Praca oryginalna *Original paper*

RNA sequence analysis and fluorescent *in situ* RNA hybridisation revealed the temporal gene expression profile of the porcine reproductive and respiratory syndrome virus in Marc-145 cells

HAIYAN YAO[1], XINXIAN WANG[1], YUNHUA WANG[1], YONGNENG LI[1], WENYING LI[1], YING YANG[1], XIAOLIN GAO[2], YINGBO LIN[3], JIANPING LIU[4], JUNLONG BI[1], LIPING WANG[1]

[1]College of Animal Veterinary Medicine, Yunnan Agricultural University, Kunming 650201, Yunnan, China
[2]College of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, Yunnan, China
[3]Department of Oncology-Pathology, Karolinska Institutet, Stockholm 17176, Sweden
[4]Department of Gastroenterology, The First Affiliated Hospital of Nanchang University, Nanchang 330006, Jiangxi, China

Received 18.01.2024 *Accepted 21.03.2024*

Yao H., Wang X., Wang Y., Li Y., Li W., Yang Y., Gao X., Lin Y., Liu J., Bi J., Wang L.
RNA sequence analysis and fluorescent *in situ* RNA hybridisation revealed the temporal gene expression profile of the porcine reproductive and respiratory syndrome virus in Marc-145 cells

Summary

The global outbreaks of porcine reproductive and respiratory syndrome (PRRS) led to a huge loss to the swine industry. Hence, there is a need to further understand how PRRSV transcribes its genes in host cells to better control PRRSV infection. In this study, we implemented RNA-seq and fluorescent in situ RNA hybridisation techniques to determine the temporal gene expression profile of PRRSV in Marc-145 cells. The expression level of PRRSV genes increases over infection time, with a tendency of the expression level from the highest to the lowest for ORF7, ORF6, PRF5, ORF4, ORF3, ORF2, ORF1b and ORF1b, which warrant further confirmations. Next, analyses of protein interaction showed that the host may react to PRRSV infection through many genes, proteins, and pathways involved in surveillance, metabolism, regulation, binding, and splicing of RNA. Furthermore, our codon usage bias (CUB) analysis proposed the hypothesis that synonym mutations by reverse genetic tools can mutate the third letters of codons to be different from the host system, to reduce viral gene transcription and protein translation. In this way, the same amino acids or protein sequences can be produced, but with much less efficacy, and as a result, the artificially synthesised virions can be used as live attenuated vaccines with synonym mutations to stimulate host defence responses and immunity against PRRSV infection. Collectively, our findings provided meaningful information to the temporal gene expression profile of PRRSV. This will assist in filling research gaps and developing strategies for better control of PRRS.

Keywords: porcine reproductive and respiratory syndrome virus, RNA-seq, temporal gene expression, fluorescent *in situ* RNA hybridisation, Marc-145 cells

Porcine reproductive and respiratory syndrome (PRRS), one of the most destructive diseases in swine, is characterised by reproductive failure in sows and respiratory constraints in weaned piglets (27), and thus causes huge economic loss and public health problems worldwide. A recent study estimated that the economic impact of PRRS in China was ¥1424.37 (Chinese Yuan) per sow based on the management system before and after the PRRS outbreaks occurring in November 2014, March 2015, December 2016, and February 2017 (49). However, currently available vaccines on the market only provide rather limited protection from infection with porcine reproductive and respiratory syndrome virus (PRRSV), mainly due to insufficient cross-protection, restored, or even enhanced virulence of attenuated vaccines (35, 39, 50). Consequently, there is an urgent need to develop some alternative vaccines by looking into the genetic and transcriptional profiles of both the virus and the host side, to effectively prevent and treat PRRS, combating the emergence of various new PRRSV strains with distinct immunopathologic phenotypes.

The PRRSV is an enveloped single-stranded positive sense RNA virus and belongs to the order *Nidovirales*, family *Arteriviridae*, genus *Arterivirus*. PRRSV infects porcine alveolar macrophages (PAMs) *in vivo*, while *in vitro* some immortalized passaged cell lines such as Marc-145 and MA-104 cells can also be infected by PRRSV (46). The PRRSV is approximately 50 to 65 nm in diameter and its genome is approximately 15 kb in length, capped at the 5'end and polyadenylated at the 3'terminus. The PRRSV genome encodes at least eight open reading frames (ORFs), with ORF1a, ORF1b, ORF2 (subdivided into ORF2a and ORF2b), ORF3, ORF4, ORF5 (subdivided into ORF5 and ORF5a), ORF6, and ORF7 from the 5'end to the 3'end, respectively (23). Each ORF partially overlaps with its neighboring ORFs. ORF1a and ORF1b encode viral replicases, accounting for approximately 80% of the viral genome. ORF2-5 encodes the proteins GP2a, GP2b, GP3, GP4, GP5 and GP5a, respectively; while ORF6 and ORF7 encode the matrix protein (M) and the viral nucleocapsid protein (N), respectively.

RNA sequencing (RNA-seq) has been widely used to study the gene expression profile of viruses (14, 29) providing a complete overview of all genes expressed during the replication cycle of a virus and allowing a better characterisation of the virus transcriptome. However, to our knowledge, there is no data on the gene expression profile of PRRSV in Marc-145 cells. Our current study provides pilot data to delineate the temporal gene expression profile of PRRSV in Marc-145 cells using RNA sequencing and fluorescent *in situ* RNA hybridisation techniques, which may deepen our understanding of PRRSV gene expression *in vitro* and pave the way toward further *in vivo* investigations for novel and potent PRRS control strategies.

Material and methods

Virus and cells. The pathogenic PRRSV YN-1 strain used in this study was isolated within our laboratory (GenBank accession No.: KJ747052.1) (54) and belongs to PRRSV genotype 2 (the North American genotype), with a viral titer of $10^{-5.3}$ $TCID_{50}$/100 µl in Marc-145 cells. The Marc-145 cells, a PRRSV permissive immortal cell model derived from the kidney of green monkey (*Chlorocebus sabaeus*), were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco, Cat. #C11995500BT), supplemented with 10% feotal bovine serum (FBS, Gibco, Cat. #10099-141), 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1 mmol/l sodium pyruvate and an antibiotics mixture (Gibco, #15140-122) of 100 IU/ml penicillin and 100 µg/ml streptomycin. Maintenance and use of PRRSV YN-1 strain and Marc-145 cells were thoroughly described in our previous studies (21, 44, 45).

Infection of Marc-145 cells with PRRSV YN-1 strain. Marc-145 cell infection with PRRSV was performed as previously reported (20, 55). First, 3×10^5 Marc-145 cells per well were seeded in 6-well plates and cultured in 2 ml of DMEM medium at 37°C and 5% CO_2 until approximately 80% confluence. The PRRSV YN-1strain (MOI = 1) was then added to Marc-145 cells. After a 1.5 h incubation to allow virus attachment to cells, the medium was replaced with fresh DMEM cell culture medium to remove unattached virus. Finally, the infected Marc-145 cells at selected time points were subjected to total RNA extraction for RNA-seq or fluorescence *in situ* RNA hybridisation (FISH), for which cells were cultured on glass slides.

Fluorescence *in situ* RNA hybridisation (FISH). Probes (Tab. 2) complementary ORF1b, ORF2, ORF3, ORF5 or ORF7 viral RNA sequences of the PRRSV YN-1 strain (Tab. 1) were designed, synthesised with fluorophore conjugation by Servicebio Co. Ltd (Wuhan, China), and applied to fluorescence *in situ* RNA hybridisation (FISH) according to the standard protocol. Briefly, infected Marc-145 cells were first fixed with 4% paraformaldehyde (PFA) for 20 min, washed with PBS three times (5 min for each), digested with proteinase K (20 µg/ml) for 8 min and

Tab. 1. Primer sequences of qPCR in this study

Target gene	Primer name	Primer sequence (5'-3')	Location of the primer (nt)
ORF1a	ORF1a-F	ACTCTCGCCCGGTGACATTG	6747-6766
	ORF1a-R	CCATCCTGCGCCTCTTCTTCT	6998-7018
ORF1b	ORF1b-F	GGCCGAATTGAGGTAGATTG	11169-11188
	ORF1b-R	CGACCACCGCAACTGATTCC	11323-11343
ORF2	ORF2-F	ACCGTTCACCCTGAGCAATT	12177-12196
	ORF2-R	ACATTCGACGCGACACCATT	12312-12331
ORF3	ORF3-F	CAGCTACACGGCCCAGTTCC	12956-12975
	ORF3-R	CAGCCATTCCAGGTGAAACC	13141-13160
ORF4	ORF4-F	CTATGCTTCCGAGATGAGTGA	13465-13485
	ORF4-R	TGATCGACCACTAAGGAGCG	13583-13602
ORF5	ORF5-F	GCTGAATGGCACAGATTGGC	13847-13866
	ORF5-R	AGGACATGCAGTTCTTCGCA	14081-14100
M	M-F	CTACACGCCAGTGATGATATATGC	14362-14563
	M-R	GTGTACTCAGCCATAGAAACCTGG	14362-14385
ORF7	ORF7-F	AATGGCCAGCCAGTCAATCA	14844-14863
	ORF7-R	TCATGCTGAGGGTGATGCTG	15185-15166

Tab. 2. Probe sequences used in this study

Target gene	Probe name	Probe sequence (5'-3')	Location of the probe (nt)	Conjugated fluorophore
ORF1b	ORF1b-1	TGTAAGGTATGTCTCCAAACCTTGT	8110-8134	Cy3
	ORF1b-2	GACGGACGCTGGAATGGTTGGCACA	8231-8255	
ORF2	ORF2-1	CTTTTCCATGATGCGGTACATTCGA	12324-12348	Cy3
	ORF2-2	TAATGTAGCCTCGCTCACCACCTGT	12344-12368	
ORF3	ORF3-1	CAGGCGGAACCATGAACCCTAGTTC	12823-12847	FAM488
	ORF3-2	CGCAGCCATTCCAGGTGAAACCAGT	13063-13087	
ORF5	ORF5-1	AACTGAAAATGAGAGCTGTTGCTGG	13798-13822	FAM488
	ORF5-2	CAGACTGCGTAAATGCTACTCAAGA	13995-14019	
ORF7	ORF7-1	TTCTTTTGCTGCCTGCCGTTGTTAT	14809-14833	Cy5
	ORF7-2	ATTGAAGGCAGTCTGGATCGACGAC	15004-15028	

again washed with PBS three times (5 min for each). Next, the pre-hybridisation was performed in the pre-hybridisation solution at 37°C for 1 h, followed by hybridisation in the hybridisation solution containing probes (500 nmol/l for each) at 40°C overnight. Next, the slides were sequentially washed with 2 × SSC (saline-sodium citrate) at 37°C for 10 min, 1 × SSC at 37°C for 2 times (5 min for each) and 0.5 × SSC at 37°C for 10 min. Subsequently, the nuclear specimens were counter stained with Hoechst 33342 (1 µg/ml) at room temperature for 10 min. Finally, the slides were subjected to image acquisition on PANNORAMIC DESK/MIDI/250/1000 in the panoramic slide scanner (3DHISTECH, Hungary) and analysed using Case Viewer 2.4 (Halo v3.0.311.314).

Library construction and RNA sequencing. Marc-145 cells were infected with PRRSV YN-1 strain (MOI = 1) in duplicates for 2, 12, 24 and 48 h, with non-infected cells as corresponding controls. Total RNA was extracted from the 16 samples with Trizol (TIANGEN, DP424, Kunming, China) and the integrity of the RNA was verified by Bioanalyzer 2100 (Agilent). Sequencing libraries were constructed using the VAHTS® Universal V6 RNA-seq Library Prep Kit for Illumina (Vazyme, NR604) according to the manufacturer's recommendations. Briefly, a total amount of 500 ng of RNA per sample was used for the capture of Poly(A) mRNAs with oligo(dT)-conjugated magnetic beads. CDNA was synthesised following fragmentation of mRNA. End repair, dA-tailing, and addition of sequencing adapters were subsequently carried out. Enrichment PCR was performed to obtain the sequencing library. Libraries were deep sequenced on the Illumina NovaSeq6000 platform with the PE150 strategy (Biolinker Technology Kunming Co. Ltd, Kunming, China).

RNA-seqprocessing. RNA-seq data were processed using a widely used pipeline. Briefly, the sequencing qualities of 16-s RNA-seq profiles were examined with FASTQC V 0.11.9 (2) and MultiQC V1.11 (9). Sequencing reads were subjected to Trimmomatic V 0.39 (3) programme for adapter and quality trimming. To distinguish the different sources, the remaining reads were separately aligned to the genomes of the PRRSV YN-1 strain (KJ747052) and *Chlorocebus sabaeus* (assembly Vero_WHO_p1.0, GenBank assembly accession: GCA_015252025.1) with STAR algorithm V 2.7.9a (7) to set up the databases with sequences and annotations of the two genomes. The reads derived from the virus were used for subsequent analysis in this study, while the reads assigned to the *Chlorocebus sabaeus* genome will be used for another study. Expression levels of the eight viral genes were quantified using feature Counts v.2.0.1 (19). Normalised FPKM (fragments per kilobase of transcript per million fragments mapped) values were used to perform PCA analysis using the FactoMineR R package V2.4 (16).

Downstream bioinformatic analyses. First, the protein-protein interaction was downloaded from HPIDB3.0 (https://hpidb.igbb.msstate.edu/index.html) (1), genes or homologs from *Sus scrofa* (pig) were used for the construction of a gene regulatory network by using the predicted interaction method psi-mi:MI:1110 (12). Trimmed networks were visualised with Cytoscape v.3.9.1 (32). Second, enriched GO terms and KEGG pathways for *Sus scrofa* genes from protein-protein networks were annotated with R package clusterProfiler (43). Significant GO terms and KEGG pathways were visualised with the dotplot function in the enrichplot package in the R package. Finally, YN-1 and *Chlorocebus sabaeus* protein coding sequences were queried to codon W (V 1.4.4, http://codonw.sourceforge.net/) for codon usage bias (CUB) analysis (4).

Results and discussion

Detection of viral transcripts from PRRSV infected Marc-145 cells. In this study, our initial objective was to investigate how PRRSV infection would affect the global transcriptomic profile of Marc-145 cells over the infection period from 2 hpi to 48 hpi (hours after infection). After mapping the sequencing reads to the *Chlorocebus sabaeus* genome, we found that for the samples from 24 hpi and 48 hpi, quite a number of reads could not be mapped to the *Chlorocebus sabaeus* genome, which is more remarkable at 48 hpi, as shown with orange bars in Fig. 1A.

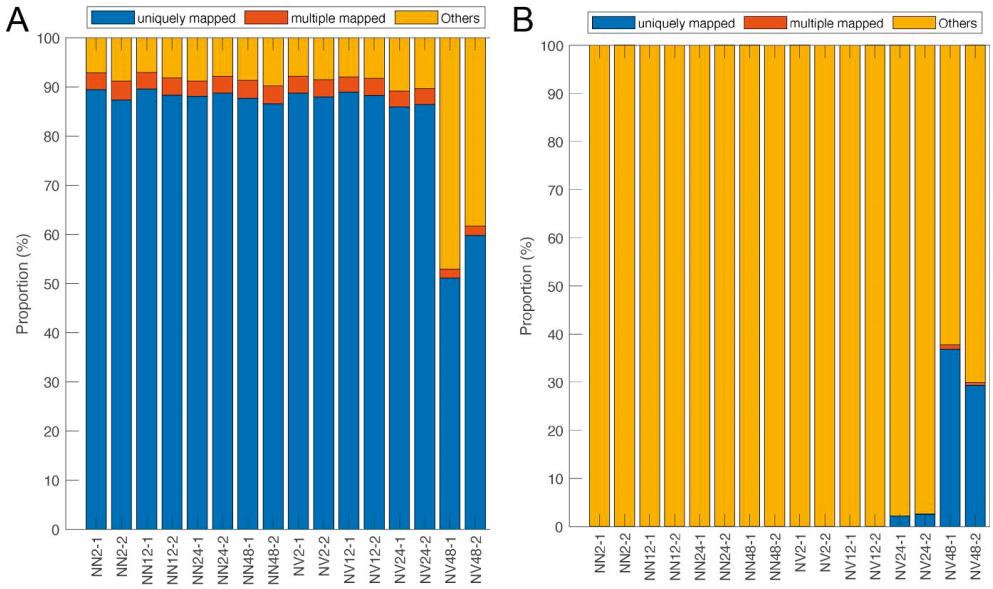

Fig. 1. Mapping of the reads to the genomes of green monkey (*Chlorocebus sabaeus*, assembly Vero_WHO_p1.0, GenBank assembly accession: GCA_015252025.1) (A) or PRRSV YN-1 strain (GenBank accession No.: KJ747052.1) (B)

Explanation: X-axis shows the 16 sequencing samples, while Y-axis shows the mapping rate of each sample to the corresponding genomes. Blue bars indicate that the reads can be uniquely mapped to the genome, red bars indicate that the reads can be mapped to multiple locations of the genome, yellow bars indicated that the reads cannot be mapped to the genome. Both types of reads shown in red or orange were excluded from analyses

Then we realised that the PRRSV genomic sequence is polyadenylated at the 3' terminus (23), which might make it possible for viral transcripts to be captured using oligo (dT)-based library construction strategies and pair end 150 bp sequencing platform. Therefore, the sequencing reads were remapped to the PRRSV genome (KJ747052). Subsequently, we found that the viral transcripts were detected by RNA-seq at 24 hpi and more prominently at 48 hpi, as shown with blue bars (Fig. 1B), with the proportions corresponding to the yellow bars (Fig. 1A). We then summarised all the mapping statistics into Fig. 3A, where we can see that the proportions of undetermined and multiple alignments of each of the 16 samples to *Chlorocebus sabaeus* genome remain at a similar alignment and that the decreased percentage of reads uniquely assigned to *Chlorocebus sabaeus* genome at 24 hpi and 48 hpi was replaced by the increased percentage of reads uniquely assigned to the genome of the PRRSV YN-1 strain. A conclusion can be drawn based on Fig. 3B that with RNA-seq technology, we could detect viral transcripts from PRRSV infected Marc-145 cells at 24 hpi, and on a much higher level at 48 hpi, with no or very few viral transcripts detected from noninfected cells or samples collected at early time points post infection.

Determination of the expression level of viral mRNA sequences in PRRSV-infected Marc-145 cells. PRRSV reads were further mapped to the eight individual ORFs (Tab. 1). As shown in (Fig. 3C), very few reads were detected as a background from the non-infection samples, while the expression of the eight viral genes gradually increased over the selected time points post infection from 2 hpi to 48 hpi, with a similar level between the duplicate of each treatment. Furthermore, hierarchical cluster analysis (Fig. 3C) showed that there was a tendency to increase the expression level of ORF1a, ORF1b, ORF2, ORF3, ORF4 to ORF5, ORF6 and ORF7, which was confirmed by the distribution of each individual viral gene in PCA (Fig. 2 and Tab. 1).

To validate the observations from RNA-seq described above, fluorescence *in situ* RNA hybridisation (FISH) was used and five viral genes were selected to exemplify the determination of the viral RNA expres-

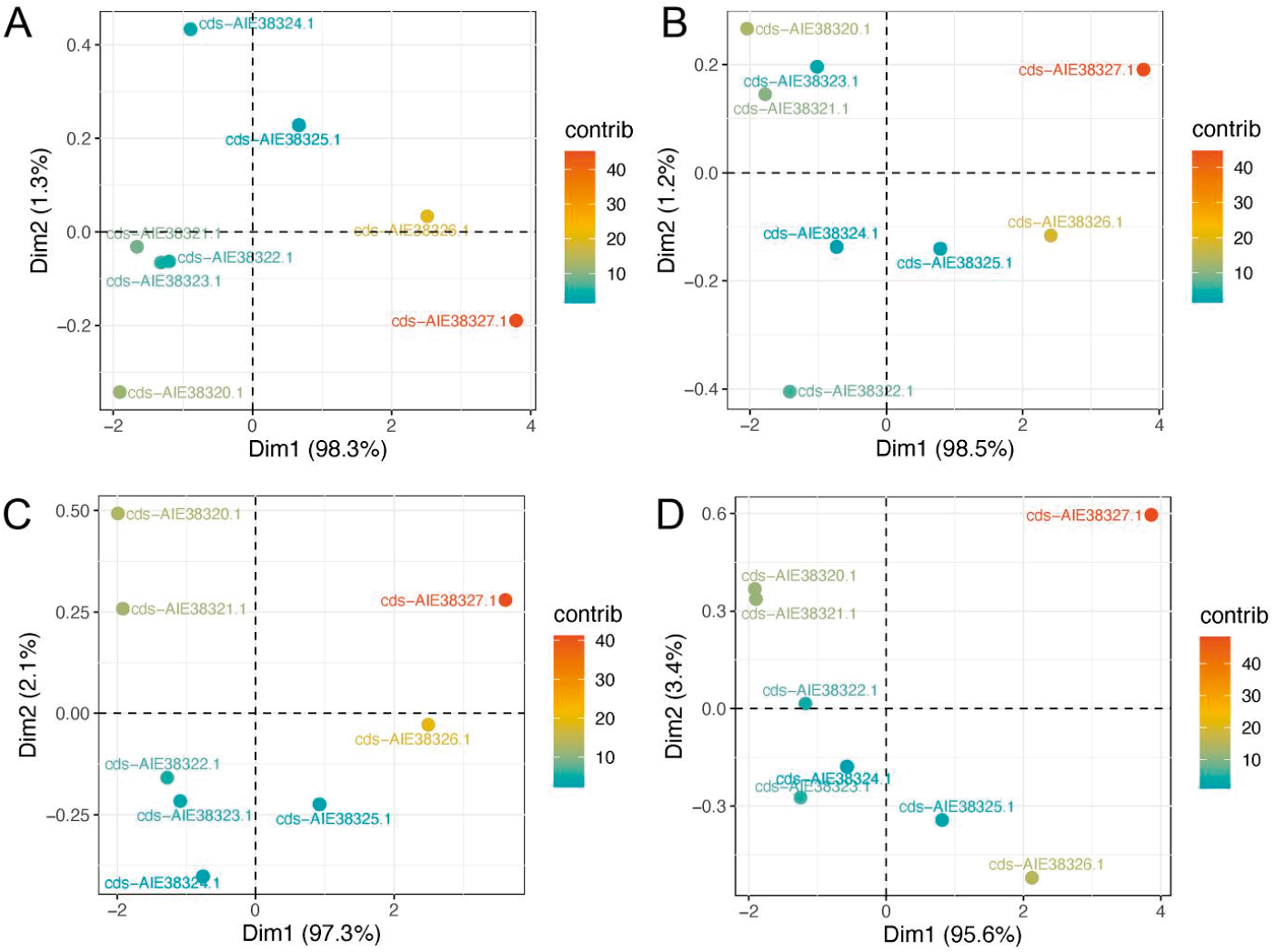

Fig. 2. Principal Component Analysis (PCA) of the RPKM of the eight viral genes
Explanation: Marc-145 cells in duplicates were infected with PRRSV YN-1 strain for 2 h (A), 12 h (B), 24 h (C) and 48 h (D), respectively. Color code indicates the RPKM amount or expression level

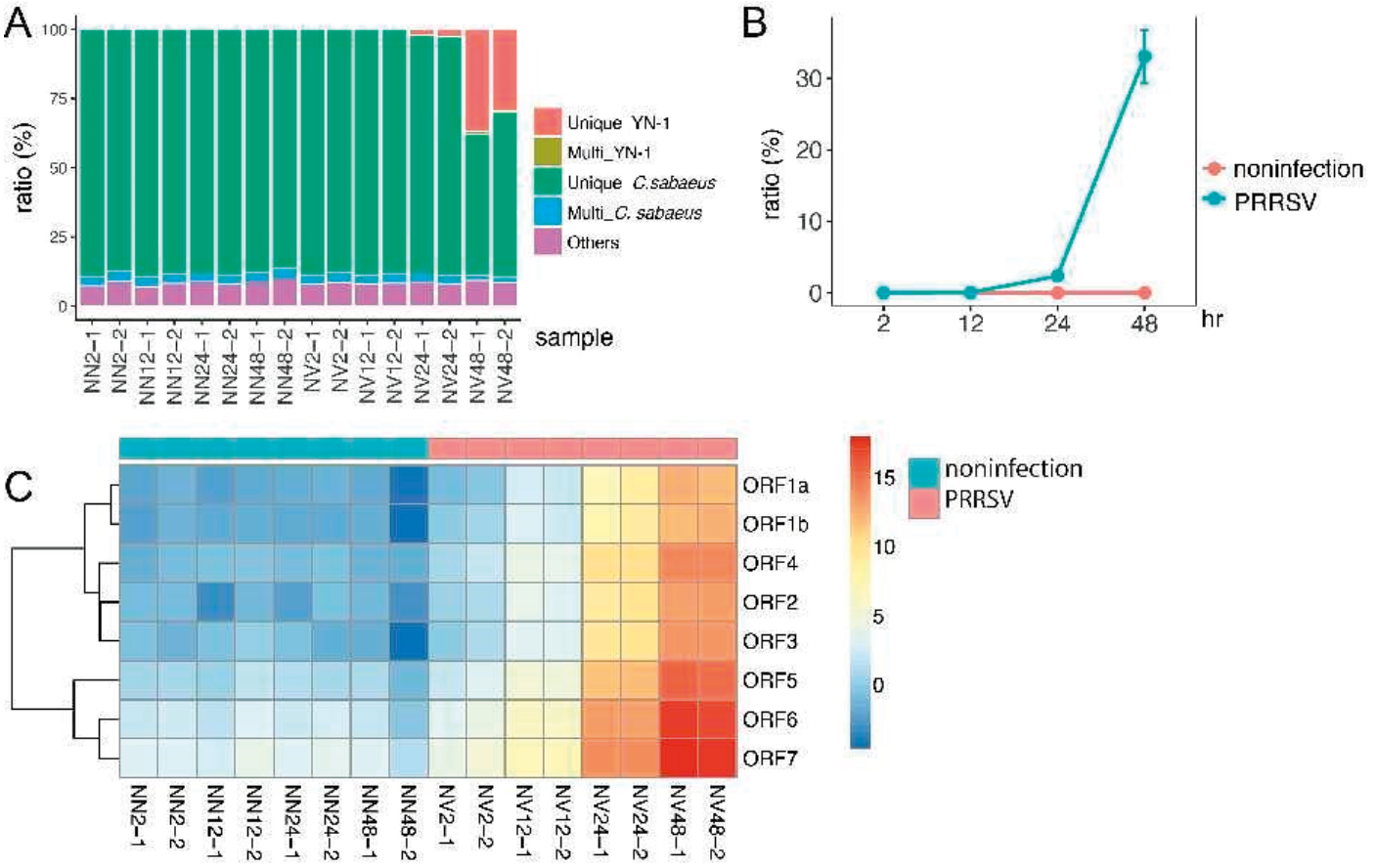

Fig. 3. Detection of viral transcripts by RNA-seq from Marc-145 cells inoculated with PRRSV in an infection-time-dependent manner

Explanation: Marc-145 cells (3×10^5cells/well) were cultured in 6-well plates overnight, followed by PRRSV infection (MOI = 1). Subsequently after a 1.5 h incubation to allow virus attachment to cells, the medium was replaced with fresh DMEM cell culture medium to remove the unattached virus. Finally, infected (NV) and corresponding noninfected control (NN) Marc-145 cells in duplicates (-1 and -2) at selected time points (2 h, 12 h, 24 h, and 48 h) were subjected to total RNA extraction and RNA sequencing analysis. A. Percentage of reads of each sample assigned to the genomes of the PRRSV YN-1 strain (GenBank accession No.: KJ747052.1) or green monkey (*Chlorocebus sabaeus*, assembly Vero_WHO_p1.0, GenBank assembly accession: GCA_015252025.1), where Marc-145 cells were originally from (unique or multiple). Other means that the reads can be mapped to either of the two genomes. The X-axis indicates the sample name, while the Y-axis shows the percentage of reads. B. Average percentage of reads from each treatment in duplicated mapped to the genomes of the PRRSV YN-1 strain at the four selected time points. C. Hierarchical clustering of the eight viral genes based on reads detected from the RNA-seq. Red blocks indicates higher expression while blue lower expression

sion levels in PRRSV infected Marc-145 cells: ORF7, which encodes the only nucleocapsid protein (N) of high immunogenicity in infected pigs (5); ORF5, which encodes a transmembrane glycosylated major envelope protein GP5 (22); ORF3, which encodes the minor glycosylated envelope viral protein GP3, one of the most variable PRRSV proteins decisive for virus assembly and infection (5); ORF2, which encodes two minor structural proteins (GP2a and E) (34); ORF1b, encoding polyproteins pp1ab, which can be processed in nsp9 to nsp12, critical for participation in virus transcription and replication (37).

Due to the limited detection channels of the slide scanner, we divided the five probes into two groups for FISH. As ORF7 or its encoded N protein has been widely applied in PRRSV research due to its high expression, the probe against ORF7 was used in both groups as a positive control for comparison. The RNA expression levels of all the five selected viral genes increased over the infection time in PRRSV infected Marc-145 cells (Fig. 4 and Fig. 5), which is consistent with our previous N expression data (44, 45). Furthermore, we can see that the expression level of ORF7 (cyan signal) is higher than that of ORF5 (red signal), followed by ORF2 (green signal) (Fig. 4), and that the expression level of ORF7 (cyan signal) is also higher than that of ORF3 (red signal), followed by ORF1b (green signal) (Fig. 5). Furthermore, under experimental conditions, we observed the signal co-localization of ORF7/ORF5/ORF2, or ORF7/ORF3/ORF1b, as highlighted by white squares; however, we also observed that ORF7 RNA did not colocalize with ORF3 or ORF1b, while ORF3 signal overlapped with ORF1b signal, shown in red squares. In addition, no nuclear viral transcripts were detected in this test.

Viral protein-host protein interaction. When viruses invade cells, the host's antiviral immune system is quickly activated to inhibit virus replication.

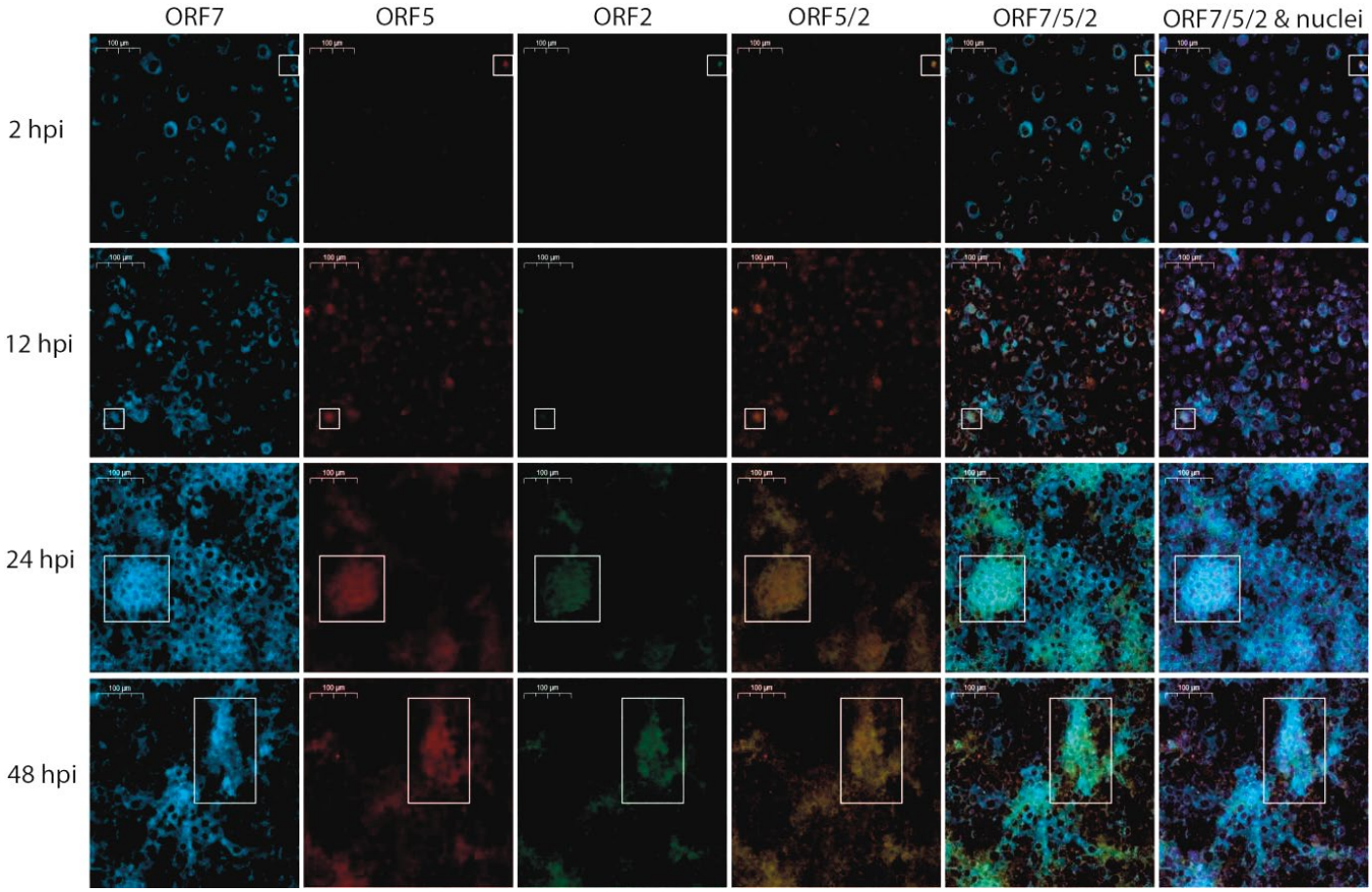

Fig. 4. Determination of the expression of three viral mRNA sequences (ORF7, ORF5, and ORF2) in Marc-145 cells by fluorescence *in situ* RNA hybridisation (FISH)

Explanation: Marc-145 cells were cultured on a glass slide overnight and infected with PRRSV YN-1 strain (MOI = 1), followed by fixation at four selected time points (2 h, 12 h, 24 h, and 48 h after infection) and subjected to FISH using the dye conjugated probe mix targeting viral mRNA sequences of ORF7, ORF5, and ORF2 (Tab. 2). Hoechst 33342 staining was performed to counter stain the nuclei. The images were acquired using PANNORAMIC DESK/MIDI/250/1000 in the panoramic slide scanner (3DHISTECH, Hungary) and analysed using CaseViewer 2.4 (Halo v3.0.311.314). Magnification is 10 × and the scale bar indicates 100 µm. Cyan signal represents the hybridisation signal from Cy5-labeled probe against the ORF7 sequence, red signal represents the hybridisation signal from Cy3-labeled probe against the ORF5 sequence, while green signal represents the hybridisation signal of the FAM488-labeled probe against the ORF2 sequence. The cells in the white squares represent cells where we detected colocalization of the three hybridisation signals

For host adaptation, viruses have evolved multiple effective strategies to manipulate the host genes and proteins to escape host antiviral responses. Therefore, the determination of virus-host protein interactions is fundamental for a better understanding of host defence against viral infections and the pathogenesis of viral infectious diseases. To understand the interactions between the PRRSV protein and host cellular factors (46), the interacting PRRSV proteins and *Sus scrofa* proteins were downloaded from HPIDB3.0 (1) and the network was visualised with Cytoscape v.3.9.1 (32) (Fig. 6) analysis of the KEGG pathway (Fig. 7A) and enriched GO terms (Fig. 7B, 7C, and 7D) for the *Sus scrofa* genes from the protein-protein network revealed that many of the *Sus scrofa* genes are involved in the surveillance, metabolism, regulation, binding and splicing of RNA and ribonucleoproteins (highlighted in red squares in Fig. 7). Interestingly, according to this database, ILF2 (interleukin-2 enhancer binding factor 2) can interact with ORF1a, ORF1b, and ORF7 or their encoding proteins, according to previous studies (31, 41).

Codon usage bias analysis of PRRSV genes. Due to differential transfer RNA supply within the cell, synonymous codons are not used with equal frequency, a phenomenon termed codon usage bias (CUB). Previous studies have demonstrated that CUB of endogenous genes transregulates the translational efficiency of other genes. CUB has evolved through mutation, natural selection, and genetic drift, and can be influenced by genome composition, GC content, gene length, position and context of codons in the genes, recombination rates, mRNA folding, and tRNA abundance and interactions. CUB is the preferential use of synonymous codons, an ubiquitous phenomenon observed in pathogens, plants, and animals. Different species have consistent and specific codon biases. Furthermore, codon bias varies not only with species,

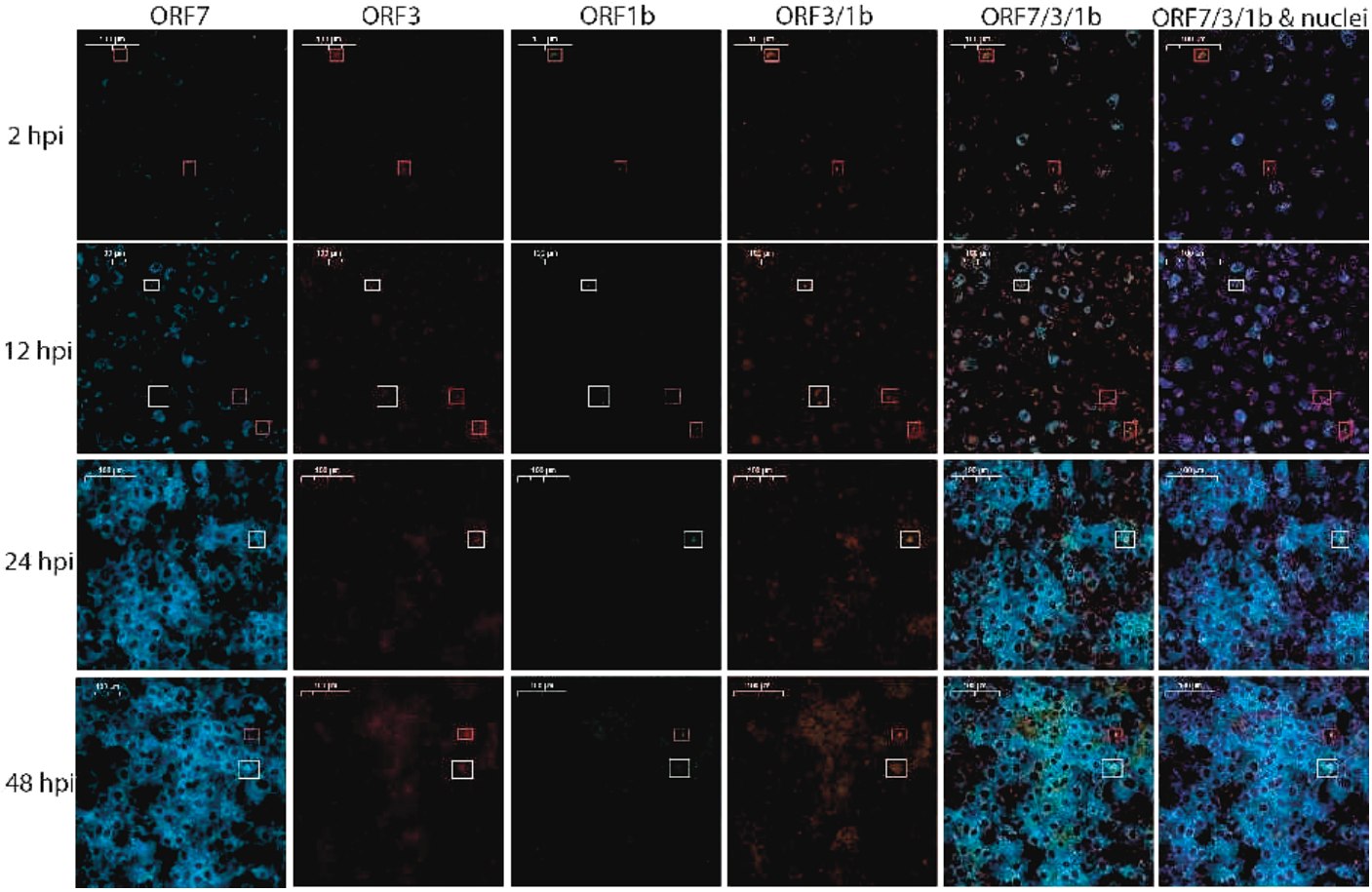

Fig. 5. Determination of the expression of three viral mRNA sequences (ORF7, ORF3, and ORF1b) in Marc-145 cells by fluorescence *in situ* RNA hybridisation (FISH)

Explanation: Except for the dye conjugated probe mix targeting viral mRNA sequences of ORF7, ORF3 and ORF1b (Tab. 2), the other experimental parts are the same as shown in Fig. 5. Cyan signal represents the hybridisation signal from Cy5-labeled probe against ORF7 sequence, red signal represents the hybridisation signal from Cy3-labeled probe against ORF3 sequence, while green signal represents the hybridisation signal from FAM488-labeled probe against ORF1b sequence. The cells in the white squares represent cells where we detected colocalization of the three hybridisation signals, while the cells in the red squares represent cells where colocalization of red and green signals was detected, but not cyan signal

family, or group within kingdom, but also between the genes within an organism. Therefore, analysis of codon usage is a good measure for understanding the genetic and evolutionary characteristics of an organism (13). However, there is no report related with the codon usage of PRRSV. Hence, the codon usage of PRRSV was analysed here to reveal the constraint factors and it could be helpful to improve the understanding of the pathogenesis of PRRSV and vaccine development to further control PRRS.

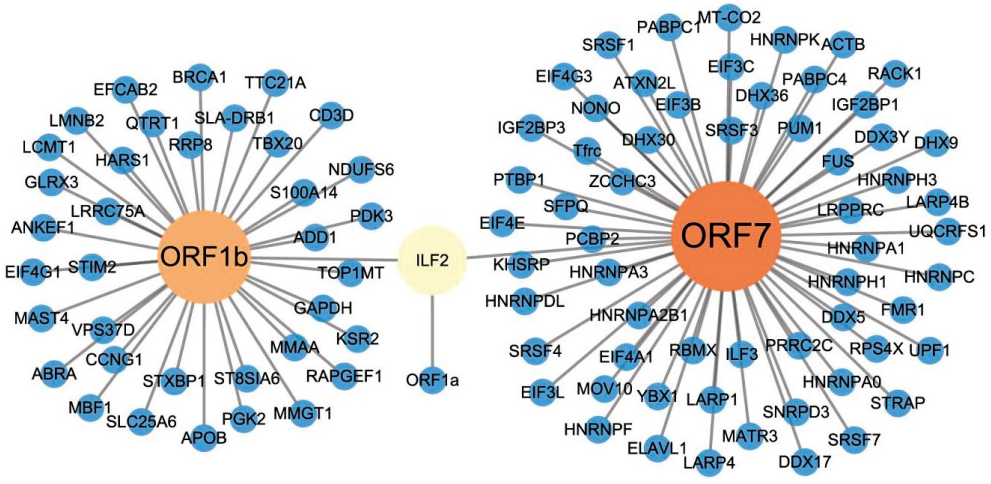

Fig. 6. Protein-protein interaction network between host proteins of *Sus scrofa* (Pig) and viral genes ORF1a, ORF1b and ORF7 or their protein products

Explanation: The data was downloaded from HPIDB3.0 and visualised with Cytoscape v.3.9.1 (see Material and methods)

CUB analyses clearly showed that PRRSV YN-1 strain used codons similar to the green monkey (*Chlorocebus sabaeus*) for protein translation, which can be reflected by analysis of CBI (codon bias index, Fig. 8A), FOP (Frequency of optimal codons, Fig. 8B) index and GC3s (GC content on the third synonymous codon position,

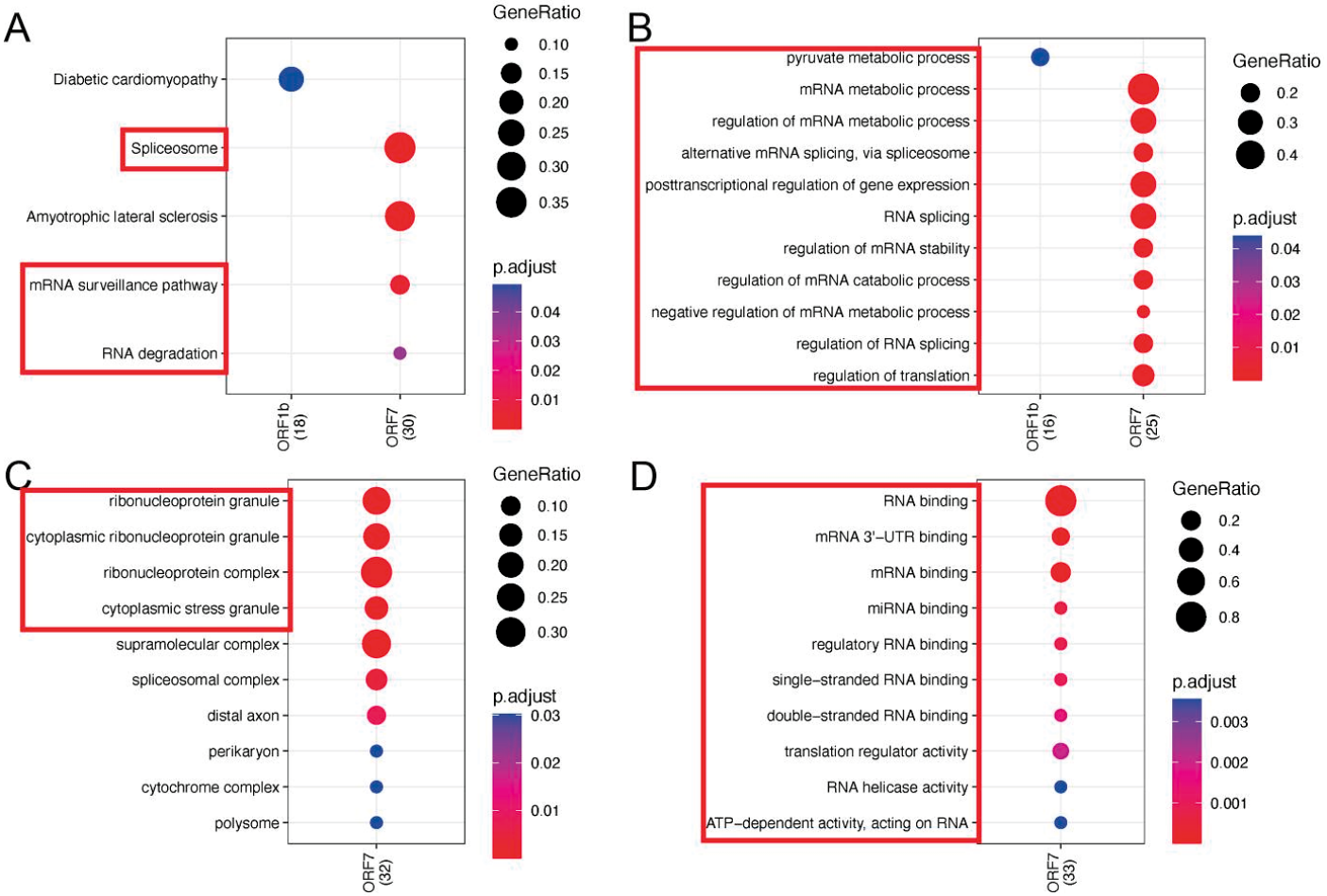

Fig. 7. KEGG and GO enrichment analyses of the known host proteins of *Sus scrofa* (Pig) that can interact with the viral genes ORF1a, ORF1b, and ORF7 or their protein products

Explanation: The input were the host proteins of *Sus scrofa* (Pig) displayed in Fig. 4. A. KEGG analysis. B. GO biological process (BP) analysis. C. GO cellular component (CC) analysis. D. GO molecular function (MF) analysis. The numbers in the brackets beside ORF1b or ORF7 indicate the numbers of proteins from *Sus scrofa* which may interact with the corresponding viral protein. The enriched items in the red squares are of high interest for further investigations

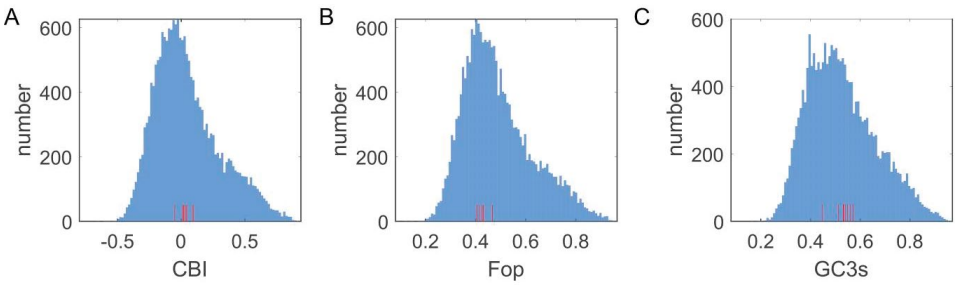

Fig. 8. Codon usage bias (CUB) analysis of PRRSV genes

Explanation: A. CBI (codon bias index) analysis. B. FOP (Frequency of optimal codons) analysis. C. GC3s (GC content on the third synonymous codon position) analysis. X-axis indicates the computed values for CBI, FOP and GC3s, respectively, by software codonW. Y-axis shows the number of mRNA coding sequences. Blue bars suggest mRNA coding sequences from green monkey (*Chlorocebus sabaeus*), while the red bars stand for mRNA coding sequences from PRRSV YN-1 strain

implying the difference in codon usage between organisms, Fig. 8C) analysis. In Fig. 8, the eight viral proteins (indicated by the red bars) are located within the peak areas for all the three analyses, suggesting PRRSV can hijack and adapt to the host translation machinery to replicate in the host cells to the best extent.

Most viruses, including PRRSV, interact with host cellular factors (including genes, RNAs and proteins) during infection. On the one hand, such an interaction is beneficial for the virus to escape from the host immune system. On the other hand, interactions regulate the host cell immune response to inhibit viral infections. RNA-seq has been widely implemented to gain a global transcriptomic landscape of the interaction between viruses and hosts. However, in PRRSV field, most of the publicly available RNA-seq studies focused on the host side, to study how the PRRSV infection altered the host transcriptome, such as the differential microRNA (miRNA) responses in porcine alveolar macrophages (PAMs) upon antiviral activation during PRRSV infection and interferon stimulation (17), a few kinetic transcriptomic analysis using bronchoalveolar lavage cells or

PAMs upon experimental infection with PRRSV-1 strains of different virulence (30, 38), integrated time-serial white blood cells transcriptomic and serum metabolomic analyses to investigate the host resistance to PRRSV (42), the transcriptional characteristics of NADC34-like PRRSV in PAMs (40), integration of transcriptome and proteome in lymph nodes to reveal the different immune responses to PRRSV in pigs (18), characterization of the recombined RNA molecules produced in PAMs during early passage infection of PRRSV by long-read direct RNA sequencing (47). However, to our knowledge, there is no report on the study and characterise the viral transcriptomic expression during PRRSV infection.

Here, we first used RNA-seq to investigate the temporal gene expression profile of PRRSV in Marc-145 cells and found that in general, viral genes that are closer to the 3' terminus showed a higher expression level than those closer to 5' end at all four time points after the infection we selected (2, 12, 24 and 48 hpi) (Fig. 1-3). To minimize the bias stemmed from RNA-seq process and to validate the observations from RNA-seq, we then performed fluorescent *in situ* RNA hybridisation to visualize the RNA expression in situ in Marc-145 cells. Indeed, Fig. 4 and Fig. 5 displayed similar tendency of a higher RNA expression level of the viral genes closer to the 3' terminus, although we cannot directly compare the hybridisation signals between these two Figs., as they were performed on separate slides. It would be of great importance to perform similar analyses in PAMs infected with various PRRSV strains of different virulence at more time points after infection, to draw a solid conclusion about whether the PRRSV gene expression profile is cell type, strain, or time dependent. Furthermore, the sensitivity and specificity of the hybridisation signal detection should be explored, together with the multiplexing capacity of more probes in a single assay, to better measure the hybridisation signals for direct comparisons.

Next, we constructed the interaction network of PRRSV proteins and green monkey proteins (Fig. 6) using the publicly available database, followed by GO and KEGG analyses of host proteins that can interact with viral proteins (Fig. 7). We found that many host proteins are ribonucleoproteins, enriched in RNA binding, metabolic, splicing, and degradation, suggesting how the host cells react to PRRSV infection. It should be noted that the database we used is far from complete and some other host factors reported to be involved in PRRSV infection are not included in this database, therefore not analysed in this study, such as porcine von Hippel-Lindau tumour suppressor (pVHL), which could stabilise PRRSV NSP1β via K63-linked ubiquitination to enhance PRRSV replication (25); exostosin glycosyltransferase 1 (EXT1), an enzyme involved in the biosynthesis of heparin sulphate that is a PRRSV receptor (15) and can interact with NSP3 and NSP5 (48); porcine TRIM26 that binds to N protein and affects PRRSV replication in PAM (53); TRIM4 that inhibits PRRSV replication through ubiquitination and degradation of the NSP2 protein in Marc-145 cells (52); porcine E3 ubiquitin ligase RNF122 that orchestrate NSP4 to promote PRRSV proliferation (36). A more comprehensive analysis of relevant porcine proteins should be performed to see which pathways could be enriched and essential for PRRSV infection, so that we can target those host genes, proteins, and pathway to develop new strategies for better PRRS control from the host side.

To our knowledge, there are some vaccines available on the market or in clinical trials for PRRSV control, exemplified by recombination PRRSV GP3 and GP5 DNA vaccines (51), modified live virus vaccine Fostera® PRRS (F-PRRS) (28) and commercial modified live PRRSV-1 vaccine „Ingelvac PRRSFLEX® EU" (8). However, the protection efficacy is rather unsatisfactory due to the high mutation and recombination rate of PRRSV strains. CUB is a critical factor determining gene expression and cellular function by influencing diverse processes such as RNA processing, protein translation and protein folding, and reflects the origin, mutation patterns and evolution of the species or genes. Investigations of codon bias patterns in genomes can reveal phylogenetic relationships between organisms, horizontal gene transfers, molecular evolution of genes and identify selective forces that drive their evolution. One of the most important applications of codon bias analysis is to design transgenes, to alter gene expression levels through codon optimization, and to develop vaccines (26). Based on the results of codon bias index, frequency of optimal codons and GC3s as shown in Fig. 8, and according to the previous study (4) that virus CUB tended to be more similar to that of symptomatic hosts than that of natural hosts, supporting a general deleterious effect of excessive CUB similarity between virus and host, synonym mutations can be achieved by reverse genetic tools to mutate the third letters of codons and thus not hijacking the most frequently used codon system from the host side, to reduce the viral gene transcription and protein translation levels. Subsequently, on the one hand, the same amino acids or even the same protein sequences can be synthesised, but with fewer virion particles and less virulence or pathogenicity, and on the other hand, artificially synthesised virions, as live attenuated vaccines with synonym mutations, can still stimulate host defence responses and immunity against PRRSV infection. This novel vaccine development strategy has been thoroughly evaluated in coronavirus (6, 33) and influenza A virus (10, 11, 24), as a proof-of-principle. Therefore, to overcome the disadvantages of currently available vaccines, at least partially, our codon usage bias (CUB) analysis may shed some new light on vaccine development to further control PRRS.

References

1. *Ammari M. G., Gresham C. R., McCarthy F. M., Nanduri B.*: HPIDB 2.0: a curated database for host-pathogen interactions. Database (Oxford). 2016 Jul 3, 2016, baw103.
2. *Andrews S.*: FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom 2010.
3. *Bolger A. M., Lohse M., Usadel B.*: Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014, 30 (15), 2114-2120.
4. *Chen F., Wu P., Deng S., Zhang H., Hou Y., Hu Z., Zhang J., Chen X., Yang J. R.*: Dissimilation of synonymous codon usage bias in virus-host coevolution due to translational selection. Nat. Ecol. Evol. 2020, 4 (4), 589-600.
5. *Dea S., Gagnon C. A., Mardassi H., Pirzadeh B., Rogan D.*: Current knowledge on the structural proteins of porcine reproductive and respiratory syndrome (PRRS) virus: comparison of the North American and European isolates. Arch. Virol. 2000, 145 (4), 659-688.
6. *Dilucca M., Forcelloni S., Georgakilas A. G., Giansanti A., Pavlopoulou A.*: Codon usage and phenotypic divergences of SARS-CoV-2 genes. Viruses 2020, 12 (5), 498.
7. *Dobin A., Davis C. A., Schlesinger F., Drenkow J., Zaleski C., Jha S., Batut P., Chaisson M., Gingeras T. R.*: STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013, 29 (1), 15-21.
8. *Duerlinger S., Knecht C., Sawyer S., Balka G., Zaruba M., Ruemenapf T., Kraft C., Rathkjen P. H., Ladinig A.*: Efficacy of a modified live porcine reproductive and Respiratory Syndrome Virus 1 (PRRSV-1) vaccine against experimental infection with PRRSV AUT15-33 in weaned piglets. Vaccines (Basel). 2022, 10 (6), 934.
9. *Ewels P., Magnusson M., Lundin S., Käller M.*: MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016, 32 (19), 3047-3048.
10. *Fan R. L., Valkenburg S. A., Wong C. K., Li O. T., Nicholls J. M., Rabadan R., Peiris J. S., Poon L. L.*: Generation of live attenuated Influenza Virus by using codon usage bias. J. Virol. 2015, 89 (21), 10762-10773.
11. *Gonin P., Mardassi H., Gagnon C. A., Massie B., Dea S.*: A nonstructural and antigenic glycoprotein is encoded by ORF3 of the IAF-Klop strain of porcine reproductive and respiratory syndrome virus. Arch. Virol. 1998, 143 (10), 1927-1940.
12. *Hermjakob H., Montecchi-Palazzi L., Bader G., Wojcik J., Salwinski L., Ceol A., Moore S., Orchard S., Sarkans U., von Mering C., Roechert B., Poux S., Jung E., Mersch H., Kersey P., Lappe M., Li Y., Zeng R., Rana D., Nikolski M., Husi H., Brun C., Shanker K., Grant S. G., Sander C., Bork P., Zhu W., Pandey A., Brazma A., Jacq B., Vidal M., Sherman D., Legrain P., Cesareni G., Xenarios I., Eisenberg D., Steipe B., Hogue C., Apweiler R.*: The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. Nat. Biotechnol. 2004, 22 (2), 177-183.
13. *Jia X., Liu S., Zheng H., Li B., Qi Q., Wei L., Zhao T., He J., Sun J.*: Non-uniqueness of factors constraint on the codon usage in Bombyx mori. BMC Genomics 2015, 16 (1), 356.
14. *Jones M., Dry I. R., Frampton D., Singh M., Kanda R. K, Yee M. B, Kellam P., Hollinshead M., Kinchington P. R., O'Toole E. A., Breuer J.*: RNA-seq analysis of host and viral gene expression highlights interaction between varicella zoster virus and keratinocyte differentiation. PLoS Pathog. 2014, 10 (1), e1003896.
15. *Jusa E. R., Inaba Y., Kouno M., Hirose O.*: Effect of heparin on infection of cells by porcine reproductive and respiratory syndrome virus. Am. J. Vet. Res. 1997, 58 (5), 488-491.
16. *Le S., Josse J., Husson F.*: FactoMineR: An R package for multivariate analysis. Journal of Statistical Software 2008, 25, 1-18.
17. *Li J., Sang E. R., Adeyemi O., Miller L. C., Sang Y.*: Comparative transcriptomics reveals small RNA composition and differential microRNA responses underlying interferon-mediated antiviral regulation in porcine alveolar macrophages. Front Immunol. 2022, 13, 1016268.
18. *Liang W., Meng X., Zhen Y., Zhang Y., Hu X., Zhang Q., Zhou X., Liu B.*: Integration of transcriptome and proteome in lymph nodes reveal the different immune responses to PRRSV between PRRSV-Resistant Tongcheng Pigs and PRRSV-Susceptible Large White Pigs. Front Genet. 2022, 13, 800178.
19. *Liao Y., Smyth G. K, Shi W.*: Feature Counts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014, 30 (7), 923-930.
20. *Liu X., Bi J., Zhao Q., Li M., Zuo Q., Wang X., Lan R., Li X., Yang G., Liu J., Yin G.*: Overexpression of RACK1 enhanced the replication of porcine reproductive and respiratory syndrome virus in Marc-145 cells and promoted the NF-κB activation via upregulating the expression and phosphorylation of TRAF2. Gene 2019, 709, 75-83.
21. *Liu X., Gao L., Zhao Q., Wang X., Yang C., Bi J., Yang R., Jin X., Lan R., Cui R., Wang X., Li W., Wang X., Yang Y., Yu X., Lin Y., Liu J., Yin G.*: Inhibition of porcine reproductive and respiratory syndrome virus by PKC inhibitor dequalinium chloride in vitro. Vet. Microbiol. 2020, 251, 108913.
22. *Mardassi H., Mounir S., Dea S.*: Molecular analysis of the ORFs 3 to 7 of porcine reproductive and respiratory syndrome virus, Québec reference strain. Arch. Virol. 1995, 140 (8), 1405-1418.
23. *Music N., Gagnon C. A.*: The role of porcine reproductive and respiratory syndrome (PRRS) virus structural and non-structural proteins in virus pathogenesis. Anim. Health Res. Rev. 2010, 11 (2), 135-163.
24. *Nogales A., Baker S. F., Ortiz-Riaño E., Dewhurst S., Topham D. J., Martínez-Sobrido L.*: Influenza A virus attenuation by codon deoptimization of the NS gene for vaccine development. J. Virol. 2014, 88 (18), 10525-10540.
25. *Pang Y., Zhou Y., Wang Y., Sun Z., Liu J., Li C., Xiao S., Fang L.*: Porcine Reproductive and Respiratory Syndrome Virus nsp1β Stabilizes HIF-1α to Enhance Viral Replication. Microbiol. Spectr. 2022 Dec 21, 10 (6), e0317322.
26. *Parvathy S. T., Udayasuriyan V., Bhadana V.*: Codon usage bias. Mol. Biol. Rep. 2022, 49 (1), 539-565.
27. *Pejsak Z., Markowska-Daniel I.*: Losses due to porcine reproductive and respiratory syndrome in a large swine farm. Comp. Immunol. Microbiol. Infect. Dis. 1997, 20 (4), 345-352.
28. *Rawal G., Angulo J., Linhares D. C. L., Mah C. K., Van Vlaenderen I., Poulsen Nautrup B.*: The efficacy of a modified live virus vaccine Fostera® PRRS against heterologous strains of porcine reproductive and respiratory syndrome virus: A meta-analysis. Res. Vet. Sci. 2022, 150, 170-178.
29. *Rozman B., Nachshon A., Levi Samia R., Lavi M., Schwartz M., Stern-Ginossar N.*: Temporal dynamics of HCMV gene expression in lytic and latent infections. Cell Rep. 2022, 39 (2), 110653.
30. *Sánchez-Carvajal J. M., Rodríguez-Gómez I. M., Ruedas-Torres I., Zaldívar-López S., Larenas-Muñoz F., Bautista-Moreno R., Garrido J. J., Pallarés F. J., Carrasco L., Gómez-Laguna J.*: Time series transcriptomic analysis of bronchoalveolar lavage cells from piglets infected with virulent or low-virulent porcine reproductive and respiratory syndrome virus 1. J. Virol. 2022, 96 (3), e0114021.
31. *Sha H., Zhang H., Chen Y., Huang L., Zhao M., Wang N.*: Research progress on the NSP9 protein of porcine reproductive and respiratory syndrome virus. Front. Vet. Sci. 2022, 9, 872205.
32. *Shannon P., Markiel A., Ozier O., Baliga N. S., Wang J. T., Ramage D., Amin N., Schwikowski B., Ideker T.*: Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003, 13 (11), 2498-2504.
33. *Sheikh A., Al-Taher A., Al-Nazawi M., Al-Mubarak A. I., Kandeel M.*: Analysis of preferred codon usage in the coronavirus N genes and their implications for genome evolution and vaccine design. J. Virol. Methods 2020, 277, 113806.
34. *Snijder E. J., van Tol H., Pedersen K. W., Raamsman M. J., de Vries A.A.*: Identification of a novel structural protein of arteriviruses. J. Virol. 1999, 73 (8), 6335-6345.
35. *Song Z., Zhang Q., Chen Y., Shen H., Yang G., Jiang P., Chen J. L., Lin L.*: The emergence of a novel recombinant porcine reproductive and respiratory syndrome virus with an amino acid insertion in GP5 protein. Microb. Pathog. 2020, 149, 104573.
36. *Sun R., Guo Y., Li X., Li R., Shi J., Tan Z., Zhang L., Zhang L., Han J., Huang J.*: PRRSV non-structural proteins orchestrate porcine E3 ubiquitin ligase RNF122 to promote PRRSV proliferation. Viruses 2022, 14 (2), 424.
37. *Van Dinten L. C., Wassenaar A. L., Gorbalenya A. E., Spaan W. J., Snijder E. J.*: Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. J. Virol. 1996, 70 (10), 6625-6633.
38. *Wang F. X., Liu X., Wu H., Wen Y. J.*: Transcriptome sequencing analysis of porcine alveolar macrophages infected with PRRSV strains to elucidate virus pathogenicity and immune evasion strategies. Virus Dis. 2021, 32 (3), 559-567.
39. *Wang J., Zhang M., Cui X., Gao X., Sun W., Ge X., Zhang Y., Guo X., Han J., Zhou L., Yang H.*: Attenuated porcine reproductive and respiratory syndrome virus regains its fatal virulence by serial passaging in pigs or porcine alveolar macrophages to increase its adaptation to target cells. Microbiol. Spectr. 2022, 10 (6), e0308422.
40. *Wang P., Ma X., Zhang R., Zhao Y., Hu R., Luo C., Zeshan B., Yang Z., Qiu L., Wang J., Liu H., Zhou Y., Wang X.*: The transcriptional characteristics of NADC34-like PRRSV in porcine alveolar macrophages. Front. Microbiol. 2022, 13, 1022481.
41. *Wen X., Bian T., Zhang Z., Zhou L., Ge X., Han J., Guo X., Yang H., Yu K.*: Interleukin-2 enhancer binding factor 2 interacts with the nsp9 or nsp2 of porcine reproductive and respiratory syndrome virus and exerts negatively regulatory effect on the viral replication. Virol. J. 2017, 14 (1), 125.

42. *Wu Q., Han Y., Wu X., Wang Y., Su Q., Shen Y., Guan K., Michal J. J., Jiang Z., Liu B., Zhou X.*: Integrated time-series transcriptomic and metabolomic analyses reveal different inflammatory and adaptive immune responses contributing to host resistance to PRRSV. Front. Immunol. 2022, 13, 960709.
43. *Wu T., Hu E., Xu S., Chen M., Guo P., Dai Z., Feng T., Zhou L., Tang W., Zhan L., Fu X., Liu S., Bo X., Yu G.*: ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb). 2021, 2 (3), 100141, doi: 10.1016/j.xinn.2021.100141.
44. *Yang C., Lan R., Wang X., Zhao Q., Li X., Bi J., Wang J., Yang G., Lin Y., Liu J., Yin G.*: Integrin β3, a RACK1 interacting protein, is critical for porcine reproductive and respiratory syndrome virus infection and NF-κB activation in Marc-145 cells. Virus Res. 2020, 282, 197956.
45. *Yang C., Zuo Q., Liu X., Zhao Q., Pu H., Gao L., Zhao L., Guo Z., Lin Y., Liu J., Bi J., Yin G.*: Small molecule screening identified cepharanthine as an inhibitor of porcine reproductive and respiratory syndrome virus infection in vitro by suppressing integrins/ILK/RACK1/PKCα/NF-κB signalling axis. Vet. Microbiol. 2021, 255, 109016.
46. *Zhang H., Sha H., Qin L., Wang N., Kong W., Huang L., Zhao M.*: Research progress in porcine reproductive and respiratory syndrome virus – host protein interactions. Animals (Basel) 2022, 12 (11), 1381.
47. *Zhang R., Wang P., Ma X., Wu Y., Luo C., Qiu L., Zeshan B., Yang Z., Zhou Y., Wang X.*: Nanopore-based direct RNA-sequencing reveals a high-resolution transcriptional landscape of porcine reproductive and respiratory syndrome virus. Viruses 2021, 13 (12), 2531.
48. *Zhang X., Dong W., Wang X., Zhu Z., He S., Zhang H., Chen Y., Liu X., Guo C.*: Correction: Exostosin glycosyltransferase 1 reduces porcine reproductive and respiratory syndrome virus infection through proteasomal degradation of nsp3 and nsp5. J. Biol. Chem. 2022, 298 (7), 102154.
49. *Zhang Z., Li Z., Li H., Yang S., Ren F., Bian T., Sun L., Zhou B., Zhou L., Qu X.*: The economic impact of porcine reproductive and respiratory syndrome outbreak in four Chinese farms: Based on cost and revenue analysis. Front. Vet. Sci. 2022, 9, 1024720.
50. *Zhang Z., Qu X., Zhang H., Tang X., Bian T., Sun Y., Zhou M., Ren F., Wu P.*: Evolutionary and recombination analysis of porcine reproductive and respiratory syndrome isolates in China. Virus Genes 2020, 56 (3), 354-360.
51. *Zhao G., Zhang J., Sun W., Xie C., Zhang H., Gao Y., Wen S., Ha Z., Nan F., Zhu X., Feng S., Cao X., Zhang Y., Zhu Y., Jin N., Lu H.*: Immunological evaluation of recombination PRRSV GP3 and GP5 DNA vaccines in vivo. Front. Cell Infect. Microbiol. 2022, 12, 1016897.
52. *Zhao M., Sha H., Zhang H., Wang R.*: TRIM4-mediated ubiquitination of NSP2 restricts porcine reproductive and respiratory syndrome virus proliferation. BMC Vet. Res. 2022, 18 (1), 208.
53. *Zhao P., Jing H., Dong W., Duan E., Ke W., Tao R., Li Y., Cao S., Wang H., Zhang Y., Sun Y., Wang J.*: TRIM26-mediated degradation of nucleocapsid protein limits porcine reproductive and respiratory syndrome virus-2 infection. Virus Res. 2022, 311, 198690.
54. *Zheng L., Li X., Zhu L., Li W., Bi J., Yang G., Yin G., Liu J.*: Inhibition of porcine reproductive and respiratory syndrome virus replication in vitro using DNA-based short antisense oligonucleotides. BMC Vet. Res. 2015, 11, 199.
55. *Zhu L., Bi J., Zheng L., Zhao Q., Shu X., Guo G., Liu J., Yang G., Liu J., Yin G.*: In vitro inhibition of porcine reproductive and respiratory syndrome virus replication by short antisense oligonucleotides with locked nucleic acid modification. BMC Vet. Res. 2018, 14 (1), 109.

Corresponding author: Liping Wang, College of Animal Veterinary Medicine, Yunnan Agricultural University, Kunming 650201, Yunnan, China; e-mail: 1620523988@qq.com