The Transcriptome Profile of a Turkey B-cell Line Upon Infection With Turkey Hemorrhagic Enteritis Virus

Abraham Quaye[†,a], Brett E. Pickett[a], Joel S. Griffitts[a], Bradford K. Berges[a], Brian D. Poole[a,*]

[a]Department of Microbiology and Molecular Biology, Brigham Young University

[†]Primary Author

[*]Corresponding Author

Corresponding Author Information

brian_poole@byu.edu

Department of Microbiology and Molecular Biology,

4007 Life Sciences Building (LSB),

Brigham Young University,

Provo, Utah

15 # ABSTRACT

INTRODUCTION

Turkey hemorrhagic enteritis virus (THEV), belonging to the family *Adenoviridae*, genus *Siadenovirus*, infects turkeys, chickens, and pheasants (1, 2). Infecting its hosts via the feco-oral route, THEV causes hemorrhagic enteritis (HE) in turkeys, a debilitating disease affecting predominantly 6-12-week-old turkey poults characterized by immunosuppression (IMS), depression, splenomegaly, intestinal lesions leading to bloody droppings, and up to 80% mortality (3–6). The clinical disease usually persists in affected flocks for about 7-10 days. However, secondary bacterial infections may extend the duration of illness and mortality for an additional 2-3 weeks due to the immunosuppressive nature of the virus, exacerbating the economic losses (5, 7). Low pathogenic (avirulent) strains of THEV have been isolated, which show subclinical infections but retain the immunosuppressive effects. Since its isolation from a pheasant spleen, the Virginia Avirulent Strain (VAS) has been used effectively as a live vaccine despite the immunosuppressive side-effects, but the vaccinated birds are rendered more susceptible to opportunistic infections and death than unvaccinated cohorts leading to significant economic losses (4, 5, 8–10).

It is well-established that THEV primarily infects and replicates in turkey B-cells of the bursa and spleen and somewhat in macrophages, inducing apoptosis and necrosis. Consequently, a significant drop in number of B-cells (specifically, IgM+ B-cells) and macrophages ensue along with increased T-cell counts with abnormal T-cell subpopulation (CD4+ and CD8+) ratios. The cell death seen in the B-cells and macrophages is generally proposed as the major cause of THEV-induced IMS as both humoral and cell-mediated immunity are impaired (5, 6, 8, 11). It is also thought that the virus replication in the spleen attracts T-cells and peripheral blood macrophages to the spleen where the T-cells are activated by cytokines from activated macrophages and vice versa. The activated T-cells undergo clonal expansion and secrete interferons: type I (IFN-α and IFN-β) and type II (IFN-γ) as well as tumor necrosis factor (TNF) while activated macrophages secrete interleukin 6 (IL-6), TNF, and nitric oxide (NO), an antiviral agent with immunosuppressive properties. The inflammatory cytokines released by T-cells and macrophages (e.g., TNF and IL-6) may also induce apoptosis in bystander splenocytes, exacerbating the already numerous apoptotic and necrotic splenocytes, culminating in IMS (8, 11) (see **Figure 1**). However, the precise molecular mechanisms of THEV-induced IMS or pathways involved are poorly understood (6). Elucidating the specific mechanisms and pathways of THEV-induced IMS is the most crucial step in THEV research as it will present a means of mitigating the IMS.

Next generation sequencing (NGS) is a groundbreaking technology that has significantly enhanced our understanding of DNA and RNA structure and function and facilitated exceptional advancements in all domains of biology and the Life Sciences (12). mRNA sequencing (RNA-seq), an NGS approach to transcriptomic studies, is a versatile, high throughput, and cost-effective technology that allows a broad scan

of the entire transcriptome, thereby uncovering the active genes and molecular pathways and processes. This technology has been leveraged in an ever increasing number of studies to elucidate active cellular processes under a wide range of treatment conditions, including the transcriptomics of viral infections (12–16). In RNA-seq studies, differentially expressed genes (DEGs) identified under different experimental conditions are key to unlocking the interesting biology or mechanism under study. Identified DEGs are typically used for functional enrichment analysis in large curated knowledgebases which connect genes to specific biological processes, functions, and pathways such as gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, shedding light on the biological question under study (17, 18).

To the best of our knowledge, no study has leveraged the wealth of information offered by RNA-seq to elucidate the molecular mechanisms and pathways leading to THEV-induced IMS. To effectively counteract the immunosupressive effect of the vaccine, it is essential to unravel the host mechanisms/pathways influenced by the virus to bring about IMS. In this study, we present the first transcriptomic profile of a THEV infection using paired-end RNA-seq in a turkey B-cell line (MDTC-RP19), highlighting key host genes, cellular/molecular processes and pathways affected during a THEV infection. Our RNA-seq yielded 149 bp long high quality (mean PHRED Score of 36) sequences from each end of cDNA fragments, which were mapped to the genome of domestic turkey (*Meleagris gallopavo*).

RESULTS

Sequencing Results

To identify the host transcriptome profile during THEV infection, MDTC-RP19 cells were THEV-infected or mock-infected in triplicates or duplicates, respectively, and collected in like manner at 4-, 12-, 24-, and 72-hours post infection (hpi). mRNAs extracted from mock- or THEV-infected cells were sequenced on the Illumina platform, yielding a total of **776.1** million raw reads (149 bp in length) across all samples (statistics for the sequencing reads obtained from each RNA library are presented in **Table 1**). After trimming off low-quality reads, the remaining **742.8** million total paired-end trimmed reads (approximately, **34.7**-**47.9** million reads per sample) were mapped to the genome of *Meleagris gallopavo* obtained from the National Center for Biotechnology Information (NCBI). The percentage of reads mapping to the host genome across all samples ranged from **32.4**-**89.2**%. Although our sequencing reads have excellent quality scores (see **Table 1**) at all time points, the DEGs identified at 4- and 72-hpi did not yield any results in the functional enrichment analyses (i.e, GO term and KEGG pathway analysis); hence, they were excluded from all subsequent analyses. In the remaining samples from 12- and 24-hpi, there is a high correlation was seen between biological replicates (**Figure 2A** and **B**)

DEGs of THEV-infected Versus Mock-infected Cells

Gene expression levels were estimated with the StringTie software (19) in Fragments per kilobase of transcript per million (FPKM) units. The analysis of DEGs was performed with the DESeq2 R package (20) which employs negative binomial distribution model for read count comparisons. Using a $P_{adjusted}$-value cutoff ≤ 0.05 as the inclusion criteria, a total of **2,343** and **3,295** genes were identified as differentially expressed at 12-hpi and 24-hpi, respectively. ~~DEG analyses results at 12 and 24-hpi are presented in **Supplementary Tables 1 and 2**, respectively~~. At 12-hpi, **1,079** genes were upregulated and **1,264** genes downregulated, whereas **1,512** genes were upregulated and **1,783** genes downregulated at 24-hpi (**Figure 2C**,and **Figure 3A-C**). The $\log_2$fold-change(FC) values at 12-hpi ranged between **-1.4** and +**1.7** for **TMEM156** (Transmembrane Protein 156) and **LIPG** (Lipase G), respectively. At 24-hpi, the $\log_2$FC values ranged between **-2.0** and +**2.6** for **C1QTNF12** (C1q And TNF Related 12) and **KCNG1** (Potassium Voltage-Gated Channel Modifier Subfamily G Member 1), respectively.

Functional Enrichment Analyses (GO, KEGG pathway, and interaction network analyses)

Gene ontology (GO) enrichment analysis was performed for 12- and 24-hpi DEGs with the gprofiler2 R package (21), which outputs results in three GO categories – cellular components, biological processes, and molecular functions. Results with $P_{adjusted}$-value ≤ 0.05 were considered as functionally enriched. GO

enrichment analyses at 12-hpi and 24-hpi showed significant intersections of results that were enriched among all three GO categories (ref tables/figures here).

At 12-hpi, The primary molecular functions of the DEGs are mainly enriched into binding (44.5%), catalytic activity (29.9%), transcription regulator activity (7.6%), molecular function regulation (7.6%), transporter activity (6.6%), molecular transducer activity (1.4%), and structural molecule activity (1.4%) (Fig. 2A). The distribution patterns of the up- regulated or down-regulated genes are similar to those of the whole DEGs (Fig. 2B and C). As for the cellular component, DEGs were roughly related to endoplasmic reticulum chaperone complex, melanosome, pigment granule, lewy body core and protein complexes (Fig. 3A and Fig. 4). Based on the biological processes, the GO terms were mainly summarized into the following categories, namely, defense response to virus, response to stress, interferon signaling pathway, regulation of viral process, immune response, metabolism, stimulus, apoptosis and protein catabolic process (Figs. 3B and 4).

~~Discuss enrichment analysis plots and tables here~~

DISCUSSION

CONCLUSIONS

MATERIALS AND METHODS

Cell culture and THEV Infection

The Turkey B-cell line (MDTC-RP19, ATCC CRL-8135) was grown as a suspension culture in 1:1 complete Leibovitz's L-15/McCoy's 5A medium with 10% fetal bovine serum (FBS), 20% chicken serum (ChS), 5% tryptose phosphate broth (TPB), and 1% antibiotic solution (100 U/mL Penicillin and 100μg/mL Streptomycin), at 41^{o}C in a humidified atmosphere with 5% CO_2. Infected cells were maintained in 1:1 serum-reduced Leibovitz's L15/McCoy's 5A media (SRLM) with 2.5% FBS, 5% ChS, 1.2% TPB, and 1% antibiotic solution. A commercially available THEV vaccine was purchased from Hygieia Biological Labs (VAS strain). The stock virus was titrated using an in-house qPCR assay with titer expressed as genome copy number (GCN)/mL, similar to Mahshoub *et al* (22). Cells were THEV-infected or mock-infected in triplicates or duplicates, respectively at a multiplicity of infection (MOI) of 100 GCN/cell, incubated at 41^{o}C for 1 hour, and washed three times with phosphate buffered saline (PBS) to get rid of free virus particles. At each time point (4-, 12-, 24-, and 72-hpi), triplicate (THEV-infected) and duplicate (mock-infected) samples were harvested for total RNA extraction.

RNA extraction and Sequencing

Total RNA was extracted from infected cells using the Thermofisher RNAqueous™-4PCR Total RNA Isolation Kit (which includes a DNase I digestion step) per manufacturer's instructions. An agarose gel electrophoresis was performed to check RNA integrity. The RNA quantity and purity was initially assessed using nanodrop, and RNA was used only if the A260/A280 ratio was 2.0 ± 0.05 and the A260/A230 ratio was >2 and <2.2. Extracted total RNA samples were sent to LC Sciences, Houston TX for poly-A-tailed mRNA sequencing. RNA integrity was checked with Agilent Technologies 2100 Bioanalyzer High Sensitivity DNA Chip and poly(A) RNA-seq library was prepared following Illumina's TruSeq-stranded-mRNA sample preparation protocol. Paired-end sequencing, generating 150 bp reads was performed on the Illumina NovaSeq 6000 sequencing system. The paired-end 150bp sequences obtained during this study and all expression data have been submitted to the Gene Expression Omnibus database, under accession no #######

Quality Control and Mapping Process

Sequencing reads were processed following a well-established protocol described by Pertea *et al* (19), using Snakemake - version 7.32.4 (23), a popular workflow management system to drive the pipeline. Briefly, raw sequencing reads were trimmed with Cutadapt - version 1.10 (24) and the quality of trimmed reads evaluated using the FastQC software, version 0.12.1 (Bioinformatics Group at the Babraham Institute, Cambridge, United Kingdom; www.bioinformatics.babraham.ac.uk), achieving an overall Mean Sequence Quality (PHRED Score) of 36. Trimmed reads were mapped the reference *Meleagris gallopavo* genome

(https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/146/605/GCF_000146605.3_Turkey_5.1/GCF_000146605.3_Turkey_5.1_genomic.fna.gz) with Hisat2 - version 2.2.1 (19) using the accompanying gene transfer format (GTF) annotation file (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/146/605/GCF_000146605.3_Turkey_5.1/GCF_000146605.3_Turkey_5.1_genomic.gtf.gz) to build a genomic index. Samtools - version 1.19.2 was used to convert the output Sequence Alignment Map (SAM) file to the more manageable Binary Alignment Map (BAM) format. The StringTie (v2.2.1) software (19), set to expression estimation mode was used to generate normalized gene expression estimates from the BAM files for genes in the reference GTF file after which the prepDE.py3 script was used to extract read count information from the StringTie gene expression files, providing an expression-count matrix for downstream DEG analysis.

DEG Analysis and Functional Enrichment Analysis

DEG analysis between mock- and THEV-infected samples was performed using the very popular DESeq2 (20), which employs a Negative Binomial distribution model for read count comparisons. Genes with $P_{adjusted}$-value ≤ 0.05 were considered as differentially expressed. The read count data are deposited at Gene Expression Omnibus (GEO) under accession number ###. The functional profiling of DEGs (GO and KEGG analyses) were performed based on GO databases and KEGG databases using the R package gprofiler2 (21) with *Meleagris gallopavo* as the reference organism. Results with $P_{adjusted}$-value ≤ 0.05 were included as functionally enriched. ~~Additionally, the DAVID analysis tool was used for KEGG pathway analysis~~. All visualization plots were made using ggplot2, pheatmap, and ggvenn R packages (25–27). Venn diagram —

Validation of DEGs by Reverse Transcriptase Quantitative PCR (RT-qPCR)

Statistical Analysis

DATA AVAILABILITY

CODE AVAILABILITY

ACKNOWLEDGMENTS

164 # REFERENCES

165 1. Harrach B. 2008. Adenoviruses: General features, p. 1–9. *In* Mahy, BWJ, Van Regenmortel, MHV (eds.), Encyclopedia of virology (third edition). Book Section. Academic Press, Oxford.

166 2. Davison A, Benko M, Harrach B. 2003. Genetic content and evolution of adenoviruses. The Journal of general virology 84:2895–908.

167 3. Gross WB, Moore WE. 1967. Hemorrhagic enteritis of turkeys. Avian Dis 11:296–307.

168 4. Beach NM. 2006. Characterization of avirulent turkey hemorrhagic enteritis virus: A study of the molecular basis for variation in virulence and the occurrence of persistent infection. Thesis.

169 5. Dhama K, Gowthaman V, Karthik K, Tiwari R, Sachan S, Kumar MA, Palanivelu M, Malik YS, Singh RK, Munir M. 2017. Haemorrhagic enteritis of turkeys – current knowledge. Veterinary Quarterly 37:31–42.

170 6. Tykałowski B, Śmiałek M, Koncicki A, Ognik K, Zduńczyk Z, Jankowski J. 2019. The immune response of young turkeys to haemorrhagic enteritis virus infection at different levels and sources of methionine in the diet. BMC Veterinary Research 15.

171 7. Pierson F, Fitzgerald S. 2008. Hemorrhagic enteritis and related infections. Diseases of Poultry 276–286.

172 8. Rautenschlein S, Sharma JM. 2000. Immunopathogenesis of haemorrhagic enteritis virus (HEV) in turkeys. Dev Comp Immunol 24:237–46.

173 9. Larsen CT, Domermuth CH, Sponenberg DP, Gross WB. 1985. Colibacillosis of turkeys exacerbated by hemorrhagic enteritis virus. Laboratory studies. Avian Dis 29:729–32.

174 10. Beach NM, Duncan RB, Larsen CT, Meng XJ, Sriranganathan N, Pierson FW. 2009. Persistent infection of turkeys with an avirulent strain of turkey hemorrhagic enteritis virus. Avian Diseases 53:370–375.

175 11. Rautenschlein S, Suresh M, Sharma JM. 2000. Pathogenic avian adenovirus type II induces apoptosis in turkey spleen cells. Archives of Virology 145:1671–1683.

176 12. Satam H, Joshi K, Mangrolia U, Waghoo S, Zaidi G, Rawool S, Thakare RP, Banday S, Mishra AK, Das G, Malonia SK. 2023. Next-generation sequencing technology: Current trends and advancements. Biology 12:997.

177 13. Pandey D, Onkara Perumal P. 2023. A scoping review on deep learning for next-generation RNA-seq. Data analysis. Functional & Integrative Genomics 23.

178 14. Wang B, Kumar V, Olson A, Ware D. 2019. Reviving the transcriptome studies: An insight into the emergence of single-molecule transcriptome sequencing. Frontiers in Genetics 10.

179 15. Choi SC. 2016. On the study of microbial transcriptomes using second- and third-generation sequencing technologies. Journal of Microbiology 54:527–536.

180 16. Mo Q, Feng K, Dai S, Wu Q, Zhang Z, Ali A, Deng F, Wang H, Ning Y-J. 2023. Transcriptome profiling highlights regulated biological processes and type III interferon antiviral responses upon crimean-congo hemorrhagic fever virus infection. Virologica Sinica 38:34–46.

181 17. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: Tool for the unification of biology. Nature Genetics 25:25–29.

182 18. Kanehisa M. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research 28:27–30.

183 19. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and ballgown. Nature Protocols 11:1650–1667.

184 20. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.

185 21. Kolberg L, Raudvere U, Kuzmin I, Vilo J, Peterson H. 2020. gprofiler2– an r package for gene list functional enrichment analysis and namespace conversion toolset g:profiler. F1000Research 9 (ELIXIR).

186 22. Mahsoub HM, Evans NP, Beach NM, Yuan L, Zimmerman K, Pierson FW. 2017. Real-time PCR-based infectivity assay for the titration of turkey hemorrhagic enteritis virus, an adenovirus, in live vaccines. Journal of Virological Methods 239:42–49.

187 23. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, Forster J, Lee S, Twardziok SO, Kanitz A, Wilm A, Holtgrewe M, Rahmann S, Nahnsen S, Köster J. 2021. Sustainable data analysis with snakemake. F1000Research 10:33.

188 24. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17:10.

189 25. Wickham H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org.

190 26. Kolde R. 2019. Pheatmap: Pretty heatmaps. https://CRAN.R-project.org/package=pheatmap.

191 27. Yan L. 2023. Ggvenn: Draw venn diagram by 'ggplot2'. https://CRAN.R-project.org/package=ggvenn.

TABLES AND FIGURES

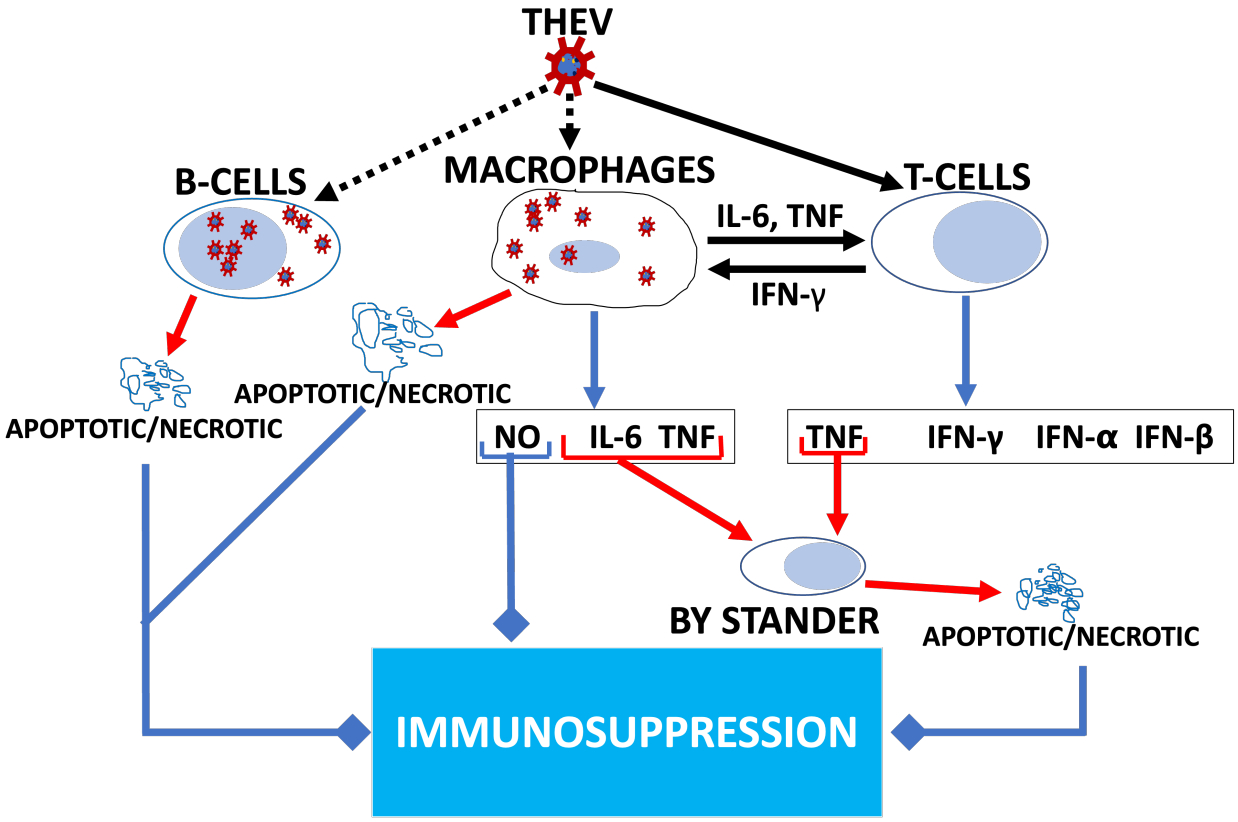

Figure 1: Model of THEV-induced immunosuppression in turkeys. THEV infection of target cells is indicated with black dotted arrows. Black unbroken arrows indicate cell activation. Red arrows indicated signals leading to apoptosis. Blue arrows indicate all cytokines released by the cell. Blue arrows with square heads indicated an event leading to IMS. Adapted from Rautenschlein *et al.* (8).

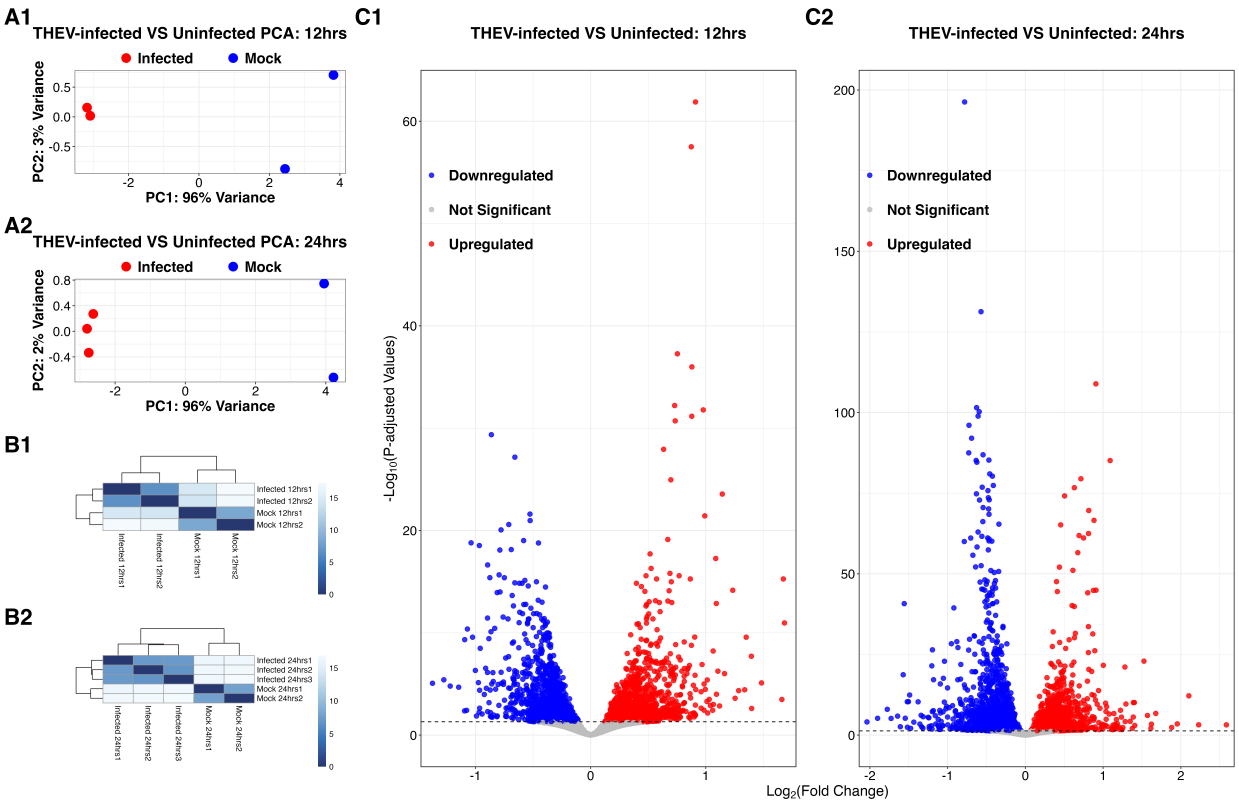

Figure 2. (A) Principal component analysis (PCA) of turkey B-cells during THEV infection. At 12-hpi **(A1)**,the results indicate that the first (PC1) and second (PC2) principal components account for 96% and 3% of the variation in the samples, respectively. Whereas PC1 and PC2 account for 96% and 2% of the variation, respectively at 24-hpi **(A2)**. **(B) Poisson distance matrices illustrating the RNA-seq library integrity within treatment (infected versus mock) groups**. The color scale represents the distances between biological replicates for both 12-hpi samples **(B1)** and 24-hpi samples **(B2)**. Dark colors represent high correlation (similarity) between the samples involved. **(C) Volcano plots of DEGs between THEV-infected versus mock-infected cells at 12- and 24-hpi**. Red, blue, and grey dots represent upregulated, downregulated, and non-significant genes, respectively for both 12-hpi samples **(C1)** and 24-hpi samples **(C2)**.

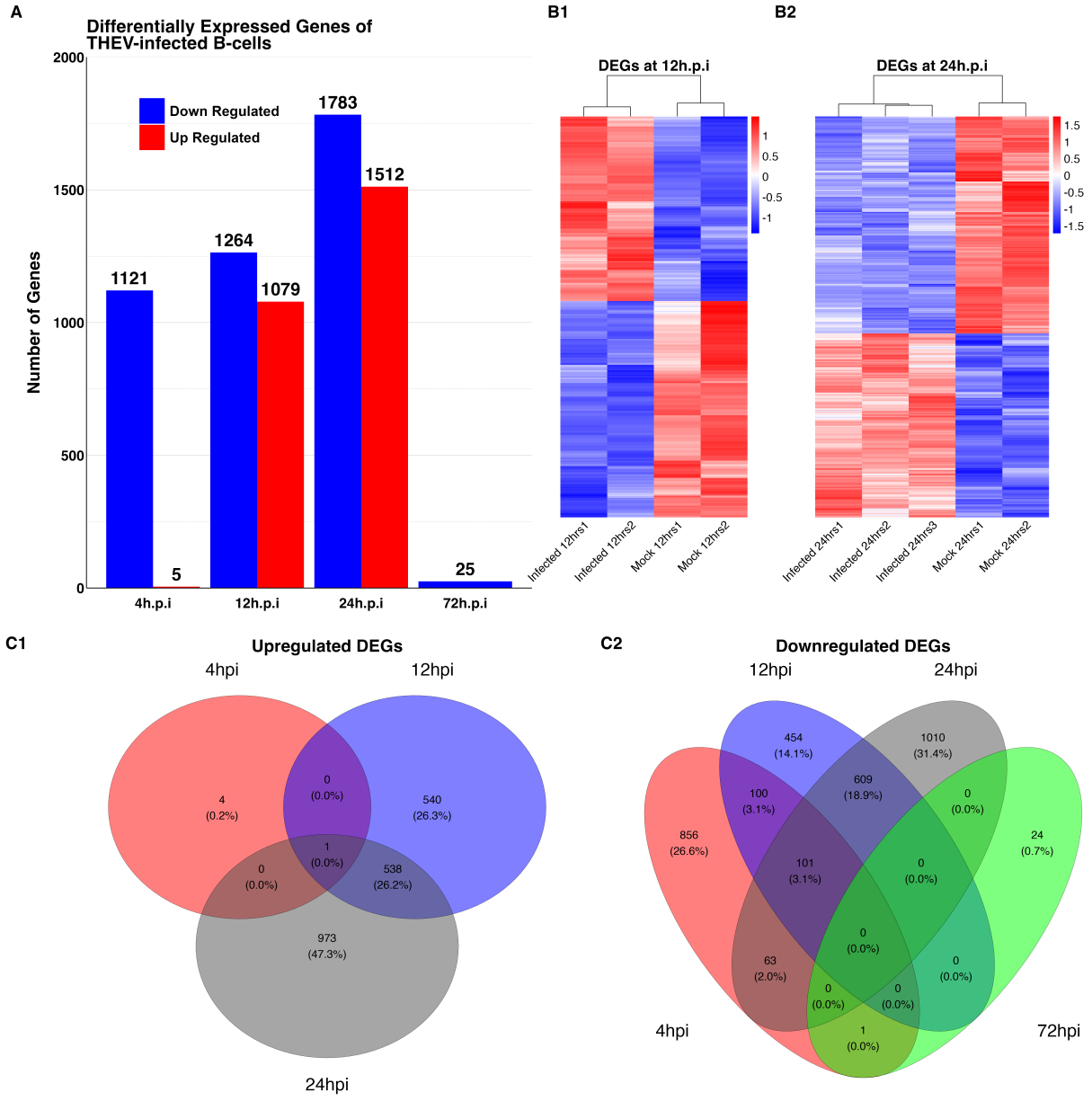

Figure 3: Differentially expressed genes (DEGs) of THEV-infected versus mock-infected samples at different time points. (A) Bar plot of number DEGs identified. Red represents upregulated genes and blue represents downregulated genes. **(B) Heatmaps of scaled expression data (Z-scores) of DEGs.** DEGs identified at 12-hpi are shown in **(B1)** and DEGs at 24-hpi in **(B2). (C) Venn diagrams showing the number of DEGs identified at different time points.** For the upregulated genes **(C1)**, the red circle represents genes at 4-hpi, the blue circle, 12-hpi, and the grey circle, 24-hpi. For the downregulated genes **(C2)**, the green circle represents genes at 72-hpi, while all the other time points retain the colors from **(C1)**.

Table 1: Summary of sequencing, quality control, and mapping processes

Sample	Raw Reads^{M}	Trimmed Reads^{M}	Mapped Reads^{M}	Uniquely Mapped Reads^{M}	Non-uniquely Mapped Reads^{M}	Q20%	Q30%	GC Content (%)
I_12hrsS1^{Inf}	40.6	39.0	34.7 (88.92%)	33.1 (84.78%)	1.6 (4.14%)	99.95	97.23	47.5
I_12hrsS3^{Inf}	38.8	37.3	33.1 (88.78%)	31.7 (84.95%)	1.4 (3.83%)	99.95	97.53	47.5
I_24hrsS1^{Inf}	42.7	41.0	36.2 (88.13%)	34.5 (84.2%)	1.6 (3.93%)	99.95	96.95	46.5
I_24hrsS2^{Inf}	42.0	40.4	35.6 (88.1%)	33.9 (83.83%)	1.7 (4.27%)	99.94	97.05	46.5
I_24hrsS3^{Inf}	40.5	38.9	34.2 (88.01%)	32.7 (84.12%)	1.5 (3.89%)	99.95	97.08	47.0
I_4hrsS1^{Inf}	39.1	37.4	33 (88.16%)	31.2 (83.43%)	1.8 (4.73%)	99.93	97.04	48.5
I_4hrsS2^{Inf}	41.3	39.6	35.3 (89.24%)	33.6 (84.92%)	1.7 (4.33%)	99.95	97.15	47.0
I_4hrsS3^{Inf}	41.5	39.8	35.5 (89.2%)	33.2 (83.29%)	2.4 (5.91%)	99.95	97.11	47.5
I_72hrsS1^{Inf}	41.2	39.8	28.3 (71.09%)	26.9 (67.7%)	1.3 (3.38%)	99.96	97.23	44.5
I_72hrsS2^{Inf}	39.3	38.0	27 (71.11%)	25.8 (67.86%)	1.2 (3.25%)	99.96	97.34	44.5
I_72hrsS3^{Inf}	39.9	37.1	28.3 (76.36%)	26.1 (70.3%)	2.2 (6.05%)	99.87	96.14	52.5
U_12hrsN1^{Mk}	42.1	40.4	35.9 (88.72%)	34.1 (84.39%)	1.7 (4.33%)	99.95	97.04	47.5
U_12hrsN2^{Mk}	41.0	39.3	34.7 (88.4%)	33.2 (84.53%)	1.5 (3.86%)	99.94	97.08	47.5
U_24hrsN1^{Mk}	38.4	37.0	32.7 (88.46%)	31.4 (84.74%)	1.4 (3.72%)	99.96	97.48	47.5
U_24hrsN2^{Mk}	39.9	38.4	34 (88.58%)	32.6 (84.96%)	1.4 (3.61%)	99.95	96.95	47.0
U_4hrsN1^{Mk}	39.4	37.9	33.7 (88.9%)	32 (84.41%)	1.7 (4.49%)	99.96	97.36	47.0
U_4hrsN2^{Mk}	37.6	34.7	22 (63.43%)	18.5 (53.18%)	3.6 (10.25%)	99.80	94.96	61.0
U_72hrsN1^{Mk}	50.3	47.9	15.5 (32.4%)	11.7 (24.5%)	3.8 (7.9%)	99.88	96.54	56.0

Sample	Raw Reads[M]	Trimmed Reads[M]	Mapped Reads[M]	Uniquely Mapped Reads[M]	Non-uniquely Mapped Reads[M]	Q20%	Q30%	GC Content (%)
U_72hrsN2[Mk]	40.5	38.9	34.5 (88.82%)	32.7 (84.14%)	1.8 (4.68%)	99.95	97.04	46.5

[M]All values for number of reads are in millions;

[Inf]These are infected samples indicated by the letter 'I' and 'S' in sample names

[Mk]These are mock-infected samples indicated by the letters 'U' and 'N' in sample names

SUPPLEMENTARY INFORMATION/MATERIALS