coli* were downloaded from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) with accession number GSE15020, GSE24217, GSE50685 (platform: GPL2112 Affymetrix Bovine Genome Array).

GEO Raw expression data files and the annotation files were downloaded and background correction and quantile normalization were used for preprocessing and normalization of the gene expression data of all the

samples. Linear models for microarray data (LIMMA) package of Bioconductor (https://www.bioconductor.org/packages/release/bioc/html/limma.html) was used for probe summarization (Ritchie et al., 2015). LIMMA package was also used to identify differentially expressed genes (DEGs) between infected and healthy bovine mammary gland's sample at 24 hour post infection with *E. coli* from three different GSEs under study.

2.2 Comparison analysis of selected DEGs

The expression of selected DEGs belongs to GSE15020, GSE24217, and GSE50685 at post infection with *E. coli*, were used to identify overlapping DEGs (ODEGs). Venn diagrams, developed through Venn diagram package in R-Studio, were utilized to observe overlapping genes among the GSE.

2.3 Regulatory and PPI network analysis

To construct the regulatory network, reported miRNA and TFs against the predicted DEGs, for their regulatory role, are screened out from available databases. The miRTarbase database was used to search miRNA, miRNA obtained from database searching were compared with those reported in the database of cattle candidate gene and genetic marker for milk production and mastitis. Transcription factors against the genes were collected from 'Regulators' database (Wang and Nishida, 2015). Predicted miRNA and TF were integrated with DEGs through Cytoscape 3.2.0 (http://www.cytoscape.org/) to construct regulatory network, i.e., miRNA-DEG-TF regulatory network.

Protein-protein interactions were predicted using the relationship between predicted ODEGs by setting the cutoff threshold as gene pair PPI score > 0.4. PPIN among the ODEGs proteins was developed through STRING, v11.2 and visualization and network statistics analysis was performed through Cytoscape v3.2.2.

2.4 Enrichment analysis for the ODEGs

To search the functional enrichment of ODEGs and biological pathways Database for Annotation, Visualisation and Integrated Discovery (DAVID) was utilized. Adjusted p-value < 0.05 and gene count >= 2 were set as cutoff criteria. Henceforth ggplot2 in 'R' was used to visualize a diagram as histogram for the enrichment term GO and KEGG along with their gene counts and p-value (Yang and Zhang, 2022).

2.5 Identification of immunogenic key genes for bovine mastitis

Plugin of Cytoscape, CytoHubba were used to rank nodes by using network features i.e., MNC, MCC, EPC, EcCentricity and degree (Zhang, 2023). These network features were made five network topologies and these five network topologies were combined to generate key immunogenic genes by Venn diagram.

2.6 Literature mining for experimentally (GCMS) characterized bioactive compounds

To list out experimentally characterized bioactive compounds, literature mining has been done from PubMed database. Screening has been done for bioactive compounds belonging to three well established medicinal plants, i.e., Aloe, Basil and Neem. In total, 147 compounds from Aloe, 91 compounds from Basil, and 191 compounds from Neem, were screened out. Further to perform structure based analysis, 3-D structures of all listed compounds were downloaded from PubChem database.

2.7 Screening of anti inflammatory compounds from Neem, Basil and Aloe through molecular docking

To screen-out bioactive anti-inflammatory compounds from all three considered plants against identified inflammatory causative targets, i.e., toll-like receptor-2 (PDBID: 3RG1) and bovine granulocyte colony stimulating factor (PDBID: 1BGC). To find out the interaction affinity between bioactive compounds and target protein molecular docking was performed through autodock vina (Hoppstädter et al., 2019).

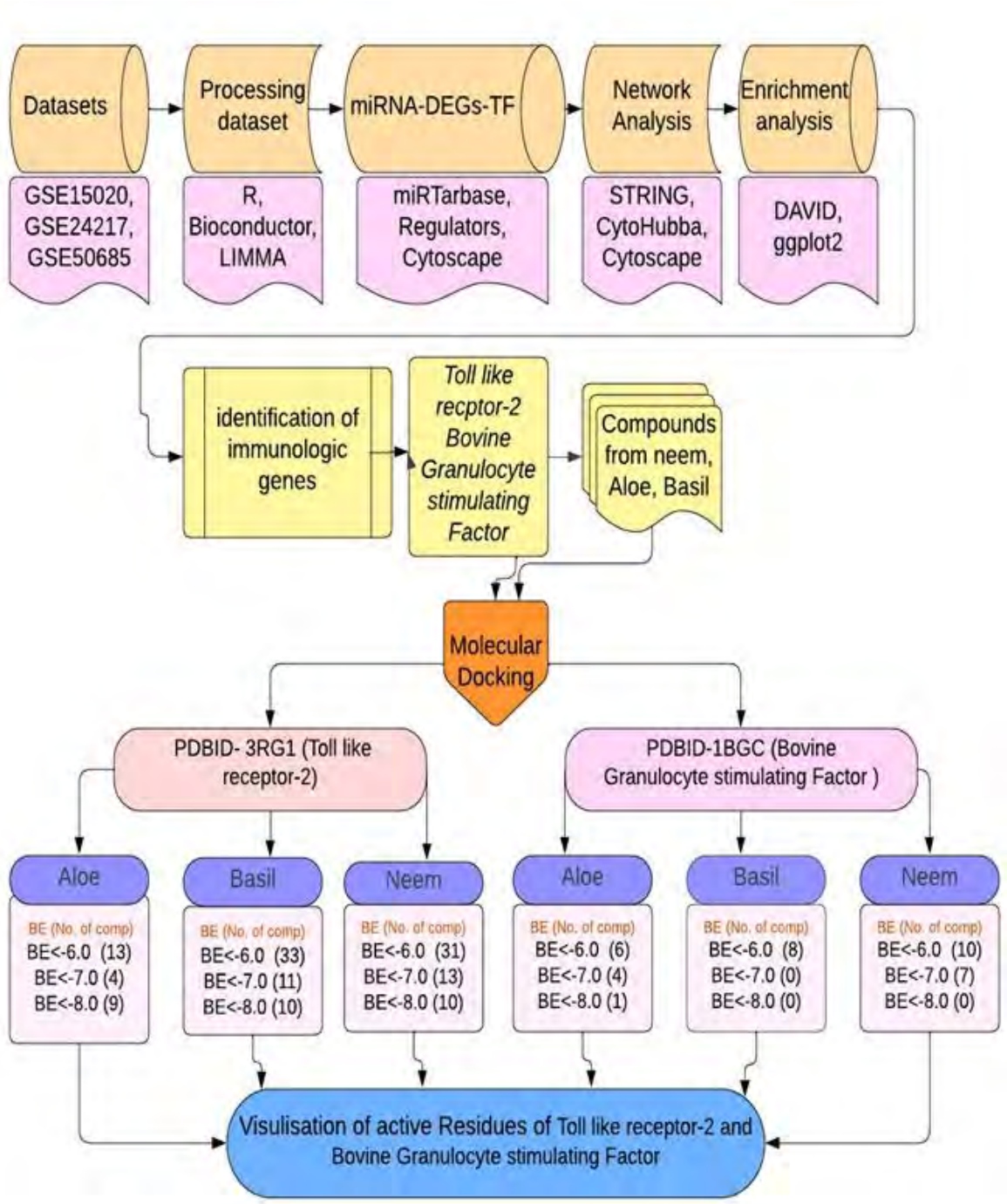


Fig. 1 Graphical summary of the work.

3 Results

3.1 Identification of differentially expressed genes (DEGs)

Genome wide gene expression analysis summarizes the involvement of genes in considered conditions. To find out list of considerable genes involved in *E. coli* infected mastitis, attempt made to identify differentially expressed genes across all three data sets. To minimize false positive gene selections, log2 and FC based selection parameters were considered for DEG identification, i.e., |log2 fold change (FC)| > 1 and FDR 0.05.

Using the cut-off criterion, in total 1177, 615, and 470 DEGs have been found from three datasets, i.e., GSE15020, GSE24217 and GSE50685, respectively (Fig. 2). The DEG expressions from the three dataset GSEs were identified 247 intersect genes among three *E. coli* infected mastitis datasets. Additionally, the 247 DEGs had similar expression patterns across the datasets (Fig. 3).

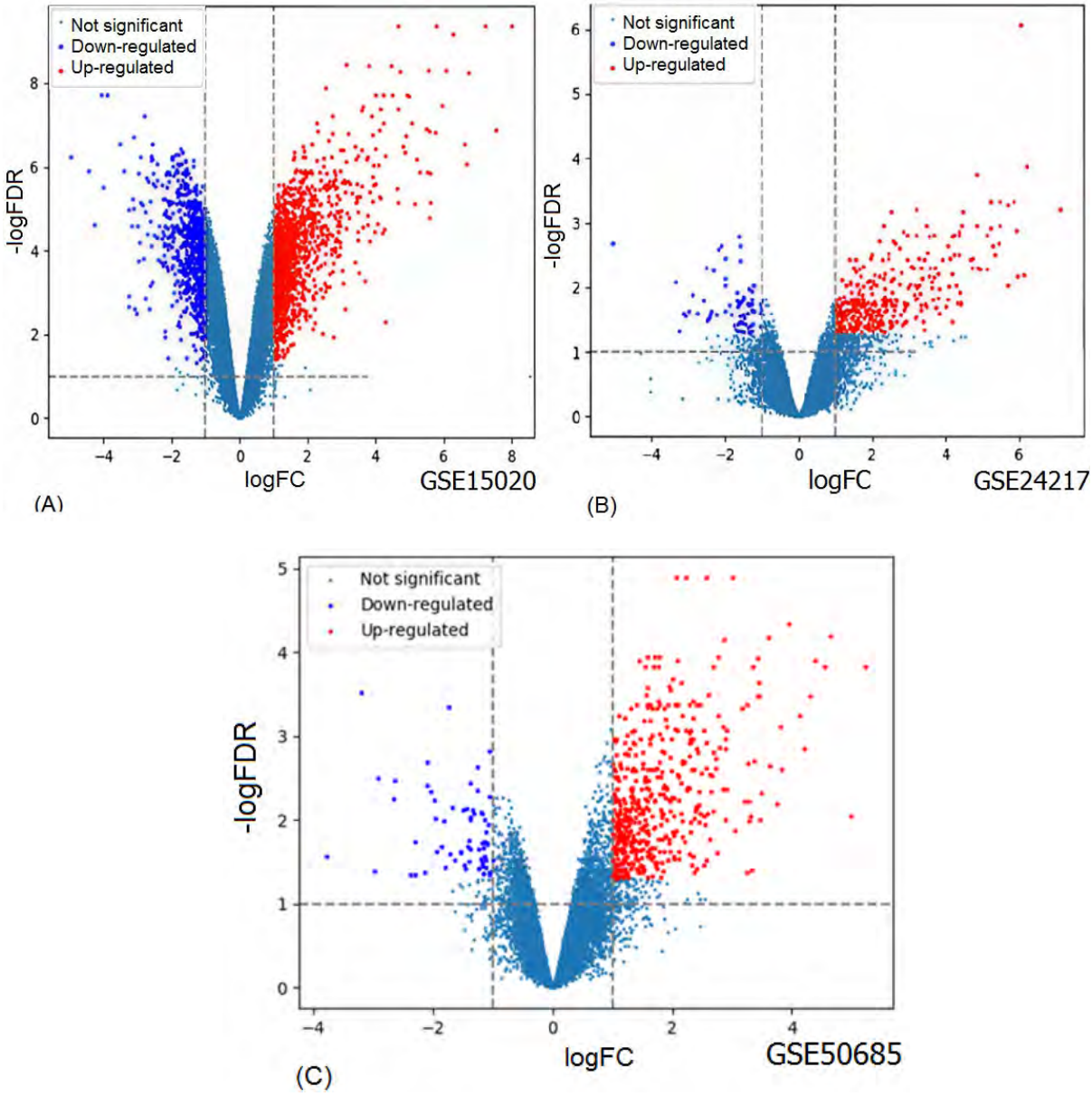


Fig. 2 Volcano plot showing expression patterns of the genes. Genes with |log2 fold change (FC)| > 1 and FDR 0.05) parameter were considered as DEGs. (A) Volcano plot of GSE15020. (B) Volcano plot of GSE24217. (C) Volcano plot of GSE50685.

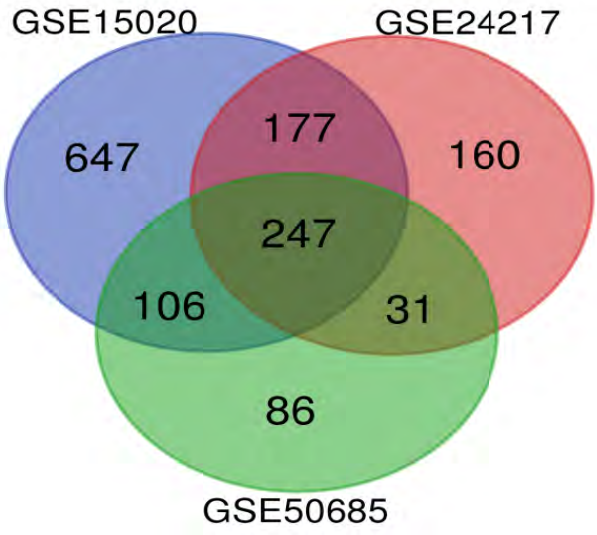


Fig. 3 Distribution of DEGs across three mastitis data sets.

3.2 Regulatory network analysis

Regulatory network provides connectivity between regulatory elements and related targets. MicroRNA and TFs were screened out for the predicted 247 mastitis related genes. Naturally, these genes are controlled by their miRNAs through post transcriptional gene regulation process, which acts on the basis of sequence complementarily. Transcription factors are also involved in gene expression regulation process during the transcription initiation event. To find out the involvement of miRNAs and TFs factors and their interaction with multi genes, gene regulatory network was developed (Fig. 4).

In the resulting regulatory network, 50 regulatory links have been observed among mi-RNA, transcription factors (TFs) and overlapping differentially expressed targeted genes and considered this association network as 'miRNA-DEG-TF' regulatory network. The 'miRNA-DEG-TF' regulatory network contained 48 mi-RNA-DEGs pairs and 4 TF-DEGs pairs. Transcription factor has been found from REGULATOR database (Table s1). Regulatory links have been established from miRTarBase database (Table s2). Multiple numbers of miRNAs demonstrate significant interactions with specific gene targets. As demonstrated in the network, 7 genes, i.e., GPAM, CD69, SOCS3, VCAM1, PLAU, CXCL2 and SRGN are linked to the miRNA bta-miR-181a (Fig. 4). bta-miR-17-5p is connected with six genes, i.e., GPAM, MPZL2, CASP4, CD69, CXCL5 and FGL2. Micro RNA bta-miR-16a is associated with 5 genes, i.e., FGL2, FASN, IL6, IL10RA, and CD69. Micro RNA bta-miR-15b is associated with 5 genes, i.e., IL10RA, IL6, CD69 FGL2, and FASN. The miRNA bta-miR-21-5 has connections to CTSC, CD69, and FGL2. Three genes, i.e., FGL2, F13A1, and BIRC3 are connected with miRNA bta-miR-155. These gene-miRNA connections shed light on putative regulatory linkages and the molecular underpinnings of many biological processes in bovine.

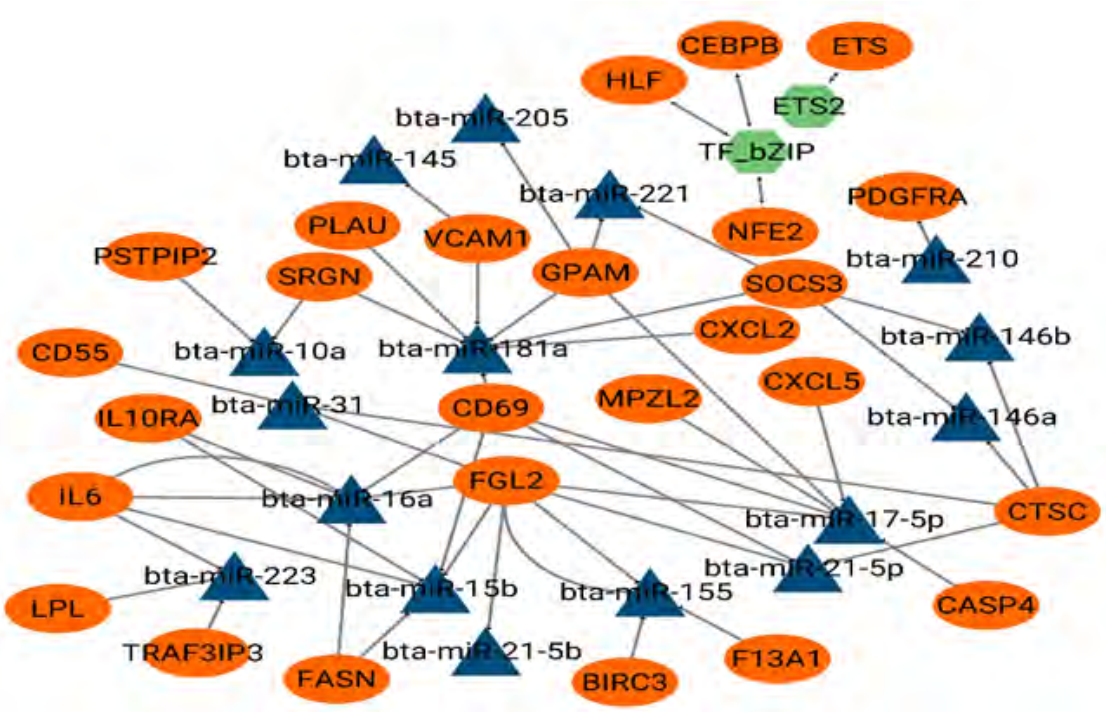
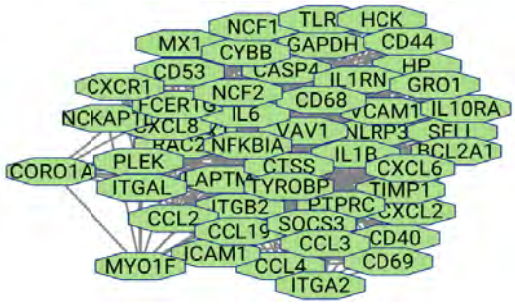


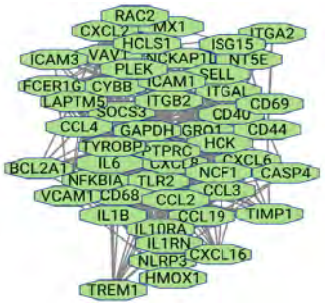
Fig. 4 The "miRNA-DEG-TF" regulatory network consists of many elements, each of which is represented by a unique set of nodes and edges. DEGs, miRNAs, and TFs are represented in this network as orange (oval), blue (triangle), and green (hexagonal) nodes, respectively. Black dotted arrows point out the TF-DEG regulatory associations, whereas black solid arrows reveal the miRNA-DEG regulatory links.

3.3 PPI network analysis ODEGs

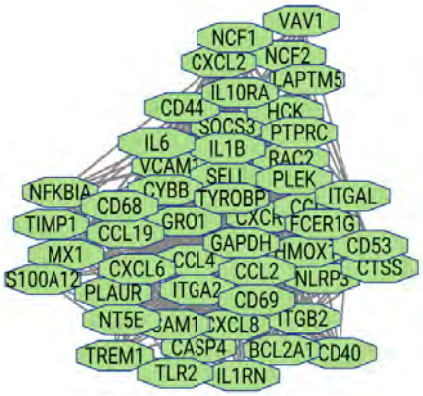
Proteins are the functional molecules of the cell and are responsible for almost all biological functions. Interaction among the proteins is one of the ways to pass the signal and act in a systemic manner. Protein-protein interaction (PPI) network provides a graphical view of interaction among the considered proteins on the basis of reported experiment and published literature. In this study, proteins of 243 predicted overlapping ODEGs was analyzed using the STRING to examine protein-protein (PPI) interactions between them. The network statistics revealed 243 nodes and 2335 significant edges (connections) with combined score greater than 0.4 and PPI enrichment p-value < 1.0e-16. A full PPI network with a total of 220 upregulated and 27 down regulated ODEGs were then built using these interactions**.** For detail analysis of protein interaction through network, STRING PPIN was imported into Cytoscape in which there are 205 nodes and their 2326 edges are observed after filtering the isolated single nodes and would go for further analysis by using CytoHubba plugin (Chin et al., 2014). Through cytoHubba, on the basis of interaction strength, top 50 ranked genes were identified. Further to avoid any false positive predictions, interaction network between these 50 genes were reconstructed through 5 different topological analysis methods (Fig. 5), i.e., i) Degree, ii) Edge Percolated Component, iii) Maximum Neighborhood Component, iv) Maximal Clique Centrality, and v) EcCentricity, based on shortest paths. Observation of all five reconstructed network resulted in 25 commonly shared genes, i.e., CYBB, LAPTM5, HCK, CCL2, VCAM1, RAC2, HMOX1, CD69, ITGB2, GAPDH, IL6, ITGAL, TLR2, VAV1, TYROBP, SELL, NCF1, CXCL8, CASP4, CD68, IL1B, ICAM1, PLEK, NLRP3, and PTPRC.



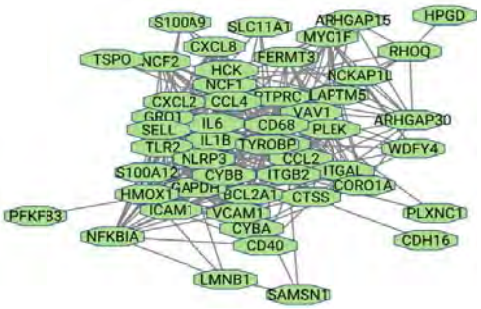
A) Maximum Neighborhood Component (MNC)



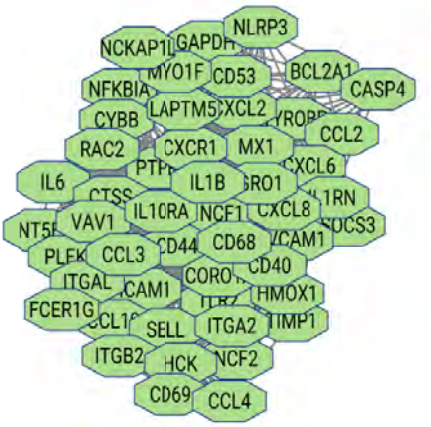
B) Maximal Clique Centrality (MCC)



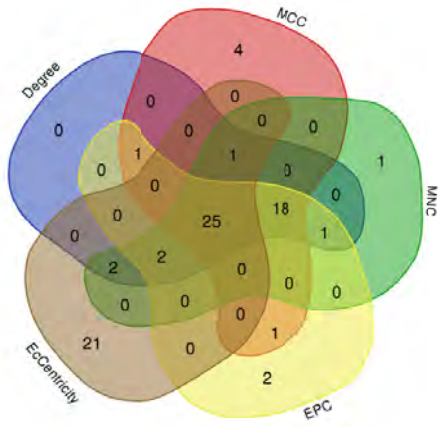
C) Edge Percolated Component (EPC)



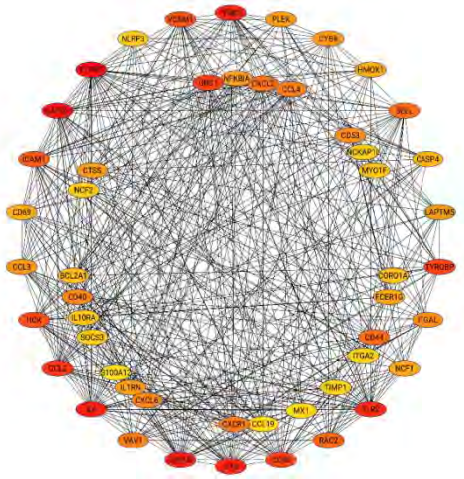
D) EcCentricity



(E) Degree



(F) 25 common genes among considering 5 topology analysis method



(G) Outer ring depicted 25 common top ranked genes.

Fig. 5 Network topologies of Top 50 ranked genes generated by considering five topological analysis methods from CytoHubba plugin. **(A)** MNC network has 50 nodes and 527 edges, **(B)** MCC network has 50 nodes interacted with 521 edges, **(C)** EPC network constructed by 50 nodes with their 535 edges, **(D)** EcCentricity network has also top 50 nodes linked with 370 edges, **(E)** Degree network also has 50 nodes interacted with 530 edges **(F)** 25 overlapping genes identified by comparing genes of five network topologies method include MNC, MCC, EPC, EcCentricity, Degree. **(G)** Most important DEG PPI network complex related to the *E. coli* mastitis module. The majority of the 530 linkages and 50 nodes in this module are connected by immunological responses.

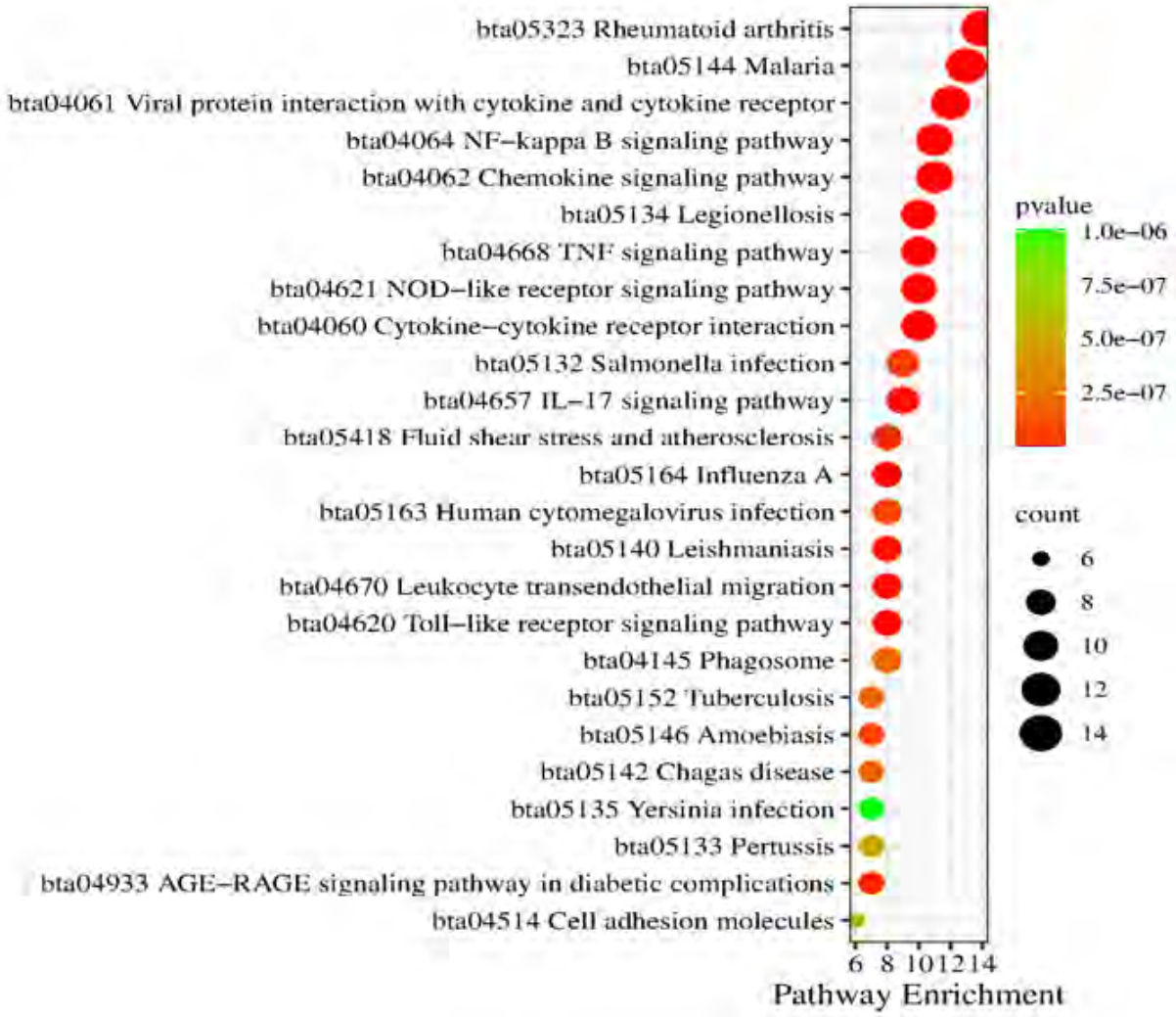
3.4 Pathway and functional enrichment analysis

To understand the involvement of genes in biological process gene enrichment was carried out. Biological processes for the DEGs stated Gene Ontology (GO) was identified (Figure 6 (A)). Through ontology analysis, it has been observed that 25 genes were engaged in the immunological response, 24 genes in the defensive response, and 21 genes in the inflammatory response among these activities. Fifteen more genes were also discovered to express themselves directly in response to the bacteria. Pathway enrichment for DEGs through KEGG enrichment analysis suggested that notably 10 DEGs were connected to the signaling route for NOD-like receptors, while 14 DEGs were connected to the chemokine signaling pathway (Table 1, Fig. 6(B)). The molecular processes underpinning the host's reaction to the infection are clarified by these discoveries, which also offer important insights into the signaling pathways that are essential.

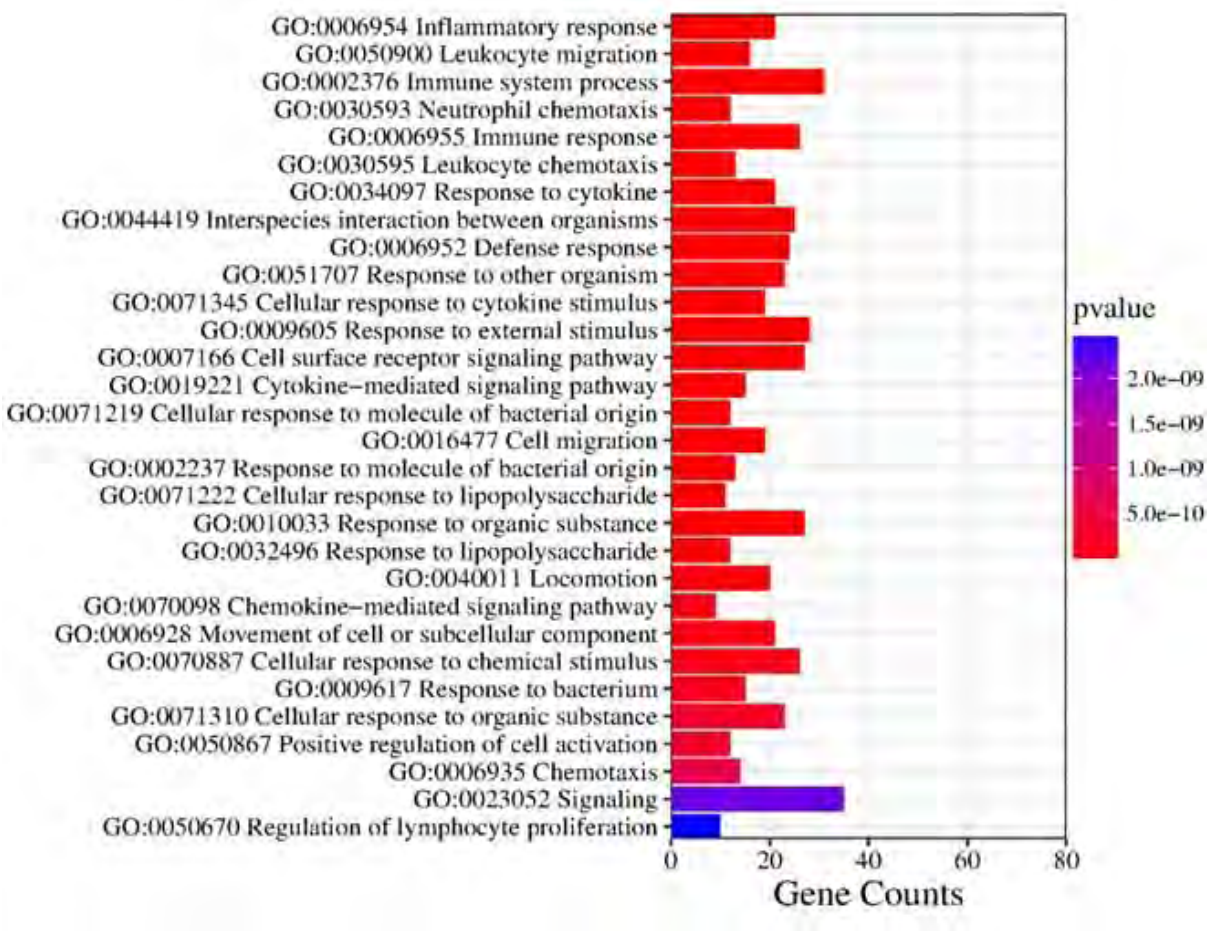
Table 1 Gene ontology and KEGG Pathway analysis for genes of selected modules.

A.	Biological process	DEGs
1	Immune system response	NLRP3, IL1B, VAV1, ITGAL, CCL4, TYROBP, TLR2, PLEK, CXCL6, CCL2, PTPRC, CCL19, IL6, CTSS, IL1RN, CXCL8, CYBB, CD40, CASP4, FCER1G, CCL3, CXCL2, MX1, VCAM1, SELL, NCKAP1L, HCK, CORO1A, GRO1, CXCR1
2	Inflammatory response	NLRP3, IL1B, CCL4, TYROBP, TLR2, CXCL6, CCL2, CCL19, IL6, IL1RN, CXCL8, CYBB, CD40, CASP4, FCER1G, CCL3, CXCL2, CD44, HCK, GRO1

3	Defense response	NLRP3, IL1B, CCL4, TYROBP, TLR2, CXCL6, CCL2, PTPRC, CCL19, IL6, IL1RN, CXCL8, CYBB, CD40, CASP4, FCER1G, CCL3, CXCL2, MX1, CD44, HCK, CORO1A, GRO1
B.	KEGG Pathways	DEGs
1	NOD-like receptor signaling pathway	NLRP3, IL1B, CCL2, IL6, NFKBIA, CXCL8, CYBB, CASP4, CXCL2, GRO1
2	Chemokine signaling pathway	NCF1, CCL4, CXCL6, CCL2, RAC2, CCL19, NFKBIA, CXCL16, CXCL8, CCL3, CXCL2, HCK, GRO1, CXCR1



(A) KEGG Pathway



(B) Functional Enrichment

Fig. 6 enrichment analysis of degs **(a)** investigation of the signaling pathways of degs linked to mastitis caused by *E. coli.* kyoto encyclopedia of genes and genome pathway enrichment of degs was found using the online tool david with gene count and adjusted p values. **(b)** investigation of the gene ontology of degs linked to mastitis caused by *E. coli*. biological process (bp) in functional enrichment of degs was performed by webtool david with gene count and adjusted p values.

3.5 Docking based Screening of bioactive compounds (Aloe, Basil and Neem) against TLR and BGCS

In this study, we examined the bioactive components of Aloe (147), Basil (91), and Neem (191), collected from published literature. Investigations were carried out to determine how well these compounds bound to the toll-like receptor-2 (PDB ID: 3RG1) and the bovine granulocyte-colony stimulating factor (PDB ID: 1BGC) proteins. Fig. 7 showed the outcomes of the docking assessments. Based on their affinity for toll-like receptor-2 (3RG1), compounds from aloe, basil, and neem were depicted by green, yellow, and blue bars, respectively. Energy thresholds were denoted by the color codes green (-6.0 kcal/mol), yellow (-7.0 kcal/mol), and blue (-8.0 kcal/mol). 13, 4, and 9 Aloe compounds for toll-like receptor-2 (3RG1) showed binding energies under -6.0 kcal/mol, -7.0 kcal/mol, and -8.0 kcal/mol, respectively. Similar results were found for basil and neem, where 2, 33, 11, and 10, 13 and 31 compounds, respectively, showed binding energies below the corresponding thresholds.

Bovine granulocyte-colony stimulating factor (1BGC) binding energies for the Aloe's 6, 4, and 1 compounds were respectively below -6.0 kcal/mol, -7.0 kcal/mol, and -8.0 kcal/mol. Only 8 compounds from basil were observed with ranges below -6.0 kcal/mol and 10, 7 and 0 compounds from Neem, revealed binding energies

ranges < -6.0 kcal/mol, -7.0 kcal/mol, and -8.0 kcal/mol for each plant, respectively. Fig. 8 exclusively showcases the details and binding affinities of compounds from the considered herbs against TLR2 and the docking results are represented in the form of minimum binding energy values ranging below -8.0 kcal/mol. Nine compounds with below -7.0 kcal/mol, affinity against the 1BGC stimulating factor protein are shown in Fig. 9.

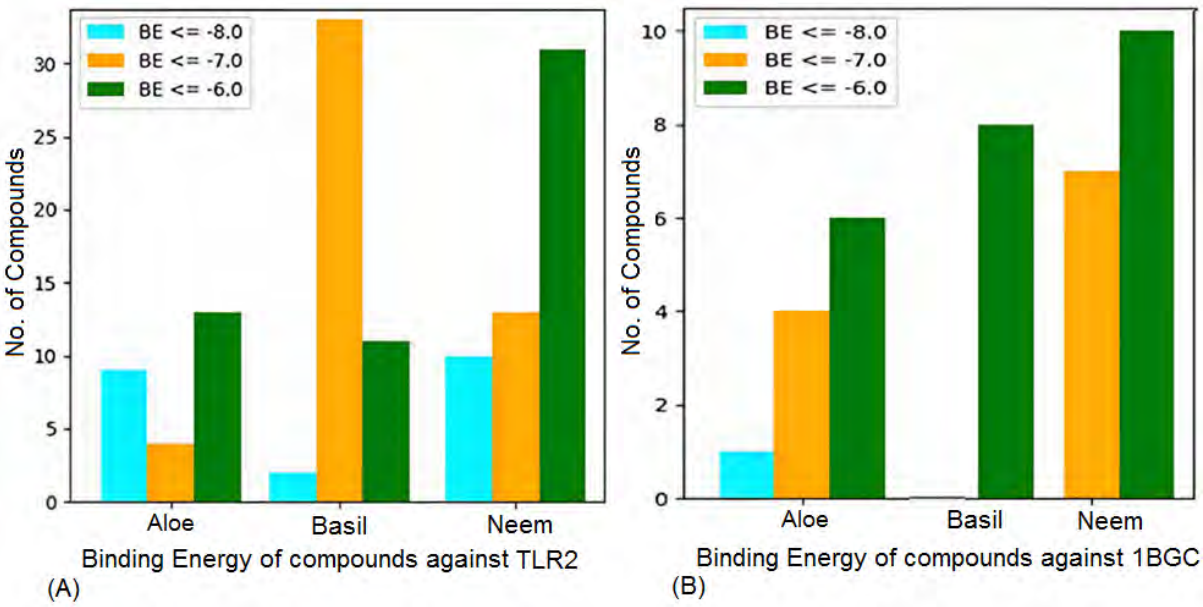


Fig. 7 The bar graph representing the number of compounds for Aloe, Basil, and Neem, while the color of each bar indicates different ranges of binding affinity i.e., -8.0 kcal/mol, -7.0 kcal/mol and -6.0 kcal/mol. The cyan color bar represents the number of compounds with a binding energy of less than or equal to -8 kcal/mol, the orange color bar represents the binding energy is less than or equal to -7 kcal/mol for the number of compounds, and the green color bar represents the number of compounds with a binding energy of less than or equal to -6 kcal/mol.

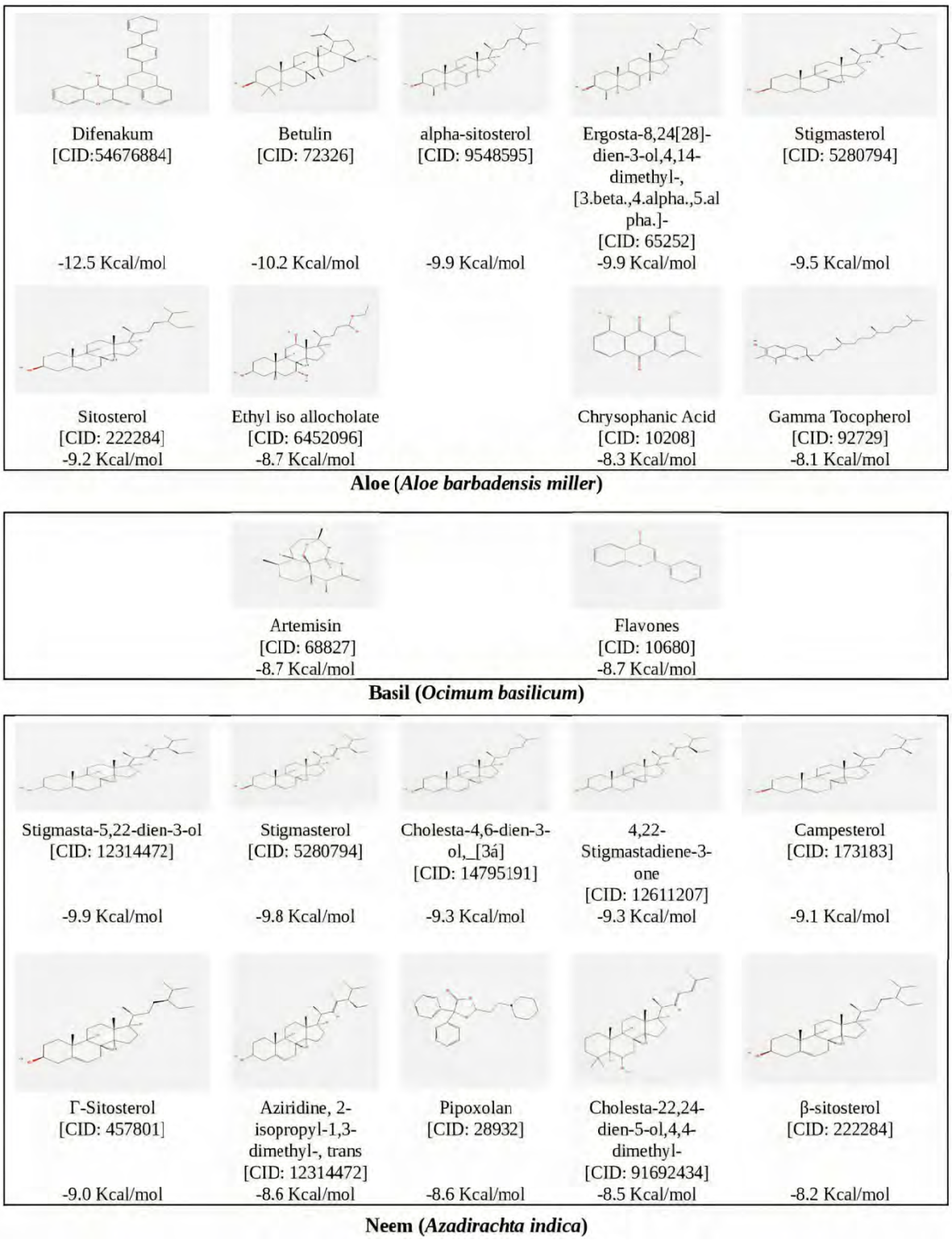


Fig. 8 Herbal compounds (A-Aloe, B-Basil, and C-Neem) with good binding affinity towards Toll-Like Receptors 2 (TLRs) 3RG1.

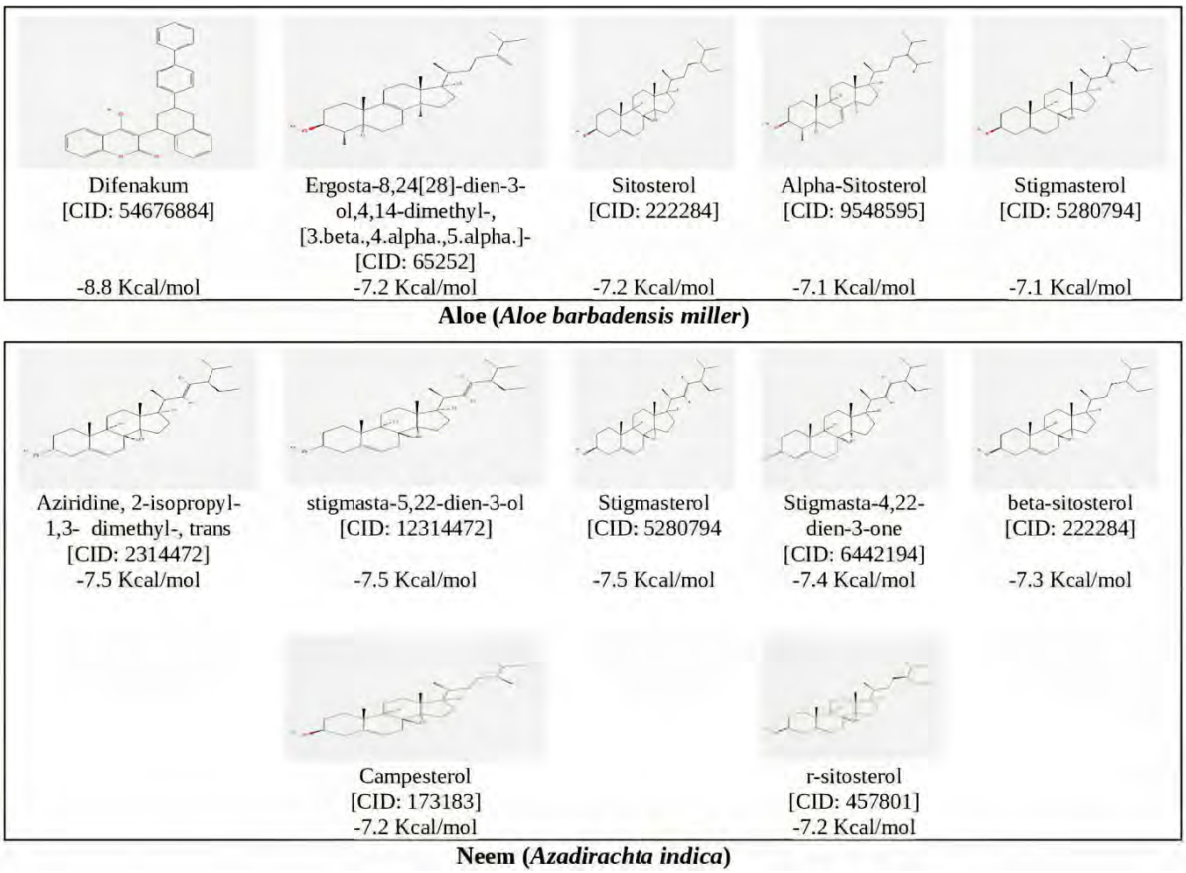


Fig. 9 Herbal compounds with good affinity towards Granulocyte-Colony Stimulating Factor (G-CSF) 1BGC.

All mentioned ligands in Figs 8-9 revealed that the binding pattern varied with the nature of the ligands. But here we studied the highest binding affinity between toll-like receptor-2 and ligands, i.e., Difenakum (Aloe), Flavone (basil), and stigmasterol (Neem). The docking results of bioactive compounds are shown in Fig. 10. The same analysis for Bovine granulocyte-colony stimulating factor (1BGC) was observed for ligands including Difenakun (Aloe) and Aziridine,2-isopropyl-1,3- dimethyl-,trans (Neem), depicted by Fig. 11.

3.5.1 Toll-like receptor-2 protein-ligand docking analysis

Protein-ligand docking analysis showed that Difenakum had an adequate binding affinity towards TLR2. There are 2 H-bonds observed between TLR2 and Difenakum (Aloe); one bond with PHE-349 and another bond with LEU-350, one pi stacking bond with PHE-322, pi cation interaction with LYS-347 and 8 hydrophobic bonds were observed with single bond include ILE-319, LUE-328, VAL-348, PHE-349, LEU-350, PRO-352, and two bonds with PHE-322, LYS-347.

In the analysis of Flavon (basil) and TLR2 docking, active residues were found to form several interactions, including hydrogen bonds, hydrophobic interactions, and pi-stacking. Specifically, two hydrogen bonds were observed with residues PHE-349 and LUE-350, while seven hydrophobic bonds formed single bonds with PHE-322, PHE-325, LYS-347, and two with ILE-319 and PHE-349. Additionally, one pi-stacking interaction was identified with PHE-322.

The investigation of the docking results between Stigmasterol from Neem and TLR2 revealed interactions involving active residues. These interactions included hydrogen bonds and hydrophobic interactions. Notably, LUE-350 formed single hydrogen bonds, while PHE-261, PHE-266, PHE-273, PHE-284, LEU-306, LEU-312,

ILE-314, PHE-325, LEU-334, LUE-328 and VAL-348 active residues bind with hydrophobically in Fig. 10.

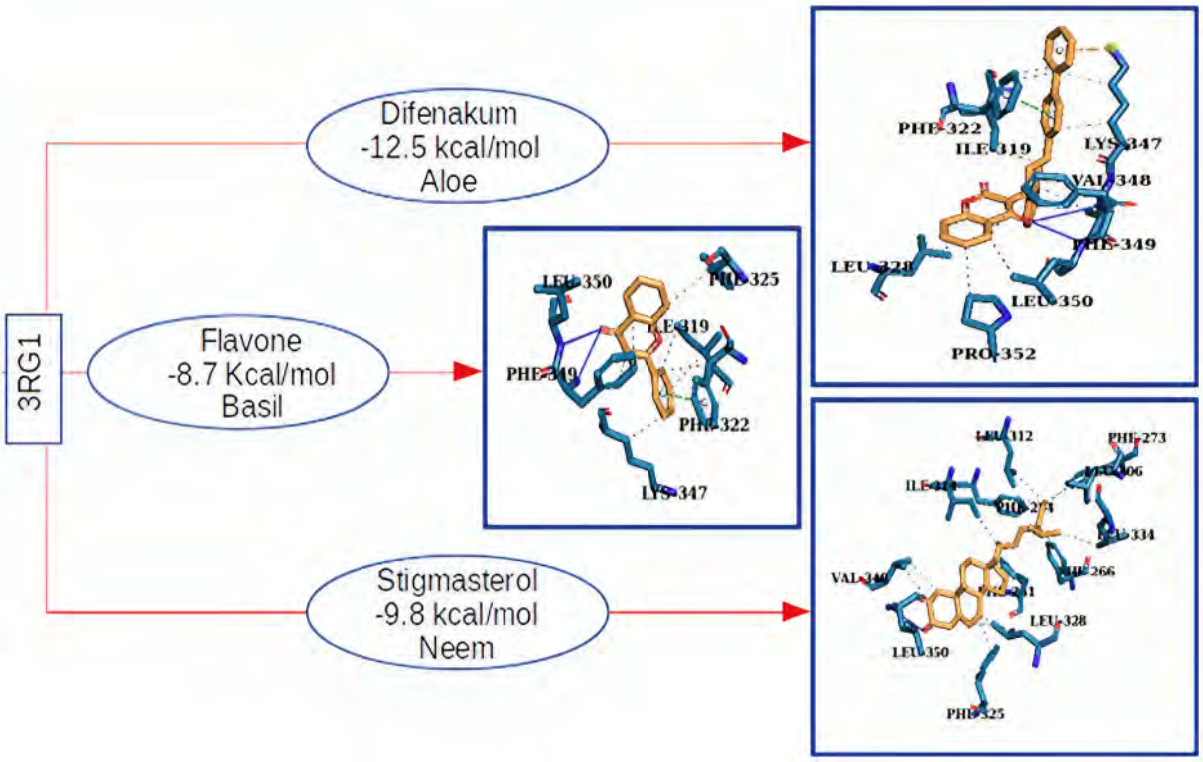


Fig. 10 Binding pocket's residues of Toll-like receptor-2 (3RG1) protein interacted with Difenakum (Aloe), Flavone (Basil), and Stigmasterol (Neem).

3.5.2 Bovine granulocyte-colony stimulating factor protein-ligand docking analysis

Bovine granulocyte-colony stimulating factor (G-CSF) is a cytokine that promotes the growth and differentiation of neutrophils. It is also known to have anti-inflammatory effects. The protein-ligand docking analysis identified that Difenakum from Aloe and Aziridine, 2-isopropyl-1,3-dimethyl-,trans from Neem, have an efficient affinity for binding to Bovine granulocyte-colony stimulating factor (PDBID: 1BGC). Difenakum from Aloe, binds with the active site residues: LEU-19, VAL-22, ARG-23, GLU-163, TYR-166, TYR-171 (Hydrophobic bonds), TYR-171 (pi-stacking)) while Aziridine,2-isopropyl-1,3-dimethyl-,trans from Neem, bind with active residues: LEU-19, GLU-163, TYR-166, ARG-170, TYR-171(Hydrophobic interations) (Fig. 11).

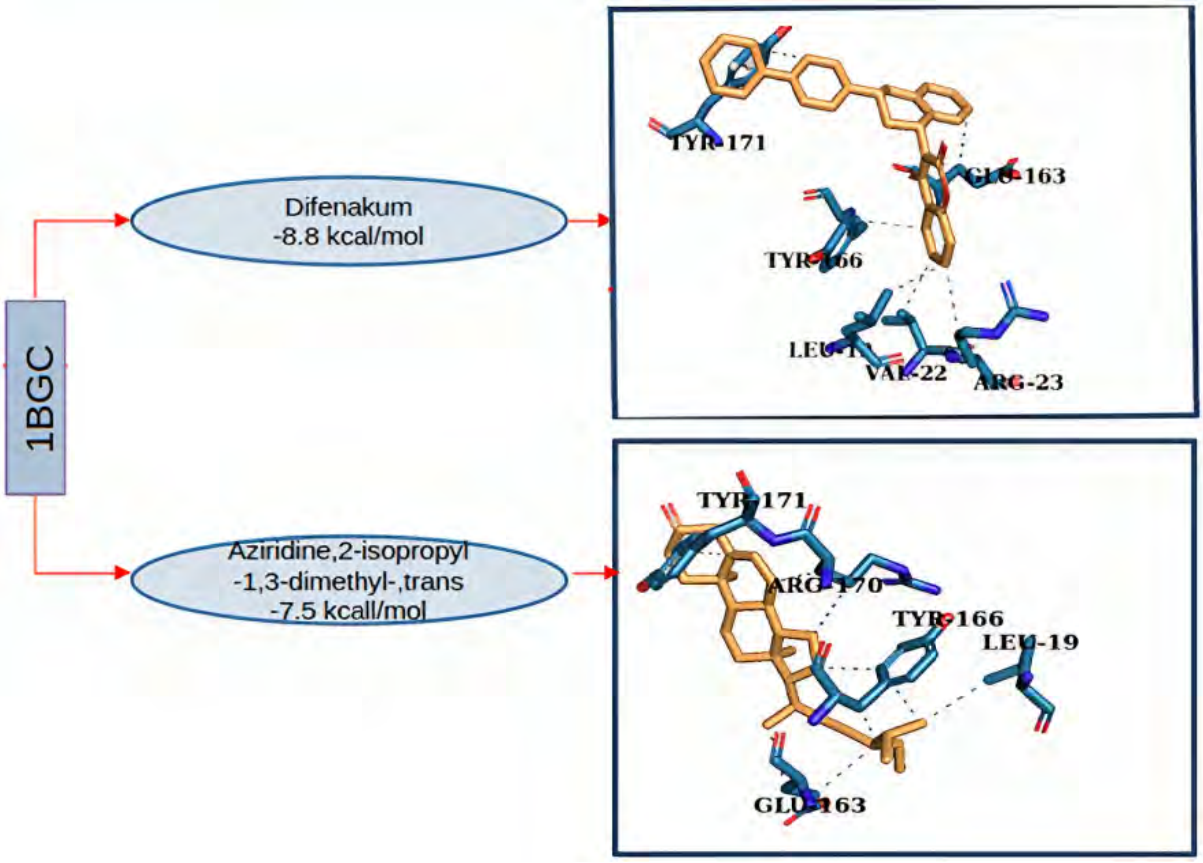


Fig. 11 Binding pocket residues of 1BGC interacted with Difenakum (Aloe) and Aziridine, 2-isopropyl-1,3-dimethyl-, trans (Neem).

4 Discussion

Mastitis in dairy cows carries significant economic importance due to its detrimental impact on the dairy industry, affecting the quality and composition of dairy products. The study identified crucial genes related to differentially expressed genes (DEGs) with a, particularly found to be enriched in the immune response, suggesting the development of an immune response against *E. coli* infection. Earlier Buitenhuis et al. (2011) also reported an increase in the expression of upregulated genes related to the immune response biological process during *E. coli*-induced mastitis. Furthermore, it is well reported that DEGs between udder quarters infected with *E. coli* and adjacent tissue quarters, along with control animals, were predominantly associated with immune response pathways (Mitterhuemer et al., 2010; Darang 2023). In both humans and animals, microRNAs (miRNAs) play a critical role in regulating inborn immunity and diverse immunity (Ahmed et al., 2017).

The miRnaTarBase database identified 541 miRNAs in total for bovine. Upon analyzing miRNA expression patterns, it was observed that 279 miRNAs and 305 miRNAs exhibited differential expression in tissues infected with *Staphylococcus aureus* and *E. coli*, respectively, when compared to control samples (Luoreng et al., 2018). In a study carried out by Santos et al., 2019, a recently identified medication therapy for infectious bovine mastitis was investigated using an approach composed of transcriptomics and bioinformatics analysis. The study demonstrated the possible efficacy of the drug treatment by targeting specific mechanisms of bovine mammary genes, including MTOR and TP53. However, 9 miRNAs were related to bovine mastitis inflammation from the databases of cattle candidate genes and genetic markers for milk production and mastitis. These miRNAs

included bta-miR-146a, bta-miR-146b, bta-miR-21-5p, bta-miR-31, bta-miR-16a, bta-miR-15b, bta-miR-181a and bta-miR-223. There were 6 DEG that were targeted by the bovine mastitis inflammation-related miRNAs. CTSC was regulated by bta-miR-146a, bta-miR-146b, bta-miR-21-5p, bta-miR-31, as well as was IL6 modulated by bta-miR-16a, bta-miR-223, bta-miR-15b. Also, bta-miR-16a and bta-miR-15b, were regulated to FASN, and CASP4, CXCL2, CXCL5 were regulated by bta-miR-17-5p, bta-miR-181a and bta-miR-17-5p respectively. In the current study, we implemented bioinformatic analysis on three microarray profile datasets to find significant key genes involved in the response to mastitis caused by bovine *E. coli*. Exclusively choosing reported microarray data from in vivo investigations with live *E. coli* infection was a unique strategy used. This made it possible for us to concentrate on certain genes and pathways linked to *E. coli* mastitis. The main inferences from previous individual study findings, such as the encouragement of immune response, inflammation, and the TNF signaling pathway, as reported by Buitenhuis et al., 2011 and Günther et al., 2011, were supported by the findings of our study. These results help to clarify the molecular processes underlying the immune system's reaction to *E. coli* mastitis in cattle, (Li et al., 2019). In the current research, the top 50 ranked genes were studied of which 25 common genes came from five topology networks. These 25 common genes were CYBB, LAPTM5, HCK, CCL2, VCAM1, RAC2, HMOX1, CD69, ITGB2, GAPDH, IL6, ITGAL, TLR2, VAV1, TYROBP, SELL, NCF1, CXCL8, CASP4, CD68, IL1B, ICAM1, PLEK, PTPRC, NLRP3, and remaining top ranked nodes came from Degree topology network i.e., GRO1, NFKBIA, CXCL2, CCL4, CD53, NCKAP1L, MYO1F, CORO1A, FCER1G, CD44, ITGA2, TIMP1, MX1, CCL19, CXR1, CXCL6, IL1RN, S100A1, SOCS3, IL10RA, CD40, BCL2A1, NCF2, CTSS, CCL3. This result is consistent with other studies, such as those by Sharifi et al., 2018, who also found the top three genes, i.e., CXCL8, CXCL2, and GRO1 were associated with *E. coli* mastitis. Han 2019 also identified CXCL2 and GRO1, IL6, NFKBIA as important genes in mastitis caused by *Staphylococcus aureus* and *E. coli* (Han, 2019). Nine genes, namely, CXCL8, CXCL2, IL10RA, CXCL6, GRO1, IL6, IL1B, ICAM1, and TLR2 are important genes for bovine mastitis in response to *E. coli* strongly recommended by Li et al., 2019. The fact that these essential genes are consistently found in investigations verifies the outcomes of the study we conducted. Toll-like receptor2 is a protein that is produced by the TLR2 gene. A class of proteins known as toll-like receptors participates in the immunological response of the mammary gland of bovine. Gram-negative bacteria produce a unique sort of cell wall component called LPS, which is specially recognized by TLR2. Experiment kits that use the TLR2 gene to check anti-inflammatory activity typically involve exposing cells to LPS in the presence or absence of a potential anti-inflammatory compound (An et al., 2021; Reid et al., 2019). Overall, the TLR is a useful tool for studying anti-inflammatory activity. Experiment kits that use this gene can help to identify new compounds that have the potential to treat inflammatory diseases (Hoppstädter et al., 2019). It is used to examine the anti-inflammatory properties of the substance curcumin. The findings demonstrated that curcumin prevented cells exposed to LPS from producing IL-1 and TNF. This implies that curcumin could be useful as an anti-inflammatory medication. Whereas Bovine granulocyte-colony stimulating factor (G-CSF) belongs to the cytokine family that plays an essential role in stimulating the production and maturation of granulocytes is a type of white blood cell, in cattle. It is produced by the number of immune cells, such as macrophages, fibroblasts, and endothelial cells, in response to infection or inflammation (Reid et al., 2021). In Figure 8-9, we identified anti-inflammatory compounds derived from Aloe, Basil, and Neem that exhibited strong binding affinity with TLR and Bovine granulocyte-colony stimulating factor proteins. We suggested, compounds from Aloe: Difenakum, Betulin, alpha-sitosterol, Ergosta-8,24[28]-dien-3-ol,4,14-dimethyl-,[3.beta.,4.alpha.,5.alpha.]-, Stigmasterol, Sitosterol, Ethyl iso allocholate, Chrysophanic Acid, Gamma Tocopherol; compounds from Basil: Artemisin, flavone and compounds from Neem; stigmasta-5,22-dien-3-ol, Stigmasterol, Cholesta-4,6-dien-3-ol,_[3á], 4,22-Stigmastadiene-3-one, Campesterol, Γ-Sitosterol, Aziridine,2-isopropyl-1,3-dimethyl-,trans, Pipoxolan,

Cholesta-22,24-dien-5-ol,4,4-dimethyl-, and β-sitosterol; it might be possible that they are found strongly recommended anti inflammatory compounds for Toll-like receptor-2 (TLR2) protein. While for Bovine granulocyte-colony stimulating factor protein: there are 7 and 5 compounds from Neem and Aloe found whose binding affinities stay below -7.0 Kcal/mol, the remaining docked compounds were not getting space in Fig. 8 - 9. Aziridine, 2-isopropyl-1,3-dimethyl-, trans, stigmasta-5,22-dien-3-ol, Stigmasterol, Stigmasta-4,22-dien-3-one, beta-sitosterol, campesterol, r-sitosterol from Neem while Ergosta-8,24[28]-dien-3-ol,4,14-dimethyl-,[3.beta.,4.alpha.,5.alpha.]-, Alpha-Sitosterol, Sitosterol and Stigmasterol from Aloe were suggested as potent compounds against the protein. It can be administered as a therapeutic agent to enhance the body's defense mechanisms and aid in the recovery process.

5 Conclusion

In this research, considering genome wide gene expression data of *E. coli* infected mastitis in total 247 DEGs were identified and studied further for their biological role. The expression analysis and molecular network based investigation pinpointed essential key genes, i.e., CYBB, LAPTM5, HCK, CCL2, VCAM1, RAC2, HMOX1, CD69, ITGB2, GAPDH, IL6, ITGAL, TLR2, VAV1, TYROBP, SELL, NCF1, CXCL8, CASP4, CD68, IL1B, ICAM1, PLEK, and NLRP3, involved in the immune response of *E. Coli* infected bovine mastitis. Further studied explore two immune response related key protein, i.e., TLR and G-CSF as potential therapeutic targets and bioactive compounds from Aloe (*Aloe barbadensis miller*), Basil (*Ocimum basilicum*) and Neem (*Azadirachta indica*), were screened against them and found 33 compounds in total with considerable affinity against the considered target proteins. The findings from this study offer valuable molecular insights of mastitis, and offer mastitis related biomarkers which will be helpful for earlier detection of bovine mastitis. Along with that, though predicted bioactive compounds need further experimental validation, they raise the hope for the effective mastitis treatment.

Acknowledgement

SV and NK thanks to UGC for UGC-CRET fellowship during this research work.

List of Abbreviations

DAVID- Database for Annotation, Visualization and Integrated Discovery
DEGs- Differentially Expressed Genes
FC- Fold Change
FDR- False Discovery Rate
GCMS- Gas Chromatography Mass Spectrometry
GEO- Gene Expression Omnibus,
LIMMA- Linear Models for Microarray
LPS- Lipopolysaccharides
ODEGs- Overlapping Differentially Expressed Genes
PAMPs- Pathogen-Associated Molecular Patterns
PPI- Protein-Protein Interaction,
PPIN- Protein-Protein Interaction Network
TF- Transcription Factor
TLRs- Toll-Like Receptors

Supplementary Data

Table s1 Details of considered mi-RNAs known for regulation of DEGs.

DEGs Symbol	Ensembl	Family	Protein	Entrez-ID
NFE2	ENSBTAG00000001562	TF_bZIP	ENSBTAP00000059395.1; ENSBTAP00000044631.1; ENSBTAP00000002043.1;	514006
PLEK	ENSBTAG00000009658	Others	-	518658
ETS2	ENSBTAG00000009214	ETS	ENSBTAP00000072496.1; ENSBTAP00000012144.5;	281148
HLF	ENSBTAG00000006618	TF_bZIP	ENSBTAP00000063281.1; ENSBTAP00000008682.5;	516069
TCF19	ENSBTAG00000014435	Others	-	514216
CEBPB	ENSBTAG00000051972	TF_bZIP	ENSBTAP00000073202.1;	338319

Table s2 Information of considered mi-RNAs known for regulation of 19 DEGs.

S.No.	Target gene	miRTarBase ID	miRNA	Entrez ID
1	IL10RA	MIRT053818 MIRT053842	bta-miR-15b bta-miR-16a	513478
2	FGL2	MIRT053816 MIRT053840 MIRT053885 MIRT053893 MIRT053949 MIRT053853 MIRT053880 MIRT053948	bta-miR-15b bta-miR-16a bta-miR-21-5p bta-miR-31 bta-miR-155 bta-miR-17-5p bta-miR-21-5p bta-miR-155	616885 616885 616885 616885 616885 511711 511711 511711
3	PLAU	MIRT053969	bta-miR-181a	281408
4	VCAM1	MIRT053908 MIRT053974	bta-miR-145 bta-miR-181a	534578 534578
5	MPZL2	MIRT053858	bta-miR-17-5p	540423
6	TRAF3IP3	MIRT054010	bta-miR-223	505371
7	LPL	MIRT054012	bta-miR-223	280843
8	F13A1	MIRT053950	bta-miR-155	617881
9	FASN	MIRT053492 MIRT053828	bta-miR-15b bta-miR-16a	281152 616885
10	CD55	MIRT053896	bta-miR-31	518609
11	SRGN	MIRT053803 MIRT053983	bta-miR-10a bta-miR-181a	509501 509501
12	CD69	MIRT053824 MIRT053848 MIRT053865 MIRT053888 MIRT053982	bta-miR-15b bta-miR-16a bta-miR-17-5p bta-miR-21-5p bta-miR-181a	281058 281058 281058 281058 281058
13	CASP4	MIRT053861	bta-miR-17-5p	338039
14	SOCS3	MIRT053915 MIRT053931 MIRT053970	bta-miR-146a bta-miR-146b bta-miR-181a	282081 282081 282081

		MIRT054001	bta-miR-221	282081
15	IL6	MIRT053807 MIRT053833 MIRT053955 MIRT054007	bta-miR-15b bta-miR-16a bta-miR-16a bta-miR-223	517016 517016 517016 517016
16	GPAM	MIRT053876 MIRT053990 MIRT054004 MIRT054016	bta-miR-17-5p bta-miR-181a bta-miR-221 bta-miR-205	497202 497202 497202 497202
17	PSTPIP2	MIRT053800	bta-miR-10a	523223
18	BIRC3	MIRT053956	bta-miR-155	514386
19	CTSC	MIRT053890 MIRT053897 MIRT053922 MIRT053938	bta-miR-21-5p bta-miR-31 bta-miR-146a bta-miR-146b	352958 352958 352958 352958
20	CXCL2	MIRT053975	bta-miR-181a	613667
21	CXCL5	MIRT053864	bta-miR-17-5p	281735

References

Ahmed FE, Gouda MM, Hussein LA, Ahmed NC, Vos PW, Mohammad MA. 2017. Role of melt curve analysis in interpretation of nutrigenomics' microRNA expression data. Cancer Genomics Proteomics, 14: 469-481. https://doi.org/10.21873/cgp.20057

Akoh OU, Mac-Kalunta OM, Amadi OK. 2021. Antimicrobial screening, vitamin assay and gcms analysis of chloroform extract of *Azadirachta indica* (Neem) leave. Journal of Chemical Society Nigeria, 46(1). DOI:10.46602/jcsn.v46i1.574

Akpuaka A, Ekwenchi MM, Dashak DA, Dildar A. 2013. Biological activities of characterized isolates of n-hexane extract of *Azadirachta indica* A. Juss (Neem) leaves. Nature and Science, 11: 141-147

An R, Gao M, Meng Y, Tong X, Chen J, Wang J. 2021. Infective mastitis due to bovine-associated Streptococcus dysgalactiae contributes to clinical persistent presentation in a murine mastitis model. Veterinary Medicine and Science, 7: 1600-1610. https://doi.org/10.1002/vms3.509

Anderson M. 1982. Factors affecting the distribution of lipoprotein lipase activity between serum and casein micelles in bovine milk. Journal of Dairy Research, 49: 51-59. https://doi.org/10.1017/s0022029900022123

Aye A, Jeon YD, Lee JH, Bang KS, Jin JS. 2019. Anti-inflammatory activity of ethanol extract of leaf and leaf callus of basil (*Ocimum basilicum* L.) on RAW 264.7 macrophage cells. Oriental Pharmacy and Experimental Medicine, 19: 217-226. https://doi.org/10.1007/s13596-019-00372-2

Ayoola AA, Ekunseitan DA, Muhammad SB, Oguntoye MA, Adejola Y. 2020. Phytochemicals analysis and GC-MS determination of ethanolic extracts of *Azadirachta indica* and *Mangifera indica* stem bark and their

biological potentials. Pacific Journal of Science and Technology, 21: 219-229

Banakar P, Jayaraj M. 2018. GC-MS analysis of bioactive compounds from ethanolic leaf extract of *Waltheria indica* Linn. and their pharmacological activities. International Journal of Pharmaceutical Sciences and Research, 9: 2005-2010. https://doi.org/ 10.13040/IJPSR.0975-8232.9(5).2005-10

Barkema HW, von Keyserlingk MA, Kastelic JP, et al. 2015. Invited review: Changes in the dairy industry affecting dairy cattle health and welfare. Journal of Dairy Science, 98: 7426-7445. https://doi.org/10.3168/jds.2015-9377

Bele AA, Khale A. 2016. Comparison of constituents in *Aloe vera* gel collected in different seasons by chromatography and spectroscopy techniques. World Journal of Pharmaceutical Research, 5: 1028-1040. https://doi.org/ 10.20959/wjpr20169-6928

Benzidia B, Barbouchi M, Hammouch H, Belahbib N, Zouarhi M, Erramli H, Hajjaji N. 2019. Chemical composition and antioxidant activity of tannins extract from green rind of *Aloe vera* (L.) Burm. F. Journal of King Saud University-Science, 31: 1175-1181. https://doi.org/10.1016/j.jksus.2018.05.022

Buitenhuis B, Røntved CM, Edwards SM, Ingvartsen KL, Sørensen P. 2011. In depth analysis of genes and pathways of the mammary gland involved in the pathogenesis of bovine *Escherichia coli*-mastitis. BMC Genomics, 12: 130. https://doi.org/10.1186/1471-2164-12-130

Caamal-Herrera IO, Muñoz-Rodríguez D, Madera-Santana T, Azamar-Barrios JA. 2016. Identification of volatile compounds in essential oil and extracts of *Ocimum micranthum* Willd leaves using GC/MS. International Journal of Applied Research in Natural Products, 9: 31-40. https://api.semanticscholar.org/CorpusID:217677094

Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. 2014. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Systems Biology, 8. https://doi.org/10.1186/1752-0509-8-S4-S11

Darang E, Pezeshkian Z, Mirhoseini SZ, Ghovvati S. 2023. Identification of key genes and potential pathways associated with mastitis induced by *E. coli*. Biochemical Genetics, 61: 202-220. https://doi.org/10.1007/s10528-022-10254-z

De Vliegher S, Fox LK, Piepers S, McDougall S, Barkema HW. 2012. Invited review: Mastitis in dairy heifers: nature of the disease, potential impact, prevention, and control. Journal of Dairy Science, 95: 1025-1040. https://doi.org/10.3168/jds.2010-4074

Günther J, Esch K, Poschadel N, et al. 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

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

Hossain MA, Al-Toubi WA, Weli AM, Al-Riyami QA, Al-Sabahi JN. 2013. Identification and characterization of chemical compounds in different crude extracts from leaves of Omani neem. Journal of Taibah University for Science, 7(4): 181-188. https://doi.org/10.1016/j.jtusci.2013.05.003

Ibeagha-Awemu EM, Lee JW, Ibeagha AE, Bannerman DD, Paape MJ, Zhao X. 2008. Bacterial lipopolysaccharide induces increased expression of toll-like receptor (TLR) 4 and downstream TLR signaling molecules in bovine mammary epithelial cells. Veterinary Research, 39: 11. https://doi.org/10.1051/vetres:2007047

Joshi RK, Hoti SL. 2014. Chemical composition of the essential oil of *Ocimum tenuiflorum* L. (Krishna tulsi)

from North West Karnataka, India. Plant Science Today, 1: 99-102. https://doi.org/10.14719/pst.2014.1.3.52

Katara P. 2014. Potential of Bioinformatics as functional genomics tool: an overview. Network Modeling and Analysis in Health Informatics and Bioinformatics, 3(1): 1-7. doi: https://doi.org/10.1007/s13721-014-0052-3

Krishna N, Vishwakarma S, Katara P. 2022. Identification and annotation of milk-associated genes from milk. Bioinformation, 18(8): 703-709. https://doi.org/10.6026/97320630018703

Kumar R, Kumar R, Prakash O, Srivastava RM, Pant AK. 2019. GC-MS analysis of the hexane extract of *Limnophila indica* (L.) Druce, its total phenolics, in-vitro antioxidant, anti-inflammatory and antifeeding activity against *Spilosoma obliqua*. Journal of Entomology and Zoology Studies, 7:970-975

Li L, Chen X, 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

Lippolis JD, Holman DB, Brunelle BW et al. 2017. Genomic and transcriptomic analysis of *Escherichia coli* strains associated with persistent and transient bovine mastitis and the role of colanic acid. Infection and Immunity, 86: e00566-17. https://doi.org/10.1128/IAI.00566-17

Loganathan T, Barathinivas A, Soorya C, Balamurugan S, Nagajothi TG, Jayakumararaj R. 2021. GCMS profile of bioactive secondary metabolites with therapeutic potential in the ethanolic leaf extracts of *Azadirachta indica*: A sacred traditional medicinal plant of India. JDDT, 11: 119-126. https://doi.org/10.22270/jddt.v11i4-S.4967

Luoreng ZM, Wang XP, Mei CG, Zan LS. 2018. Comparison of microRNA profiles between bovine mammary glands infected with *Staphylococcus aureus* and *Escherichia coli*. International Journal of Biological Science, 14: 87-99. https://doi.org/10.7150/ijbs.22498

Mitterhuemer S, Petzl W, Krebs S, et al. 2010. *Escherichia coli* infection induces distinct local and systemic transcriptome responses in the mammary gland. BMC Genomics, 11: 138. https://doi.org/10.1186/1471-2164-11-138

Palii AP, Kovalchuk YO, Boyko YA, Bondaruk YV, Diachuk PV, Duka TM, Paliy AP. 2020. Impact of various milking equipment on incidence of mastitis in dairy herd. Ukrainian Journal of Ecology, 10: 160-165. https://doi.org/10.15421/2020_224

Rai P, Chatrath H. 2021. The comparison of gc-ms data of leaves of *Aloe vera* (*A. barbadensis* Mill) plant grown in different soil combinations with chemical lab wastes. Green Chemistry and Technology Letters, 6: 25-30. https://doi.org/10.18510/gctl.2020.623

Raina AP, Kumar A, Dutta M. 2013. Chemical characterization of aroma compounds in essential oil isolated from “Holy Basil” (*Ocimum tenuiflorum* L.) grown in India. Genetic Resources and Crop Evolution, 60: 1727-1735. https://doi.org/10.1007/s10722-013-9981-4

Ramya R. 2022. GC-MS analysis of bioactive compounds in ethanolic leaf extract of *Hellenia speciosa* (J. Koenig) S.R. Dutta. Applied Biochemistry and Biotechnology, 194: 176-186. https://doi.org/10.1007/s12010-021-03742-2

Rana L, Tewari G, Pande C, Prakash O, Bisht M, Bhatt S, Rana K. 2022. Identification of the aroma compounds of *Ocimum americanum* as a function of growth stages and their in vitro antioxidant and anti-inflammatory potential. Journal of Essential Oil Bearing Plants, 25: 403-418. https://doi.org/10.1080/0972060X.2022.2069479

Reid C, Beynon C, Kennedy E, O'Farrelly C, Meade KG. 2021. Bovine innate immune phenotyping via a standardized whole blood stimulation assay. Scientific Reports, 11: 17227. https://doi.org/10.1038/s41598-021-96493-3

Ritchie ME, Phipson B, Wu D, et al. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research, 43: e47. https://doi.org/10.1093/nar/gkv007

Santos EMS, Almeida AC, Santos HO, Cangussu AR, Costa KS, Alves JN, Aguiar RWS. 2019. Mechanism of *Brassica oleracea* performance in bovine infectious mastitis by bioinformatic analysis. Microbial Pathogenesis, 129: 19-29. https://doi.org/10.1016/j.micpath.2019.01.029.

Selvaraju R, Sakuntala P, Jaleeli KA. 2021. Comparative analysis of chemical compounds and trace elements in the leaves of few Ocimum species plants by using different spectroscopic methods. Romania Journal of Biophysics, 31: 65-78

Sharifi S, Pakdel A, Ebrahimi M, Reecy JM, Fazeli Farsani S, 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

Tanamati F, Stafuzza NB, Gimenez DFJ, et al. 2019. Differential expression of immune response genes associated with subclinical mastitis in dairy buffaloes. Animal, 13: 1651-1657. https://doi.org/10.1017/S1751731118003324

Trott O, Olson AJ. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31: 455-461. https://doi.org/10.1002/jcc.21334

Verma Y, Yadav A, Katara P. 2020. Mining of cancer core-genes and their protein interactome using expression profiling based PPI network approach. Gene Reports, 18: 100583, https://doi.org/10.1016/j.genrep.2019.100583

Wang K, Nishida H. 2015. REGULATOR: a database of metazoan transcription factors and maternal factors for developmental studies. BMC Bioinformatics, 16: 114. doi: 10.1186/s12859-015-0552-x

Wang R, Lai L, Wang S. 2002. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. Journal of Computer-Aided Molecular Design, 16: 11-26. https://doi.org/10.1023/a:1016357811882

Yadav A, Vishwakarma S, Krishna N and Katara P. 2020. Integrative Omics: Current Status and Future Directions. Recent Trends in Computational Omics: Concepts and Methodology. Nova Science Publishers, USA

Yang S, Zhang WJ. 2022. Systematic analysis of olfactory protein-protein interactions network of fruitfly, *Drosophila melanogaster*. Archives of Insect Biochemistry and Physiology, 110(2): e21882

Zhang WJ. 2023. netAna: A tool for network analysis. Network Biology, 13(4): 192-212