**Explor Anim Med Res.**
**Vol. 15, Issue 2, 2025**
**DOI: 10.52635/eamr/15.2.219-225**

ISSN 2277-470X (Print), ISSN 2319-247X (Online)
Website: www.animalmedicalresearch.org

***Research Article***

# INTER-SPECIES SNPS FROM BUFFALO LIVER TRANSCRIPTOME INDICATE DIVERSITY IN IMMUNE PATHWAYS BETWEEN BUFFALO AND CATTLE

Chaitanya Kumar Thota Venkata[1*], Sudhakar Singh[2], Gowdar Veerapa Vedamurthy[2],
Dheer Singh[2], Suneel Kumar Onteru[2]

*Received 16 November 2025, revised 27 October 2025*

**ABSTRACT: The duration of negative energy balance (NEB), a physiological adaptation in females during early postpartum, and its coping mechanisms vary among animals, breeds, and species because of genetic differences. However, genetic variations that may influence the NEB differences between cattle and buffaloes were not reported. Therefore, the present study aims to identify such genetic differences between cattle and buffaloes by identifying inter-species single nucleotide polymorphisms (SNPs) by mapping the available liver RNA-seq data earlier obtained from three early postpartum buffaloes and three heifers to the cattle genome (UMD 3.1) using the SNP eff tool. The liver transcriptome data were chosen, because the liver is the major metabolic organ responsible for balancing metabolic milieu to meet different physiological demands during early postpartum. Using the identified interspecies SNPs, intragenus SNPs (Buffalo) were derived from six different Murrah buffaloes and mapped them to the *Bubalus bubalis* (ASM312139*v*1) genome. Further, functional annotation and biological pathway analyses were performed for those genes harbouring the putative SNPs. We identified a total of 1,40,056 interspecies and 188 intragenus SNPs. Functional annotation and pathway analyses revealed that most of the putative interspecies SNPs between the cattle (*Bos taurus* UMD 3.1) and buffalo (*Bubalis bubalus*) were in the protein metabolism genes (>1900) predominantly involved in the immune system (>1800 genes). Similarly, most of the putative intragenus SNPs among buffaloes were in the genes of lipid metabolic pathways. In conclusion, our study suggests that cattle and buffaloes might differ in protein metabolism genes involved in immunity at the deoxyribonucleic acid (DNA) level.**

**Keywords:** Single Nucleotide Polymorphisms; Liver transcriptome; Functional annotation; Cattle genome; Buffalo**.**

## INTRODUCTION

Understanding genetic variation among species and breeds of dairy animals is utmost important for selection of efficient animals that could balance the energy needs during transition period. During transition period, the energy demands for milk production surpass the energy intake, leading to negative energy balance (NEB), which is common in high yielding dairy cows and buffaloes. However, the genetic variations between cows and buffaloes influencing NEB were not reported. Therefore, the present study is targeted to identify such genetic differences.

Liver is the principal organ that balances the metabolic milieu during postpartum to meet the milk production and reproduction demands. Particularly, liver is the gluconeogenic organ in ruminants for supplying the inputs for lactose production [1]. During NEB, the liver induces lipolysis in adipose tissue through angiopoietin like-4 production [2, 3]. It also controls ovarian follicular development through molecules like IGF-1 [4].

Therefore, liver transcriptome during early postpartum has tremendous importance to explore the genetic variants in its expressed gene repertoire for understanding the

---

[1]*Department of Veterinary Biochemistry, College of Veterinary Science, Sri Venkateswara Veterinary University, Tirupati, Andhrapradesh-517502 India.*
[2]*Molecular Endocrinology, Functional Genomics and Systems Biology lab Animal Biochemistry Division ICAR-National Dairy Research Institute, Karnal-132001 India.*
**Corresponding author. e-mail: meechaitanya@yahoo.co.in*

genetic variation influencing NEB. Hence, the present study aims to identify interspecies and intragenus SNPs from the liver transcriptome data of the early postpartum buffaloes and to annotate their functional importance through bioinformatics approaches.

## MATERIALS AND METHODS

### Buffalo liver transcriptome data

The liver transcriptome data from our previous study [5] were analysed in the present study to identify interspecies and intragenus SNPs. The liver transcriptome data were obtained by performing RNA-seq on the liver biopsy samples (NDRI Institutional Animal Ethics Committee approval no. 95/16) collected from three early postpartum buffaloes and three heifers. Liver tissues were collected by biopsy from all six animals. The liver biopsies were performed on heifers at the diestrous stage (10th day of estrous cycle) and on lactating animals on the 15th and 30th day of early postpartum. Briefly, the animals were fasted overnight prior to biopsy. On the day of biopsy, the animals were washed very cleanly with water, and a standard protocol was followed to collect liver biopsy samples [5]. From the lactating buffaloes, transcriptome data were obtained twice, once on the 15th day and another on the 30th day. Overall, the present study utilized nine transcriptome datasets [5] to obtain the interspecies and intragenus SNPs as mentioned in the table below.

| Animal No | Day of collection/No of samples |
|---|---|
| Control Animals (7203, 7204, 7205) | 10th day of oestrous cycle, Total three samples once from each animal |
| Lactating buffaloes (7206, 7207, 7208) | 15th and 30th day, Total six samples, twice from each animal |

### Identification of interspecies SNPs

All the transcriptome datasets were mapped to the cattle reference genome using SNP eff software integrated to GATK software (https://software.broadinstitute.org/gatk/) [6] to obtain interspecies SNPs.

### Verification and finalization of interspecies SNPs obtained from the liver transcriptome data of lactating animals

As two liver transcriptome data sets (15th and 30th day of early postpartum) were available from one lactating animal, two gene lists containing interspecies SNPs were obtained by the SNP-eff software for each lactating animal. The two gene lists from each lactating animal were compared by interactivenn tool (http://www.interactivenn.net/) [7] to identify the unique and common list of genes containing interspecies SNPs between these two datasets. Later, the unique and common list of genes were pooled for each animal to avoid duplicity of the gene name. For each gene in this pooled list of each lactating animal, its corresponding nucleotide base was extracted from two arrays of the gene lists along with their interspecies SNPs, which were obtained from the transcriptome data of the same animal on the 15th and 30th day of lactation by using VLOOKUP function in the MS-Excel software. The unique genes had an interspecies SNP with a single nucleotide base difference from the cattle genome. This strategy gave us the list of the unique nucleotide base for each common gene containing interspecies SNPs for one lactating animal based on transcriptome datasets of the 15th and 30th days of lactation. The total list of unique genes and common genes with their corresponding interspecies SNPs were considered together as interspecies SNPs specific to that lactating animal. Likewise, the list of genes containing interspecies SNPs were verified and finalized for each lactating animal.

### Derivation of intragenus SNPs

The list of interspecies SNPs obtained from six Murrah buffaloes were compared to get intragenus SNPs. This comparison was done by using the functions available in MS-Excel software. For this, the list of genes containing interspecies SNP nucleotide base (NB) of all individual six animals under study were arranged in an ascending order in different excel sheets. Then, the common genes containing interspecies SNPs between all the six animals were identified using the "VLOOKUP" function by a process developed in our lab. various functions available in MS-Excel. The NBs from six animals were compared to derive intragenus SNPs.

### Mapping of the putative SNPs to the buffalo genome

As the intragenus and interspecies SNPs were identified from the liver transcriptome data of six Murrah buffaloes based on cattle genome (UMD 3.1) using SNPeff software, the fifty bases of the flanking sequences at both the upstream and downstream of the SNP position were extracted from the *Bos taurus* (UMD 3.1.1) genome to identify their syntenic positions in the buffalo genome. These sequences containing putative SNPs were aligned to the *Bubalus bubalis* (ASM312139*v*1) genome by the NCBI local Blast software (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [8] by considering the sequences containing putative SNPs as the query sequence and the *Bubalus bubalis* (ASM312139*v*1) genome as a subject sequence.

### Functional annotation of the genes containing interspecies SNPs

The genes containing inter-genus SNPs were functionally annotated initially by analysing the data with CPDB (consensus pathway database: http://cpdb.molgen.mpg.de) [9]. The pathways with highest numbers of candidate genes were considered for further analysis by the STRING (https://string-db.org) [10] and CYTOSCAPE software version 3.6.1. The .tsv file obtained on the network analysis containing node 1 and node 2 genes and their interaction score were further analysed by the CYTOSCAPE software (https://www.cytoscape.org/php) [11] to visualize the networks. In the CYTOSCAPE, cluster one plugin was used to obtain the significant ($P<0.05$) sub networks. Functional annotation of the list of genes in each sub network was carried by the clustering option in the DAVID (Database for annotation, visualization, and integrated discovery (https://www.david.ncifcrf.gov/) software [12].

### Functional annotation of the genes containing intragenus SNPs

Functional importance of the genes having the intragenus SNPs were analysed using CPDB. Further, synonymous, or non-synonymous nature of intragenus SNPs present in the coding regions were analysed. For this, the sequence of fifty bases upstream and the fifty bases downstream to the SNP position from *Bubalus bubalis* genome was blasted against non-redundant protein sequences of *Bubalus bubalis* using the blast X software (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx & PAGE_TYPE= Blast Search & LINK_LOC=blasthome). The triplet codons with amino acids for the two allelic forms of SNPs were determined to know whether the SNPs were synonymous or non-synonymous. Further, the point allele mutation scores for non-synonymous SNPs identified were determined using PAM 250 log odds substitution matrix.

## RESULTS AND DISCUSSION

### Intragenus and interspecies SNPs identified from buffalo liver transcriptome data

A total of 1,40,056 interspecies SNPs were found in 13801 genes based on *Bos taurus* genome. By considering an average gene size of 1200 base pairs (bp) in eukaryotes, the 13801 genes would cover nearly 16,561,200 bp of the genome. As 1,40,056 locations showed interspecies SNPs between cattle and buffaloes, the variation between these two species could be 0.84% in the genes specific to early postpartum liver transcriptome in buffaloes. However, the variation at

**Table 1. Intragenus SNPs present in coding regions.** [Eight SNPs are present in coding regions of the genes out of 163 intragenus SNPs, of these six SNPs are nonsynonymous and two are synonymous. PAM score determined for non synonymous SNPs shows LMPTP gene non synomous SNP with maximum score].

| Sl. No. | Gene | Intra-genus SNP and type of SNP | Codon change with SNP | Amino acid change with SNP | PAM Score |
|---|---|---|---|---|---|
| 1 | Interferon alpha/beta receptor 1 | A/T and Non synonymous SNP | TTA/TTT | Leucine/ Phenylalanine | 2 |
| 2 | SGTA (small glutamine rich tetratricopeptide repeat containing protein alpha isoform X2) | C/T and Synonymous SNP | CCC/CCT | Proline | NA |
| 3 | Low molecular weight phosphotyrosine phosphatase isoform x2 | G/C/A and Non synonymous SNP | TGC/TTC/TAC | Cysteine/ Phenylalanine | -4 |
| 4 | Ribosome binding protein 1 isoform X2 | C/G and Non Synonymous SNP | CGT/TGT | Alanine/ Threonine | 1 |
| 5 | Fatty acyl amino acid synthase / hydrolase | G/T/C and Non synonymous SNP | TCG/TCT/TCC | Alanine/Serine /Proline | 1 |
| 6 | Complement C4 like | A/G and Non synonymous SNP | GG/CAG | Arginine/Gluta mine | 1 |
| 7 | HSP1A | C/A and Synonymous SNP | CGG/AGG | Glycine | NA |
| 8 | LST1 | A/G/C and Non synonymous SNP | AAG/GAG/CAG | Lysine/Glutamic acid | 0 |

genomic level between the two species was reported to be 3 percent [13] indicating a lesser variation in the coding region of the genome than in the non-coding regions between the cattle and buffaloes. Although these SNPs were distributed throughout the genome, much variation was on the chromosome No. 3, which is syntenic to the buffalo Chromosome No. 6 (Fig. 1a). From these 1,40,056 interspecies SNPs, 188 intra-genus

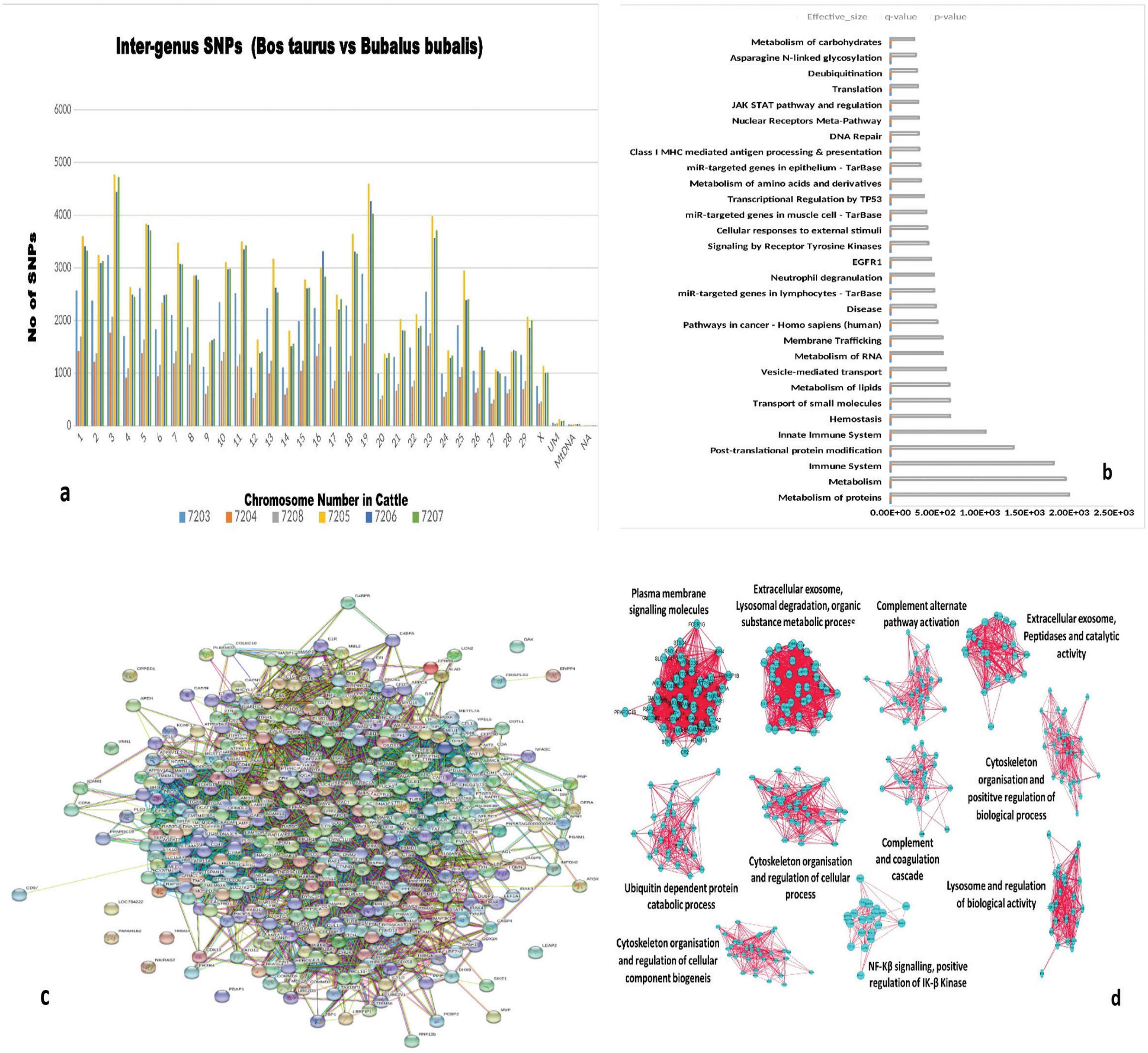

**Fig. 1. [a) Histogram of the distribution of inter-species SNPs on the cattle chromosomes.** The six color bars under each chromosome indicate the number of inter-species SNPs identified from six animals. The SNPs were distributed among all the chromosomes with the major variation on the chromosome number 3. **b) Pathway analysis for the genes containing inter-species SNPs (Cattle *vs.* Buffalo).** The Y-axis shows the pathways annotated to be enriched for the genes containing inter genus SNPs by CPDB database. The X-axis indicates the number of genes (effective size), P-values and q values related to the annotated pathways. The highest number of genes were enriched in the pathways involved in the metabolism of proteins, immune system, post-translational modification, and innate immune system. **c) Protein-protein interaction network for the genes containing inter-species SNPs being annotated to be involved in innate immunity.** Maximum number of genes containing inter-genus SNPs were annotated to be involved in innate immunity by CPDB database. This network was designed among these genes by using protein-protein interaction network algorithm available in the STRING software. **d) Eleven significant sub networks ($p<0.05$) among the genes containing inter-species SNPs being annotated to be involved in innate immunity.]**

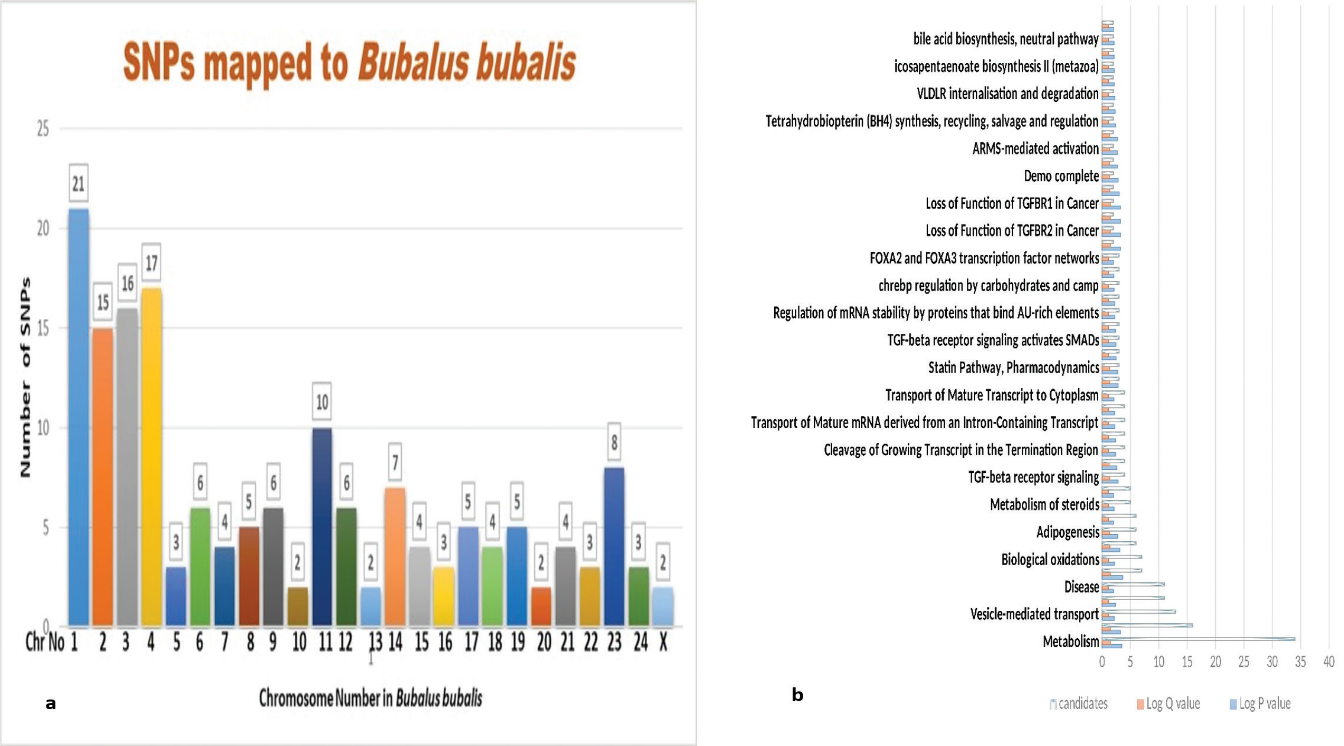

**Fig. 2. [a) Histogram of the distribution of the identified intra-genus SNPs on the buffalo chromosomes.** The SNPs were distributed among all the chromosomes with the major variation is on the first four chromosomes. **b) Pathway analysis for the genes containing intra-genus SNPs (Buffaloes).** The Y-axis shows the pathways annotated to be enriched for the genes containing inter genus SNPs by CPDB database. The X-axis indicates the number of genes (effective size); ,p-values and q values related to the annotated pathways. The highest number of genes were in the pathways involved in the metabolism (lipid metabolism and lipid associated metabolisms), vesicle mediated transport and, disease and signalling].

buffalo SNPs were identified in 156 genes. Among these 188 SNPs, 163 has been successfully mapped to different chromosomes of Bubalus bubalis (breed Mediterranean ASM312139v1) (Fig. 2a). The remaining SNPs were indicated as unplaced or not matched or shown many matches with the buffalo genome. Maximum number of intragenus SNPs were mapped to the chromosomes, 1, 2, 3 and 4 in buffaloes (Fig. 2a). The variation in the coding region between cattle and buffaloes appear to be high on cattle chromosomes 3 (syntenic position on chromosome number 6 in buffaloes) and 19 (syntenic position on chromosome number 3 in buffaloes).

### Bioinformatics analysis of the identified intragenus and interspecies SNPs

Pathway analysis of the genes containing interspecies SNPs found that majority of the genes were annotated to be involved in top five pathways, such as metabolism of proteins (1999), metabolism (1960), immune system (1827), post-translational modification (1378), gene expression (1364), RNA polymerase II transcription (1228) and innate immunity (1065) (Fig. 1b). Variation in innate immune genes between cattle and buffalo may explain their differential susceptibility to certain diseases. The SNPs in the innate immune genes may affect their differential expression of these genes between the two species. For instance, differential expression of immune genes among the two species was reported to understand the less susceptibility of water buffalo to *Schistosoma japonicum* than yellow cattle [14]. As the innate immunity pathway was annotated to be functionally defined pathway by CPDB analysis, the gene list from this pathway was used for the network analysis by the STRING and CYTOSCAPE software. This network analysis resulted in eleven significant ($p<0.05$) sub networks (Fig. 1c and Fig. 1d). DAVID functional annotation of the genes in these sub-networks revealed that majority of these genes were involved in cytoskeletal organization, protein catabolism, and membrane specific signalling involved in innate immunity. This observation based on liver transcriptome generated interspecies SNPs suggest that the two species vary at the basic molecular signalling response repertoire involved in the defence mechanism against infections.

Pathway analysis of the successfully mapped 163 intragenus SNPs in 156 genes found that the SNPs were found majorly in the genes involved in metabolic pathways, vesicle mediated transport, and disease signalling (Fig. 2b). Among the genes involved in the metabolic pathways, the genes related to lipid metabolism explains the importance of their SNPs and the liver in handling lipids. During NEB, the liver and its lipid metabolizing genes play an important role in meeting the energy demand for milk synthesis [1], especially in high yielders by initiating lipolysis in adipose tissue [2, 3] and adapting liver's metabolism [15]. In specific, the genes involved in metabolism of steroids and adipogenesis have shown much of the variation accounting to metabolism. This further endorses the importance of these variations during early lactation their adaptability and susceptibility to NEB [16]. Therefore, variations imposed by gene polymorphisms in lipid metabolic pathways in the liver may affect the animal's ability to sustain the stress caused by the NEB during early post-partum and resume their normal physiological activities like reproduction. These differences in animal's capability will also influence their differential susceptibility to post-partum diseases. Similarly, among the genes annotated to be involved in vesicle mediated transport and disease signalling, the genes related to the TGF (Tumor growth factor) beta signalling and EGFR1 (Epidermal growth factor receptor 1) signalling showed variation among buffaloes. Vesicle mediated transport and signalling are important during cross-talk among the liver, mammary, adipose, and reproductive organs, which manage the reciprocal control of metabolic activities and immune adaptation among these organs [17] during post-partum.

Among eight intragenus SNPs found in the coding regions of the genes, six SNPs in the genes, Interferon alpha/beta receptor 1 (INFAR1B), Low molecular weight phosphotyrosine phosphatase isoform x2 (LMPTP), Ribosome binding protein 1 isoform X2 (RBP1), Leukocyte specific transcript 1 (LST1), Complement 4 like and N-fatty-acyl-amino acid synthase/hydrolase (PM20D1) are non-synonymous for the two allelic forms (Table 1). Among these six SNPs, the maximum point mutation score (-4) for amino acid change was observed in LMPTP as per PAM 250 log odds substitution matrix. The polymorphisms in this gene were reported to be associated with the conception season and insulin resistance in humans [18, 19]. Insulin resistance during NEB may affect lipolysis and lipid mobility, which are very crucial for regaining reproductive efficacy. Hence, future studies are needed to explore its association with reproductive efficiency in Murrah buffaloes.

The sub-networks were obtained by the CYTOSCAPE (cluster one plugin) software by considering the interaction scores among the genes in the PPI network derived by the STRING software. The sub-networks explain that the genes containing inter-genus SNPs between cattle and buffaloes show much variation in cytoskeleton organization, protein catabolism and membrane specific signalling involved in innate immune response. The pathways will explain the differences in the susceptibility of the two species for various diseases. Further, plasma membrane signalling, and cytoskeletal organisation pathways variation will explain how these two species will interact with and deal with disease causing agents. This helps in explaining buffaloes being more tolerant and resistant to certain diseases than cattle (Fig. 1)..

## CONCLUSION

The identified interspecies SNPs based on early postpartum buffalo liver transcriptome data suggests that cattle and buffalo vary in their innate immunity during this period. Likewise, intragenus SNPs in buffalo indicate variation in the expressed genes of metabolism and in particular lipid metabolism during early lactation.

## ACKNOWLEDGEMENT

The authors thank Director, ICAR-National Dairy Research Institute for providing the infrastructure to carry out this study, and National Agricultural Science Fund, India for the financial assistance to this study (Grant No. NASF/GTR-5005).

### Ethics approval

The study was performed in guidelines and norms of NDRI Institutional Animal Ethics Committee with approval no. 95/16.

## REFERENCES

1. Drackley JK, Overton TR, Douglas GN. Adaptations of glucose and long-chain fatty acid metabolism in liver of dairy cows during the periparturient period. J Dairy Sci. 2001; 84: E100-E112, DOI: 10.3168/jds.S0022-0302(01)70204-4.

2. Loor JJ, Everts RE, Bionaz M, Dann HM, Morin DE, *et al.* Nutrition-induced ketosis alters metabolic and signaling gene networks in liver of periparturient dairy cows. Physiol Genomics. 2007; 32(1): 105-116, DOI: 10.1152/physiolgenomics.00188.2007.

3. McCarthy SD, Waters SM, Kenny DA, Diskin MG,

Fitzpatrick R, *et al*. Negative energy balance and hepatic gene expression patterns in high-yielding dairy cows during the early postpartum period: A global approach. Physiol Genomics. 2010; 42(3): 188-199, DOI: 10.1152/physiolgenomics.00118.2010.

4. Leroy JL, Van Soom A, Opsomer G, Goovaerts IG, Bols PE. Reduced fertility in high-yielding dairy cows: Are the oocyte and embryo in danger? Part II. Mechanisms linking nutrition and reduced oocyte and embryo quality in high-yielding dairy cows. Reprod Domest Anim. 2008; 43(5): 623-632, DOI: 10.1111/j.1439-0531.2007.00961.x.

5. Singh S, Golla N, Sharma D, Singh D, Onteru SK. Buffalo liver transcriptome analysis suggests immune tolerance as its key adaptive mechanism during early postpartum negative energy balance. Funct Integr Genomics. 2019; 19(5): 759-773, DOI: 10.1007/s10142-019-00676-1.

6. Van der Auwera GA, O'Connor BD. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra (1st edn.). 2020; O'Reilly Media.

7. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC bioinformatics. 2015; 16: 169, DOI: 10.1186/s12859-015-0611-3.

8. Sayers EW, Beck J, Bolton EE, Brister JR, Chan J, *et al*. Database resources of the National Centre for Biotechnology Information in 2025. Nucleic Acids Res. 2025; 6:53, DOI: 10.1093/nar/gkae979.

9. Herwig R, Hardt C, Lienhard M, Kamburov A. Analyzing and interpreting genome data at the network level with ConsensusPathDB. Nat protocols. 2016; 11(10): 1889-907.

10. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, *et al*. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023; 6:51(D1): D638-D646, DOI: 10.1093/nar/gkac1000.

11. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13: (11): 2498-2504.

12. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, *et al*. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022; 5:50(W1): W216-W221, DOI: 10.1093/nar/gkac194.

13. Moaeen-ud-Din M, Bilal G. Sequence diversity and molecular evolutionary rates between buffalo and cattle. J Anim Breed Genet. 2015; 132(1): 74-84, DOI:10.1111/jbg.12100.

14. Yang J, Fu Z, Hong Y, Wu H, Jin Y, *et al*. The differential expression of immune genes between water buffalo and yellow cattle determines species-specific susceptibility to *Schistosoma japonicum* infection. PLoS One. 2015; 10(6): e0130344, DOI: 10.1371/journal.pone.0130344.

15. Veshkini A, Hammon HM, Sauerwein H, Tröscher A, Viala D, *et al*. Longitudinal liver proteome profiling in dairy cows during the transition from gestation to lactation: Investigating metabolic adaptations and their interactions with fatty acids supplementation via repeated measurements ANOVA-simultaneous component analysis. J Proteomics. 2022; 252: 104435, DOI: 10.1016/j.jprot.2021.104435.

16. Anhe GF, Bordin S. The adaptation of maternal energy metabolism to lactation and its underlying mechanisms. Mol Cell Endocrinol. 2022; 553: 111697.

17. Bu D, Bionaz M, Wang M, Nan X, Ma L, *et al*. Transcriptome difference and potential crosstalk between liver and mammary tissue in mid-lactation primiparous dairy cows. PLoS One. 2017; 12(3): e0173082, DOI: 10.1371/journal.pone.0173082.

18. Bottini N, Bottini E, Gloria-Bottini F, Mustelin TBottini N, Bottini E, *et al.* Low-molecular-weight protein tyrosine phosphatase and human disease: in search of biochemical mechanisms. Arch Immunol Ther Exp. (Warsz.). 2002; 50(2): 95-104.

19. Stanford SM, Diaz MA, Ardecky RJ, Zou J, Roosild T, *et al*. Discovery of orally bioavailable purine-based inhibitors of the low-molecular-weight protein tyrosine phosphatase. J Med Chem. 2021; 64(9): 5645-5653, DOI: 10.1021/acs.jmedchem.0c02126.

**Cite this article as:** Venkata CKT, Singh S, Vedamurthy GV, Singh D, Onteru SK. Inter-species SNPs from buffalo liver transcriptome indicate diversity in immune pathways between buffalo and cattle. Explor Anim Med Res. 2025; 15(2), DOI: 10.52635/eamr/15.2.219-225.