Identification and Validation of Differentially Expressed Genes Via-s-vis Exploration of the Modular Pathways in Diseased Versus Healthy Nili Ravi Water Buffalo

Priyabrata Behera¹, Simarjeet Kaur¹, Shiva R Sethi² and Chandra Sekhar Mukhopadhyay*²

¹Department of Animal Genetics and Breeding, Guru Angad Veterinary and Animal Science University, Ludhiana, Punjab, India
²College of Animal Biotechnology, Guru Angad Veterinary, and Animal Science University, Ludhiana, Punjab, India

Abstract

Peripheral blood mononuclear cells (PBMCs) were isolated from 3 groups of she-buffaloes (Tuberculosis, Metritis, and Healthy control) was sequenced by RNA-Seq (using Illumina HiSeq 2500 platform). The pre-processed reads, obtained from transcriptome sequencing, were aligned to the Bostaurus genome using the Hisat-2 program. Gene expression was studied using the String Tie program. A total of 31982 transcripts were identified. Comparisons of the entire 3 groups' revealed 176 differentially expressed genes (DEGs) in TB vs. healthy groups and 162 DEGs in metritis vs. healthy groups. Analysis of gene ontology and pathways (molecular function and biological processes) identified certain pathways like cytokine activity, Wnt signaling, PI3K-Akt signaling, MAPK signalling (between TB and healthy groups) and cAMP signaling, Wnt signaling, TGF-beta signaling, MAPK signaling, PI3K-Akt signaling, etc. between metritis-positive and healthy buffaloes. Network analysis identified the immune-related genes contributing to the system biology related to the disease-resistance in Nili Ravi buffalo. Besides, five differentially expressed genes have been validated using SYBR-green chemistry of qPCR. In the future, these key genes could be studied in detail to explore their potential to be promising biomarkers for selecting breeding animals with higher tolerance against these economically devastating diseases.

Keywords: Transcriptome sequencing • RNA-seq • qPCR • Differentially expressed genes • Pathway analysis • Disease-resistance

Introduction

Buffalo has been an integral part of livestock farming and agriculture in Asia for over 5000 years and played a pivotal role in overall agro-socio-economic development through contributing meat, milk, hide, and draft power. India treasures about 57.3% of the world buffalo population (which is 109.85 million as per the 2019 national livestock census of India) [1]. Buffalo contributes 55% of the total milk production (185.4 million tonnes) in India [2]. Nili Ravi, an important bubaline breed in its natural breeding tract Punjab (India and Pakistan), is considered as a producer of milk with a high-fat content (7-8.5%), meat, and remarkable ability to utilize poor quality fodder.

Scientific knowledge on the incidence of clinical diseases in dairy animals allows identifying factors and assists in formulating the priorities of breeding and disease-prevention programs [3]. Genetic predisposition plays an important role in developing resistance to infectious diseases like brucellosis, hemorrhagic septicemia (HS), tuberculosis (TB), mastitis, rinderpest, Johne's disease (JD), etc. Besides, some of these diseases have a zoonotic impact on public health. A study on the occurrence of TB in cattle and humans (in India) reported that 15.7% of humans had Mycobacterium tuberculosis, 26.8% of cattle had M. bovis. However, 8.7% of humans had M. bovis TB and 36.7% of cattle had ‘mixed infections’ [4]. TB has no geographical boundaries and infection occurs in a diverse group of animals [5]. Similarly, postpartum metritis is one of the most important disorders in bovines, causing high economic losses due to prolonged days-open as well as inter-calving periods, leading to involuntary culling. The incidence rate of uterine infection in she-buffalo (24.7) is found to be much higher than in cows [6].

Global transcriptome analysis helps unravel all the transcripts available in a cell type or tissue. The massive parallel sequencing approach enables us to determine the differential expression profile of the genes between the experimental groups and to identify the significantly associated modular genes [7-9]. The detailed systems biology and related genes associated with metritis and tuberculosis are not known in buffalo [10,11]. Research endeavors directed towards discovering the underlying genes contributing to susceptibility/tolerance of Nili Ravi buffaloes to infectious diseases have been done been undertaken. Increasing evidence suggests that mRNA expression plays an important role in host-pathogen interactions through modulation of both innate and acquired immune responses [12,13]. The gene and miRNA expression profile in the PBMCs alters due to diseases [14] like TB and metritis. Considering the gap in knowledge on systems biology associated with disease resistance in Nili Ravi buffaloes, the present study was designed to discover the mRNA repertoire and explore the modular genes and pathways associated with resistance/susceptibility to tuberculosis (TB) and metritis in this economically important breed of water buffalo.

Materials and Methods

Sample collection and experimental design

Six adult female Nili Ravi water buffaloes (aged between 2 to 4 years),
maintained under similar management in the organized institutional dairy farm of Guru Angad Dev Veterinary and Animal Sciences University (GADVASU), Ludhiana, India were included in the study. These animals were divided into 3 groups (two biological replicates each): Tuberculosis positive (TB), Metritis positive (Met), and Healthy control (Ctrl) animals. Peripheral blood samples (10 ml) were aseptically collected from the jugular vein in tubes containing 500 µl of 0.5M EDTA and immediately brought to the laboratory to process for PBMCs isolation by density gradient centrifugation method [15,16]. We have used PBMC for comparison of RNA expression because the diseases have manifestations in the PBMCs due to their involvement in innate immunity. The gene expression profile in the PBMCs alters due to diseases like TB and metritis [17,18]. This research work was certified by the Institutional Animal Ethics Committee (IAEC). All the experiments and protocols were conducted at the College of Animal Biotechnology, GADVASU.

RNA extraction

The total RNA from all 6 samples was isolated using the mirVanaRNA isolation kit (Ambion, Life Technologies, USA). The RNA samples with OD (260/280) between 2.0 and 2.2 (assessed by Nanodrop, Thermo Fisher) were selected for further processing.

Transcriptomic sequencing

The total RNA samples were outsourced to Agri-Genome Labs Private Limited, Kochi, Kerala, India by maintaining proper cold-chain (using dry ice) during transit, for next-generation sequencing (NGS) using Illumina Hiseq 2500 platform (100 bp paired-end reads) and preliminary biocomputational analysis of the raw data.

Analysis of sequenced data

The raw data were pre-processed using the FASTX tool. About ~94% of the total reads of all the samples passed ≥30 Phred score (Table 1). The quality of the reads was checked for each of the samples using sequence quality score distribution, average base composition, average base quality, GC% distribution, check for over-represented sequences as well as quality score distribution, average base composition, average base quality, quality of the reads was checked for each of the samples using sequence analysis of the raw data.

The quality passed reads were subjected to adapter trimming and then the pre-processed reads were aligned to the taurine genome (UMD3.1) [19] downloaded from Ensembl database (ftp://ftp.ensembl.org/pub/release-89/). The raw data were pre-processed using the FASTX tool. About ~94% of the total reads of all the samples passed ≥30 Phred score (Table 1). The quality of the reads was checked for each of the samples using sequence quality score distribution, average base composition, average base quality, GC% distribution, check for over-represented sequences as well as biasing of kmers, and read-length distribution of forward and reverse reads (Table 2). The raw sequence reads in FASTQ format have been published through NCBI SRA (https://www.ncbi.nlm.nih.gov/sra/PRJNA514883) with the experiment accession numbers SRX2528354 to SRX2528359.

The quality passed reads were subjected to adapter trimming and then the pre-processed reads were aligned to the taurine genome (UMD3.1) [19] downloaded from Ensembl database (ftp://ftp.ensembl.org/pub/release-89/). The raw data were pre-processed using the FASTX tool. About ~94% of the total reads of all the samples passed ≥30 Phred score (Table 1). The quality of the reads was checked for each of the samples using sequence quality score distribution, average base composition, average base quality, GC% distribution, check for over-represented sequences as well as biasing of kmers, and read-length distribution of forward and reverse reads (Table 2). The raw sequence reads in FASTQ format have been published through NCBI SRA (https://www.ncbi.nlm.nih.gov/sra/PRJNA514883) with the experiment accession numbers SRX2528354 to SRX2528359.

Identification of differentially expressed genes (DEGs)

The average FPKM values of two biological replicates belonging to each of the experimental groups were calculated. The non-coding RNA genes (viz. miRNA, snRNA, and srRNA) were removed from the list of 31982 genes for each of the experimental groups. The differentially expressed coding mRNA genes were identified between the healthy groups with each of the diseased groups. The common and unique protein-coding genes among the experimental groups being compared and were represented by the Venn diagram using online web-tool IneractiVenn (http://www.interactivenn.net/) and R Program (v.3.2.0) package Venn Diagram v.1.6.9s. The fold change of expression was calculated by taking the logarithm (with base 2) of the ratio of expression values between the two groups (healthy versus diseased) being compared. The genes with fold change were considered as differentially expressed genes. Expression heatmaps (on the color scale bar; white-orange-red; representing the low-medium-high expression, respectively) of the mRNAs for the experimental samples were constructed using the WGCNA package [20] of R program (v.3.2.0).

Validation of selected DEGs using qPCR

Five differentially expressed genes were selected for validation of the obtained transcriptomic analysis results (primer details given in Table 3). The expression of these genes in the whole blood or leukocytes was checked from the Genecards database (www. genecards.org). The primers were designed using the available mRNA sequences in NCBI Nucleotide (https://www.ncbi.nlm.nih.gov/nuccore/) using primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) online tool. The specificity and sensitivity of the designed primers were checked by IDT Oligo Analyzer (v.3.1.1) (https://eu.idtdna.com/pages). Expression of selected differentially expressed genes was quantified (in three technical replicates) by real-timePCR (Bio-Rad) using a miScript SYBR-Green PCR kit (Qiagen, Germany) and beta-actin as an endogenous control.

Total RNA was isolated (for validation using qPCR) from PBMCs extracted from fresh blood samples (three biological replicates of each of the 3 experimental groups) using PureLink® RNA Minikit (Ambion) and then reverse-transcribed into cDNAs bymiScript II RT Kit (Qiagen, Germany)

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>Total Reads</th>
<th>QC Passed</th>
<th>QC Passed%</th>
<th>Aligned read-count</th>
<th>Aligned %</th>
<th>Unaligned Read-count</th>
<th>Unaligned %</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB_1</td>
<td>54149470</td>
<td>54114736</td>
<td>99.93</td>
<td>37084829</td>
<td>68.53%</td>
<td>17029807</td>
<td>31.47</td>
</tr>
<tr>
<td>TB_2</td>
<td>54476460</td>
<td>54417428</td>
<td>99.9</td>
<td>37388446</td>
<td>68.67%</td>
<td>17048890</td>
<td>31.33</td>
</tr>
<tr>
<td>Metritis_1</td>
<td>50614338</td>
<td>50574744</td>
<td>99.92</td>
<td>32590365</td>
<td>64.44%</td>
<td>17984379</td>
<td>35.56</td>
</tr>
<tr>
<td>Metritis_2</td>
<td>53517678</td>
<td>53418886</td>
<td>99.81</td>
<td>36688778</td>
<td>68.88%</td>
<td>18731108</td>
<td>31.32</td>
</tr>
<tr>
<td>Healthy_1</td>
<td>53294154</td>
<td>53238892</td>
<td>99.89</td>
<td>36191799</td>
<td>67.98%</td>
<td>17047093</td>
<td>32.02</td>
</tr>
<tr>
<td>Healthy_2</td>
<td>54652348</td>
<td>54600404</td>
<td>99.94</td>
<td>37434037</td>
<td>68.56%</td>
<td>17163637</td>
<td>31.44</td>
</tr>
</tbody>
</table>

Table 1. Quality scores of the 6 RNA samples subjected to transcriptomic sequencing.

Table 2. Quality checking and detail of read-counts of the 6 experimental samples subjected to paired-end sequencing using Illumina Hiseq 2500 platform.
Table 3. Detail of the primers used for validation of expression profile of the selected differentially expressed genes by real-time PCR (SYBR Green chemistry).

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5’-3’)</th>
<th>Length</th>
<th>NCBI Accession Number</th>
<th>Tm (°C)</th>
<th>GC%</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM7-F</td>
<td>ctcagcgctcttcctcttg</td>
<td>22</td>
<td>XM_006074849.1</td>
<td>60.09</td>
<td>54.55</td>
</tr>
<tr>
<td>TOMM7-R</td>
<td>ttagcggcacagctaaac</td>
<td>22</td>
<td>XM_006062969.1</td>
<td>59.96</td>
<td>50</td>
</tr>
<tr>
<td>SPTSSB-F</td>
<td>agctaggagagatcttggcgctg</td>
<td>22</td>
<td>XM_006062969.1</td>
<td>59.96</td>
<td>50</td>
</tr>
<tr>
<td>SPTSSB-R</td>
<td>cgctcaatctctccagcagga</td>
<td>22</td>
<td>59.9</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>MSHA-F</td>
<td>cgaaaaccttggcgacgtgatt</td>
<td>22</td>
<td>60.29</td>
<td>45.45</td>
<td></td>
</tr>
<tr>
<td>MSHA-R</td>
<td>tctctctactctctactctcttg</td>
<td>22</td>
<td>59.89</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

Genes for validation: Tuberculosis-positive vs. Healthy control groups

Genes for validation: Metritis-positive vs. Healthy control groups

MRPL33-F    | aagtggttggagttgctggc | 22     | XM_006046139.1        | 59.9    | 45.45 |
MRPL33-R    | ctagtctctcatagctggtctt | 22     | 60.09                | 50      |
GPC3-F      | ttcagccgcaagagggac | 22     | XM_006074811.1        | 60.09   | 50  |
GPC3-R      | ttccacctctgcgccttttg | 22     | 59.96                | 45.45   |

Endogenous control beta-actin

ACTB-F      | aagatattgccgcgctcgtg | 20     | NM_001290932.1        | 61.56   | 55   |
ACTB-R      | catcccccagctacagtc | 20     | 62.01                | 65      |

mRNAs were commonly detected in all three groups. The Fragments per Kilobase of transcript per Million mapped reads (FPKM) values of each of the six samples for all the 31982 genes have been tabulated in Supplementary Table S1. While comparative analysis between the diseased and healthy groups revealed uniquely expressed mRNAs that are likely to be associated with disease resistance or susceptibility viz., 411 mRNAs specific to Tuberculosis-positive group and 468 mRNAs to the healthy control group, 373 and 469 uniquely expressed mRNAs in Metritis-positive group and healthy control group, respectively. The number of unique genes of TB, metritis, and healthy groups were 231, 193, and 273, respectively. The detailed annotations of all 31982 genes, as well as the lists of all the differentially expressed vis-à-vis, uniquely expressed protein-coding genes are available in the Supplementary Tables S2 (TB vs. Healthy comparison) and S3 (Metritis vs. Healthy comparison). In total, 162 differentially expressed (up- and down-regulated) mRNA genes (DEGs) were identified in 176 DEGs (mRNA) in TB vs. healthy group (Supplementary Table S2a) and metritis vs. healthy group (Supplementary Table S3a). The expression profile of the protein-coding genes has been represented in the heat map (Figure 2). It indicated that treatment group 2 (i.e. metritis samples) along with a healthy control group exhibited the most similar mRNA expression profiles, while TB group exhibit a distinguished mRNA expression profile.

Identification of differentially expressed genes

The differentially expressed protein-coding genes were identified as up and down-regulated based on log2 (FPKM) >2 and <-2, respectively. The detailed annotations of all 31982 genes, as well as the lists of all the differentially expressed vis-à-vis, uniquely expressed protein-coding genes are available in the Supplementary Tables S2 (TB vs. Healthy comparison) and S3 (Metritis vs. Healthy comparison). In total, 162 differentially expressed (up- and down-regulated) mRNA genes (DEGs) were identified in 176 DEGs (mRNA) in TB vs. healthy group (Supplementary Table S2a) and metritis vs. healthy group (Supplementary Table S3a). The expression profile of the protein-coding genes has been represented in the heat map (Figure 2). It indicated that treatment group 2 (i.e. metritis samples) along with a healthy control group exhibited the most similar mRNA expression profiles, while TB group exhibit a distinguished mRNA expression pattern.

Results and Discussion

Uniquely and differentially expressed genes

A total of 31982 transcripts (Figure 1A) were identified which included 15772 protein-coding mRNAs (inclusive of multi-copy of some genes) and 16210 non-coding RNAs (miRNA, misc_RNA, Mt_rRNA, Mt_tRNA, rRNA, snoRNA, and snRNA), processed-pseudogenes, pseudogenes, and other unidentified transcripts. The Venn diagram (Figure 1B) depicts that 8132
which regulates the assembly and stability of the TOM complex [29]. It is genes revealed that the Translocase of Outer Mitochondrial Membrane TB as compared to the healthy group and SPTSSB was almost same in MSH4 was down-regulated (fold-change 1.21 and 0.83, respectively) in results of the qPCR analysis showed that TOMM7 was up-regulated while high bacterial load and weight loss leading to reduced survival [28]. The of the transcriptome (Supplementary Table S1). TB is characterized by DEGs (Figure 3). The selected 5 differentially expressed genes (3 genes were analyzed for fold-change using the ddCt method [27]. The FPKM values of MSH4 and TOMM7 genes showed up-regulation (mean FPKM values 0.19 and 15.67, respectively) and SPTSSB was found to be down-regulated (mean FPKM value 0.00) in Tuberculosis samples as compared to the healthy ones (0.02, 0.00 and 0.71, respectively, for MSH4, TOMM7, and SPTSSB), as obtained from the Illumina next generations sequencing of the transcriptome (Supplementary Table S1). TB is characterized by high bacterial load and weight loss leading to reduced survival [28]. The results of the qPCR analysis showed that TOMM7 was up-regulated while MSH4 was down-regulated (fold-change 1.21 and 0.83, respectively) in TB as compared to the healthy group and SPTSSB was almost same in both experimental groups. Ontology study and functional analysis of these genes revealed that the Translocase of Outer Mitochondrial Membrane 7 (TOMM7) gene encodes a subunit-peptide of the translocase enzyme, which regulates the assembly and stability of the TOM complex [29]. It is well known that host-mitochondria are snipped-targets of several bacteria including *Mycobacterium* [30,31]. The MutS Homolog 4 (MSH4) gene encodes a meiosis-specific protein that forms a heterodimer with MSH5 to bind to a Holliday Junction and thereby provoke ADP-ATP exchange. Mutation of this protein is associated with altered or mismatched DNA-binding during Meiosis-I thus affecting the cell-cycle. Host-cell death is caused by the host-pathogen interaction to prevent host immunity to overcome the *Mycobacterium*-infection and promoting host cell necrosis [32]. It has been propounded that *Mycobacterium* promotes programmed necrosis of host-cells and thwarts host cell apoptotic signalling [33]. The up-regulation of TOMM7 and down-regulation of the MSH4 gene in the TB-positive samples are logical and also validated in our qPCR study. The Serine Palmityltranseferase Small Subunit B (SPTSSB) gene encodes a peptide that acts as a small SPT subunit associated with SPT activity and also involved in conferring acyl-CoA preference to the SPT catalytic heterodimer of SPTLC1 [34]. Although this gene was selected based on its expression in the blood cells (so that we can easily detect in the PBMCs) available at the Gene Cards database (www.genecards.org), it seems SPTSSB gene has no direct association with resistance or susceptibility to *Mycobacterium* infection. So the TB-positive sample and the healthy sample didn't show a similar pattern of expression as evident in the NGS data.

The other two genes that were selected for validation of Metritis-positive vs. healthy groups were GPC3 (down-regulated in Metritic samples) and MRPL33 (up-regulated in Metritic samples), as presented in the transcriptome sequencing data. The MRPL33 has shown up-regulation in the Metritic sample in the qPCR validation experiment. Mitochondrial Ribosomal Protein L33 (MRPL33) is associated with mitochondrial protein synthesis, viral gene translation in infected cells, and organelle biogenesis [35]. In our validation study, the MRPL33 was found to be more expressed (Fold-change 1.18) in the PBMCs of metritis-positive samples. Perhaps, the bacterial infection has stimulated the expression of this gene. Finally, the GPC3h as show almost the same level of expression in both groups (Fold-change 0.99). This gene encodes Glycianic 3 protein that forms the membrane-associated protein core of cell surface heparin sulphate proteoglycan. This protein is associated with apoptosis-induction, modulation of growth in mesodermal tissues, and interactions between IGF2 and its receptor [36]. Thus this gene can also be associated with inflammatory responses in the uterine cells. In a recent study on genome-wide association study for identifying the DEGs expressed in the endometrium of Holstein Cows, Stephen et al. identified GPC3 to be expressed poor-fertile cows (logFC=1.61, P=0.03) [37]. There is a difference in transcriptomics expression profile (revealed by NGS) and qPCR results of the selected 5 genes. This could be attributed to using new blood samples for qPCR. Research reports suggest that these genes are involved in the disease regulation pathways [38].

### Gene ontology and functional classification

Gene ontology (GO) for biological processes, molecular functions, cellular component, pathway, and protein class of differentially and uniquely expressed genes in the comparison of diseased vs. healthy sample groups was studied using DAVID, Pather D Band String DB. During a comparison between the tuberculosis-positive and the metritis-positive animals with the healthy group, in total, 1055 and 1004 protein-coding DEGs were identified, respectively, for functional annotations using DAVID. The Ensembl Gene Ids were uploaded and 896 (for TB vs. Healthy) and 821 (Metritis vs. Healthy) genes from Bostaurus species were selected for further analysis.

### Tuberculosis-positive vs. healthy groups

Each of the enriched GO term with a Bonferroni p-value less than 0.01 enriched gene ontology (GO) terms belonged to the following categories:749 (83.6%), 616 (68.8%) and 651 (72.7%) genes were incorporated into the cellular component (CC), molecular function (MF) and biological process (BP), respectively. Hypergeometric distribution was used to find the significance of enrichment and each enriched GO term (Bonferroni p-value <0.01) was adjusted for multiple testing [39]. The hypergeometric distribution calculates the likelihood that the annotation of GO terms to the section of interacting genes is significantly higher than that of the fraction of all the genes in the taurine genome (http://ctdbase.org/)
In total 41 genes were identified contributing to these pathways.

The Protein-Ids of these genes were extracted using g:Profiler and then subjected to String DB for generating protein-protein interaction (PPI) network (PPI enrichment p-value <1.0, e-16 and mean local clustering coefficient=0.458). Available literature indicates that these pathways are very critical in susceptibility or resistance to tuberculosis in humans and mice.

Panther analysis of the DEGs (between TB vs. healthy) identified a total of 2255 biological process hits were predicted which was categorized into 31 groups (including cellular process, metabolic process, biological regulation, response to a stimulus, immune system process, etc) (Figure 4).

It is evident from the protein-protein interaction network that insulin-like growth factor 1 (IGF1) forms a strong network (high combined score) with several proteins, like DCN, TEK, SLC24, CSF3, CSF2, MAPK10, IL13, etc. (Supplementary Table S8). Besides, CXCL14, GJA1, DCN, CCL11, CCL25 reveals strong interaction with several genes (Figure 5). The co-expression data shows that many genes with immune-related direct functions (IGF1 with DCN, SLC24, TEK; CSF1 and CSF2, etc.) are co-expressed. Network analysis using String DB identified the immune-related genes (viz, CSF2, CSF3, CXCL14, etc.) contributing to the system biology related to the disease in Nili Ravi buffalo. The gene interaction network of DEGs was studied with a list of combined scores of gene pairs. The combined network enables a topological characterization of the reliability strength of gene associations. A detailed and thorough investigation of the ontology and functional analysis of the 52 genes incorporated in the protein-protein interaction study revealed that these genes are actively involved in chemokine related functions, immunity, binding of molecules, etc. For example, Colony-stimulating factor 2 and 3 (CSF2 and 3: involved in cytokine activity; granulocyte-macrophage colony-stimulating factor receptor binding; growth factor activity), Interleukin-13 (IL-13: cytokine activity; interleukin receptor binding), CCL11, CCL25 reveals strong interaction with several genes (Figure 5).

In total 41 genes were identified contributing to these pathways. The Protein-Ids of these genes were extracted using g: Profiler and then subjected to String DB for generating protein-protein interaction (PPI) network (PPI enrichment p-value <1.0, e-16 and mean local clustering coefficient=0.458). Available literature indicates that these pathways are very critical in susceptibility or resistance to tuberculosis in humans and mice.
Figure 4. Pie chart of gene ontology of the biological process of the differentially and uniquely expressed protein-coding genes between tuberculosis-positive and healthy groups.

Figure 5. Protein-protein interaction network analyzed by selecting the genes that are intricately involved in the important pathways contributing to resistance or susceptibility to tuberculosis.
infection and this pathway is critical to confer innate immunity against 
M. bovis BCG stimulation [58].

Functional annotation clustering was performed by DAVID using 829

taurine genes (selected by DAVID) under classification stringency and
default threshold parameters. The results showed 61 clusters with the
highest Enrichment Score of 4.93 (Supplementary Table S12).

Panther analysis of the DEGs for Metritis-positive and healthy control
groups detected 1341 process hits (falling in 13 categories) (Figure 6).

The protein-protein interaction network was constructed using String
DB with 29 genes (protein-Ids converted by g:Profiler) selected from

**Figure 6.** Pie chart of gene ontology of the biological process of the differentially and uniquely expressed protein-coding genes between metritis-positive and healthy groups.

mechanisms, viz. activation of the innate inflammatory response which further induces the NF-κB and MAPK pathways to affect the release of pro-inflammatory cytokines. Besides, TNF-α can contribute to the production of pro-inflammatory cytokines by activating the intracellular IkBα and JNK signaling pathways through TLRs [61-63].

Conclusion

In this study, we have identified the differentially expressed (Log2 FPKM ratio >2 and <2) mRNA genes between disease (Tuberculosis, Metritis) versus healthy (control) Nili Ravi buffaloes. The gene ontology and pathway analysis associated with the diseased and control group revealed that cytokine activity, transforming growth factor-beta receptor binding, Wnt signaling, Cytokine-cytokine receptor interaction, PI3K-Akt signaling, MAPK signaling and GMP-PKG signaling pathways related genes play important role in TB susceptibility. Similarly, the important pathways associated with metritis-susceptibility are CAMP signaling, Neuroactive ligand-receptor interaction, Wnt signaling, TGF-beta signaling, MAPK signaling, PI3K-Akt signaling. Further studies aimed at targeting these modular genes could be used effectively in the future for marker-assisted selection for better tolerance against these economically devastating diseases in buffaloes.

Acknowledgments

The work is supported by the state-sponsored research project ONBS-01 (One Nucleus Breeding System to improve Sahiwal Cattle and Nili Ravi Buffaloes in the State of Punjab).

References

2. https://www.dahd.nic.in/sites/default/files/Tables%20of%20BAH&FS%202017%20(1).pdf


