Genome-wide analysis of long non-coding RNA expression profile in porcine circovirus 2-infected intestinal porcine epithelial cell line by RNA sequencing

Manxin Fang 1, Yi Yang 1, Naidong Wang 1, Aibing Wang 1, Yanfeng He 1, Jiaoshun Wang 1, You Jiang 1, Zhibang Deng 1

1 Hunan Agricultural University, Hunan Provincial Key Laboratory of Protein Engineering in Animal Vaccines, Changsha, Hunan, China

Corresponding Author: Zhibang Deng
Email address: zbangd@hunau.edu.cn

Porcine circovirus-associated disease (PCVAD), which is induced by porcine circovirus type 2 (PCV2), is responsible for severe economic losses. Recently, the role of noncoding RNAs, and in particular, microRNAs (miRNAs), in PCV2 infection has received great attention. However, the role of long noncoding RNA (lncRNA) in PCV2 infection is unclear. Here, for the first time, we describe the expression profiles of lncRNAs in an intestinal porcine epithelial cell line (IPEC-J2) after PCV2 infection, and analyze the features of differently expressed lncRNAs and their potential target genes. After strict filtering of approximately 150 million reads, we identified 13,520 lncRNAs, including 199 lncRNAs that were differentially expressed in non-infected and PCV2-infected cells. Furthermore, trans analysis found lncRNA-regulated target genes enriched for specific GO terms (P < 0.05), such as DNA binding, RNA binding, and transcription factor activity, which are closely associated with PCV2 infection. In addition, we analyzed the predicted target genes of differentially expressed lncRNAs, including SOD2, TNFAIP3, and ARG1, all of which are involved in infectious diseases. Our study identifies many candidate lncRNAs involved in PCV2 infection and provides new insight into the mechanisms underlying the pathogenesis of PCVAD.
Genome-wide analysis of long non-coding RNA expression profile in porcine circovirus 2-infected intestinal porcine epithelial cell line by RNA sequencing

Manxin Fang\textsuperscript{a}, Yi Yang\textsuperscript{a}, Naidong Wang\textsuperscript{a}, Aibing Wang\textsuperscript{a}, Yanfeng He\textsuperscript{a}, Jiaoshun Wang\textsuperscript{a}, You Jiang\textsuperscript{a}, Zhibang Deng\textsuperscript{a,*}

\textsuperscript{a} Hunan Provincial Key Laboratory of Protein Engineering in Animal Vaccines, Hunan Agricultural University, Changsha, Hunan 410128, China

* To whom correspondence should be addressed: Zhibang Deng, PhD.

E-mail: zbangd@hunau.edu.cn
Abstract

Porcine circovirus-associated disease (PCVAD), which is induced by porcine circovirus type 2 (PCV2), is responsible for severe economic losses. Recently, the role of noncoding RNAs, and in particular, microRNAs (miRNAs), in PCV2 infection has received great attention. However, the role of long noncoding RNA (lncRNA) in PCV2 infection is unclear. Here, for the first time, we describe the expression profiles of lncRNAs in an intestinal porcine epithelial cell line (IPEC-J2) after PCV2 infection, and analyze the features of differently expressed lncRNAs and their potential target genes. After strict filtering of approximately 150 million reads, we identified 13,520 lncRNAs, including 199 lncRNAs that were differentially expressed in non-infected and PCV2-infected cells. Furthermore, trans analysis found lncRNA-regulated target genes enriched for specific GO terms ($P < 0.05$), such as DNA binding, RNA binding, and transcription factor activity, which are closely associated with PCV2 infection. In addition, we analyzed the predicted target genes of differentially expressed lncRNAs, including $SOD2$, $TNFAIP3$, and $ARG1$, all of which are involved in infectious diseases. Our study identifies many candidate lncRNAs involved in PCV2 infection and provides new insight into the mechanisms underlying the pathogenesis of PCVAD.
1. Introduction

Infectious diseases threaten pig production, which is an important source of meat (Karuppannan & Opriessnig 2017). Porcine circovirus type 2 (PCV2) in the family Circoviridae is one of the most important pathogens affecting the pig population (Todd et al. 1991). This small, non-enveloped, and circular DNA virus is the causative agent of porcine circovirus-associated disease (PCVAD), which can manifest as PCV2-systemic disease, porcine dermatitis, nephropathy syndrome, porcine respiratory disease complex, PCV2-enteric disease, reproductive failure, and acute pulmonary edema (Meng 2013; Segales 2012; Segales et al. 2005). Increasing evidence indicates that PCVAD is a disease that causes considerable economic damage (Lekcharoensuk et al. 2004; Segales 2012).

Recent studies showed that noncoding RNAs play important regulatory roles in PCVAD (Hong et al. 2015; Wang et al. 2017). However, studies of PCV2-induced PCVAD have mainly focused on microRNA rather than on long non-coding RNA (lncRNA). The lncRNAs are defined as non-protein-coding transcripts greater than 200 nucleotides in length (Batista & Chang 2013). They play significant roles in cellular activities, such as genome regulation, and cell growth, differentiation, and apoptosis (Batista & Chang 2013; Yang et al. 2013). Furthermore, many viruses, such as influenza virus, enterovirus, and porcine reproductive and respiratory syndrome virus (PRRSV), can alter the expression of lncRNAs (Winterling et al. 2014; Yin et al. 2013; Zeng et al. 2018). Dysregulation of lncRNAs may also lead to diseases in pigs (Gao et al. 2017; Zhou et al. 2014).

No studies have examined the role of swine lncRNAs in PCV2-associated PCVAD. Enteritis is a common clinical manifestation of PCV2 infection because the intestinal mucosa is the initial site of PCV2 infection (Chae 2005; Kim et al. 2004). The intestinal porcine epithelial cell line IPEC-J2 is a non-transformed columnar epithelial cell line that was isolated from the neonatal piglet mid-jejunum by Helen Berschneider and colleagues (Orlando et al. 1989). IPEC-J2 is a well-validated model for studying the processes involved in pathogenic infections in the porcine intestinal epithelium (Arce et al. 2010; Koh et al. 2008; Skjolaas et al. 2007). Thus, in the current study, we...
evaluated the expression profiles of lncRNAs in an intestinal porcine epithelial cell line (IPEC-J2) after PCV2 infection by RNA sequencing and validated the results by quantitative real-time polymerase chain reaction (qRT-PCR). Moreover, Gene Ontology (GO) and pathway analyses were conducted to identify the biological roles of lncRNAs that were differentially expressed after PCV2 infection. To our knowledge, this is the first study to describe the aberrant lncRNA expression profile in response to PCV2 infection in IPEC-J2.
2. Materials and Methods

2.1. Cell culture and virus

IPEC-J2 cells lines free from porcine circovirus (Guangzhou Jennio Biotech Co., Ltd., China) were used in the present study. Cells were cultured in DMEM-F12 medium supplemented with 5% FBS. All cells were maintained at 37°C in a humidified incubator containing 5% CO₂. Here, IPEC-J2 cells were cultured on 6-well plastic tissue culture plates (Corning, Inc., Corning, NY, USA) at a density of 3 × 10⁵/well. The virus PCV2b (GenBank accession number: KJ867555) used in this work was provided by Hunan Provincial Key Laboratory of Protein Engineering in Animal Vaccines and stored at −80°C. The infectious titer of the PCV2 virus prepared from IPEC-J2 cells was 10⁴.⁷ TCID₅₀/ml.

2.2. Virus infection

PCV2 infection was performed as previously described (Yan et al. 2014). IPEC-J2 cells were grown to approximately 85% confluence and washed twice with phosphate-buffered saline (PBS). Next, cells were infected with PCV2 at 3×10².⁵ TCID₅₀/ml. After 1 h of adsorption, infected cells were cultured in fresh medium supplemented with 2% FBS. Uninfected cells were used as a negative control. Both the PCV2-infected and uninfected cells were harvested 36 h post-infection and used for total RNA extraction.

2.3. Library preparation and Illumina sequencing

Total RNA was extracted from each cell group using Trizol (Invitrogen, Carlsbad, USA). The RNA amount and purity of each sample was quantified by NanoDrop ND-1000 (NanoDrop, Wilmington, DE, USA) (Table 1). Moreover, the RNA integrity was assessed by Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA) (Table 1). In this study, RNA with a RIN (RNA integrity number) = 10 was used for library preparation. After reverse transcription, the purified first-strand cDNA was subjected to PCR amplification. The preparation for library and sequencing were performed according to previous studies (Wang et al. 2016) by LC-Bio (Hangzhou, China).
Finally, RNA sequencing was assessed by the HiSeq 4000 (Illumina, San Diego, CA, USA) on a 150 bp paired-end run.

2.4. Quality control

First, raw reads in the FASTA format were processed through in-house Perl scripts. During this step, clean reads (clean data) were obtained by discarding reads either containing adapter or over 10% poly-N sequences, and low quality reads (>50% of bases whose Phred scores were <5%). Meanwhile, Q20, Q30, and GC content were calculated for the clean data. All downstream analyses were carried out based on clean reads with high quality.

2.5. Mapping to the reference genome

Clean reads were aligned to the porcine reference genome (Sus scrofa 10.2) downloaded from the Genome website using TopHat v2.0.9 run with default parameters (Trapnell et al. 2012).

2.6. Transcriptome assembly

The mapped reads of each cell group were assembled using both StringTie (v1.3.3) and Cufflinks (v2.1.1) based on default parameters as previously described (Pertea et al. 2015; Trapnell et al. 2010).

2.7. Identification of lncRNAs

In addition to lncRNAs identified by BLAST searches, we also predicted novel lncRNAs from clean reads (Fig.1). Novel putative lncRNAs had to meet three requirements that have also been described previously: 1) transcripts should be longer than 200 bp and contain more than one exon; 2) the fragments per kilobase of transcript per million mapped reads (FPKM) score should be more than 0.5; 3) transcripts should have no coding potential. (Ran et al. 2016; Weng et al. 2017). In addition, the FRKM was transformed by the following formula:

$$\text{FPKM} = \frac{\text{Total exon Fragments}}{\text{Mapped reads (millions) \times \text{exon length (kb)}}}$$
To distinguish protein-coding and non-coding sequences effectively, Coding-Non-Coding-Index (CNCI v2) profiles were analyzed independent of known annotations according to the following parameters: length $\geq 200$ bp and exon $\geq 1$ and score $\leq 0$ (Meng 2013; Segales et al. 2005; Sun et al. 2013; Todd et al. 1991). The Coding Potential Calculator (CPC 0.9r2) mainly assesses the extent and quality of the open reading frame (ORF) in transcript sequences and compares it to sequences in a protein sequence database to differentiate coding and non-coding transcripts (Kong et al. 2007). All transcripts with CPC scores $\leq -1$ were removed. This filtering process yielded transcripts without coding potential that formed our candidate set of lncRNAs and were used for subsequent analysis.

2.8. Distribution of lncRNAs along each chromosome

Based on the location with respect to protein-coding genes, lncRNAs are classified into three types: intergenic lncRNAs (lincRNAs), intronic lncRNAs, and anti-sense lncRNAs (Harrow et al. 2012; St Laurent et al. 2015). After comparing with known mRNAs via the class_code module in cuffcompare, the putative lncRNAs were separated into these three classes (Yan et al. 2014). Next, all these kinds of lncRNAs were mapped to the porcine genome separately to determine their chromosomal distribution. Briefly, lncRNAs were aligned by short blast, followed by best hit analysis in short 500 kb segments. To evaluate their chromosomal distribution, we used the start sites of lncRNAs in the chromosomes counted in the pig reference genome (Susscrofa10.2).

2.9. Target gene prediction

Transcripts without coding potential constituted our candidate set of lncRNAs. Next, coding genes were searched within a range of 10–100 kb upstream or downstream of each candidate lncRNA for the cis target gene(Yan et al. 2014; Zhang et al. 2017a). For trans interactions, we determined the level gene expression based on Pearson’s correlations. Briefly, the Pearson’s correlation coefficients (R) between lncRNAs and mRNAs were calculated using the R statistical package, and lncRNA target genes were predicted with $R \geq 0.95$(Li et al. 2014).
2.10. Quantification of gene expression levels

To determine gene expression levels, we calculated the FPKMs for both lncRNAs and coding genes in each cell group using Cuffdiff (v2.1.1) software run with default parameters as previously described (Trapnell et al. 2010).

2.11. Differential expression analysis

To determine differential expression in digital transcript or gene expression datasets, we utilized a model based on a negative binomial distribution provided by Cuffdiff software. In addition, transcripts or genes in biological replicates with $P < 0.05$ were classed as differentially expressed. For non-biological replicates, transcripts or genes with $P < 0.05$ and an absolute value of $\log_2$ (fold change) more than 1 were assigned as the threshold for significant differential expression. Coverage signals used to generate heatmaps were obtained using heatmaply (a R package) based on the z-score obtained following dimensionality reduction of FPKM values (Galili et al. 2018).

2.12. Validation of RNA-Seq data by qRT-PCR

To further confirm the reliability of the RNA-seq data, qRT-PCR assays were performed. Briefly, single-stranded cDNA was generated using the RevertAid kit (Fermentas Life Science, Burlington, ON, Canada) with random primers. Real-time PCR was conducted using the SYBR Green q-PCR SuperMix (Bio-Rad, Hercules, CA, USA). The primers used are listed in Table 2. Subsequently, Ct values were acquired with manual thresholds using the 7500 System SDS software (ABI, USA). Levels of lncRNA expression were normalized to the level of GAPDH expression according to the $\Delta\DeltaCT$ method. The lncRNA expression levels between different groups were compared using $2^{-\Delta\DeltaCT}$. $P$ values $< 0.05$ were considered statistically significant.

2.13. Gene Ontology (GO) and KEGG pathway enrichment analysis

Analysis of GO enrichment for differentially expressed lncRNAs-regulated target genes was..
performed using the GOseq R package as previously described (Young et al. 2010). For KEGG pathway enrichment of differentially expressed IncRNA-regulated target genes (Young et al. 2010), the KOBASE software program was used as described previously (Mao et al. 2005).

3. Results

3.1. Reads and mapping of RNA-seq in IPEC-J2 cells

Total RNAs without rRNA for 6 samples including mock (IPEC_A1-A, IPEC_A2-A, IPEC_A3-A) and PCV2-infected IPEC-J2 cells (IPEC_B1-B, IPEC_B2-B, IPEC_B3-B) were sequenced. After discarding low-quality and adapter sequences, approximately 100 million clean reads were obtained for each cell group. The percentage of clean reads was approximately 98% (Table S1). Next, clean reads were mapped to the pig reference genome (Susscrofa10.2) using TopHat; > 69% of the clean reads were mapped (Table S2). More than 58% of reads were uniquely mapped, 42% were non-splice reads, and approximately 16% were splice reads (Table S2). Additionally, more than 60% of the reads were located on exons, whereas approximately 13% of the reads were on introns; the remaining reads were in intergenic regions (Fig. 2A-F, Table S3).

3.2. Assessing the quality of RNA-seq data

FPKM values were used to measure gene expression levels in IPEC-J2 cells. Based on the distribution profile for all transcripts shown in Fig. 3A, the patterns of expression between the control group (IPEC_A1-A, IPEC_A2-A, IPEC_A3-A) and PCV2 infection group (IPEC_B1-B, IPEC_B2-B, IPEC_B3-B) were similar. The FPKM density in the control and PCV2 infection groups was consistent (Fig. 3B). Importantly, the RNA-Seq Pearson correlation coefficients of transcript levels were greater than 0.84 in the control group and greater than 0.94 in the PCV2 infection group, indicating the rationality of the experimental design between these two groups and similarity of expression within the groups (Fig. 3C).

3.3. Identification of IncRNAs in IPEC-J2 cells
Reads were assembled using StringTie and selected lncRNAs were spliced using Cufflinks. After these rigorous selections, 13,520 novel lncRNAs were identified, including 10,975 lincRNAs, 2,182 intronic lncRNAs, and 301 anti-sense lncRNAs (Fig. S1A, Table S4). To further evaluate whether these lncRNAs had coding potential, we predicted the protein coding potential using CPC and CNCI. The CPC and CNCI scores in the control and PCV2 infection groups were similar, showing only slight differences (Fig. S1B, C). In total, 14,001 non-coding transcripts were predicted by CPC and 13,520 non-coding transcripts were determined using CNCI. Detailed information for the predicted lncRNAs is shown in Table S5. Moreover, lncRNAs were evenly distributed on each chromosome in both the control and PCV2 infection groups (Fig. S2).

3.4. Feature comparison of transcripts

To determine the differences between lncRNAs and mRNAs in IPEC-J2 cells, we compared their transcript structure, sequence conservation, and expression levels. The results showed a contrasting distribution tendency of exon number between mRNAs and lncRNAs (Fig. 4A). The distribution of lncRNAs was enriched on one exon term, which differed from that of mRNAs (Fig. 4A). Additionally, a large proportion of mRNAs were longer than 1000 base pairs, whereas lncRNAs were generally shorter (Fig. 4B). As expected, most of the lncRNAs contained a comparatively shorter ORF (mean = 61.41) compared to the ORFs in mRNAs (Fig. 4C). Compared to mRNAs, the expression level of lncRNAs was generally lower (Fig. 4D, E).

3.5. Characteristics of lncRNA expression levels between the control and PCV2 infection groups

The tentative lncRNAs were quantified by Cuffdiff software using the read count and FPKM analyses. The graphs in Fig. 5A show the lncRNA expression levels in IPEC-J2 cells after PCV2 infection; differentially expressed lncRNAs are also shown. There were 132 up-regulated and 67 down-regulated lncRNAs ($P < 0.05$) (Fig. 5B, Table S6). The heat map (Fig. 5C) indicates the differentially expressed lncRNAs ($P < 0.05$) between the control group and the PCV2 infection group. Nine lncRNAs, MSTRG.19762.1, MSTRG.1454.1, MSTRG.641.1, MSTRG.31692.1,
MSTRG.385.1, MSTRG.2965.1, MSTRG.15360.1, MSTRG.22503.1, and IMSTRG.5484.1, were selected for validation by qRT-PCR, and the results were consistent with the differential expression observed in the heat map (Fig. 6A-I). Particularly, MSTRG.31692.1 was down-regulated by 20-fold (Fig. 6I).

3.6. Prediction and functional analysis of differentially expressed IncRNA trans-regulated target genes

All reads were assembled into 31,836 transcripts as either mRNAs or IncRNAs. Briefly, there were 18,316 mRNAs, including 623 mRNAs predicted for differentially expressed IncRNA target genes. Of these 623 mRNAs, the expression of 373 mRNAs was altered by >2-fold. Genes can be regulated by IncRNAs in cis or in trans, and both regulatory mechanisms may play an important role in pathological and biological processes in pigs. Of the 373 differentially expressed genes, 362 were IncRNA trans-regulated target genes. Thus, we focused on differentially expressed IncRNA trans-regulated target genes in this study. The list of differentially expressed IncRNA target genes is shown in Table S7.

To further evaluate the relationship between IncRNAs and target genes, we mapped the nine validated IncRNAs and their target genes into the IncRNA-mRNA regulatory network (Fig. 7, Table S8). This interaction network was delineated using the Cytoscape software (v3.7.0). From this network, we inferred that IncRNAs may play a central role in PCV2 infection, as they regulate numerous target genes. Additionally, some of these target genes, including SOD2, TNFAIP3, ARG1, SERPINB2, VLDLR, HSPA5, and LCN2, are associated with infectious diseases, suggesting that IncRNAs respond to PCV2 infection by regulating these genes.

Next, differentially expressed target genes of the IncRNAs in trans were subjected to GO enrichment. The histogram in Fig. 8A shows the number of genes for a term distributed across biological processes, cellular components, and molecular functions. The histogram also shows the enriched GO terms of molecular function, including metal ion binding, DNA binding, ATP binding, nucleotide binding, RNA binding, zinc ion binding, nucleic acid binding, identical protein
binding, transcription factor activity, sequence-specific DNA binding, and calcium ion binding (Fig. 8A); detailed information is shown in Table S9.

Finally, we performed KEGG pathway enrichment for the differentially expressed lncRNA target genes. The 44 critical pathways with low P values (P < 0.05) and 501 genes are shown in Fig. 8B. Briefly, most enriched pathways were related to the TNF signaling pathway, *Salmonella* infection, Pertussis, MAPK signaling pathway, cytokine-cytokine receptor interaction, Influenza A, and Amoebiasis (Table S10).

4. Discussion

PCV2 is associated with PMWS and other porcine diseases that have a major negative impact on the global pig industry. In our efforts to characterize non-coding RNAs that may be involved in porcine diseases, we identified 13,520 novel porcine lncRNAs. Similarly, a large number of novel pig lncRNAs was also identified by another group, who predicted 12,867 novel lncRNAs in porcine alveolar macrophages after infection of the HP-PRRSV GSWW15 strain and the North American strain FL-12 (Zhang et al. 2017b). However, these numbers are low compared with those in other animals, and further characterization of porcine lncRNAs is required. The lncRNAs identified here share universal characteristics with other mammals including fewer exons, shorter length and lower expression level than protein-coding genes (Cabili et al. 2011; Ravasi et al. 2006; Ulitsky & Bartel 2013). Thus, although further characterization is required, the present study provides a reference for studying lncRNAs in other species.

There is strong evidence that lncRNAs play a clear role in viral infection. For example, lncRNA NRAV (negative regulator of antiviral response) modulates antiviral responses by suppressing the initiation of interferon-stimulated gene transcription (Ouyang et al. 2014). In addition, the upregulation of lncRNA-CMPK2 contributes to the negative regulation of the interferon response (Kambara et al. 2014). To date, the roles of lncRNAs in viral-host interactions during PCV2 infection are still unclear. To our knowledge, this is the first report that characterizes the expression profiles of lncRNAs in the non-transformed columnar epithelial cell line IPEC-J2 after PCV2
infection. Post PCV2 infection, we identified 199 differentially expressed lncRNAs, which appear to modify genes associated with viral infection, such as SOD2, TNFAIP3, ARG1, SERPINC2, VLDLR, HSPA5, and LCN2 (Table S11). For example, TNFAIP3 is involved in influenza A virus infection (Maelfait et al. 2012), whereas ARG1 suppresses arthritogenic alphavirus infection (Burrack et al. 2015). Therefore, PCV2 infection-associated lncRNAs may modulate viral infection through regulating these targeted genes.

The lncRNAs predicted as regulators of genes related to infectious diseases include MSTRG.4625, MSTRG.8436, MSTRG.4146, MSTRG.5886, MSTRG.5870, MSTRG.4146 and MSTRG.4592 (Table S11). However, since every targeted gene is regulated by several lncRNAs (Table S11), further studies are required to determine the mechanism of combinatorial control.

Our functional enrichment analysis revealed that processes including DNA binding, transcription factor activity and identical protein binding are closely associated with PCV2 infection. Consistent with this, PCV2 induces the activation of transcription factor nuclear factor kappa B (NF-kappa B) by increasing DNA binding activity (Han et al. 2017; Wei et al. 2008). Another study revealed that PCV2 protein ORF4 induces apoptosis by binding to mitochondrial adenine nucleotide translocase 3 (mtANT3)(Lin et al. 2018). These data suggest that lncRNAs may respond to PCV2 infection through regulating DNA binding, transcription factor activity and identical protein binding.

The most enriched pathways in PCV2 infection include the TNF signaling pathway, Salmonella infection, Pertussis, MAPK signaling pathway, cytokine-cytokine receptor interaction, Influenza A, and Amoebiasis. Salmonella infection often occurs concurrently with PCV2-associated disease (Takada-Iwao et al. 2011); thus, lncRNAs involved in PCV2 infection may also play roles in Salmonella infection. However, no studies have directly demonstrated the involvement of the other pathways we identified in PCV2 infection. A previous study has demonstrated that PCV2 vaccination may protect piglets against PCV2 infection through inducing TNFα production (Koinig et al. 2015). Notably, these pathways, such as the TNF signaling pathway, MAPK signaling pathway, and cytokine-cytokine receptor interactions, play major roles in the host inflammatory response.
response to numerous infectious diseases e, suggesting that swine lncRNAs may contribute to host inflammation during PCV2 infection. Further studies should focus on exploring the underlying mechanisms by which swine lncRNAs affect PCV2 infection. PCV2 is associated with PMWS and other porcine diseases that have a major negative impact on the global pig industry. In our efforts to characterize non-coding RNAs that may be involved in porcine diseases, we identified 13,520 novel porcine lncRNAs. Similarly, a large number of novel pig lncRNAs was also identified by another group, who predicted 12,867 novel lncRNAs in porcine alveolar macrophages after infection of the HP-PRRSV GSWW15 strain and the North American strain FL-12 e. However, these numbers are low compared with those in other animals, and further characterization of porcine lncRNAs is required. The lncRNAs identified here share universal characteristics with other mammals including fewer exons, shorter length and lower expression level than protein-coding genes. Thus, although further characterization is required, the present study provides a reference for studying lncRNAs in other species.

The lncRNAs predicted as regulators of genes related to infectious diseases include MSTRG.4625, MSTRG.8436, MSTRG.4146, MSTRG.5886, MSTRG.5870, MSTRG.4146 and MSTRG.4592 (Table S11). However, since every targeted gene is regulated by several lncRNAs (Table S11), further studies are required to determine the mechanism of combinatorial control.

Our functional enrichment analysis revealed that processes including DNA binding, transcription factor activity and identical protein binding are closely associated with PCV2 infection. Consistent with this, PCV2 induces the activation of transcription factor nuclear factor kappa B (NF-kappa B) by increasing DNA binding activity (Han et al. 2017; Wei et al. 2008). Another study revealed that PCV2 protein ORF4 induces apoptosis by binding to mitochondrial adenine nucleotide translocase 3 (mtANT3)(Lin et al. 2018). These data suggest that lncRNAs may respond to PCV2 infection through regulating DNA binding, transcription factor activity and identical protein binding.

The most enriched pathways in PCV2 infection include the TNF signaling pathway, Salmonella infection, Pertussis, MAPK signaling pathway, cytokine-cytokine receptor interaction, Influenza
A, and Amoebiasis. *Salmonella* infection often occurs concurrently with PCV2-associated disease (Takada-Iwao et al. 2011); thus, lncRNAs involved in PCV2 infection may also play roles in *Salmonella* infection. However, no studies have directly demonstrated the involvement of the other pathways we identified in PCV2 infection. A previous study has demonstrated that PCV2 vaccination may protect piglet against PCV2 infection through inducing TNF-α production (Koinig et al. 2015). Notably, these pathways, such as the TNF signaling pathway, MAPK signaling pathway, and cytokine-cytokine receptor interactions, play major roles in the host inflammatory response to numerous infectious diseases (Benedict et al. 2003; Gong et al. 2011; Li et al. 2011; Maegraith & Harinasuta 1953), suggesting that swine lncRNAs may contribute to host inflammation during PCV2 infection. Further studies should focus on exploring the underlying mechanisms by which swine lncRNAs affect PCV2 infection.

**Conclusions**

In summary, the current study reveals candidate lncRNAs associated with PCV2 infection and the cellular signaling pathways that they modulate. These findings could be valuable in designing novel potential strategies to identify the molecular mechanisms underlying PCV2-associated diseases. In turn, this will facilitate the development of antiviral strategies against PCV2 that will benefit animals in the global pig industry.

**References**

Arce C, Ramirez-Boo M, Lucena C, and Garrido JJ. 2010. Innate immune activation of swine intestinal epithelial cells (IPEC-J2 and IPI-2I) in response to LPS from Salmonella typhimurium. *Comp Immunol Microbiol Infect Dis* 33:161-174. 10.1016/j.cimid.2008.08.003

Batista PJ, and Chang HY. 2013. Long noncoding RNAs: cellular address codes in development and disease. *Cell* 152:1298-1307. 10.1016/j.cell.2013.02.012

Benedict CA, Banks TA, and Ware CF. 2003. Death and survival: viral regulation of TNF signaling pathways. *Curr Opin Immunol* 15:59-65.
of Arthritogenic Alphavirus Infection by Suppressing Antiviral T Cells. PLoS Pathog 11:e1005191.

10.1371/journal.ppat.1005191

Cabil MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, and Rinn JL. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25:1915-1927.

10.1101/gad.17446611

Chae C. 2005. A review of porcine circovirus 2-associated syndromes and diseases. Vet J 169:326-336.

10.1016/j.tvjl.2004.01.012

Galili T, O’Callaghan A, Sidi J, and Sievert C. 2018. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. Bioinformatics 34:1600-1602. 10.1093/bioinformatics/btx657

Gao PF, Guo XH, Du M, Cao GQ, Yang QC, Pu ZD, Wang ZY, Zhang Q, Li M, Jin YS, Wang XJ, Liu H, and Li BG. 2017. LncRNA profiling of skeletal muscles in Large White pigs and Mashen pigs during development. J Anim Sci 95:4239-4250. 10.2527/jas2016.1297

Gong J, Shen XH, Chen C, Qiu H, and Yang RG. 2011. Down-regulation of HIV-1 infection by inhibition of the MAPK signaling pathway. Virol Sin 26:114-122. 10.1007/s12250-011-3184-y

Han J, Zhang S, Zhang Y, Chen M, and Lv Y. 2017. Porcine circovirus type 2 increases interleukin-1beta and interleukin-10 production via the MyD88-NF-kappa B signaling pathway in porcine alveolar macrophages in vitro. J Vet Sci 18:183-191. 10.4142/jvs.2017.18.2.183

Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigo R, and Hubbard TJ. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 22:1760-1774. 10.1101/gr.135350.111

Hong JS, Kim NH, Choi CY, Lee JS, Na D, Chun T, and Lee YS. 2015. Changes in cellular microRNA expression induced by porcine circovirus type 2-encoded proteins. Vet Res 46:39. 10.1186/s13567-015-0172-5

Kambara H, Niazi F, Kostadinova L, Moonka DK, Siegel CT, Post AB, Carnero E, Barriocanal M, Fortes P, Anthony DD, and Valadkhan S. 2014. Negative regulation of the interferon response by an interferon-induced long non-coding RNA. Nucleic Acids Res 42:10668-10680. 10.1093/nar/gku713

Karuppannan AK, and Opriessnig T. 2017. Porcine Circovirus Type 2 (PCV2) Vaccines in the Context of Current Molecular Epidemiology. Viruses 9. 10.3390/v9050099

Kim J, Ha Y, Jung K, Choi C, and Chae C. 2004. Enteritis associated with porcine circovirus 2 in pigs. Can J Vet Res 68:218-221.

Koh SY, George S, Brozel V, Moxley R, Francis D, and Kaushik RS. 2008. Porcine intestinal epithelial cell lines as a new in vitro model for studying adherence and pathogenesis of enterotoxigenic Escherichia coli. Vet Microbiol 130:191-197. 10.1016/j.vetmic.2007.12.018

Koinig HC, Talker SC, Stadler M, Lading A, Graage R, Ritzmann M, Hennig-Pauka I, Gerner W, and Saalmuller A. 2015. PCV2 vaccination induces IFN-gamma/TNF-alpha co-producing T cells with a potential role in protection. Vet Res 46:20. 10.1186/s13567-015-0157-4

Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, and Gao G. 2007. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res 35:W345-349.

10.1093/nar/gkm391
Lekcharoensuk P, Morozov I, Paul PS, Thangthumniyom N, Wajjawalku W, and Meng XJ. 2004. Epitope mapping of the major capsid protein of type 2 porcine circovirus (PCV2) by using chimeric PCV1 and PCV2. *J Virol* 78:8135-8145. 10.1128/JVI.78.15.8135-8145.2004

Li L, Eichten SR, Shimizu R, Petsch K, Yeh CT, Wu W, Chettoor AM, Givan SA, Cole RA, Fowler JE, Evans MM, Scanlon MJ, Yu J, Schnable PS, Timmermans MC, Springer NM, and Muehlbauer GJ. 2014. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biol* 15:R40. 10.1186/gb-2014-15-2-r40

Li Y, Zhou H, Wen Z, Wu S, Huang C, Jia G, Chen H, and Jin M. 2011. Transcription analysis on response of swine lung to H1N1 swine influenza virus. *BMC Genomics* 12:398. 10.1186/1471-2164-12-398

Lin C, Gu J, Wang H, Zhou J, Li J, Wang S, Jin Y, Liu C, Liu J, Yang H, Jiang P, and Zhou J. 2018. Caspase-Dependent Apoptosis Induction via Viral Protein ORF4 of Porcine Circovirus 2 Binding to Mitochondrial Adenine Nucleotide Translocase 3. *J Virol* 92. 10.1128/JVI.00238-18

Maegraith B, and Harinasuta C. 1953. Experimental amoebiasis in the guinea pig. *Trans R Soc Trop Med Hyg* 47:582-583.

Maelfait J, Roose K, Bogaert P, Sze M, Saelens X, Pasparakis M, Carpentier I, van Loo G, and Beyaert R. 2012. A20 (Tnfaip3) deficiency in myeloid cells protects against influenza A virus infection. *PLoS Pathog* 8:e1002570. 10.1371/journal.ppat.1002570

Mao X, Cai T, Olyarchuk JG, and Wei L. 2005. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21:3787-3793. 10.1093/bioinformatics/bti430

Meng XJ. 2013. Porcine circovirus type 2 (PCV2): pathogenesis and interaction with the immune system. *Annu Rev Anim Biosci* 1:43-64. 10.1146/annurev-animal-031412-103720

Orlando RC, Powell DW, Croom RD, Berschneider HM, Boucher RC, and Knowles MR. 1989. Colonic and esophageal transepithelial potential difference in cystic fibrosis. *Gastroenterology* 96:1041-1048.

Ouyang J, Zhu X, Chen Y, Wei H, Chen Q, Chi X, Qi B, Zhang L, Zhao Y, Gao GF, Wang G, and Chen JL. 2014. NRAV, a long noncoding RNA, modulates antiviral responses through suppression of interferon-stimulated gene transcription. *Cell Host Microbe* 16:616-626. 10.1016/j.chom.2014.10.001

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, and Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 33:290-295. 10.1038/nbt.3122

Ran M, Chen B, Li Z, Wu M, Liu X, He C, Zhang S, and Li Z. 2016. Systematic Identification of Long Noncoding RNAs in Immature and Mature Porcine Testes. *Biol Reprod* 94:77. 10.1095/biolreprod.115.136911

Ravasi T, Suzuki H, Pang KC, Katayama S, Furuno M, Okunishi R, Fukuda S, Ru K, Frith MC, Gongora MM, Grimmond SM, Hume DA, Hayashizaki Y, and Mattick JS. 2006. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res* 16:11-19. 10.1101/gr.4200206

Segales J. 2012. Porcine circovirus type 2 (PCV2) infections: clinical signs, pathology and laboratory diagnosis. *Virus Res* 164:10-19. 10.1016/j.virusres.2011.10.007

Segales J, Allan GM, and Domingo M. 2005. Porcine circovirus diseases. *Anim Health Res Rev* 6:119-142.

Skjolaas KA, Burkey TE, Dritz SS, and Minton JE. 2007. Effects of Salmonella enterica serovar Typhimurium, or serovar Choleraesuis, Lactobacillus reuteri and Bacillus licheniformis on chemokine and cytokine expression in the swine jejunal epithelial cell line, IPEC-J2. *Vet Immunol Immunopathol* 115:299-308. 10.1016/j.vetimm.2006.10.012

St Laurent G, Wahlestedt C, and Kapranov P. 2015. The Landscape of long noncoding RNA classification. *Trends Genet*
Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, and Zhao Y. 2013. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res* 41:e166. 10.1093/nar/gkt646

Takada-Iwao A, Nakanishi M, Souma J, Chikata S, Okuda Y, Imai Y, and Sato S. 2011. Porcine circovirus type 2 potentiates morbidity of Salmonella enterica serovar Choleraesuis in Cesarean-derived, colostrum-deprived pigs. *Vet Microbiol* 154:104-112. 10.1016/j.vetmic.2011.06.036

Todd D, Niagro FD, Ritchie BW, Curran W, Allan GM, Lukert PD, Latimer KS, Steffens WL, 3rd, and McNulty MS. 1991. Comparison of three animal viruses with circular single-stranded DNA genomes. *Arch Virol* 117:129-135.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, and Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28:511-515. 10.1038/nbt.1621

Ulitsky I, and Bartel DP. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* 154:26-46. 10.1016/j.cell.2013.06.020

Wang X, Xu X, Wang W, Yu Z, Wen L, He K, and Fan H. 2017. MicroRNA-30a-5p promotes replication of porcine circovirus type 2 through enhancing autophagy by targeting 14-3-3. *Arch Virol* 162:2643-2654. 10.1007/s00705-017-3400-7

Wang Y, Xue S, Liu X, Liu H, Hu T, Qiu X, Zhang J, and Lei M. 2016. Analyses of Long Non-Coding RNA and mRNA profiling using RNA sequencing during the pre-implantation phases in pig endometrium. *Sci Rep* 6:20238. 10.1038/srep20238

Wei L, Kwang J, Wang J, Shi L, Yang B, Li Y, and Liu J. 2008. Porcine circovirus type 2 induces the activation of nuclear factor kappa B by IkappaBalpha degradation. *Virology* 378:177-184. 10.1016/j.viro.2008.05.013

Weng B, Ran M, Chen B, He C, Dong L, and Peng F. 2017. Genome-wide analysis of long non-coding RNAs and their role in postnatal porcine testis development. *Genomics* 109:446-456. 10.1016/j.ygeno.2017.07.001

Winterling C, Koch M, Koeppel M, Garcia-Alcalde F, Karlas A, and Meyer TF. 2014. Evidence for a crucial role of a host non-coding RNA in influenza A virus replication. *RNA Biol* 11:66-75. 10.4161/rna.27504

Yan M, Zhu L, and Yang Q. 2014. Infection of porcine circovirus 2 (PCV2) in intestinal porcine epithelial cell line (IPEC-J2) and interaction between PCV2 and IPEC-J2 microfilaments. *Virol J* 11:193. 10.1186/s12985-014-0193-0

Zeng N, Wang C, Liu S, Miao Q, Zhou L, Ge X, Han J, Guo X, and Yang H. 2018. Transcriptome Analysis Reveals Dynamic Gene Expression Profiles in Porcine Alveolar Macrophages in Response to the Chinese Highly Pathogenic Porcine Reproductive and Respiratory Syndrome Virus. *Biomed Res Int* 2018:1538127. 10.1155/2018/1538127
Zhang G, Duan A, Zhang J, and He C. 2017a. Genome-wide analysis of long non-coding RNAs at the mature stage of sea buckthorn (Hippophae rhamnoides Linn) fruit. *Gene* 596:130-136. 10.1016/j.gene.2016.10.017

Zhang J, Sun P, Gan L, Bai W, Wang Z, Li D, Cao Y, Fu Y, Li P, Bai X, Ma X, Bao H, Chen Y, Liu Z, and Lu Z. 2017b. Genome-wide analysis of long noncoding RNA profiling in PRRSV-infected PAM cells by RNA sequencing. *Sci Rep* 7:4952. 10.1038/s41598-017-05279-z

Zhou ZY, Li AM, Adeola AC, Liu YH, Irwin DM, Xie HB, and Zhang YP. 2014. Genome-wide identification of long intergenic noncoding RNA genes and their potential association with domestication in pigs. *Genome Biol Evol* 6:1387-1392. 10.1093/gbe/evu113

**Figure legends**

**Figure 1. The pipeline for identifying putative IncRNAs.** Briefly, 55,463 transcripts were assembled and 13,520 novel IncRNAs were identified.

**Figure 2. Region distribution of raw reads.** The summary of the region distribution of raw reads in the genome from the control groups IPEC_A1-A (A), IPEC_A2-A (B), IPEC_A3-A (C) and PCV2 infection groups IPEC_B1-B (D), IPEC_B2-B (E), IPEC_B3-B (F).

**Figure 3. Reads and results of mapping RNA identified by deep sequencing.** (A) The FPKM distribution is shown as a box plot. (B) FPKM density distribution for all transcripts. (C) Pearson correlation coefficients for all samples.

**Figure 4. Comparison between IncRNAs and mRNAs.** LncRNA and mRNA transcripts compared by exon number (A), length (B), ORF length (C), and expression level (D and E).

**Figure 5. Characteristics of IncRNA expression levels between PCV2-infected and control groups.** All IncRNA expression levels are shown; differentially expressed IncRNAs are shown in red (up-regulated) or blue (down-regulated) (A). The number of differentially expressed IncRNAs (B). Heat map showing the expressed IncRNAs (P < 0.05) in the two groups. Colors from dark blue to orange stand for z-score passed the dimensionality reduction of FPKM value filter and reveal increasing RNA levels in each group (C).
Figure 6. Validation of RNA-Seq data by qRT-PCR. Left Y-axis shows the FPKM values of the selected lncRNAs (A-I) using RNA-seq, whereas the right Y-axis shows the relative expression levels of selected lncRNAs (A-I) using qPCR. * indicates \( P < 0.05 \).

Figure 7. Regulatory lncRNA-mRNA network based on validated lncRNAs and target genes. View of the lncRNA-mRNA regulatory network according to nine validated lncRNAs and their target genes.

Figure 8. Functional analysis of target genes regulated by differentially expressed lncRNAs in trans. For trans interactions, the GO enrichment histogram (A), GO terms (B), and KEGG pathway enrichment scatter plot (B) are shown.
Figure 1

The developed pipeline for identifying putative IncRNAs.

Briefly, 55,463 transcripts were assembled and 13,520 novel IncRNAs were identified.

- 55,463 transcripts were assembled by StringTie and Cufflinks
- 55,211 transcripts were filtered
- 37,174 transcripts were annotated
- 36,224 transcripts had protein-coding potential
- 13,520 transcripts were identified with CPC (score < -1) and CNCL (score < 0)
Figure 2

Region distribution of raw reads.

The summary of the region distribution of raw reads in the genome from the control groups IPEC_A1-A (A), IPEC_A2-A (B), IPEC_A3-A (C) and PCV2 infection groups IPEC_B1-B (D), IPEC_B2-B (E), IPEC_B3-B (F).
Figure 3

Reads and mapping results of RNA deep sequencing.

(A) The FPKM distribution is shown as a box plot. (B) FPKM density distribution for all transcripts. (C) Pearson correlation coefficients for all samples.
Figure 4

Comparison between lncRNAs and mRNAs.

LncRNA and mRNA transcripts compared by exon number (A), length (B), ORF length (C), and expression level (D and E).
Figure 5

Characteristics of lncRNA expression levels between PCV2 infection and control groups.

All lncRNA expression levels are shown; differentially expressed lncRNAs are shown in red (up-regulated) or blue (down-regulated) (A). Number of differentially expressed lncRNAs (B). Heat map showing the expressed lncRNAs (P < 0.05) in the two groups. Colours from dark blue to orange reveal increasing RNA levels in each group (C).
Figure 6

Validation of RNA-Seq data by qRT-PCR.

Left Y-axis shows the FPKM values of the selected IncRNAs (A-I) using RNA-seq, whereas the right Y-axis shows the relative expression levels of selected IncRNAs (A-I) using qPCR. * indicates P < 0.05.
Figure 7

LncRNA-mRNA regulatory network between validated IncRNAs and target genes.

View of IncRNA-mRNA regulatory network according to nine validated IncRNAs and their target genes.
Figure 8

Functional analysis of target genes regulated by differentially expressed IncRNAs in _trans_.

For _trans_ interactions, the GO enrichment histogram (A), GO terms (B), and KEGG pathway enrichment scatter plot (B) are shown.
Table 1 (on next page)

| Quality of total RNA |
|----------------------|

| Sample name | Conc.(μg/μL) | Total(μg) | 260/280 | RIN | 28S/18S | QC Evaluation |
|-------------|--------------|-----------|---------|-----|---------|---------------|
| IPEC-A1-A   | 2.04         | 61.32     | 2.06    | 10  | 2.1     | A             |
| IPEC-A2-A   | 1.98         | 59.31     | 2.08    | 10  | 2.0     | A             |
| IPEC-A3-A   | 2.03         | 60.82     | 2.06    | 10  | 2.0     | A             |
| IPEC-B1-B   | 1.96         | 58.93     | 2.08    | 10  | 1.9     | A             |
| IPEC-B2-B   | 1.85         | 55.62     | 2.08    | 10  | 2.0     | A             |
| IPEC-B3-B   | 2.22         | 66.48     | 2.03    | 10  | 1.9     | A             |
Table 2 (on next page)

Primer list
### Table 2 Primer list

| Gene          | ID     | Primer sequence       | Product length |
|---------------|--------|-----------------------|----------------|
| MSTRG.19762.1 | F      | CGACGACAAAACGAGAGTCA  | 196            |
|               | R      | AATTCTTGAAAAAGCGGCTGA |                |
| MSTRG.1454.1  | F      | CACCTTCTCCATGTCCCAT  | 207            |
|               | R      | CATGCTGCTTTATTGCCAAA |                |
| MSTRG.641.1   | F      | TGCTTCGGTCTCCCTTCTA  | 202            |
|               | R      | TTGGGATCCTCGACATTCTC |                |
| MSTRG.31692.1 | F      | CGTGAGAGATGCCATTCAGA | 215            |
|               | R      | AGGACTACCCCTCACGGGAGT|                |
| MSTRG.385.1   | F      | TCCGACTAGGAACCATGAGG | 173            |
|               | R      | TCCCCAGCTAGGGGTCTAAT  |                |
| MSTRG.2965.1  | F      | CTCAGTGGGTTAGGGTCCA  | 161            |
|               | R      | GTTTCTGGCTGCACATACG  |                |
| MSTRG.15360.1 | F      | ATAAGGTTGGGCTGCAGG   | 172            |
|               | R      | TCCCTCAGCATATGGAGTT  |                |
| MSTRG.22503.1 | F      | AACCAACTC GGTTTCCTG  | 239            |
|               | R      | CCTATCGCCTTTCTCTGTGC |                |
| MSTRG.5484.1  | F      | GAGCCGCATCTGCTACCTAC | 207            |
|               | R      | ACACGGTTCCCAGTGAGTT  |                |
| GAPDH         | 396823 | TCGGAGTGAAACGGATTTCG | 189            |
|               | R      | TGACAAGCTTCGGCTTCC   |                |