



www.bioinformatics.net
Volume 18(8)

Research Article

Received July 2, 2022; Revised August 31, 2022; Accepted August 31, 2022, Published August 31, 2022

DOI: 10.6026/973206300186703

Declaration on Publication Ethics:

The author's state that they adhere with COPE guidelines on publishing ethics as described elsewhere at <https://publicationethics.org/>. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Declaration on official E-mail:

The corresponding author declares that lifetime official e-mail from their institution is not available for all authors

License statement:

This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

Comments from readers:

Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article without open access charges. Comments should be concise, coherent and critical in less than 1000 words.

Edited by P Kanguane

Citation: Krishna *et al.* Bioinformatics 18(8): 703-709 (2022)

Identification and annotation of milk associated genes from milk somatic cells using expression and RNA-seq data

Neelam Krishna, Shraddha Vishwakarma & Pramod Katara*

Computational Omics Lab, Centre of Bioinformatics, University of Allahabad, Prayagraj - 211002, India; *Corresponding author

Institution URL:

<https://www.allduniv.ac.in/>

Author contacts:

Pramod Katara - E-mail: pmkatara@gmail.com; pkatara@allduniv.ac.in

Abstract:

It is of interest to identify and annotate milk associated genes using expression profiling and RNA-Seq data from milk somatic cells. RNA-Seq data was pre-processed and mapping was done to identify differentially expressed genes (DEG). The functional insights about the up and down regulated genes were gleaned using the protein-protein interaction Network in the STRING database followed by CytoHubba

analysis in Cytoscape. Gene ontology, annotation and pathway enrichment was completed using ShinyGO, David tool and QTL analysis. These analysis shows that 21 genes are linked with the secretion of milk.

Keywords: Casein, Somatic cell, Lactation, Milk Traits, Gene expression

Background:

Cow milk represents important nutritional support to human health; in particular, it is a rich source of all nutrients such as protein, fat, and minerals that are required for human growth. Studies based on clinical evidence also support the consumption of cow milk and dairy products for better health; it is also reported to be beneficial for cardio metabolic health [1]. Milk composition has a dynamic nature which comprises immunoglobulins, hormones, growth factors, cytokines, nucleotides, peptides, polyamines, enzymes, and other bioactive peptides [2]. The composition and quality of milk vary with the stage of lactation, age, breed, nutrition, and health status of the udder. The health of the udder is one of the major factors which affect the quality of raw milk [3]. Milk Somatic cells (SCs) in a cow are a mixture of milk-producing cells and immune cells, i.e., WBC (White blood cells), they are associated with udder health and milk production. It is a known process that somatic cells are released into the cow milk to defend against udder infections, i.e., mastitis, these cells fight with infection and also repair the tissue damage. Milk SC is influenced by various factors including cow productivity, health, parity, lactation stage, and breed [4]. Change in any kind of environmental conditions and increased stress significantly increases the amount of SC in milk. In developed countries, milk somatic cell counts (SCCs) are in practice as a marker, mainly to monitor the udder-infection and as an indicator of raw milk quality [5]. Somatic cells are also in practice to measure the shelf-life of pasteurized fluid milk. It is now well established that there is a direct correlation between the number of somatic cells and udder infection, which is directly associated with milk quality. There is a need to observe the expression profiles of the genes belonging to somatic cells of the udder, which allows understanding of the involvement of genes and their regulation process, and their associations with milk quality and production rate [6]. Milk protein and fats are major constituents of milk and these traits are genetically influenced by associated genes. The use of expression profiling and genome wide gene expression is common now a day's [7], dairy science also reported studies based on expression profiling [8]. Therefore, the analysis of expression patterns of genes of interest is a key step to understanding the production and quality rate. During the lactation period expression of milk somatic cells may help to identify more essential genes that are involved in milk production and in its formation. Mammary transcriptome in dairy cows revolved around the cattle's health and milk production [9]. Availability of gene expression data, i.e., microarray and RNA-seq, and Bioinformatics resources for functional genomics make it feasible to perform functional genomics in a more advanced and precise manner [10]. Considering the facts current research aims to find out genes that affect the milk production and composition during the early and late days of lactation, for that differential expression of the gene and network analysis has been done. The expression level of genes will facilitate understanding of particular genes that enhance milk production

and genes that decline milk production. QTLdb is used to identify genes that are directly involved in milk production and composition traits [11, 12]. Therefore, it is of interest to identify and annotate milk associated genes using expression profiling and RNA-Seq data from milk somatic cells.

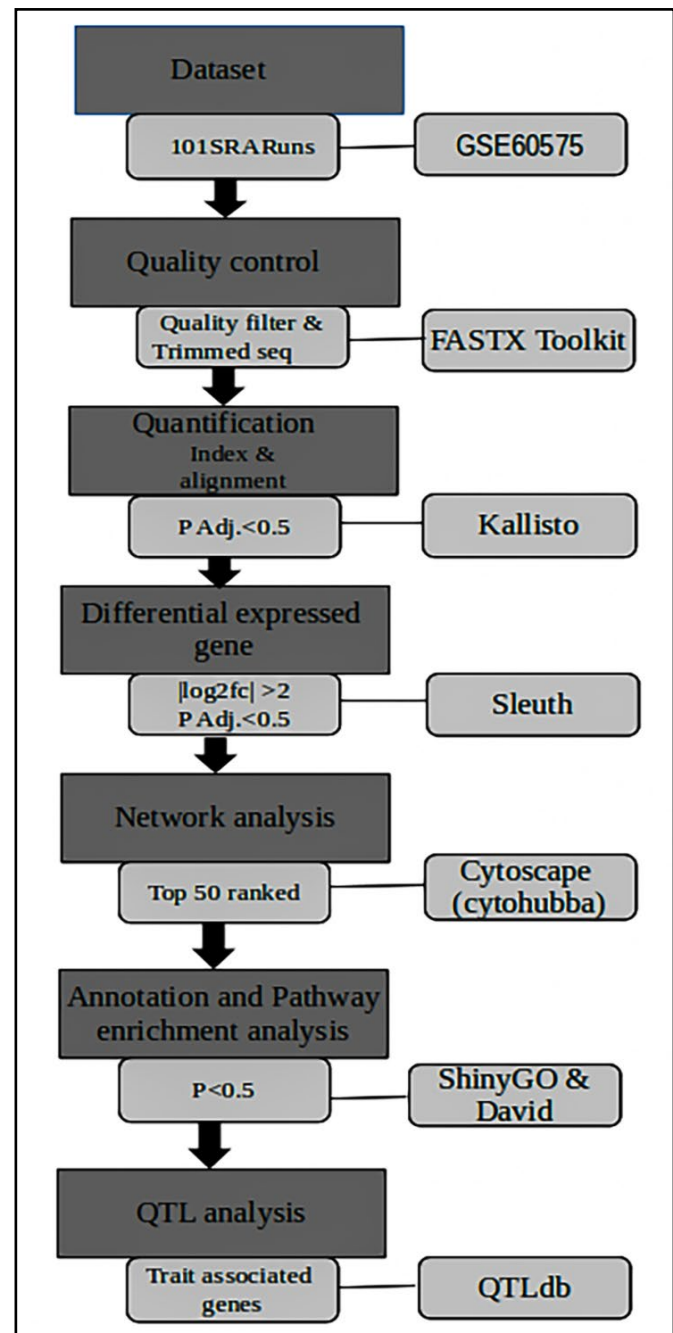


Figure 1: Schematic representation of the considered methodology

Material and methodology:**Data collections:**

To perform the expression profiling RNA-seq data of Holstein cows were retrieved from the GEO database (GSE60575) [13]. It includes 21 biological samples (SRP045641; Bio project ID PRJNA258561) and is extracted from somatic cells belonging to Holstein bovine milk [14]. This data contains 101 SRA runs which are paired-end layout reads. For the analysis purpose, data has been divided into 2 different groups: High yield and low yield, based on their lactation periods from early days of lactation to late days of lactation.

Quality control and Quantification:

FASTX Toolkit was used to check the raw reads quality and pre-processing. Raw reads were filtered based on the Phred score Q33 by the FASTQ quality filter and end-trimming raw reads based on the quality by FASTQ Quality Trimmer [15]. Pre-processed data was used for quantifying transcripts abundance from RNA-Seq data using the Kallisto program [16].

Differentially expressed genes and Network analysis:

Analysis of differentially expressed genes between two conditions, i.e. high yield of milk and low yield of milk, using p-value < 0.05 and $|\log_2fc| > 2$ by Sleuth was performed. Fold change is estimated and overall expression strength is indicated by the identification of the coefficient [17]. Further, network analysis of identified DEGs was performed by STRING (v11.5) [18] and Cytoscape [19]. Cytohubba, a plugin of Cytoscape was used to find out top-ranked genes and central elements of the biological network on the basis of the rank and their connectivity through the MCC algorithm $MCC(v) = \sum C \in S(v) (|C| - 1)!$ Where $S(v)$ is the collection of maximal cliques which contain v , and $(|C| - 1)!$ is the product of all positive integers less than $|C|$ [20].

Gene enrichment and Pathways analysis:

Highly interacted top-ranked genes obtained from Cytohubba were used for gene ontology (GO) analysis which was carried out with ShinyGO [21] and functional annotation tool DAVID [22] with P-value cut-off 0.05. Further for QTL analysis, differentially expressed genes were mapped with QTLdb [23]. Identification of QTL genes was observed and their involvement was checked in milk production and composition traits (Figure 1).

Results and Discussion:**RNAseq data and differentially expressed gene:**

Collected raw reads from 21 SRA experiments were pre-processed with FASTX, and then mapped with Bos_taurus_UMD_3.1.1 genome by pseudo alignment tool, Kallisto. Moreover Bos_taurus_UMD_3.1.1 reference assembly was aligned approximately up to 85.5% and quantified in total 76,341 transcripts identified from Kallisto. Further data was normalized and resulted in 14,433 differential expressed genes by setting up a threshold p-value < 0.05 between high yielding and low yielding milk production. Out of these genes, 1231 elevated expression of genes > +2 FC (fold change) and 360 down-regulated genes < -2 FC were identified by setting up threshold values for $|\log_2fc| > 2$.

PPIs Network:

The Protein-Protein Interaction Network provides the connection between proteins which helps us to understand the role of proteins in a systemic manner. Predicted up (1231) and down-regulated (360) genes were used to construct the PPI networks. PPI network has been observed through a string database [24], separately for upregulated and down-regulated genes and network characteristics of both the network have been examined (Table 1).

For further analysis, these PPI networks have been imported to a network analyzer and visualization tool, i.e., Cytoscape. Cytohubba plugin has been used to identify the most interactive, top-ranked 50 nodes for upregulated and down-regulated PPIs networks (Table 4). Upregulated most interactive, top-ranked nodes are represented by blue color while pink color nodes are portrayed with top-ranked 50 down-regulated genes (Figure 2). The molecular network displayed that these proteins have high connectivity among them and make a network motif. The network also suggested that top-ranked genes consist of more casein protein-related genes; Caseins protein is a major constituent of milk, which is highly interacted [25].

In Figure 2, yellow nodes (shown in the blue and pink colour top most ranked 50 genes upregulated and down-regulated PPI networks) have been identified as the QTL gene that is mapped by the QTLdb database for milk yield, milk protein, and milk fat traits in dairy cattle. In the network, these QTL genes are well connected to each other. The genes, i.e., LTF, GLYCAM1, CSN1S1, CSN1S2, CSN2, CSN3, PIGR, and PAEP are made compact structures altogether and singly connected to the rest of the blue color network and signified as milk trait module by analyzing the network [26]. Eight genes, i.e., GNAS, PTPN11, SELL, STAT1, STAT6, STAT5B, IL10RA, and CD44 are shown as yellow color nodes among pink color nodes network (Figure 2b) for down-regulated top most 50 gene PPI network. These nodes well interact in a circular module of a network; it's not made a compact structure while it is part of a well compact, circular structure of the network. In both networks, yellow color nodes are QTL genes that are signified as milk production and composition traits associated genes. All top ranked genes are further used for annotation and pathways enrichment analysis based on fold change enrichment.

Functional annotations and pathway enrichment:

To find out the additional information regarding predicted up and down-regulated genes and their functional annotation and pathway enrichment were performed. Annotation results suggest that upregulated genes, i.e., PRPF39, HNRNPUL1, RBM39, HNRNPA2B1, SRRM1, HNRNPD, RPL35A, SRSF2, and PCBP2 are mainly involved in the necessary function, such as RNA processing, RNA metabolism, and Ribonucleoprotein complex. CSN3 prevents casein precipitation in milk; LALBA activates LS (Lactose synthase) to synthesize lactose and forms a major carbohydrate component of milk [27, 28]. PAEP act as the Primary component of whey protein, and LTF transferrins are iron binding transport proteins and are associated with the binding of an anion. Other three genes, namely, CSN1S1, CSN1S2, and CALB1 play crucial roles in the capacity of milk to transport calcium phosphate and calcium-binding [29, 30].

Down-regulated genes such as CASC3, POU2AF1, STAT6, EIF3CL, ADAR, STAT1, SMG5, PDCD11, STAT5B, and BCL6 control the innate immune response, Macromolecule metabolic process, and Cytokine-mediated signalling pathways during late days of lactation [31]. The genes, i.e., RPS27L, RPL34, RPL7, RPL8, RPL6, RPS24, and SAFB are involved in major biological Reactome pathways such as Metabolism of proteins, Translation, and

Metabolism of RNA. Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC) pathways maintains the expression of the genes and reduces the error in gene expression [32]. Down-regulated genes, i.e., PTPN11, STAT6, CSF2RB, and STAT1 are involved in 3 Reactome pathways (Table 2), which mainly belong to the immune system signalling process [33, 34].

Table 1: Network characteristics for upregulated and down-regulated genes while most interactive, top-ranked 50 genes from upregulated and down-regulated genes given by string

S. No.	PPIs	Upregulated genes	Down-regulated genes	Most interactive upregulated genes (50)	Most interactive down-regulated network (50)
1	Number Of Nodes	1116	324	50	50
2	Number Of Edges	5357	619	348	167
3	Average Node Degree	9.6	3.82	13.9	6.68
4	Avg. Local Clustering Coefficient:	0.238	0.319	0.809	0.536
5	Expected Number Of Edges	4433	534	63	36
6	PPI enrichment p-value	< 1.0e-16	0.000178	< 1.0e-16	< 1.0e-16

Table 2: Major biological reactome pathways in which highly interacted up and down-regulated genes involved

S. No.	Category	Term	Count	P-Value	Benjamini
A. Biological reactome pathways shared by upregulated genes					
1	R-BTA-156827	L13a-mediated translational silencing of Ceruloplasmin expression	4	8.50E-05	1.90E-04
2	R-BTA-392499	Metabolism of proteins	5	1.20E-02	2.50E-02
3	R-BTA-72613	Eukaryotic Translation Initiation	5	1.20E-05	3.90E-05
4	R-BTA-72689	Formation of a pool of free 40S subunits	5	6.00E-06	3.90E-05
5	R-BTA-72737	Cap-dependent Translation Initiation	5	1.20E-05	3.90E-05
6	R-BTA-72766	Translation	5	1.80E-05	5.30E-05
7	R-BTA-8953854	Metabolism of RNA	5	3.60E-05	9.10E-05
8	R-BTA-927802	Nonsense-Mediated Decay (NMD)	5	1.10E-05	3.90E-05
9	R-BTA-975956	Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	5	8.30E-06	3.90E-05
B. Biological reactome pathways shared by down-regulated genes					
1	R-BTA-449147	Signalling by interleukins	3	9.2E-3	8.2E-1
2	R-BTA-1280215	Cytokine signalling in immune system	3	3.2E-2	1.0E-0
3	R-BTA-512988	Interleukin-3, Interleukin-5 and GM-CSF signalling	2	3.8E-2	1.0E-0

Table 3: Genes associated with milk yield and composition

S. No.	Gene name	Traits (QTL ID)
A Upregulated genes		
1	GLYCAM1 (5)	Milk fat yield (175859); Milk yield (175950); Milk protein yield (176227)
2	CSN2 (6)	Average daily milk yield (281728); Milk rennet coagulation time (136223); Milk casein percentage (136227); Milk yield (13567); Milk casein content (136222); Milk fat yield (136219); Milk protein percentage (13570)
3	CSN3 (6)	Milk yield (10468); Milk kappa-casein percentage (13565); Milk protein yield (31639); Milk alpha-S1-casein percentage (13562); Milk alpha-lactalbumin percentage (36287); Somatic cell count (31646); Milk casein yield (36290); Length of productive life (31641); Milk fat yield (10470)
4	LTF (22)	Milk urea nitrogen content (95740); Milk lactose content (95724); Clinical mastitis (30827); Milk solids percentage (95739)
5	PAEP (11)	Milk yield (12419); Milk alpha-lactalbumin content (13551); Milk beta-lactoglobulin protein content (20583); Milk whey protein content (12423); Milk beta-casein percentage (13555); Milk kappa-casein content (12404); Milk solids percentage (238160); Somatic cell count (238175); Milk alpha-S1-casein percentage (12431); Milk alpha-lactalbumin percentage (13551); Milk protein yield (13551); Milk casein content (12422); Milk alpha-casein content (238161)
6	CSN152 (6)	Milk beta-casein content (20562); Milk protein content (20590)
7	CSN151 (6)	Milk yield (3855); Milk alpha-casein content (20558); Milk alpha-casein to beta-casein ratio (20596)
B Down-regulate genes		
1	PTPN11 (17)	Milk fat yield (181826)
2	CD44 (15)	Lactation persistency (125214)
3	GNAS (13)	Milk protein percentage (16139); Milk protein yield (16137); Somatic cell count (16143); Milk fat percentage (16138); Milk yield (16136)
4	IL10RA (15)	Somatic cell score (125376); <i>M. paratuberculosis</i> susceptibility (13115)
5	SELL (16)	Milk yield (281485)
6	STAT5B (19)	Milk yield (14140); Milk protein yield (14141)
7	STAT6 (5)	Yield grade (11476)
8	STAT1 (2)	Milk arachidonic acid content (122306); Milk protein percentage (9984); Somatic cell score (95253); Milk trans-6/8-Octadecenoic acid content (122320); Milk trans-16-Octadecenoic acid content (122328); Milk cis-8,11,14-Eicosatrienoic acid content (122341); Milk protein yield (9985); Milk yield (136183); Milk fat percentage (9983)

Table 4: List of top 50 up and down-regulated genes

Upregulated Genes	Down-regulated Genes
-------------------	----------------------

Gene Names	LOC510852, HNRNPM, HNRNPA3, HNRNPA2B1, HNRNPU, SF3B1, SRSF5, HNRNPD, HNRNPH1, SRSF2, SNRPA, SFPQ, TRA2B, HNRNPR, TRA2A, RPL7, RPL8, RPL6, RPL34, RPL35A, RPS24, RPS27L, HNRNPH3, MRPS5, MRPL16, BTF3, EIF5B, HNRNPUL1, SFRS13A, SRRM1, SRRM2, NACA, PCBP2, PAEP, CSN1S1, CSN1S2, CSN2, CSN3, LTF, PIGR, YTHDC1, RBM39, ZFAND4, GLYCAM1, LPO, PRPF3, RBM22, SAFB, PRPF39.	HNRNPA3, HNRNPD, ILF3, HNRNPR, SFPQ, FUS, HNRNPUL1, RBMX, CD19, SELL, CD44, STAT1, EFTUD2, BCL6, VCAM1, CD3E, ITGAL, STAT6, STAT5B, PTPN11, NOS2, CSF2RB, IRF8, PDCD11, EMR1, PABPC4, MX1, ATP5A1, ARR2, GNB1, GNAS, IL10RA, SNRNP, CASC3, RUVBL1, FES, MAVS, MRPL16, ADAR, POLDIP3, PRKCB, SP1, VARS, POU2AF1, NCOR1, SMG5, HDAC6, EIF3CL, CNOT7, AARS.
------------	--	--

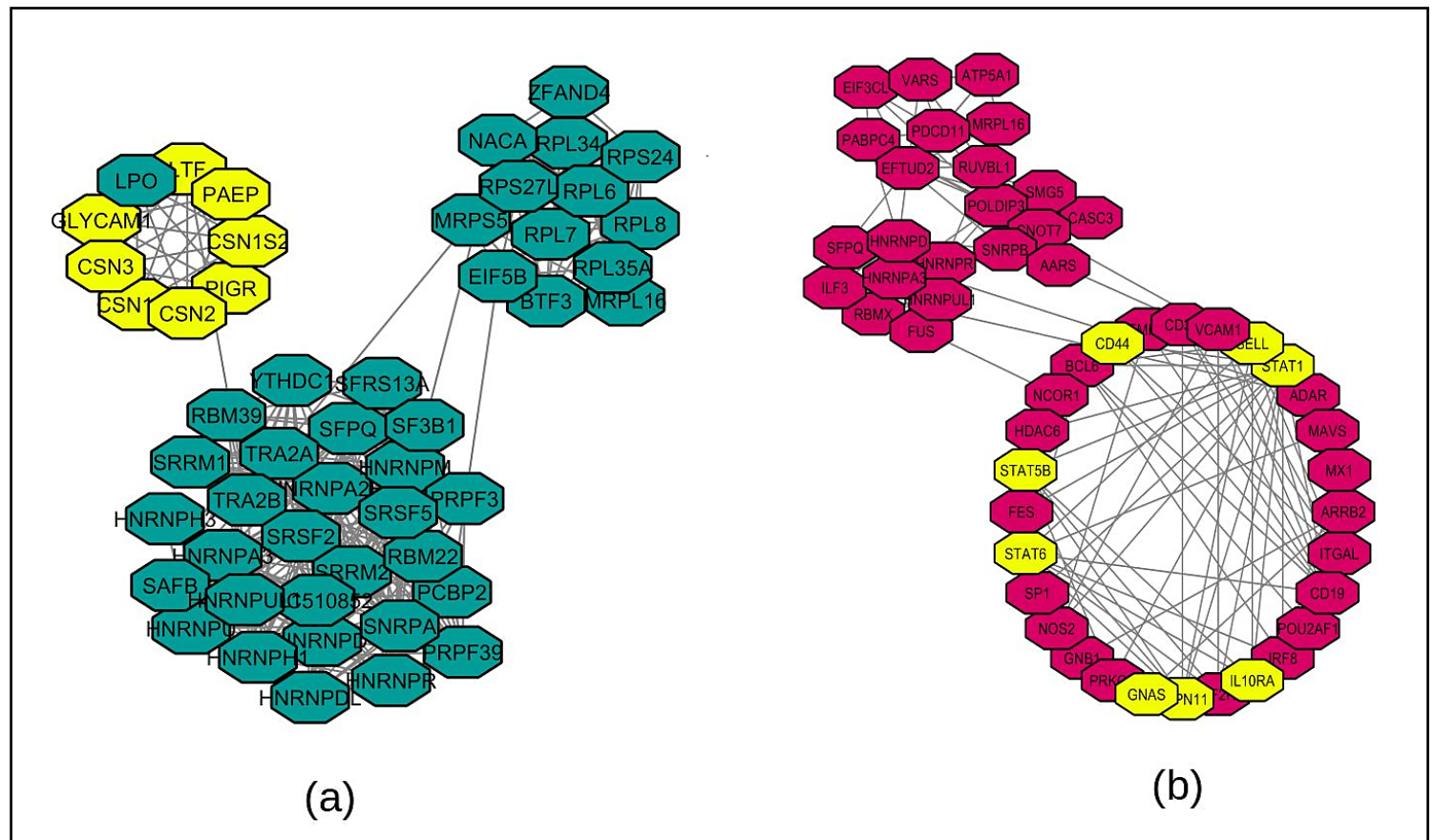


Figure 2: Gene network interaction of top ranked genes, yellow colour nodes showed the important genes that are associated with production and composition traits are interacted with each other (a) Blue colour nodes depicted with upregulated genes and (b) Pink colour nodes portrayed with down regulated genes.

Identification of genes associated with milk production and composition:

In total, 15 candidate functional genes have been identified, which are known to affect milk production and composition traits in which 7 up-regulated genes and 8 down-regulated genes are associated with the total 63 traits of milk quality and quantity. Comparative analysis with previously reported genetic data including QTL mapping was used for the assessment of the position of DEG's. The physical position of each DEG with the position of known QTLs that have been shown to be associated with the milk yield, milk protein, and milk fat traits in dairy cattle from the QTLdb database has been examined [35]. Through QTL mapping it has been observed that in total 67,849 QTL are involved in milk, out of the 5,532 QTL are involved in milk production, 39,536 QTL are involved in milk composition fat, 20,108 QTL belongs to milk composition protein, and 2,673 belong to other milk

nutrients [36, 37]. Highly expressed and interacted genes are also associated with the protein and fat composition of milk, which are influenced by nutrition, stage of lactation, and breed of cattle [38], GLYCAM1 and CSN2 play a major role in the milk fat composition and CSN1S1, CSN1S2, CSN2, CSN3 are involved in protein composition [39]. LTF is mainly involved in the composition of Milk lactose content, Milk urea nitrogen content, and Milk alpha-S1-casein percentage [40]. Milk solids percentage and milk somatic counts are influenced by the expression of PAEP gene (Table 3), that is well reported in previous studies [41].

In down-regulated gene group, genes such as GNAS and PTPN11 are reported to be important for oocyte maturation and embryo development, and also act as growth factors [42]. SELL gene involved in mastitis, biological process reported [43]. STAT1 and STAT6 are involved in calculating yield grade and maintaining cow

body weight and health and STAT5B have an almost fatalistic attitude towards milk fat and protein composition, which directly affect milk quality during the late days of lactation [44]. Negative expression of IL10RA protects the cow from *M. paratuberculosis* susceptibility and it is also reducing the number of somatic milk cells which directly affects the milk production rate [45]. The presence of CD44 affects lipid metabolism and lactation persistence [46]. Available reports justify that the expression level of upregulated genes helps to enhance the milk production and composition rate in the early days of lactation and down-regulated genes degrade the milk quality and quantity in the late days of lactation [47].

Conclusion:

Analysis shows that some genes are highly upregulated and are linked with milk yield and its composition. Data shows that 6 genes namely RPS27L, RPL34, RPL7, RPL8, RPL6 and RPS24 are found to be involved in the secretion of milk proteins. Data further shows that 9 down-regulated genes (PTPN11, CD44, GNAS, IL10RA, SELL, STAT5B, STAT1, STAT6, and PAEP) are linked to affect milk yield and its composition.

Conflict of interest:

No conflict of interest exists

Acknowledgements:

The authors (NK and SV) thanks to UGC for the UGC-CRET fellowship, they also thanks to the Centre of Bioinformatics, University of Allahabad for providing the computational facility for this work.

References:

- [1] Hirahatake KM *et al. Adv Nutr.* 2020 **11**:533. [PMID: 31904812]
- [2] Haug A *et al. Lipids Health Dis.* 2007 **6**:25. [PMID: 17894873]
- [3] Lemay DG *et al. PLoS One.* 2013 **8**:e67531. [PMID: 23861770]
- [4] Alhussien MN *et al. Vet World.* 2018 **11**:562. [PMID: 29915493]
- [5] Petzer IM *et al. Onderstepoort J. Vet. Res.* 2017 **84**:e1. [PMID: 28397516]
- [6] Li N *et al. Dairy Sci Technol.* 2014 **94**:517. [PMID: 25309683]
- [7] Kumari R & Katara P, *Bioinformation* 2021 **17**:608. [PMID: 35173382]
- [8] Yang J *et al. Anim Genet.* 2016 **47**:174. [PMID: 26692495]
- [9] Ahlawat S *et al. Mol Biol Rep.* 2020 **47**:7029. [PMID: 32880836]
- [10] Wickramasinghe S *et al. BMC Genomics* 2012 **13**:45. [PMID: 22276848]
- [11] Zhou C *et al. Front Genet.* 2019 **10**:672. [PMID: 31456815]
- [12] Chin CH *et al. BMC Syst Biol.* 2014 **8**:S11. [PMID: 25521941]
- [13] <https://www.ncbi.nlm.nih.gov>
- [14] Seo M *et al. Sci. Rep.* 2016 **6**:24375. [PMID: 27071914]
- [15] <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- [16] Du Y *et al. G3 (Bethesda).* 2020 **10**:1775. [PMID: 32220951]
- [17] Yi L *et al. Genome Biol.* 2018 **19**:53. [PMID: 29650040]
- [18] <https://string-db.org/>
- [19] Shannon P *et al. Genome Res.* 2003 **13**:2498. [PMID: 14597658]
- [20] Grisart B *et al. Proc Natl Acad Sci U S A.* 2004 **101**:2398. [PMID: 14983021]
- [21] <http://bioinformatics.sdstate.edu/go/>
- [22] <https://david.ncifcrf.gov/summary.jsp>
- [23] <https://www.animalgenome.org/cgi-bin/QTLdb/BT/index>
- [24] Szklarczyk D *et al. Nucleic Acids Res.* 2019 **47**:D607. [PMID: 30476243]
- [25] Rijnkels M *et al. J Mammary Gland Biol Neoplasia* 2002 **7**:327. [PMID: 12751895]
- [26] Li C *et al. Sci Rep.* 2016 **6**: 26813. [PMID: 27254118]
- [27] Głab TK & Boratyński J, *Top Curr Chem. (Cham)* 2017 **375**:71. [PMID: 28712055]
- [28] <https://www.uniprot.org/uniprot/P00711>.
- [29] <https://www.uniprot.org/uniprot/P02662>.
- [30] <https://www.uniprot.org/uniprot/P02663>.
- [31] Huang C *et al. Nat Immunol.* 2013 **14**:380. [PMID: 23455674]
- [32] Chang YF *et al. Annu Rev Biochem.* 2007 **76**:51. [PMID: 17352659]
- [33] Zhao Y *et al. BMC Systems Biology.* 2019 **13**:8. [PMID: 30642337]
- [34] Li HS & Watowich SS. *Immunol Rev.* 2014 **261**:84. [PMID: 25123278]
- [35] Li Q *et al. BMC Genet.* 2020 **21**:72. [PMID: 32646377]
- [36] Kolbehdari D *et al. J Anim Breed Genet.* 2009 **126**:216. [PMID: 19646150]
- [37] Coppieters W *et al. Mamm Genome* 1998 **9**:540. [PMID: 9657851]
- [38] Cui X *et al. BMC Genomics* 2014 **15**:226. [PMID: 24655368]
- [39] Meier S *et al. Front Genet.* 2019 **10**:1129 [PMID: 31781175]
- [40] Groenenda MA *et al. Gene* 1995 **158**:189. [PMID: 7607540]
- [41] Kolenda M *et al. Anim Biosci.* 2021 **34**:1283. [PMID: 33677915]
- [42] Idrees M *et al. Cells* 2019 **8**:1272. [PMID: 31635340]
- [43] Dusza M *et al. Czech J. Anim. Sci.* 2018 **63**:256. [PMID: 33902174]
- [44] Rincon G *et al. Anim Genet.* 2009 **40**:878. [PMID: 19519791]
- [45] Mallikarjunappa S *et al. BMC genetics* 2020 **21**:121 [PMID: 33138773]
- [46] Jiang P *et al. J Dairy Sci.* 2020 **103**:6661. [PMID: 32359993]
- [47] Lemay DG *et al. PLoS One* 2013 **8**:e67531. [PMID: 23861770]

