



## Expression profiling of host long non-coding RNAs under ORF virus infection

ARUNA PANDEY\*, SHIKHA SAXENA, RAJA ISHAQNABI KHAN, RAVI KUMAR GANDHAM and PRAMOD W RAMTEKE

ICAR-Indian Veterinary Research Institute, Izatnagar, Uttar Pradesh 243 122 India

### ABSTRACT

The present study aimed at gaining insights into the expression profile of long non-coding RNAs (lncRNAs) and coding genes of Orf virus (ORFV) infected oral mucosal tissues of sheep using RNA-Seq technology. Specifically, an expression profile and lncRNA-mRNA interaction network was inferred from a large-scale gene expression data set of sheep mucosal tissues on 0, 3, 7 and 15 days after ORFV infection. RNA-Seq profiles were obtained from the Gene Expression Omnibus (GEO) database. We found that 45, 64 and 45 lncRNAs and 1796, 2182 and 1550 coding genes were differentially expressed at early (T3), intermediate (T7) and late (T15) stages of ORFV infection in sheep mucosal tissues, respectively. Functional analysis revealed that differentially expressed long non-coding RNAs (DELncRNAs) regulate immune processes by regulating the expression level of differentially expressed coding genes (DEGs) under ORFV infection.

**Keywords:** Apoptosis, lncRNAs, ORFV, RNA-Seq, Sheep

Orf also known as contagious ecthyma, is a highly contagious viral disease of sheep and goats, caused by Orf virus of the genus *Parapoxvirus*, diagnosed by lesions on the mouth and oral mucosa (Gelaye *et al.* 2016). Several genes are involved in ORFV infection. CBP gene has a critical role in virulence and pathogenesis of ORFV (Fleming *et al.* 2017). The dysregulation of genes and their involvement in immune response and apoptosis were identified in ORFV infected sheep mucosal tissues using RNA-Seq approach (Jia *et al.* 2017). Although remarkable efforts have been done to understand ORFV pathogenesis, the mechanism by which ORFV causes fatal disease remains unclear.

Long non-coding RNAs are mRNA like transcripts known to be associated with animal diseases (Zhang *et al.* 2017, Zhao *et al.* 2018). It has also been reported that lncRNAs modulate innate immune responses during viral infection (Wang and Zheng 2018). lncRNA NEAT1 supports IL-8 production in response to influenza virus and herpes simplex virus infection (Imamura *et al.* 2014). Similarly, lncRNA NRAV suppress expression of immune response genes against influenza A virus infection (Ouyang *et al.* 2014). Studies have explored differentially expressed host lncRNAs under virus infection in mammals such as mouse and pig by RNA-Seq analysis (Zhao *et al.* 2018, Fang *et al.* 2019).

However, the lncRNAs expression profile in sheep infected with ORFV is not known. To the best of our knowledge, this is the first report that focuses on the expression profile of lncRNAs in mucosal tissues of sheep

after ORFV infection. Known lncRNAs and coding genes expression profile in three oral mucosal stages of ORFV infection in sheep was indentified in this study. Differentially expressed lncRNAs were used to predict trans target genes. The predicted trans targets were further used for the construction of lncRNA-targets correlation interaction network. KEGG pathway and gene ontology (GO) analysis was performed to investigate the functions associated with target genes.

### MATERIALS AND METHODS

**Data extraction:** RNA-Seq data of 8 oral mucosal tissue samples from sheep at 0, 3, 7 and 15 days post ORFV infection was obtained from Gene Expression Omnibus (GEO) database under accession number GSE95203. All procedures regarding tissue collection, the characteristics of oral mucosal samples, RNA extraction, and RNA-Seq were described previously (Jia *et al.* 2017). Briefly, 6 male sheep (4 months old) were used for the experiment; tissues from 3 sheep were pooled as one RNA-seq sample for each time point. To identify genes expressed in mucosal tissues at different stages of viral infection, cDNA libraries were constructed and sequenced on the IlluminaNextseq 500 system platform to generate a 151 bp paired end reads.

**Insilico analysis of RNA-Seq data:** The raw data was further processed using prinseq-lite software (Schmieder and Edwards 2011) to remove reads of low quality (mean phred score <25) and short length (<50) for downstream analysis. Filtered reads were then mapped independently to reference genome ([ftp://ftp.ensembl.org/pub/release-95/fasta/ovis\\_aries/dna/](ftp://ftp.ensembl.org/pub/release-95/fasta/ovis_aries/dna/)) using Bowtie 2 v2.3.4.3 (Langmead and Salzberg 2012). The mapped reads of each sample were

\*Corresponding author e-mail: arunabioinfo@gmail.com

then quantified separately against the lncRNAs GTF (gene transfer format) and coding genes GTF (obtained by filtering the GTF downloaded from Ensembl) using RSEM v1.2.12 (Li and Dewey 2011). The expression analysis for lncRNAs and coding genes was done using edgeR (Robinson *et al.* 2010) in R/Bioconductor environment. DElncRNAs and DEGs were considered with fold change  $\geq 2$  and P value  $\leq 0.05$ . DElncRNAs and DEGs expressions are displayed in the genome view with Circos plot.

**Target gene prediction and functional analysis:** To predict the targets of lncRNAs, a trans acting gene analysis was performed. The trans acting correlation of lncRNAs and coding genes was used to identify each other through the expression level. The expressed correlation between lncRNAs and coding genes was calculated with custom R scripts (Pearson correlation  $\geq 0.95$  or  $\leq 0.95$ ).

**Co-expression network construction:** Co-expression network was constructed according to the following steps: (i) lncRNAs and target genes that were upregulated or downregulated with fold change  $>4.0$  and P value  $<0.05$  were retained; (ii) lncRNA-genes, interaction was obtained based on expression correlation coefficient (Pearson correlation  $\geq 0.95$  or  $\leq 0.95$  was identified in this study). Cytoscape v3.6.0 was then used to display the network (Shannon *et al.* 2003).

**Gene Ontology (GO) and KEGG pathway analysis:** To identify biological processes associated with target genes of lncRNAs, a functional annotation enrichment analysis for Gene Ontology was conducted using g: Profiler (Reimand *et al.* 2011). In addition the KEGG pathway analysis was performed by KOBAS server (Wu *et al.* 2006). P value  $< 0.05$  was considered to indicate a statistically significant difference.

## RESULTS AND DISCUSSION

**RNA-Seq data analysis:** A total of 45, 64 and 45 DElncRNAs (fold change  $\geq 2$  and P value  $\leq 0.05$ ) were detected in the comparisons of T0 vs T3, T0 vs T7 and T0 vs T15, respectively (Supplementary Table 1). Among these DElncRNAs, 9, 28, and 17 lncRNAs were upregulated and 36, 36 and 28 were downregulated in T0 vs T3, T0 vs T7 and T0 vs T15 comparisons, respectively. Of these, 12 DElncRNAs including ENSOARG00000026674, ENSOARG00000025633, ENSOARG00000026363, ENSOARG00000026027, ENSOARG00000026131, ENSOARG00000026988, ENSOARG00000025403, ENSOARG00000025951, ENSOARG00000025486, ENSOARG00000026510, ENSOARG00000026846 and ENSOARG00000026451 were found common among all three comparisons; whereas 20, 28 and 16 lncRNAs were exclusively detected in T0 vs T3, T0 vs T7 and T0 vs T15 comparisons, respectively (Fig. 1A). Of these 12 common DElncRNAs, 7 DElncRNAs, viz. ENSOARG000000-25633, ENSOARG00000026027, ENSOARG000000-26363, ENSOARG00000026451, ENSOARG000000-26510, ENSOARG00000026674 and ENSOARG000000-26846 were identified with a fold change value of more

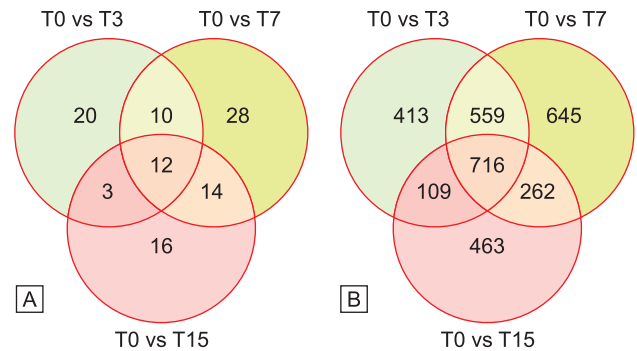


Fig. 1. Venn diagrams of DElncRNAs and DEGs showing comparisons of three stages of mucosal tissues in ORFV infected sheep. (A) Number of DElncRNAs for three comparisons (T0 vs T3, T0 vs T7 and T0 vs T15). (B) Numbers of DEGs for three comparisons (T0 vs T3, T0 vs T7 and T0 vs T15).

than  $\pm 2$ . Comparisons of coding gene expression levels among the three time-points revealed that there were 1797 DEGs (fold change  $\geq 2$  and P value  $\leq 0.05$ ) between T0 vs T3, 2182 DEGs between T0 vs T7 and 1550 DEGs between T0 vs T15 (Supplementary Table 2). Among these DEGs, 938, 1097 and 664 were upregulated and 859, 1085 and 886 were downregulated in T0 vs T3, T0 vs T7 and T0 vs T15 comparisons, respectively. Further, 706 genes were found common among all three comparisons, whereas 413, 645 and 463 genes were exclusively differentially expressed in T0 vs T3, T0 vs T7 and T0 vs T15 comparisons, respectively (Fig. 1B). In addition, most of these differentially expressed RNA transcripts were downregulated in ORFV infected sheep (Fig. 2). The difference in lncRNAs expression profile in various stages of mucosal tissues of sheep, suggested that lncRNAs might play significant biological roles in ORFV infection.

Moreover, similar to the distribution pattern of DEGs (Fig. 2A), the differentially expressed lncRNAs in all three-time points were not equally distributed among all chromosomes. The chromosome 1, 3 and 3 had the highest number of DElncRNAs and chromosome 9 had the least number of DElncRNAs in T0 vs T3, T0 vs T7 and T0 vs T15 comparisons, respectively, while chromosome 26 did not show lncRNAs expression in all three comparisons (Fig. 2B). This indicated that ORFV infection could lead to dynamic transcriptomic changes in different stages of mucosal tissues of sheep.

**Target genes prediction and network construction:** In the present study, the trans role of lncRNAs in coding genes was examined based on its expression correlation coefficient (Pearson correlation  $\geq 0.95$  or  $\leq 0.95$ ). To better explore the relationship of lncRNAs with target genes, we selected 108 DEGs and 7 DElncRNAs that were common in T0 vs T3, T0 vs T7 and T0 vs T15 comparisons, for generating the co-expression network (Fig. 3). The network displayed that downregulated lncRNA ENSOARG00000026027 governed the expression of NR4A2 gene. To our knowledge, prior to this analysis, no reports existed concerning the association between NR4A2 and lncRNA.

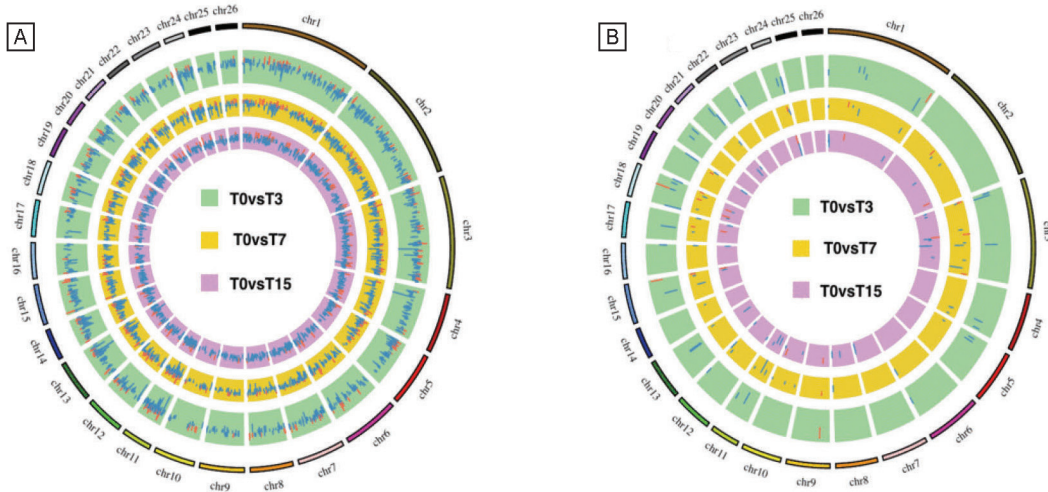


Fig. 2. Circos plot showing lncRNAs and coding genes expression profile in ORFV infected sheep mucosal tissues. Each circle represents three different stages of mucosal tissues (from inner to outer: T0 vs T15, T0 vs T7 and T0 vs T3). (A) Distribution of coding genes along each chromosome. The upregulated genes are shown by red colour and downregulated genes are shown by blue colour. (B) Distribution of lncRNAs along each chromosome. The upregulated lncRNAs are shown by red colour and downregulated lncRNAs are shown by blue colour.

NR4A2 is a nuclear receptor known to participate in cell differentiation and apoptosis (Shi *et al.* 2017). A previous study demonstrated that NR4A2 inhibits apoptosis via convergence with Wnt and mitogen-activated protein kinase (MAPK) pathways (Nordzell *et al.* 2004, Kitagawa *et al.* 2007). Similarly, based on the lncRNA-gene interaction

network, PLAUR is a predicted trans-target of ENSOARG00000026451. PLAUR or urokinase receptor is known to prevent apoptosis (Gilder *et al.* 2018, Zhou *et al.* 2018). The correlation between ENSOARG000000-26027 and NR4A2; ENSOARG00000026451 and PLAUR, suggests that host cells might regulate apoptosis via

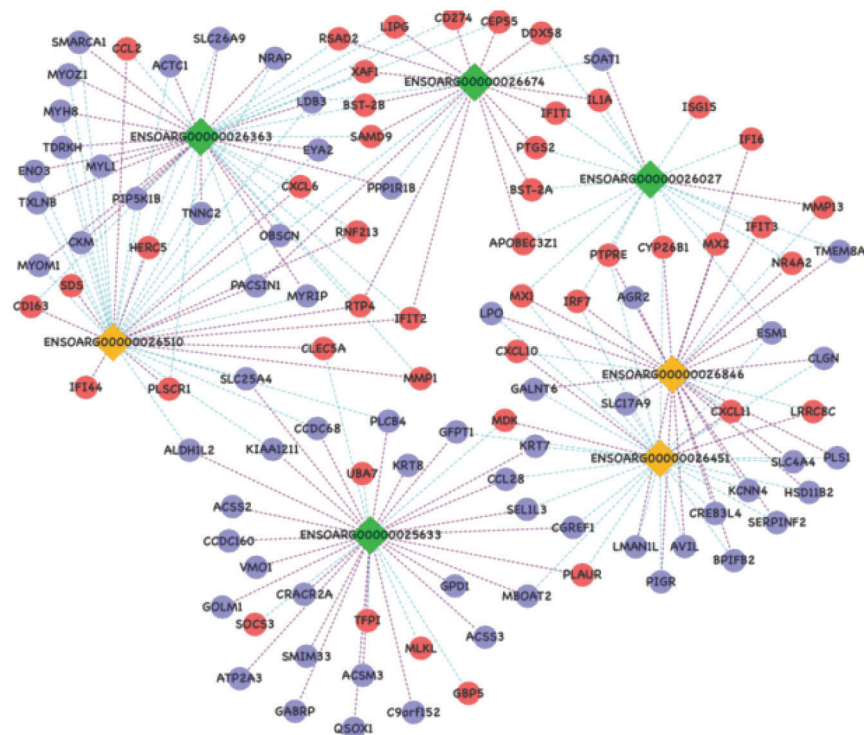


Fig. 3. lncRNA-gene co-expression network for T0 vs T15, T0 vs T7 and T0 vs T3. In this network, upregulated genes are displayed as red circles, downregulated genes are displayed as light purple circles, upregulated lncRNAs are displayed as yellow diamonds, and downregulated lncRNAs are displayed as green diamonds. The positive correlation between lncRNAs and their corresponding trans-targets shown by magenta arrow lines, whereas cyan blue dashed lines represent negative correlation between lncRNAs and their corresponding trans-targets.

lncRNAs ENSOARG00000026027 and ENSOARG00000026451 by activating anti-apoptotic genes after ORFV infection.

In addition, the downregulated lncRNA ENSOARG-00000026674 was positively correlated with RSAD2, IFIT1, IFIT2 and CD274. CD274 or the programmed death receptor 1 ligand/programmed death receptor 1 (PDL-1) pathway work as an immune checkpoint receptor (Yamauchi and Moroishi 2019). It is also reported that CD274 is negatively regulated by human antisense lncRNA-NKX2-1-AS1 (Kathuria *et al.* 2018). Studies also suggest, that several ISGs including IFIT-1 and IFIT-2 are known to have antiviral activity (Mears and Sweeney 2018). RSAD2 or viperin also reported to be participated in innate antiviral immune response (Dumbrepatil *et al.* 2019). Moreover, in the network, the upregulated lncRNA ENSOARG000000-26846 shows a positive correlation with genes- MX1 and IRF7. MX1 is reported to inhibit influenza virus by interfering with functional viral ribonucleoprotein complex assembly (Verhelst *et al.* 2012). Studies in the human umbilical vein endothelial cells suggest that transcription of lncRNA NEAT1 was induced through activation of RIG-I/IRF7 pathway during Hantavirus infection (Ma *et al.* 2017). It is also known that silencing of lncRNA#32 remarkably reduced the level of IRF7, resulting in sensitivity to encephalomyocarditis virus (EMCV) infection (Liu and Ding 2017). Furthermore, the downregulated lncRNA ENSOARG00000026363 negatively correlated with chemokine ligands -HERC5 and CCL2, and interferon induced gene IFIT2. Hence suggested, that lncRNAs ENSOARG000000-26674, ENSOARG000000-26846 and ENSOARG00000026363 could be a new host factor target for developing antiviral therapy by inducing the expression of immune genes against ORFV infection.

Additionally, in the network, the downregulated lncRNAs ENSOARG00000025633 governed the expression of downregulated KRT8. It is reported that increased expression of KRT8 enhances HBV virus replication (Zhong *et al.* 2014). Hence, lncRNA ENSOARG00000025633 might suppress ORFV replication, by regulating the expression of downregulated KRT8. Likewise, another upregulated lncRNA ENSOARG00000026510 target PLCB4, a WNT signaling modulator (Timmermans-Sprang *et al.* 2019), which indicated that lncRNA ENSOARG00000026510 might participate in WNT signaling pathways via its trans-regulation of PLCB4. Furthermore, our lncRNA-targets co-expression network, also suggests that one lncRNA can regulate the function of many genes.

**GO and KEGG pathway analysis:** Genes that were affected by DElncRNAs in a trans-acting manner was subjected to functional analysis. We found that the targets of DElncRNAs were highly enriched in biological processes like defense response to other organism, apoptosis, programmed cell death, response to type I interferon etc. It suggested that lncRNAs induced by ORFV infection might regulate the immune responses against ORFV. The top 5

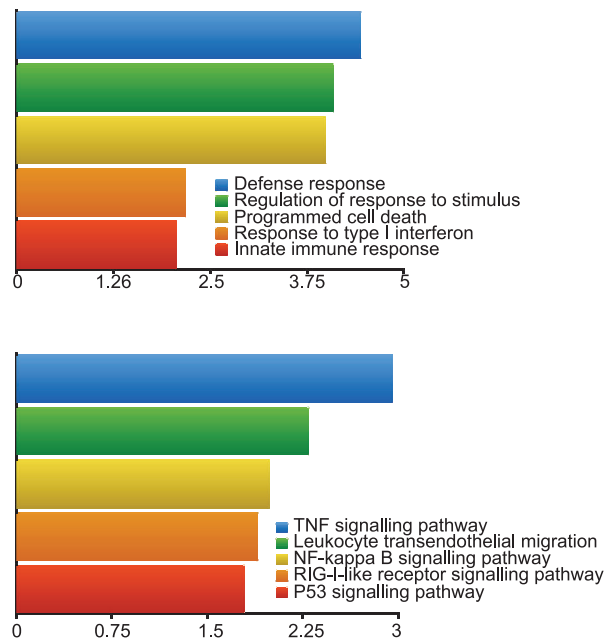


Fig. 4. Annotation of differentially expressed lncRNAs. (A) The top five biological process gene ontology for targets regulated by DElncRNAs. (B) The top five significantly enriched KEGG pathways for targets regulated by DElncRNAs.

significant GO biological terms are presented in Fig. 4A. Further, KEGG analysis predicted that DElncRNAs targets were enriched in 46 pathways. The top five KEGG pathways are shown in Fig. 4B. The targets of DElncRNAs were involved in important signaling pathways, such as TNF signaling pathway, NF-κB signaling pathway, RIG I like receptor signaling pathway, etc., which suggested that lncRNAs take part in host immune response against virus infection through various pathways.

In conclusion, the present study is the first report about the expression profile of lncRNAs upon ORFV infection in sheep. The expression of lncRNAs varied at different stages of ORFV infection. Furthermore, these dysregulated lncRNAs might have key roles in regulating immune responses post ORFV infection and exert important biological effects.

#### ACKNOWLEDGEMENTS

The authors are thankful to the Director, ICAR-Indian Veterinary Research Institute (IVRI), Izatnagar, for providing necessary laboratory facilities for carrying out the research work successfully.

#### REFERENCES

Dumbrepatil A B, Ghosh S, Zegalia K A, Malec P A, Hoff J D, Kennedy R T and Marsh E N. 2019. Viperin interacts with the kinase IRAK1 and the E3 ubiquitin ligase TRAF6, coupling innate immune signaling to antiviral ribonucleotide synthesis. *Journal of Biological Chemistry* **294**(17): 6888–98.  
 Fang M, Yang Y, Wang N, Wang A, He Y, Wang J, Jiang Y and Deng Z. 2019. Genome-wide analysis of long non-coding RNA expression profile in porcine circovirus 2-infected intestinal

- porcine epithelial cell line by RNA sequencing. *Peer Journal* 7: e6577.
- Fleming S B, McCaughan C, Lateef Z, Dunn A, Wise L M, Real N C and Mercer A A. 2017. Deletion of the chemokine binding protein gene from the parapoxvirus Orf virus reduces virulence and pathogenesis in sheep. *Frontiers in Microbiology* 8: 46.
- Gelaye E, Achenbach J E, Jenberie S, Ayelet G, Belay A, Yami M, Loitsch A, Grabherr R, Diallo A and Lamien C E. 2016. Molecular characterization of orf virus from sheep and goats in Ethiopia, 2008–2013. *Virology Journal* 13(1): 34.
- Gilder A S, Natali L, Van Dyk D M, Zalfa C, Banki M A, Pizzo D P, Wang H, Klemke R L, Mantuano E and Gonias S L. 2018. The urokinase receptor induces a mesenchymal gene expression signature in glioblastoma cells and promotes tumor cell survival in neurospheres. *Scientific Reports* 8(1): 2982.
- Imamura K, Imamachi N, Akizuki G, Kumakura M, Kawaguchi A, Nagata K, Kato A, Kawaguchi Y, Sato H and Yoneda M. 2014. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Molecular Cell* 53(3): 393–406.
- Jia H, Zhan L, Wang X, He X, Chen G, Zhang Y, Feng Y, Wei Y, Zhang Y and Jing Z. 2017. Transcriptome analysis of sheep oral mucosa response to Orf virus infection. *PLoS ONE* 12(10): e0186681.
- Kathuria H, Millien G, McNally L, Gower A C, Tagne J B, Cao Y and Ramirez M I. 2018. NKX2-1-AS1 negatively regulates CD274/PD-L1, cell-cell interaction genes, and limits human lung carcinoma cell migration. *Scientific Reports* 8(1): 14418.
- Kitagawa H, Ray W J, Glantschnig H, Nantermet P V, Yu Y, Leu CT, Harada S-i, Kato S and Freedman L P. 2007. A regulatory circuit mediating convergence between Nurr1 transcriptional regulation and Wnt signaling. *Molecular and Cellular Biology* 27(21): 7486–96.
- Langmead B and Salzberg S L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9(4): 357.
- Li B and Dewey C N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12(1): 323.
- Liu W and Ding C. 2017. Roles of lncRNAs in viral infections. *Frontiers in Cellular and Infection Microbiology* 7: 205.
- Ma H, Han P, Ye W, Chen H, Zheng X, Cheng L, Zhang L, Yu L, Wu X and Xu Z. 2017. The long noncoding RNA NEAT1 exerts antihantaviral effects by acting as positive feedback for RIG-I signaling. *Journal of Virology* 91(9): e02250–16.
- Mears H and Sweeney T. 2018. Better together: the role of IFIT protein-protein interactions in the antiviral response. Society for General Microbiology.
- Nordzell M, Aarnisalo P, Benoit G, Castro D S and Perlmann T. 2004. Defining an N-terminal activation domain of the orphan nuclear receptor Nurr1. *Biochemical and Biophysical Research Communications* 313(1): 205–11.
- Ouyang J, Zhu X, Chen Y, Wei H, Chen Q, Chi X, Qi B, Zhang L, Zhao Y and Gao G F. 2014. NRAV, a long noncoding RNA, modulates antiviral responses through suppression of interferon-stimulated gene transcription. *Cell Host and Microbe* 16(5): 616–26.
- Reimand J, Arak T and Vilo J. 2011. g: Profiler—a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Research* 39(suppl\_2): W307–15.
- Robinson M D, McCarthy D J and Smyth G K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1): 139–40.
- Schmieder R and Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27(6): 863–64.
- Shannon P, Markiel A, Ozier O, Baliga N S, Wang J T, Ramage D, Amin N, Schwikowski B and Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* 13(11): 2498–504.
- Shi X, Li W, Liu H, Yin D and Zhao J. 2017. The ROS/NF- $\kappa$ B/NR4A2 pathway is involved in H<sub>2</sub>O<sub>2</sub> induced apoptosis of resident cardiac stem cells via autophagy. *Oncotarget* 8(44): 77634.
- Timmermans-Sprang E, Collin R, Henkes A, Philipsen M and Mol J A. 2019. P-cadherin mutations are associated with high basal Wnt activity and stemness in canine mammary tumor cell lines. *Oncotarget* 10(31): 2930.
- Verhelst J, Parthoens E, Schepens B, Fiers W and Saelens X. 2012. Interferon-inducible protein Mx1 inhibits influenza virus by interfering with functional viral ribonucleoprotein complex assembly. *Journal of Virology* 86(24): 13445–55.
- Wang Z and Zheng Y. 2018. lncRNAs regulate innate immune responses and their roles in macrophage polarization. *Mediators of Inflammation*.
- Wu J, Mao X, Cai T, Luo J and Wei L. 2006. KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Research* 34(suppl\_2): W720–W24.
- Yamauchi T and Moroishi T. 2019. Hippo pathway in mammalian adaptive immune system. *Cells* 8(5): 398.
- Zhang J, Sun P, Gan L, Bai W, Wang Z, Li D, Cao Y, Fu Y, Li P and Bai X. 2017. Genome-wide analysis of long noncoding RNA profiling in PRRSV-infected PAM cells by RNA sequencing. *Scientific Reports* 7(1): 4952.
- Zhao P, Liu S, Zhong Z, Jiang T, Weng R, Xie M, Yang S and Xia X. 2018. Analysis of expression profiles of long noncoding RNAs and mRNAs in brains of mice infected by rabies virus by RNA sequencing. *Scientific Reports* 8(1): 11858.
- Zhong Q, An X, Yang Y X, Hu H D, Ren H and Hu P. 2014. Keratin 8 is involved in hepatitis B virus replication. *Journal of Medical Virology* 86(4): 687–94.
- Zhou J, Kwak K J, Wu Z, Yang D, Li J, Chang M, Song Y, Zeng H, Lee L J and Hu J. 2018. PLAUR confers resistance to Gefitinib through EGFR/P-AKT/Survivin signaling pathway. *Cellular Physiology and Biochemistry* 47(5): 1909–24.