한국가축위생학회지 제42권 제2호 (2019)
Korean J Vet Serv, 2019, 42(2), 73-83
ISSN 1225-6552, eISSN 2287-7630
https://doi.org/10.7853/kjvs.2019.42.2.73

Korean Journal of
Veterinary Service
Available online at http://kjves.org

< Original Article >

Computational approaches for prediction of protein-protein interaction between Foot-and-mouth disease virus and *Sus scrofa* based on RNA-Seq

Tamina Park[1,2], Myung-gyun Kang[1], Jinju Nah[3], Soyoon Ryoo[3], Sunghwan Wee[3], Seung-hwa Baek[1,4], Bokkyung Ku[3], Yeonsu Oh[5], Ho-seong Cho[6], Daeui Park[1,2,4]*

[1]*Department of Predictive Toxicology, Korea Institute of Toxicology, Daejeon 34114, Korea*
[2]*Department of Human and Environmental Toxicology, University of Science and Technology, Daejeon 34113, Korea*
[3]*Animal and Plant Quarantine Agency, Gimcheon 39660, Korea*
[4]*Center for Convergent Research of Emerging Virus Infection, Korea Research Institute of Chemical Technology, Daejeon 34114, Korea*
[5]*Department of Veterinary Pathology, College of Veterinary Medicine & Institute of Veterinary Science, Kangwon National University, Chuncheon 24341, Korea*
[6]*College of Veterinary Medicine and Bio-safety Research Institute, Chonbuk National University, Iksan 54596, Korea*

(Received 25 April 2019; revised 24 May 2019; accepted 25 May 2019)

Abstract

Foot-and-Mouth Disease (FMD) is a highly contagious trans-boundary viral disease caused by FMD virus, which causes huge economic losses. FMDV infects cloven hoofed (two-toed) mammals such as cattle, sheep, goats, pigs and various wildlife species. To control the FMDV, it is necessary to understand the life cycle and the pathogenesis of FMDV in host. Especially, the protein-protein interaction between FMDV and host will help to understand the survival cycle of viruses in host cell and establish new therapeutic strategies. However, the computational approach for protein-protein interaction between FMDV and pig hosts have not been applied to studies of the onset mechanism of FMDV. In the present work, we have performed the prediction of the pig's proteins which interact with FMDV based on RNA-Seq data, protein sequence, and structure information. After identifying the virus-host interaction, we looked for meaningful pathways and anticipated changes in the host caused by infection with FMDV. A total of 78 proteins of pig were predicted as interacting with FMDV. The 156 interactions include 94 interactions predicted by sequence-based method and the 62 interactions predicted by structure-based method using domain information. The protein interaction network contained integrin as well as STYK1, VTCN1, IDO1, CDH3, SLA-DQB1, FER, and FGFR2 which were related to the up-regulation of inflammation and the down-regulation of cell adhesion and host defense systems such as macrophage and leukocytes. These results provide clues to the knowledge and mechanism of how FMDV affects the host cell.

Key words : FMDV, *Sus scrofa*, RNA-Seq, Virus-host interaction, Protein-protein interaction

INTRODUCTION

Foot-and-mouth disease (FMD) is a highly contagious trans-boundary disease, which causes huge economic losses. It has important implications for international trade in animal products and for economic impacts of the livestock industry. FMDV, the causing agent, is a non-enveloped virus with a single stranded RNA genome and belongs to the family *Picornaviridae*, genus *Aphthovirus*. The RNA genome of approximately 8.4 kilo nucleotides in length is enclosed within an icosahedral capsid which is composed of 60 copies of four different structural

*Corresponding author: Daeui Park, Tel. +82-42-860-8844, Fax. +82-42-610-8157, E-mail. daeui.park@kitox.re.kr

proteins such as viral protein (VP) 1, VP2, VP3, and VP4. FMDV also includes eight non-structural proteins (L, 2A, 2B, 2C, 3A, 3B, 3C, and 3D) (Gao et al, 2016).

It infects cloven hoofed (two-toed) mammals such as cattle, sheep, goats, pigs and various wildlife species. The FMDV initiates infection via a variety of sites. The transmission is formed by the physical contact with excretions or secretions from infected animals, indirect contact via contaminated all classes of fomites and airborne transmission (Alexandersen et al, 2003). Sometimes, the airborne transmission is speculated to be a key role in some outbreaks (Christensen et al, 2005).

Several studies have shown initial events between FMDV and the receptor of host cells when the infection was established. FMDV has been shown to have highly conserved triplet arginine-glycine-aspartate (Arg-Gly-Asp, RGD) motif in the protruding G-H loop in the VP1 for integrin receptor recognition (Baxt and Becker, 1990). Until now 24 integrins were found in vertebrates, and FMDV utilizes four different types of integrin receptors (Neff et al, 1998; Jackson et al, 2004). Eventually, the FMDV-integrin complex induces viral internalization. Like most RNA viruses, FMDV is introduced into cells using endocytic mechanisms of host cells, of which clathrin-mediated endocytosis is the most potent. Another studies showed that FMDV was mainly absorbed by micropinocytosis to some extent (Han et al, 2016). After FMDV enters the susceptible host cells, the virus first encounters the immune responses. Because viruses depend on the host for survival, growth, and transmission, FMDV exploits host cell machinery to evade host's immune system and defense mechanisms.

To identify the virus replication and persistence mechanism and to control rapid FMDV transmission, it is essential to understand the virus-host protein-protein interactions (PPIs). PPIs are important not only to understand the mechanism of the infection, but can also present new strategies to prevent and treat FMD. Because of the importance of PPI, several researches have made efforts to identify interactions with hosts in many viruses. Despite these efforts, however, there is little known interaction information between FMDV and host proteins because FMDV has few experimentally validated data compared to other viruses such as human immunodeficiency virus type 1, eptein-barr virus and influenza A virus. Therefore, we predicted PPIs between FMDV and its host, *Sus scrofa*, to look for novel virus-host interactions. The computational PPI prediction has been developed in variety of ways using sequence (Shen et al, 2007), domain (Park et al, 2001; Itzhaki, 2011), protein structure (Franzosa and Xia, 2011), and so on. We used both the protein sequence information of already known virus-host interaction and the protein structure information based on the three-dimensional structure domain. The concept of "Interologs" applies to the prediction method of the PPI to extend the interaction network. (Walhout et al, 2000). Interologs are described as an orthologous pair of proteins that interact in different organisms. In order to further analyze the PPIs, additional RNA sequencing (RNA-Seq) analysis was performed using publicly available data about FMDV. The RNA-Seq is a high throughput sequencing technique widely used for transcriptome profiling because the RNA-Seq is more robust and comprehensive transcriptome analyses than microarrays (Rai et al, 2018). It is also good to identify differentially expressed gene transcripts. The gene set enrichment analysis and signature analysis were performed with differentially expressed gene transcripts (DEGs) obtained through the RNA-Seq data to understand how FMDV changes gene expressions in infected host cell. These results can be used to understand the pattern of FMDV infection by observing changes in gene expression in host cell infected with FMDV, and to discover the biological significance for the development of therapeutic agents and vaccines.

MATERIALS AND METHODS

Raw RNS-Seq data from the publicly available archive

The RNA-Seq of FMDV were retrieved from the sequence read archive (SRA, https://www.ncbi.nlm.nih.gov/sra). The SRA is part of the International Nucleotide Sequence Database Collaboration. The RNA-Seq of FMDV-host was used as SRA accession number SRP066333

which includes total six SRA files such as SRR2924891, SRR2924899, SRR2924902, and SRR2924909. These SRA files were performed with Illumina HiSeq 2000 to describe mRNA profiles during FMDV infection. The isolation source is pig kidney cell lines (PK-15) and the strain of FMDV used in cell infection is serotype Asia1 jiangsu strain. PK-15 cells are conventional cell lines for research on FMDV. The four SRA files have difference in harvest time of FMDV infected PK-15 cells. The SRR2924891 was the control of PK-15 cells uninfected with FMDV, the SRR2924899 was harvested at 3 hours post-infection (hpi), SRR2924902 was harvested at 6 hpi, and SRR2924909 was harvested at 12 hpi. The SRA files were preferentially downloaded with SRA toolkit (https://www.ncbi.nlm.nih.gov/sra/docs/toolkitsoft/).

RNA-Seq analysis for identifying differential expressed genes

The downloaded SRA files were converted to FASTQ files with fastq-dump packaged in SRA toolkit 2.6.2 and checked the quality using fastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) version 0.11.5. The sequencing reads in FASTQ files were aligned to the whole pig genome sequence for *Sus scrofa* (NCBI version 11.1) using Spliced Transcripts Alignment to a Reference (STAR) 2.4.1c (Dobin et al, 2013). The pig genome annotation general feature format 3 (GFF3) file was used for mapping because the raw sequence reads downloaded from SRA used pig kidney cell for sampling. The next step was to use featureCounts with the Binary Alignment/Map (BAM) file from the alignment result. The featureCounts is useful program in bioinformatics study for counting reads called read summarization and assigning fragment to the gene (Liao et al, 2014). Finally, the DEGs were identified with a fold change (fc) >2 and p-value <0.05 using EdgeR, which is a Bioconductor software package (https://www.bioconductor.org/) for counting data with an overdispersed *Poisson* model and an empirical *Bayes* procedure (Robinson et al, 2010). The overall procedures are based on the Icahn School of Medicine at Mount Sinai paper described in the F1000Research (Wang and Ma'ayan, 2016).

Data preparation for PPI prediction

To establish PPI network between FMDV and *Sus scrofa*, we have applied the concepts of interologs of DEGs because we could not find any direct interactions between DEGs and *Sus scrofa*. The domain enhanced lookup time accelerated BLAST (DELTA-BLAST), which can yield better homology detection, was selected for the method to iteratively search homologs and orthologous (Boratyn et al, 2012). We first constructed three blast databases using a Viruses.STRING database version 10.5 (Cook et al, 2018), the Structural Classification of Proteins - extended (SCOPe) database 2.07 (Fox et al, 2014), and Swiss-Prot database from UniProtKB (Boutet et al, 2016). The protein sequences of DEGs were retrieved from the *Sus scrofa* protein fasta file version 11.1 from NCBI. A total of eight FMDV protein sequences were used from the NCBI polyprotein sequence (accession number: ABM66095.1). The total eight proteins comprising Peptidase C28 (1-201), VP4_2 (202-285), Rhv (311-475, Rhv_311), Rhv (533-687, Rhv_533), rhv_like (723-859), RNA helicase (1210-1312), Pepticase C3 (1649-1833), and RdRp 1 (1872-2321) were compared with whole genome sequence, and they were aligned to Lab, VP4, VP2, VP3, VP1, 2C, 3C, and 3D, respectively.

Sequence-based interaction prediction

To construct interaction network with proteins which is significantly expressed in FMDV infected cell, the prediction of sequence-based interaction was performed with DEGs. We first found each interologs of FMDV polyproteins and DEGs using DELTA-BLAST. The DELTA-BLAST was iterated five times and the last iteration yield results of DELTA-BLAST were adopted. The cutoff value of blast results were sequence identity $>80\%$ or *E*-value $<10^{-70}$ (Yu et al, 2004). The PPIs between FMDV proteins and DEGs were predicted with the Viruses. STRING database, which contains lots of information about virus-host interactions and virus-virus interactions. The Viruses.STRING database provides not only information about the proteins interacting with each other, but also a combined score indicating the proba-

bility of interaction. The combined score ＞700 was set as the threshold for significant interaction pair. The network of PPIs was visualized with Cytoscape program which is an open source platform (http://www.cytoscape.org/).

Structure-based interaction prediction

Structure-based interaction prediction is based on the domain-domain and protein-protein interaction information for known three-dimensional protein structure which is determined by techniques such as X-ray and NMR. Therefore, structure-based interactions are predicted depending on information obtained from protein structure database. The protein structure database used the SCOPe database which contains the protein domain classified as family and superfamily. The proteins classified hierarchically in SCOPe rely on structural similarity obtained from known 3D structures. All possible pairs of domain-domain interactions were identified as having at least 5 residues in contact within 5 Å distances in protein structure (Park et al, 2001). The method to check interactions between any two domains was to calculate the Euclidean distance to see if they are located within a certain distance threshold. Similar to the sequence-based interaction described above, the interologs of FMDV-host interaction were predicted using the DELTA-BLAST based on the domain-domain interaction information. The predicted network in this way was appended to the previously visualized network as Cytoscape program.

Functional enrichment analysis

In one way, the representative pathway was analyzed by performing an Ingenuity Pathway Analysis (IPA, http://www.ingenuity.com) to determine the significant biological pathways associated with DEGs which were the result of RNA-Seq. IPA program is effective for searching critical pathways through many biological and chemical information proven in a variety of sources, including literatures. As a next step, the pig's DEGs interacting directly with the FMDV proteins were subjected to Gene Ontology (GO) analysis using ClueGO which generated a functional network of organized GO/pathway terminology (Bindea et al, 2009). In ClueGO, kappa statistics was used to link the terms of the network.

RESULTS

RNA-Seq analysis

Prior to establishing the network between FMDV and *Sus scrofa*, the DEGs were preferentially identified through RNA-Seq analysis. Since direct interaction between the FMDV proteins and all pig proteins can lead to too much false-positives, it is necessary to predict the interaction using only DEG to reduce this. The RNA-Seq analysis was performed according to the methods described in MATERIALS AND METHODS with a total of four samples of control (SRR2924891), 3 hpi (SRR2924899), 6 hpi (SRR2924902), and 12 hpi (SRR2924902) downloaded from SRA. The results of RNA-Seq analysis show that the expression in host cells infected with FMDV differs over time. A total of 40 genes were differentially expressed in 3 hpi (22 up-regulated genes and 18 down-regulated genes). The total 40 DEGs, which were changed in 3hpi, could be considered the initial response in FMDV-infected pig host cells. Among the total 40 DEGs, the LOC100525442 (log_2fc: 5.04, *P*-value: 0.02), LOC100152003 (log_2fc: 4.64, *P*-value: 0.04), IFN-ALPHA-9 (log_2fc: −5.36, *P*-value: 0.01), MIR30C-2 (log_2fc: −5.04, *P*-value: 0.02), MIR378-1 (log_2fc: −4.64, *P*-value: 0.04), and LOC100523076 (log_2fc: −4.64, *P*-value: 0.04) genes appear to be important. It is well known that FMDV is highly sensitive to the action of interferon. One example is that the leader proteinase of FMDV elicits the inhibition of the type I interferon mechanisms, including interferon alpha, and evades the host innate immune responses (Wang et al, 2011). LOC100525442, LOC100152003, LOC100523076 are genes encoding polyadenylate-binding protein 1-like 2, olfactory receptor 2S2-like, and olfactory receptor 11H12-like, respectively. To the next, a total of 1674 genes were differentially expressed in 6 hpi (637 up-regulated genes and 1037 down-regulated genes), and total 1931 genes were identified as DEGs in 12 hpi (707 up-regulated genes and 1224 down-regulated

genes). The DEGs were grouped into biological pathways and gene ontologies by performing the functional enrichment analysis.

The functional enrichment analysis

The canonical pathway analysis with gene set was first performed using IPA, the functional analysis tool. The canonical pathways in IPA are well-organized metabolic and cell signaling pathways, and include information from diverse journal articles, text books, and HumanCyc (IPA's KEGG pathway substitution). In 6 hpi, functional enrichment analysis demonstrated that the up-regulated genes were significantly enriched in 'Th1

Table 1. The top 10 canonical pathways by significantly up-regulated genes in 6 hpi

Ingenuity canonical pathways	-log (*P*-value)	Molecules
Wnt/β-catenin signaling	3.03	WNT7A, SFRP5, WNT6, KREMEN2, RARG, ACVR2B, SOX13, ACVR1C
Th1 and Th2 activation pathway	2.22	LTA, IL10RA, ACVR2B, CD8A, ACVR1C, NOTCH1, ICOSLG/LOC102723996
Role of IL-17A in psoriasis	2.02	S100A9, S100A8
Notch signaling	2.01	DTX4, HES7, NOTCH1
Osteoarthritis pathway	1.91	MATN3, S1PR2, S100A9, COL10A1, GREM1, S100A8, NOTCH1
Crosstalk between dendritic cells and natural killer cells	1.69	IL3RA, LTA, IFNB1, ACTA1
Th1 Pathway	1.68	LTA, IL10RA, CD8A, NOTCH1, ICOSLG/LOC102723996
Role of cytokines in mediating communication between immune cells	1.60	IFNB1, IFNL1, IL17F
Proline degradation	1.46	PRODH2
Apelin liver signaling pathway	1.44	APLN, COL10A1

Table 2. The top 10 canonical pathways by significantly down-regulated genes in 6 hpi

Ingenuity canonical pathways	-log (*P*-value)	Molecules
Role of macrophages, fibroblasts and endothelial cells in rheumatoid arthritis	5.53	CXCL8, IL1A, PLCB2, CAMK4, PIK3R1, IL15, VEGFC, CEBPB, IL6, CREB5, PLCH1, CEBPG, APC, VEGFA, TLR4, JUN, NFAT5, CSF1, OSM, ATF4, FZD5, TLR3, PDGFD, CSF2
Hepatic fibrosis/hepatic stellate cell activation	5.19	CXCL8, IL1A, CD40LG, CTGF, KLF6, SMAD7, VEGFC, IL6, CXCL9, VEGFA, TLR4, IGF2, EDN1, CSF1, PDGFD, SERPINE1, A2M
Osteoarthritis pathway	5.04	TIMP3, CXCL8, SMAD9, DDIT4, PTHLH, SMAD7, VEGFC, GDF5, HES1, CEBPB, CREB5, VEGFA, TLR4, ATF4, MEF2C, FZD5, PTGS2, JAG1
HMGB1 signaling	4.88	CXCL8, IL1A, LIF, IL17C, PIK3R1, RHOJ, IL6, TLR4, JUN, KAT2B, RND3, OSM, CSF2, SERPINE1
Altered T Cell and B Cell signaling in rheumatoid arthritis	4.01	TLR4, CD40LG, IL1A, PRTN3, CSF1, IL15, TLR3, IL6, IL23A, CSF2
Adipogenesis pathway	3.77	NR1D2, KAT2B, RUNX1T1, EGR2 ,DDIT3, FOXO1, SMAD9, TXNIP, KLF5, FZD5, CEBPB, NOCT
Hematopoiesis from pluripotent stem cells	3.64	CD247, CXCL8, IL1A, LIF, CSF1, IL6, CSF2
Role of cytokines in mediating communication between immune cells	3.37	CXCL8, IL1A, IL20, IL15, IL6, IL23A, CSF2
Crosstalk between dendritic cells and natural killer cells	3.35	TLR4, CD40LG, TLN2, IL15, TNFSF10, TLR3, IL6, SF2, ACTC1
ILK Signaling	3.25	PIK3R1, ACTN2, ACTN3, VEGFC, RHOJ, RICTOR, CREB5, VEGFA, JUN, RND3, ATF4, PTGS2, ITGB6, ACTC1

and Th2 Activation Pathway' (LTA, IL10RA, ICOSLG/LOC102723996, ACVR2B, CD8A, ACVR1C, NOTCH1) and 'Role of IL-17A in Psoriasis' (S100A9, S100A8) (Table 1), whereas the down-regulated genes were enriched in 'Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis' (CXCL8, IL1A, PLCB2, CAMK4, PIK3R1, IL15, VEGFC, CEBPB, IL6, CREB5, PLCH1, CEBPG, APC, VEGFA, TLR4, JUN, NFAT5, CSF1, OSM, ATF4, FZD5, TLR3, PDGFD, CSF2) (Table 2). The results of the IPA analysis suggested that the host defense system such as macrophage was down-regulated by FMDV infection. And, the inflammation related pathways such as 'inflammatory and immune response' and 'cytokine and chemokine secretions' were the most significantly down-regulated in 6 hpi after the host cells infected by FMDV.

In 12 hpi, the up-regulated genes involved in 'Th1 and Th2 Activation Pathway' (CD3G, SOCS1, IL18, DLL1, LTA, TNFRSF4, IKZF1, GFI1, IL27, TBX21, IL24, IL25) and engaged in 'Role of Hypercytokinemia/

Table 3. The top 10 canonical pathways by significantly up-regulated genes in 12 hpi

Ingenuity canonical pathways	-log (*P*-value)	Molecules
Pathogenesis of multiple sclerosis	5.41	CXCL10, CXCL11, CCL4, CCL5
Th1 and Th2 activation pathway	5.08	CD3G, SOCS1, IL18, DLL1, TNFRSF4, LTA, IKZF1, GFI1,IL27, TBX21, IL24, IL25
Role of hypercytokinemia/hyperchemokinemia in the pathogenesis of influenza	4.64	CXCL10, IL18, CCL4, CCL2, IFNB1, CCL5
Interferon signaling	3.94	IFIT1, IFIT3, SOCS1, IFNB1, ISG15
Role of cytokines in mediating communication between immune cells	3.10	IL18, IFNB1, IL27, IL24, IL25
Th2 pathway	3.00	CD3G, DLL1, TNFRSF4, IKZF1, GFI1, TBX21, IL24, IL25
Agranulocyte adhesion and diapedesis	2.90	CXCL10, CXCL11, IL18 ,MYH2, CCL4, CCL2, CCL3L1, CCL5, ACTA1
Notch signaling	2.78	DLL1, CNTN1 ,HEY2, HES7
Differential regulation of cytokine production in macrophages and T helper cells by IL-17A and IL-17F	2.77	CCL4, CCL2, CCL5
Communication between Innate and adaptive immune cells	2.72	CXCL10, IL18, CCL4, CD79B, IFNB1, CCL5

Table 4. The top 10 canonical pathways by significantly down-regulated genes in 12 hpi

Ingenuity Canonical Pathways	-log (*P*-value)	Molecules
Leukocyte extravasation signaling	2.84	CLDN15, MMP16, ACTN2, ACTN3, FGFR2, RAPGEF4, TIMP4, ARHGAP9, GAB1, WAS, VAV3, RASGRP1, FER, CLDN2, CLDN22
Regulation of the epithelial-mesenchymal transition pathway	2.72	MAP2K6, ESRP2, WNT2B, FGFR2, APC, FGF21, GAB1, FGF18, HGF, ZEB2, APH1B, PDGFD, FZD2, WNT1
Role of macrophages, fibroblasts and endothelial cells in rheumatoid arthritis	2.51	MAP2K6, CXCL8, WNT2B, IL15, FGFR2, NFATC4, CREB5, APC, PLCH1, NFATC1, CEBPE, GAB1, TLR6, VEGFD, SFRP5, TLR3, PDGFD, FZD2, WNT1
Neuropathic pain signaling in dorsal horn neurons	2.17	GAB1, GRM3, CAMK1G, FGFR2, TAC1, GRM4, KCNN2, GRIA4, PLCH1
Hepatic fibrosis/hepatic stellate cell activation	2.03	COL19A1, CXCL8, CD40LG, COL4A3, HGF, SMAD7, FGFR2, VEGFD, MYH11, CXCL9, PDGFD, COL28A1
Osteoarthritis pathway	2.00	CXCL8, FGF18, SMAD9, GLIS1, SMAD7, VEGFD, GDF5, SP7, CREB5, C1QTNF4, GLI1, FZD2, PPARGC1A
NAD salvage pathway III	1.79	NMNAT3, NMRK2
Amyotrophic lateral sclerosis signaling	1.78	CAPN8, CACNA1S, CACNA1D, GAB1, FGFR2, CAPN3, VEGFD, GRIA4,
Role of JAK2 in hormone-like cytokine signaling	1.77	GHR, EPOR, SOCS2, PRLR
FGF signaling	1.76	MAP2K6, FGF21, FGF18, GAB1, HGF, FGFR2, CREB5

hyperchemokinemia in the Pathogenesis of Influenza' (CXCL10, IL18, CCL4, CCL2, IFNB1, CCL5) (Table 3). The down-regulated genes were enriched in 'Leukocyte Extravasation Signaling' (MAP2K6, ESRP2, WNT2B, FGFR2, APC, FGF21, GAB1, FGF18, HGF, ZEB2, APH1B, PDGFD, FZD2, WNT1), and 'Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis' (MAP2K6, CXCL8, WNT2B, IL15, FGFR2, NFATC4, CREB5, APC, PLCH1, NFATC1, CEBPE, GAB1, TLR6, VEGFD, SFRP5, TLR3, PDGFD, FZD2, WNT1) (Table 4). The results at 12 hpi showed that the downregulation of host defense mechanisms and the up-regulation of inflammation were similar with the 6 hpi. It is anticipated that the intrusion of FMDV caused intracellular inflammation and host defense mechanisms to be regulated. This is a meaningful evidence of the general phenomenon that occurs when FMDV enters the host cell.

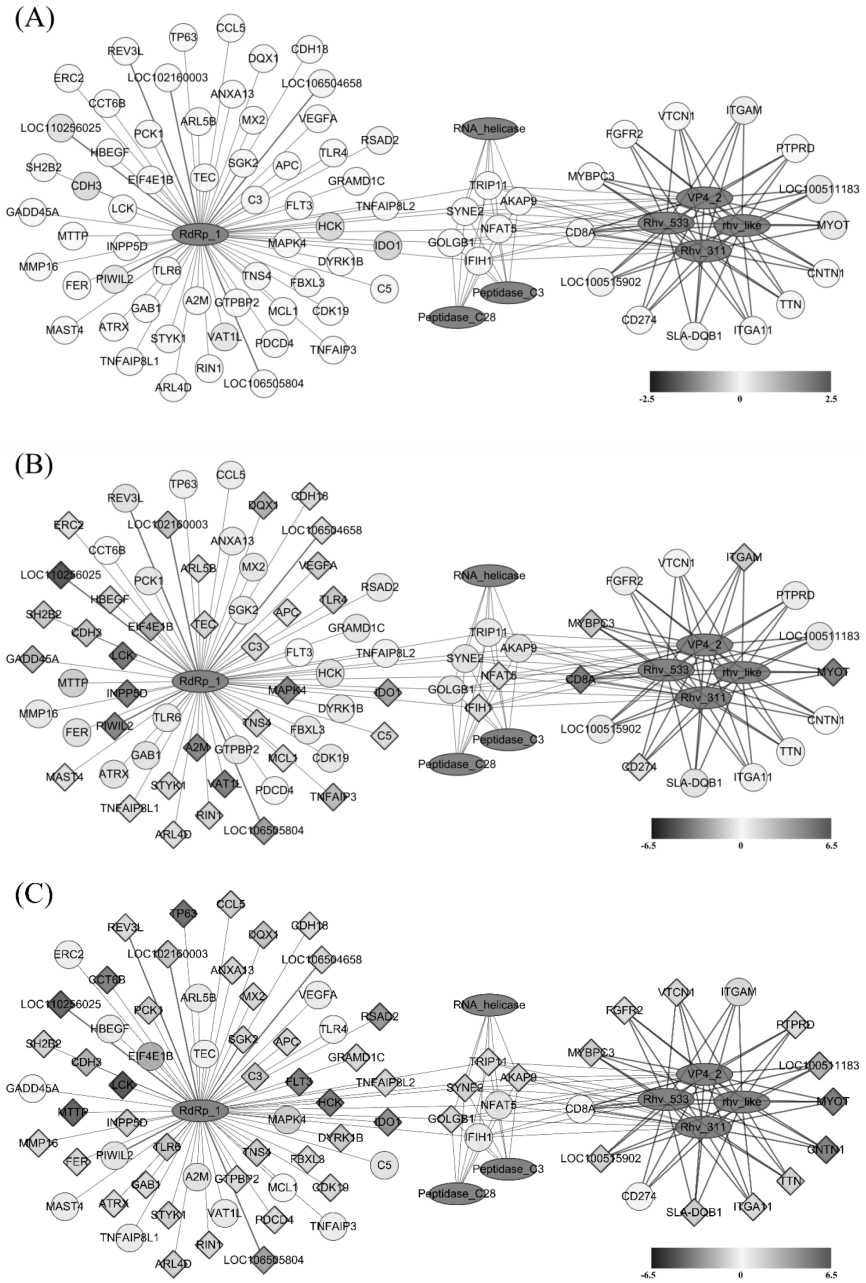

Fig. 1. The network of interaction between FMDV proteins and DEGs in 3 hpi (A), 6 hpi (B), 12 hpi (C). The green circle indicates FMDV proteins and the others are DEGs among swine (*Sus scrofa*) proteins. The purple border of the diamond indicates DEGs that were significantly expressed over each FMDV infection time (6 hpi and 12 hpi). A total of 156 direct interactions were predicted. The lines indicates the predicted interaction. The thick lines are predicted with structure-based methods and the others are predicted with sequence-based methods.

The predicted PPIs network

To comprehend the impacts of FMDV in host cell, the virus-host interactions have been predicted with DEGs. And further, the network was constructed to display the change of expression in 3 hpi, 6 hpi, and 12 hpi (Fig. 1). As a result of investigating pig's proteins interacting with a total of eight FMDV proteins, it was predicted that 156 direct interactions with 78 DEGs of pig's proteins. We used two methods, which are sequence-based and structure-based, in predicting interactions. The sequence-based method predicted 94 interactions of proteins with many functional studies in other species such as RNA dependent RNA polymerase (RdRp), helicase and peptidase. On the other hand, since the structure-based method uses domain information of the 3D structure, the 62 interactions of FMDV surface proteins such as VP4_2, rhv_like and Rhv were predicted. In addition to confirming the Integrin Subunit Alpha M (ITGAM), a subunit of cell surface protein that plays an important role in internalization of FMDV into the host cell, we also observed that expression was increased at 3 hpi, 6 hpi, and 12 hpi (log_2fc: 0.19, 1.96, and 1.45, respectively). We focused on the genes whose expression patterns persisted over time, including the up-regulated ITGAM, IDO1, VTCN1, C3, and MYBPC3 as well as the down-regulated SLA-DQB1, C5, MYOT, CDH3, CDH18, and STYK1. In gene ontology analysis, these genes were also identified as important genes.

The functional enrichment analysis (ClueGO)

To better understand the pathway, we performed the different functional enrichment analysis using ClueGO (Fig. 2). We screened the 78 key genes that were predicted to interact with FMDV among DEGs. As it is shown in Fig. 1, there are 42 gene ontology terms that is including 11 representative pathways merged as redundant group with >50.0% gene overlap. This functional enrichment analysis demonstrated that the DEGs were abundant in 'negative regulation of leukocyte activation' (GOID: 0002695), 'cytokine production involved in inflammatory response' (GOID: 0002534), 'cell adhesion molecules' (GOID: 04514), 'protein tyrosine kinase activity' (GOID: 0004713), etc.

Applying the expression data of the RNA-Seq analysis, Serine/Threonine/Tyrosine Kinase 1 (STYK1, log_2fc: −0.05, −1.46, −1.76), V-set domain-containing T-cell activation inhibitor 1 (VTCN1, log_2fc: −0.04, 0.15, 1.38), indoleamine 2,3-dioxygenase (IDO1, log_2fc: 0.56, 3.62, 4.41), the expressions of Cadherin 3 (CDH3, log_2fc: 0.48 in 3 hpi, −2.20 in 6 hpi, and −2.94 in 12 hpi), SLA class II histocompatibility antigen, DQ haplotype C beta chain (SLA-DQB1, log_2fc: −0.08, −0.89,

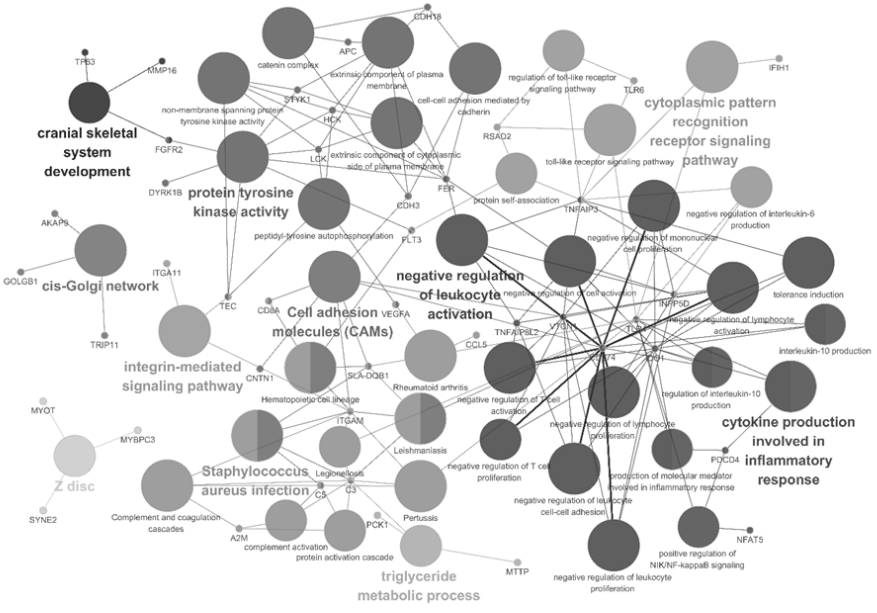

Fig. 2. The linkage of 42 gene ontology terms and 43 related genes (red). The kappa score was set to 0.4 (default) and the network only showed pathways with p-value less than 0.05. The 11 representative terms (total 11 colors) which are merged as redundant groups with >50.0% sharing genes were highlighted.

−1.85), Tyrosine-protein kinase Fer (FER, log_2fc: −0.02, −1.00, −1.49), and Fibroblast Growth Factor Receptor 2 (FGFR2, log_2fc: 0.002, −0.49, −1.20) suggested that the FMDV infection could lead to increased inflammatory responses and decreased cell adhesion and leukocytes. In particular, the presence of the 'integrin-mediated signaling pathway' (GOID: 0007229) was highly consistent with previous studies on FMDV that integrin is the major key proteins used for internalization of FMDV. These results were visually well represented in Fig. 2, supporting the results of IPA.

DISCUSSION

FMD is an infectious disease that is a global problem, and efforts are being made to control it. As part of this effort, we analyzed the network and forecasted the effect of FMDV in the host cell. The proteins known to interact between FMDV and *Sus scrofa* are BECN1 (Gladue et al, 2012). RTN1, RTN3, RTN4, KHDRBS1, CHUK, IKBKB, and RELA were additionally predicted for each their putative homologs to interact with FMDV proteins from Viruses.STRING database. There is little experimental or predictive researches on interaction of FMDV. By this necessity, we had predicted 78 pig's proteins interacting with seven FMDV proteins using sequence information and 3D structure information. The prediction accuracy of the network was improved by applying for expression information of DEGs that varied significantly from the early infection stage (3 hpi) to the late infection stage (12 hpi). Even though we tried to reduce the false-positive, additional experiments need to verify that the interaction of the proposed genes is really valid.

Additionally, the result of RNA-Seq supported that the significant change of biological pathways occurs when FMDV enters the host cell. In the early stage by FMDV infection, interferon alpha 9 was up-regulated in host cell. For the late stage, the FMDV infection resulted in increased inflammation and the downregulation of host defense systems such as macrophage and leukocytes in host cell. Zhang et al, suggested differentially expressed genes such as IL6, CXCL2, CCL4, CCL20, FOS, and NFKBIA which are involved in the host responses to FMDV infection. The research also performed the RNA-seq analysis of PK-15 cells infected with FMDV (Zhang et al, 2018). Looking at the gene expression patterns of CCL2, CCL4, CCL5, CXCL10, and CXCL11, their results could support our conclusion that increased inflammation was the major response against FMDV infection.

Also, FMDV is known to block viral countermeasures by modulating the immunity of host pathways. The major targets representing viral evasion strategies are Toll-like receptors (TLRs) and TLR adaptor TRIF which are virus recognition molecules of host innate immune system (Lester and Li, 2014). Previous picornavirus studies such as coxsackievirus B, enterovirus 71, and hepatitis A virus have been shown to inhibit TLR3-adapted TRIF which is cleaved by FMDV-coding protease (Rodríguez Pulido and Sáiz, 2017). Our results showed that TLR3, TLR4, and TLR6 were down-regulated in host cell after FMDV infection. It seems to be the resistance of FMDV to antiviral immunity of the host. Viral infection affects a large number of proteins in its host. It needs to focus significant targets because a total of proteins make highly huge networks and it is hard to understand the pathogenicity. We added interaction network information to signaling pathway information and finally highlighted genes that are thought to play an important role in FMDV infection. The genes interacting with FMDV were finally presented and these genes could be used as diagnostic markers or drug targets for the development of therapeutic agents.

ACKNOWLEDGEMENTS

This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPET) through Animal Disease Management Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (116097033SB010) and the National Research Council of Science & Technology grant by the Korea government (CRC-16-01-KRICT).

REFERENCES

Alexandersen S, Zhang Z, Donaldson A. & Garland AJ. 2003. The Pathogenesis and Diagnosis of Foot-and-Mouth Disease. Journal of Comparative Pathology, 129(1), 1-36. doi:10.1016/S0021-9975(03)00041-0

Baxt B. & Becker Y. 1990. The effect of peptides containing the arginine-glycine-aspartic acid sequence on the adsorption of foot-and-mouth disease virus to tissue culture cells. Virus genes, 4(1), 73-83. http://www.ncbi.nlm.nih.gov/pubmed/2168107. Accessed 31 December 2018

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics (Oxford, England), 25(8), 1091-3. doi:10.1093/bioinformatics/btp101

Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ. & Madden TL. 2012. Domain enhanced lookup time accelerated BLAST. Biology direct, 7, 12. doi:10.1186/1745-6150-7-12

Boutet E, Lieberherr D, Tognolli M, Schneider M, Bansal P, Bridge AJ, et al. 2016. UniProtKB/Swiss-Prot, the Manually Annotated Section of the UniProt KnowledgeBase: How to Use the Entry View (pp. 23-54). Humana Press, New York, NY. doi:10.1007/978-1-4939-3167-5_2

Christensen LS, Normann P, Thykier-Nielsen S, Sørensen JH, Stricker K. de & Rosenørn S. 2005. Analysis of the epidemiological dynamics during the 1982-1983 epidemic of foot-and-mouth disease in Denmark based on molecular high-resolution strain identification. Journal of General Virology, 86(9), 2577-2584. doi:10.1099/vir.0.80878-0

Cook H, Doncheva N, Szklarczyk D, von Mering C, Jensen L, Cook HV, et al. 2018. Viruses.STRING: A Virus-Host Protein-Protein Interaction Database. Viruses, 10(10), 519. doi:10.3390/v10100519

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England), 29(1), 15-21. doi:10.1093/bioinformatics/bts635

Fox NK, Brenner SE. & Chandonia JM. 2014. SCOPe: Structural Classification of Proteins—extended, integrating SCOP and ASTRAL data and classification of new structures. Nucleic Acids Research, 42(D1), D304-D309. doi:10.1093/nar/gkt1240

Franzosa EA. & Xia Y. 2011. Structural principles within the human-virus protein-protein interaction network. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.1101440108

Gao Y, Sun SQ. & Guo HC. 2016. Biological function of Foot-and-mouth disease virus non-structural proteins and non-coding elements. Virology Journal, 13(1), 107. doi:10.1186/s12985-016-0561-z

Gladue DP, O'Donnell V, Baker-Branstetter R, Holinka LG, Pacheco JM, Fernandez-Sainz I. et al. 2012. Foot-and-mouth disease virus nonstructural protein 2C interacts with Beclin1, modulating virus replication. Journal of virology, 86(22), 12080-90. doi:10.1128/JVI.01610-12

Han SC, Guo HC, Sun SQ, Jin Y, Wei YQ, Feng X, et al. 2016. Productive Entry of Foot-and-Mouth Disease Virus via Macropinocytosis Independent of Phosphatidylinositol 3-Kinase. Scientific Reports, 6(1), 19294. doi:10.1038/srep19294

Itzhaki Z. 2011. Domain-domain interactions underlying herpesvirus-human protein-protein interaction networks. PLoS ONE. doi:10.1371/journal.pone.0021724

Jackson T, Clark S, Berryman S, Burman A, Cambier S, Mu D. et al. 2004. Integrin alphavbeta8 functions as a receptor for foot-and-mouth disease virus: role of the beta-chain cytodomain in integrin-mediated infection. Journal of virology, 78(9), 4533-40. doi:10.1128/JVI.78.9.4533-4540.2004

Lester SN. & Li K. 2014. Toll-like receptors in antiviral innate immunity. Journal of molecular biology, 426(6), 1246. doi:10.1016/J.JMB.2013.11.024

Liao Y, Smyth GK. & Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics, 30(7), 923-930. doi:10.1093/bioinformatics/btt656

Neff S, Sá-Carvalho D, Rieder E, Mason PW, Blystone SD, Brown EJ. & Baxt B. 1998. Foot-and-mouth disease virus virulent for cattle utilizes the integrin alpha(v)beta3 as its receptor. Journal of virology, 72(5), 3587-94. http://www.ncbi.nlm.nih.gov/pubmed/9557639. Accessed 2 January 2019

Park J, Lappe M. & Teichmann SA. 2001. Mapping protein family interactions: intramolecular and intermolecular protein family interaction repertoires in the PDB and yeast. Journal of Molecular Biology, 307(3), 929-938. doi:10.1006/JMBI.2001.4526

Rai MF, Tycksen ED, Sandell LJ. & Brophy RH. 2018. Advantages of RNA-seq compared to RNA microarrays for transcriptome profiling of anterior cruciate ligament tears. Journal of orthopaedic research : official publication of the Orthopaedic Research Society, 36(1), 484-497. doi:10.1002/jor.23661

Robinson MD, McCarthy DJ. & Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England), 26(1), 139-40. doi:10.1093/bioinformatics/btp616

Rodríguez Pulido M. & Sáiz M. 2017. Molecular Mechanisms of Foot-and-Mouth Disease Virus Targeting the Host Antiviral Response. Frontiers in cellular and infection microbiology, 7, 252. doi:10.3389/fcimb.2017.00252

Shen J, Zhang J, Luo X, Zhu W, Yu K, Chen K, et al. 2007. Predicting protein-protein interactions based only on sequences information. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.0607879104

Walhout AJ, Sordella R, Lu X, Hartley JL, Temple GF, Brasch MA, et al. 2000. Protein interaction mapping in C. ele-

gans using proteins involved in vulval development. Science (New York, N.Y.), 287(5450), 116-22. doi:10.1126/SCIENCE.287.5450.116

Wang D, Fang L, Li P, Sun L, Fan J, Zhang Q, et al. 2011. The leader proteinase of foot-and-mouth disease virus negatively regulates the type I interferon pathway by acting as a viral deubiquitinase. Journal of virology, 85(8), 3758-66. doi:10.1128/JVI.02589-10

Wang Z. & Ma'ayan A. 2016. An open RNA-Seq data analysis pipeline tutorial with an example of reprocessing data from a recent Zika virus study. F1000Research, 5, 1574. doi:10.12688/f1000research.9110.1

Yu H, Luscombe NM, Lu HX, Zhu X, Xia Y, Han JDJ, et al. 2004. Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. Genome research, 14(6), 1107-18. doi:10.1101/gr.1774904

Zhang T, Chen H, Qi L, Zhang J, Wu R, Zhang Y. & Sun Y. 2018. Transcript Profiling Identifies Early Response Genes against FMDV Infection in PK-15 Cells. Viruses, 10(7), 364. doi:10.3390/v10070364