HUE UNIVERSITY
INSTITUTE OF BIOTECHNOLOGY

NGUYỄN TRẦN TRUNG

CHARACTERISTICS OF STRUCTURAL PROTEIN ENCODING GENES OF PORCINE PARVOVIRUS (PPV) IN CENTRAL PROVINCES OF VIETNAM.

SUMMARY OF DOCTORAL DISSERTATION

Major: Biology
Code number: 9420101

HUẾ - 2024

The study was conducted in: The institute of Biotechnology, Hue university

Supervisors:

1. ASSOC.PROF. PhD. Trần Quốc Dung

Department of Biology

University of Education, Hue University

2. PhD. Nguyễn Thị Diệu Thuý

Institute of Biotechnology,

Vietnam Academy of Science and Technology

Reviewer 1:

Reviewer 2:

Reviewre 3:

The dessertation will be defended at the Thesis Judging Committee at Hue University, at the Hue University office, at …….hour…… on……. month……. year……..

The dessertation is available at:

1. Vietnam's National library
2. Hue University Learning resource center.
3. Library of Institute of Biotechnology, Hue University

INTRODUCTION

1. RATIONALE FOR THE STUDY

Reproductive quality of sows plays an important role in providing good breeding stock for pig production. In countries with a developed pig industry, including Vietnam, sows often have problems related to fertility decline. Porcine Parvovirus (PPV) is considered to be the cause of Sow Fertility Deficiency Syndrome (SMEDI) (Stillbirth, Mummification, Embryonic Death and Infertility).

In theory, PPV is a single-stranded DNA virus, which used the host's DNA replication machinery to replicate its genetic material, so the virus is thought to have a low rate of genetic variation. However, studies in recent years have shown that the level of nucleotide substitutions of PPV is comparable to that of RNA viruses (about 10-4 substitutions/site/year). As of September 2022, eight PPV genotypes (1-8) have been confirmed. In recent years, the pig industry has been greatly affected by the emergence of many diseases such as blue ear disease, foot-and-mouth disease, and acute diarrhea, so the risk of infection with pathogens is very high. When an animal is infected, the immune system rapidly declines and there is a good opportunity for other pathogens that are in a latent state to erupt. This also means an increased risk of genetic recombination between strains/subtypes of viruses. Although PPV has appeared and existed for a long time on a worldwide scale, and vaccines have been used effectively for 3 decades, but considering the current situation, it is absolutely essential to control, monitor and evaluate the genetic changes and molecular epidemiology of this virus. Recent publications have confirmed that there are 4 PPV genotypes (1-4) circulating in pigs raised in our country. Currently, there are no scientific studies on genetic characteristics, and genome’s characteristics of PPV isolated on pigs

raised in central Vietnam. On the scientific and practical basis mentioned above, we conducted the project : "Characteristics of structural protein encoding genes of Porcine parvovirus (PPV) in Central provinces of Vietnam". The results of the study are an important scientific basis to evaluate the molecular evolution of PPV, contributing to the control of PPV-induced fertility decline syndrome in pigs.

2. OBJECTIVES OF THE STUDY

2.1. Overall objectives

Determining changes in the predicted nucleotide and amino acid sequences of PPV strains circulating in pigs in some Central provinces through surveying the genetic characteristics of PPV virus **2.2. Detailed objectives**

- Isolation of PPV genotypes (1-4) circulating in the central provinces.

- Determination of nucleotide sequences of genes encoding structural proteins of PPV genotypes isolated in the central provinces of Vietnam.

- Analysis and evaluation of genetic changes at the nucleotide and inferrd amino acid levels of genes encoding structural proteins of PPV isolated in the central provinces of Vietnam.

3. CONTENT OF THE STUDY

This study focuses on addressing the following main issues:

- Collecting samples, and extracting total DNA

- Isolating and screening PPV genotypes

- Amplifying and determining the nucleotide sequence of the gene encoding the structural protein of PPV

- Analyzing, evaluating changes in nucleotide and inferred amino acid sequences, inferring genes encoding structural proteins and building genetic genealogical trees of PPV strains

4. SCOPE OF THE STUDY

- Sampling locations: 7 central provinces (Quang Binh, Quang Tri, Thua Thien Hue, Da Nang, Quang Nam, Quang Ngai, Binh Dinh).

- Research location: Department of Animal Gene Technology, Institute of Biotechnology, Vietnam Academy of Science and Technology.

- Research period: 2019-2023.

5. NEW FINDINGS OF THE DESSERTATION

- Determining the prevalence and co-infection rates of 4 PPV genotypes (1-4) in domestic pigs in 7 central provinces.

- Sequencing and analyzing gene sequences encoding structural proteins of 03 PPV2 strains (GenBank code: OL913365-OL913367) and genome sequences of 02 PPV4 strains (GenBank code: MT434668-MT434669) isolated in Central Vietnam.

- Analyzing the genetic relationship of PPV2 and PPV4 strains isolated in Central Vietnam with PPV strains circulating in the region and in the world.

6. SCIENTIFIC AND PRACTICAL IMPLICATIONS OF THE DESSERTATION

- Results of sequence analysis, evaluation of genetic variation, genetic taxonomy of PPV strains isolated in central Vietnam will be valuable information in molecular epidemiology studies on PPV strains, effectively supporting the management and control of Swine Fertility Reduction Syndrome, contributing to minimizing economic losses to the pig industry not only in the central provinces but in Vietnam as well.

- Providing updated scientific data to facilicate research and teaching students specialized in Biology, Biotechnology, Animal Husbandry and Veterinary Medicine at universities.

CHAPTER 1. OVERVIEW

1.1. CURRENT SITUATION OF PIG FARMING AND DISEASES IN CENTRAL PROVINCE OF VIETNAM

Pig farming ranks first in the livestock industry in our country. However, the tropical monsoon climate, which is hot and humid with a lot of rain in the Central region, has provided many types of pig pathogens, including PPV, with favorable conditions to survive and caused significant damage to the pig industry.

1.1. INTRODUCTION ABOUT PPV AND PPV-CAUSED DISEASES IN PIGS.

1.2.1. PPV Classification

PPV (Porcine Parvovirus) belongs to:

Family: Parvoviridae

Sub-family: Parvovirinae

Genus: *Protoparvovirus*

Species: *Ungulate protoparvovirus*

Sub-species: Porcine parvovirus (PPV)

1.2.2. Biological characteristics.

1.2.2.1. Morphology

PPV is a non-enveloped virus, with small size, round shape, virion diameter about 18-26 nm, 20-sided symmetry structure.

1.2.2.2. Capsid protein

The capsid of PPV is a spherical shell consisting of 60 copies of VP1 or VP2 arranged in a symmetric polyhedral structure (Chapman and Rossmann, 1993).

1.2.2.3. Genome

The genetic material of PPV is ssDNA with a molecular size ranging from 5-6 kb. The sequence on the two-terminal ends of the PPV gives rise to a

hairpin structure (Bern and Hauswirth, 1983). The PPV genome consists of two main open reading frames: ORF1 encoding the nonstructural proteins NS1, NS2 (Bergeron et al., 1993) and ORF2 encoding the capsid proteins. ORF3 (PPV4) is a small open reading frame, located between ORF1 and ORF2, encoding the nonstructural protein NS3.

1.2.3. PPV Symptoms of PPV-caused diseases

Although PPV infects pigs of all ages, the disease mainly occurs in sows with the main clinical symptops of impaired fertility, characterized by miscarriage, premature birth, embryonic death, and mummification.

1.2.4. Epidemiology

The worldwide circulation of PPV has been confirmed (Cadar et al., 2012; Truyen and Streck, 2012). PPV is resistant to environmental factors and immune to many common disinfectants (Brown, 1981). PPV infects pigs through three main routes: food, water and insemination.

1.2.5. Interactions between viruses and cells

PPV enters cells by all three pathways: (i) endocytosis, (ii) macrocytosis, and (iii) a third pathway of entry (presumptive). PPV DNA and proteins enter the host genome and are edited by the host's genetic machinery into double-stranded DNA. After each time the host cell duplicates, the viral DNA is also duplicated.

1.2.6. Disease mechanisms

PPV infection induces cytopathogenic effect (CPE) or mitochondrial-mediated apoptosis (Zhang et al., 2019). The actual outcome of PPV infection largely depends on the virus strain and cell type (Zhang et al., 2015).

1.2.7. Diagnosis of PPV

PPV can be detected by hematologic methods including SN, MDCF, HI, HA (Mengeling, 1972; Morimoto et al., 1972; Joo et al., 1976) or immunofluorescence assays and ELISA assays. Currently, PCR and Real-

time PCR are the most useful techniques for detecting PPV in a variety of samples (Milek et al., 2019).

1.2.8. Prevention of PPV

Inactivated vaccines and attenuated vaccines are traditional vaccines, which have evolved into routine injections, creating a long-lasting immune response that helps protect the pig herd. Currently, VLP vaccine (VLP-virus-like particle) and recombinant vaccine based on VP2 gene sequence of PPV are currently under research and development to gradually replace traditional vaccines.

1.2. CURRENT SITUATION OF PPV RESEARCH IN THE WORLD AND IN VIETNAM

1.3.1. The current situation of PPV research in the world

1.3.1.1. Circulation

PPV has been detected in pig herds in most countries regardless of the sex, age and health status of pigs. The prevalence of PPV may depend on the type of sample collected in each study (Streck et al., 2013).

1.3.1.2. Genetic variation in genes encoding structural proteins

The structural protein of PPV plays an important role in the induction of various disease properties. Comparing the genome sequences between the mild strain NADL-2 used as a vaccine and the virulent strain Kresse showed that they differ only at 8 nucleotide positions and 6/8 of these positions lead to changes in amino acids. New phenotypes with amino acid substitutions in structural proteins have altered the antigenic properties of the virus. The characterization of molecular characteristics and nucleotide polymorphisms, and phylogenetic analysis based on gene sequences encoding structural proteins are needed for new strategies to control diseases caused by PPV.

1.3.1.3. PPV evolution

In the 80s and 90s of the last century, it was assumed that PPV had a low level of genetic variation comparable to that of the host. However, recent studies have shown that the nucleotide variation rate of PPV is comparable to that of RNA viruses. Besides, the discovery of eight PPV genotypes (1965-2022) completely changed the view of the evolution and immunology of PPV, showing that the virus is much more diverse than previously predicted.

1.3.1.4. Detection of other PPV genotypes

Sequence comparison of genes encoding structural proteins of PPV showed that genetic variation in the VP molecule leads to the formation of different PPV subtypes. Eight PPV genotypes have been confirmed (1965-2022) including: PPV1, PPV2, PPV3, PPV4, PPV5, PPV6, PPV7 and PPV8 (Palinski et al., 2016; Ni et al., 2014; Xiao et al., 2013); Cheung et al., 2010; Huang et al., 2010; Lau et al., 2008).

1.3.2. Current situation of PPV research in Vietnam

In Vietnam, SMEDI Syndrome has been of concern since the 1990s, there is no specific treatment and vaccination can be ignored, so in reality, the disease still occurs continuously and is common in most localities. in the country. There are still no specific statistics on the harmful effects of SMEDI in Vietnam in general and the Central provinces in particular. Serological testing has been used to confirm the prevalence of PPV in pigs (Pham Hung, 1999; Ho Dinh Chuc, 1995). Recent studies have used PCR to detect the prevalence of PPVs in pig herds in some provinces in the country, mainly in the North and the South (Giap et al., 2020; Thuy et al. , 2021). Up to now, there have been no published scientific works on infection rate and molecular characteristics of genes encoding structural proteins of PPVs isolated in Central Vietnam.

Chapter 2. SUBJECTS, MATERIALS AND METHODS

2.1. Subjects

PPV strains (1-4) circulated in pigs of slaughter age in seven central provinces.

2.2. Materials

A total of 146 samples (117 lung samples and 29 blood samples) were collected from commercial pigs raised in seven central provinces at slaughterhouses. Each province collects an average of 5 slaughterhouses, each slaughterhouse collects no more than 5 samples; equipment, tools, chemicals, primers to detect PPV genotypes, primers to amplify genes encoding PPV genotypes and reference sequences.

2.3. Research Methods

2.3.1. Sampling method

A total of 146 samples including blood and lungs were collected at slaughterhouses (4 slaughterhouses/province, no more than 5 samples/slaughterhouse). After collection, samples were kept at 4°C and transferred to the laboratory for storage at -20°C prior to total DNA extraction.

2.2.2. Total DNA extraction

DNA extraction method: According to routine laboratory method, using Proteinase K; Phenol/Chloroform.

2.2.3. PCR

The PCR method uses total DNA as a template, and requires the optimization of PCR conditions to specifically amplify DNA fragments with the corresponding primer pairs for each DNA fragment.

2.2.4. Electrophoresis

Prepare the agarose gel, apply the electrophoresis sample, stain the gel with Ethidium Bromie, then observe and capture the electrophoresis image.

2.2.5. DNA sequencing

DNA sequences were determined using an automatic sequence analyzer ABI-3100 Avabp Genetic Analyzer of Macrogen company, Korea according to Sanger's principle.

2.2.6. DNA sequence processing and analysis

DNA sequences of isolated PPV strains were processed with BioEdit software v.7.0.9.0 (Hall, 1999); and compared using GenDoc 2.6 software and BLAST tool. Skew value was calculated according to the formula AT skew = (A-T)/(A+T) and GC skew = (G-C)/G+C) (Perna and Kocher, 1995). Nucleotide sequences of reference strains and study strains were aligned and arranged using the ClustalW Multiple alignment tool (Thompson et al., 1994).

2.2.7. Phylogenetic tree construction

Phylogenetic tree was built using MEGA X software (Kumar et al., 2018) with Maximum Likelihood inference method, Tajima-Nei model and bootstrap test repeated 1000 times.

2.2.8. Data analysis

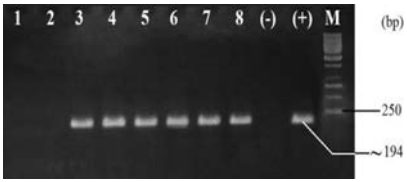
The prevalence rates of the studied genotypes and the composition rates of nucleotides were statistically processed using Excel v2010 software.

CHAPTER 3. STUDY RESULTS

3.1. DISCOVERY OF PPV (1-4) CIRCULATING ON PIGS IN SEVEN CENTRAL PROVINCES

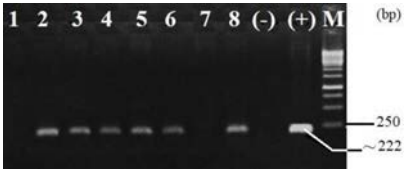
3.1.1. Discovery of PPV (1-4) circulating on pigs in 7 central provinces.

PPV detection PCR products (1-4) are presented in Figures: 3.1; 3.2; 3.3 and 3.4. Sequencing results of gene fragments detecting PPV (1-4) were compared by BLAST tool.



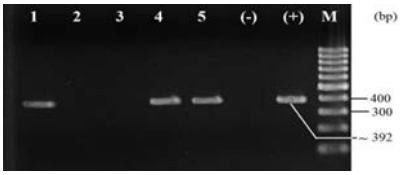
M: GeneRuler 1 kb DNA Ladder (Thermo Sciebpific)
1-2: Negative samples
3-8: Positive samples
(-): Negative control
(+): Positive control

Figure 3.1. Electrophoresis results of PCR product to detect PPV1



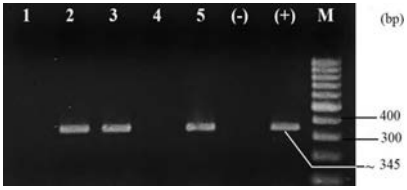
M: GeneRuler 1 kb DNA Ladder (Thermo Sciebpific)
1, 7: Negative samples
2-6, 8: Positive samples
(-):Negative control
(+):Positive control

Figure 3.2. Electrophoresis results of PCR product to detect PPV2



M: GeneRuler 100 bp DNA Ladder (Thermo Sciebpific)
2, 3: Negative samples
1, 4, 5: Positive samples
(-):Negative control
(+):Positive control

Figure 3.3. Electrophoresis results of PCR product to detect PPV3



M: GeneRuler 100 bp DNA Ladder (Thermo Sciebpific);
1, 4:Negative samples
2, 3, 5: Positive samples
(-):Negative control
(+):Positive control

Figure 3.4. Electrophoresis results of PCR product to detect PPV4 (345 bp)

3.1.2. PPV (1-4) infection rate results

The results of infection rates of PPV (1-4) in pigs raised in seven central provinces are presented in Table 3.1.

Table 3.1. PPV (1-4) infection rates in pigs in 7 central provinces

Provinces	n	Infection rate of PPV genotypes (1-4) (%)				Infection rate of at least one PPV (%)
		PPV1	PPV2	PPV3	PPV4	
Quảng Bình	17	0	17 (**100**)	0 (0)	0 (0)	17 (100,0)
Quảng Trị	22	18 (81,8)	10 (45,5)	3 (**13,6**)	6 (27,3)	19 (86,4)
Thừa Thiên Huế	28	24 (**85,7**)	20 (71,4)	3 (10,7)	6 (21,4)	25 (89,3)

Đà Nẵng	14	7 (50,0)	0 (0)	0 (0)	0 (0)	7 (50,0)
Quảng Nam	16	0	6 (37,5)	0 (0)	0 (0)	6 (37,5)
Quảng Ngãi	24	15 (62,0)	18 (75)	2 (8,3)	0 (0)	18 (75,0)
Bình Định	25	13 (52,0)	11 (44)	0 (0)	0 (0)	17 (68,0)
Tổng	146	77 (**52,7**)	82 (**56,2**)	8 (**5,5**)	12 (**8,2**)	109 (**74,7**)

Data from Table 3.1 show that PPV is present in pigs raised in the seven provinces in the study (7/7). PPV was detected in a total of 109/146 samples studied, accounting for 74.7%; specifically, the rate of PPV infection (1-4) being: 52.7%; 56.2%; 5.5% and 8.2% respectively.

3.1.2. PPV genotypes (1-4) co-infection rate results

Results of co-infection with PPV genotypes (1-4) in pigs in the Central provinces showed that PPV1/PPV2 co-infection accounted for the highest rate (34.2%) and only one sample was co-infection with 4 PPV genotypes. (1-4). In terms of locality, Thua Thien Hue is the province where all types of co-infection with 2/3/4 PPV strains were detected.

3.2. DNA SEQUENCES ENCODING STRUCTURAL PROTEIN OF PPV2 AND PPV4 STRAINS CIRCULATING ON PIGS IN CENTRAL PROVINCES

3.2.1. DNA Sequences encoding structural proteins of PPV2 strains

The VP DNA sequences of three PPV2 strains isolated in this study with molecular size 2,493 bp were submitted to the GenBank database with access codes OL913365-OL913367, respectively. Nucleotide composition of VP gene sequences of PPV2 strains isolated from central provinces is presented in Table 3.2.

Table 3.2. Nucleotide composition and skew value in VP DNA sequences of PPV2 strains isolated in Central Vietnam

Genotype	Strain	Sequence	Size (bp)	A (%)	T (%)	G (%)	C (%)	A+T (%)	AT-skew	G+C (%)	GC-skew
PPV2	QN03	*VP gene*	2.493	25,3	20,8	29,6	24,2	46,1	0,097	53,9	0,1
	HU10			25,3	20,8	29,6	24,3	46,1	0,098	53,9	0,1
	QB05			25,3	20,4	29,9	24,4	45,7	0,108	54,3	0,099

3.2.2. Genome sequences of PPV4 strains

3.2.2.1. Genome sequences of PPV4 strains isolated from Central Vietnam

The genome molecular size of two PPV4 strains isolated in this study is 5,367 bp, including 3 open reading frames ORF1, ORF2 and ORF3 and was submitted to GenBank database with access codes MT434668 and MT434669. The nucleotide composition of genome sequence of PPV4 strains isolated from central provinces is presented in Table 3.3.

Table 3.3. The nucleotide composition and skew value in genome sequence of PPV4 strains isolated in Central Vietnam.

Genotype	Strain	Sequence	Size (bp)	A (%)	T (%)	G (%)	C (%)	A+T (%)	AT-skew	G+C (%)	GC-skew
PPV4	QT02	Genome	5.387	33,0	25,2	22,5	19,3	58,2	0,135	41,8	0,078
	QT20			33,0	25,2	22,5	19,3	58,2	0,135	41,8	0,075

3.2.2.2. VP gene sequences of PPV4 strains isolated from Central Vietnam

The molecular size of the VP gene of the two PPV4 strains isolated in this study is 2,187 bp; located in the open reading frame ORF2. The nucleotide composition and the AT and GC skew values in the VP gene sequences of PPV4 strains isolated from the central provinces are presented in Table 3.4.

Table 3.4. Nucleotide composition và skew values in VP gene sequences of PPV4 strains isolated in Central Vietnam.

Genotype	Strain	Sequence	Size (Sbp)	A (%)	T (%)	G (%)	C (%)	A+T (%)	AT-skew	G+C (%)	GC-skew
PPV4	QT02	Gen *VP*	2187	34,3	24,7	20,3	20,7	59,0	0,163	41,0	0,162
	QT20			34,3	24,8	20,2	20,7	59,1	-0,001	40,9	-0,001

3.2.2.3. DNA sequences of PPV4 strains isolated from Central Vietnam

ORF3 open reading frame sequence length of the two PPV4 strains isolated in this study was 615 bp, without insertion/deletion mutations in the coding region. The nucleotide composition and AT and GC skew values in ORF3 sequences of PPV4 strains isolated from central provinces are

presented in Table 3.5.

Table 3.5. Nucleotide composition and skew values in ORF3 sequences of PPV4 strains isolated in Central Vietnam

Genotype	Strain	Sequence	Size (bp)	A (%)	T (%)	G (%)	C (%)	A+T (%)	AT-skew	G+C (%)	GC-skew
PPV4	QT02	ORF3	615	34,6	26,2	21,0	18,2	60,8	0,139	39,2	0,141
	QT20			34,8	26,2	20,8	18,2	61,0	0,071	39,0	0,067

3.2.3. Sequence of the NS1/VP1 gene region of PPV3 strains

The sequence of the NS1/VP1 gene region of the two PPV3 strains isolated in this study is 995 bp in size (Including: *NS1* gene segment of 229 bp and *VP1* gene segment of 766 bp, coding for structural proteins VP1). No nucleotide addition/loss mutations were detected in the coding gene region.

3.3. CHARACTERIZATION OF GENES ENCODING STRUCTURAL PROTEINS OF PPV ISOLATED IN CENTRAL PROVINCES

3.3.1. Chacterization of genes encoding structural proteins of PPV2 strains isolated in central provinces

3.3.1.1. Characterization of the VP gene nucleotide sequence of PPV2

The sites of nucleotide substitution mutations and amino acid inferred from the VP gene sequences of three PPV2 strains isolated in Central Vietnam are different from those of the reference strains shown in Table 3.6.

Table 3.6. The sites of nucleotide substitution mutations and amino acid inferred from the VP gene sequences of three PPV2 strains isolated in Central Vietnam compared to reference strains.

Strain	Sites of bp substitution	Sites of aa substitution
PPV2-QN03 (OL913365)	444: T→A	-
	734: C→T	245: S→F
	820: A→C	274: K→Q
	1794: T→A	-

Data from Table 3.6 shows nucleotide variation points on the VP gene sequence of 3 PPV2 strains isolated in the Central region (PPV2 QN03, PPV2-HU10 and PPV2- QB05), including 4 nucleotide substitution positions

only detected in strain PPV2-QN03 (OL913365) was completely different from the reference strains.

3.3.1.2. Characterization of amino acid sequences inferred from the VP gene of PPV2

The results of comparing the inferred amino acid sequences of the two PPV2 strains isolated in this study with the referenced PPV2 strains on GenBank are presented in Figure 3.5.

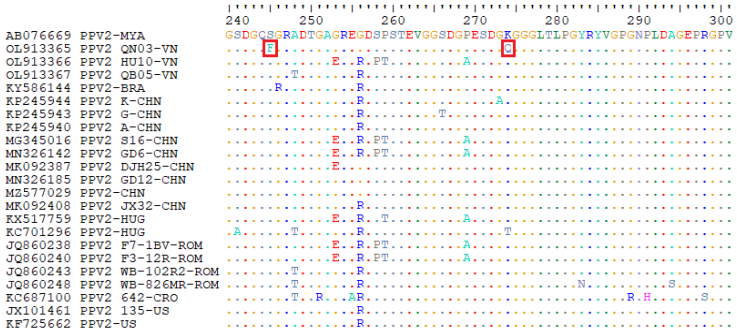


Figure 3.5. Comparison of amino acid sequences inferred from the VP gene of PPV2 strains isolated in Central Vietnam compared with reference strains

As shown in Figure 3.17, according to the inferred amino acid sequence comparison, two substitution sites were detected (245: S→F, 274: K→Q) in the inferred amino acid sequence of strain PPV2-QN03.

Results of comparing amino acid substitution mutations at important sites in the inferred amino acid sequence of PPV2 strains isolated in Vietnam with the results of previous studies (Sun et al., 2015; Cadar et al. cs, 2013) are presented in Table 3.7.

Table 3.7. Important sites of substitution in the amino acid sequence inferred from the *VP* gene of PPV2 strains isolated in Central Vietnam compared with reference strains

Country	Site / Strain	Year of isolation	A		B			C					
			245	274	269	442	714	349	437	598	689	784	796
Myamar	AB076669	2001	S	K	P	D	Q	S	D	D	T	M	Q
Vietnam	OL913365 QN03	2019	F	Q	P	D	S	S	D	E	T	M	Q

	OL913366 HU10		S	K	A	K	N	S	D	D	T	I	E
	OL913367 QB05		S	K	P	H	N	S	Y	E	S	M	Q
Brazil	KY586144	2017	S	K	P	D	Q	S	D	D	S	M	Q
China	KP245944	2014	S	K	P	D	G	R	E	E	S	I	E
	KP245943		S	K	P	D	G	R	E	E	S	I	E
	KP245940		S	K	P	D	G	R	E	E	S	I	E
	MK092387	2018	S	K	P	D	N	S	D	D	T	M	Q
	MK092408		S	K	P	D	N	S	D	D	T	M	Q
	MN326142	2019	S	K	A	D	S	S	D	D	T	M	Q
	MN326185		S	K	P	D	S	S	D	D	T	M	Q
	MG345016	2017	S	K	A	D	S	S	D	D	T	I	E
	MZ577029	2021	S	K	P	D	S	S	D	D	T	M	Q
Hungary	KX517759	2016	S	K	A	D	S	S	D	D	T	M	Q
	KC701296	2013	S	T	P	D	N	S	D	D	S	M	Q
Romania	JQ860238	2012	S	K	A	K	N	S	D	D	T	I	E
	JQ860240		S	K	A	D	N	S	D	D	T	I	E
	JQ860243		S	K	P	H	N	S	Y	E	S	M	Q
	JQ860248		S	K	P	D	N	S	Y	D	S	M	Q
Cromania	KC687100	2013	S	K	P	D	Q	S	D	D	T	M	Q
U.S.A	JX101461	2012	S	K	P	D	G	R	E	E	S	I	E
	KF725662	2013	S	K	P	N	S	S	D	D	S	M	Q

A: Mutation positions in Vietnamese strains different from reference strains; B: The mutation site under positive selection pressure (Cadar et al., 2013; C: Important mutation sites determining genetic subtypes (Sun et al., 2015).

The results of mutations detected at the sites of the specific motifs in the VP gene sequence of PPV2 according to Cadar's research on PPV2 strains isolated in central Vietnam are presented in Table 3.8.

Table 3.8. Results of mutations at the sites of the characteristic sequence motifs in the VP gene sequence of PPV2 (Cadar et al., 2013)

Characteristic sequence motifs	Site of codon khởi đầu motif	Vị trí codon xuất hiện đột biến	Chủng PPV2 xuất hiện đột biến
Polyadenylation signal (PAS) (AATAAA)	66	66 (AA**G**AAA)	PPV2-HU10
	454	454 (A**C**TAAA)	PPV2-QB05
SP1 binding site (GGGCGG)	36, 165, 329	329 (GGGC**A**G)	PPV2-QN03

3.3.1.3. Level of nucleotide/amino acid similarity inferred from the VP gene of PPV2

The level of similarity in nucleotide/amino acid sequences inferred from VP gene sequences of PPV2 strains isolated in Central Vietnam (2,493 bp and 831 aa) compared with strains isolated in the region and the world is presented in Table 3.9

Table 3.9. The ratio of nucleotide and amino acid homology inferred (%) from VP gene sequences of PPV2 strains isolated in Central Vietnam compared with reference strains

Strain \ Size		Nucleotide (2.493 bp)	Amino acid (831 aa)
Vietnam		95,8-97,0	95,3-97,5
Myamar		96,1-98,1	95,4-98,7
China	Group A	95,2-99,7	94,5-99,6
	Group B	93,7-94,6	93,2-95,5
Europe		94,9-100	93,3-100
U.S.A		93,7-96,7	93,8-97,4

3.3.1.4. PPV2 phylogenetic relationships based on VP gene sequences

The phylogenetic tree of the PPV2 strains isolated in this study and the reference strains was built based on the VP gene sequence of PPV2 with a molecular size of 2,493 bp (Figure 3.6).

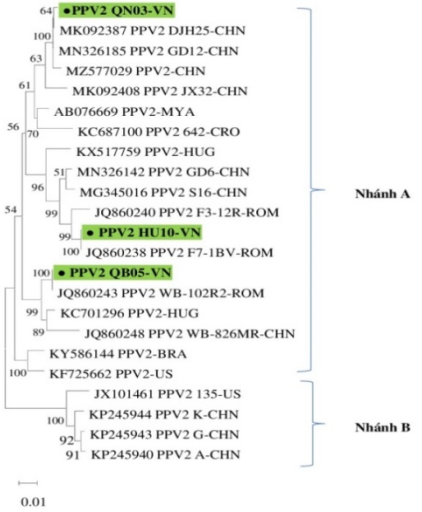


Figure 3.6. Phylogenetic tree PPV2 based on VP gene nucleotide sequences (2,493 bp) obtained from some central provinces of Vietnam and reference gene sequences.

3.3.2. Characterization of genes encoding structural proteins of PPV4 strains isolated in central provinces

3.3.2.1. Characterization of the PPV4 genome

The sites of nucleotide substitution mutations leading to substitutions in amino acid sequences inferred from ORF1 sequences of PPV4 strains isolated in the Central region are shown in Figure 3.7.

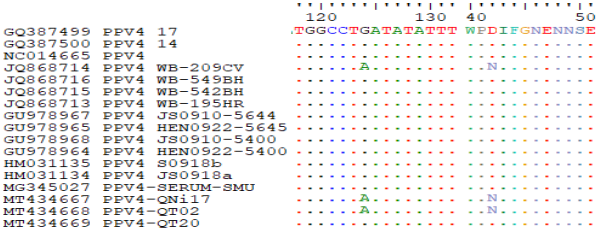
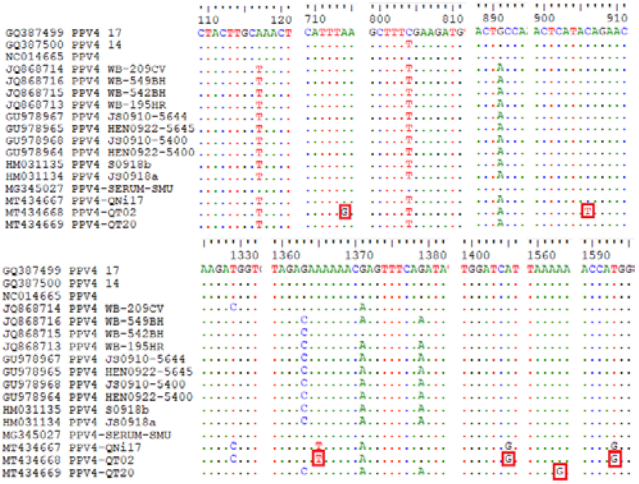


Figure 3.7. Nucleotide and amino acid substitution mutation positions in the ORF1 open reading frame of the PPV4 genome sequence.

3.3.2.2. Characterization of the VP gene of PPV4

Characterization of the VP gene nucleotide sequence of PPV4

The results of comparing VP gene nucleotide sequences of PPV4 strains isolated in Central Vietnam and reference strains are presented in Figure 3.8.



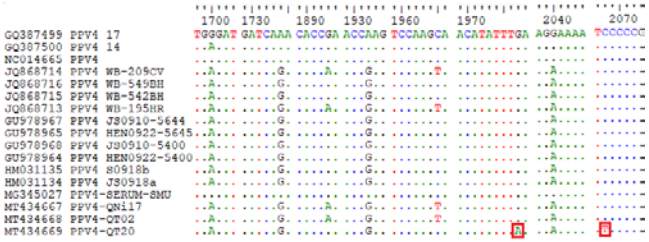
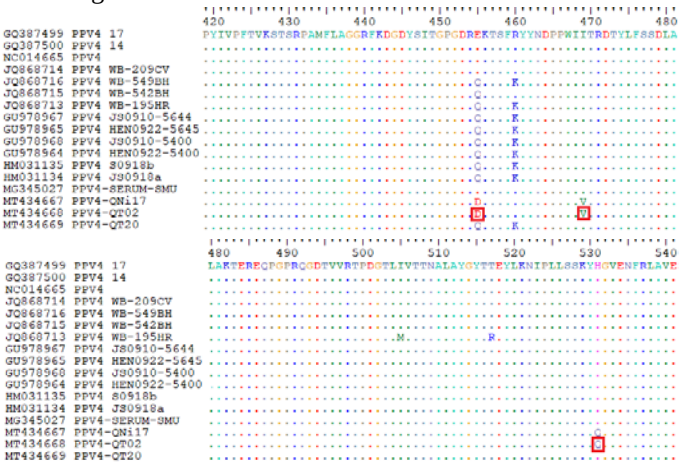


Figure 3.8. The results of comparing VP gene nucleotide sequences of PPV4 strains isolated in Central Vietnam and reference strains

Table 3.10. Sites of substitution in the amino acid sequence inferred from the *VP* gene of PPV2 strains isolated in Quang Tri compared with reference strains

Strain	Sites of bp subtitution	Sites of aa subtitution
PPV4-QT02	714: A→G	-
	906: C→T	-
	1365: A→T	455: E→D
	1405: C→G	469: I→V
	1593: A→G	531: H→Q
PPV4-QT20	1563: A→G	-
	1977: G→A	-
	2068: C→T	690: P→S

The results of comparing inferred amino acid sequences of two PPV4 strains isolated in this study with referenced PPV4 strains on GenBank are presented in Figure 3.9.



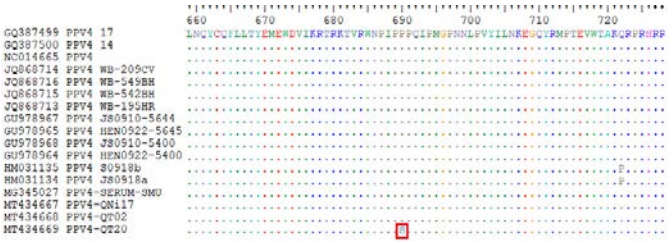


Figure 3.9. Results of comparison of amino acid sequences inferred from the VP gene of PPV4 strains isolated in central Vietnam compared with reference strains.

Results of comparing amino acid substitutions at potential sites in amino acid sequences inferred from the VP gene of PPV4 strains isolated in Quang Tri province with the results of previous studies (Sun et al., 2015; Cadar, 2015 et al., 2013) are presented in Table 3.11.

Table 3.11. Important sites of substitutions in amino acid sequences inferred from the VP gene of PPV4 strains isolated compared with reference strains

Country	Site / Strain	Year of isolation	A		B								
			531	690	178	416	455	460	469	567	583	722
U.S.A	GQ387499	2010	H	P	S	E	E	R	I	G	T	Q
	GQ387500	2010	H	P	S	K	E	R	I	E	A	Q
	NC014665	2018	H	P	S	E	E	R	I	G	T	Q
Vietnam	MT434667 QNi17	2020	Q	P	S	K	D	R	V	E	T	Q
	MT434668 QT02		Q	P	S	K	D	R	V	E	T	Q
	MT434669 QT20		H	S	S	K	Q	K	I	E	T	Q
Romania	JQ868713	2012	H	P	S	K	E	R	I	E	T	Q
	JQ868714		H	P	S	K	Q	K	I	E	T	Q
	JQ868715		H	P	S	K	Q	R	I	E	T	Q
	JQ868716		H	P	S	K	Q	K	I	E	T	Q
China	GU978964	2010	H	P	P	K	Q	K	I	E	T	Q
	GU978965		H	P	S	K	Q	K	I	E	T	Q
	GU978967		H	P	P	K	Q	K	I	E	T	Q
	GU978968		H	P	S	K	Q	K	I	E	T	Q
	HM031134		H	P	S	K	Q	K	I	E	T	P
	HM031135		H	P	S	K	Q	K	I	E	T	P
	MG345027	2018	H	P	S	E	E	R	I	G	T	Q

A: Mutation sites in Vietnamese strains are different from reference strains; B: Important mutation sites (Sun et al., 2015; Cadar et al., 2013).

3.3.2.3. Characterization of ORF3 of PPV4

The results of comparing the nucleotide and amino acid sequences inferred from the ORF3 open reading frame of the PPV4 strains isolated in Quang Tri province with the reference strains showed that only a single substitution mutation was detected in the nucleotide sequence in Quang Tri. strain PPV4-QT02 (240: A→G).*3.3.2.4. Level of nucleotide/amino acid similarity of PPV4 isolated in central Vietnam*

The results of comparing the similarity of nucleotide and inferred amino acid sequences of ORF1, ORF2, ORF3 and genome sequences between PPV4 strains isolated in Central Vietnam compared with strains in the region and the world is presented in Table 3.12.

Table 3.12. The percentage of nucleotide and amino acid similarity (%) of ORF1, ORF2, ORF3 and genomic sequences between PPV4 strains isolated in Central Vietnam compared with reference strains

Sequence	Molecular size	Sources of PPV4 strains				
		Cetral Vietnam	Northern Vietnam	China	U.S.A	Romania
ORF1	1797 bp	99,6	99,7-99,8	99,1-99,6	99,1-99,2	99,1-99,8
	598 aa	99,4	99,6-99,8	98,6-99,6	98,4-98,9	99,1-100
ORF2	2.187 bp	99,1	99,3-99,7	98,9-99,6	98,9-99,3	99,2-99,7
	728 aa	99,3	99,3-100	99,3-99,8	99,3-99,4	99,1-99,8
ORF3	615 bp	99,8	99,6-99,8	99,1-100	99,6-99,8	99,8-100
	204 aa	100	100	99,0-100	100	100
Genome	5.367 bp	99,3	99,4-99,6	98,7-99,4	98,8-99,0	99,0-99,5
	1500 aa	98,7	98,8-99,3	98,0-98,9	98,0	98,2-99,0

3.3.2.5. *Phylogenetic relationship of PPV4 strains*

The phylogenetic tree built based on genomic sequences (5,367 bp) and genes encoding structural protein (2,187 bp) are presented in Figure 3.24 and Figure 3.25.

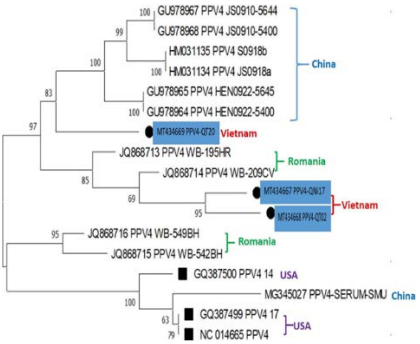


Figure 3.10. Phylogenetic tree PPV4 was built based on the almost complete genome sequence of PPV4 (5,367 bp) obtained from Vietnam and reference sequences.

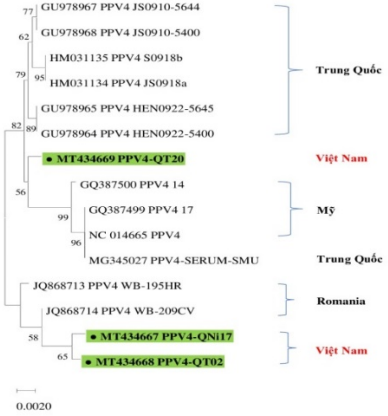


Figure 3.11. The PPV4 phylogenetic tree was built based on the complete capsid protein-coding gene sequences of PPV4 strains (2,187 bp) isolated in Central Vietnam and reference sequences.

3.3.3. Characterization of genes *NS1/VP1* gene rigion of PPV3 strains isolated in central provinces

The sequence of the NS1/VP1 gene region (995 bp) of the two strains PPV3-QN16 and PPV3-QT4 has seven nucleotide positions that are completely different from the reference strains (39: A→T; 131: G→C; 137 : T→G and 542: T→C; 930: T→G; 932: G→A and 949: G→A), leading

to deduced amino acid substitutions at positions 234 (L→W) and 235 (D→N). The phylogenetic tree built based on the sequence of the gene encoding the VP1 structural protein of PPV3 (305 bp) shows that PPV3 strains isolated from Quang Ngai and Quang Tri have a close relationship with PPV3 strains from China and USA.

CONCLUSION AND RECOMMENDATIONS

CONCLUSION

1. The circulation of PPV (1-4) was detected in pig herds in seven central provinces of Vietnam using PCR to amplify the specific nucleotide fragment of each genotype. The infection rates of PPVs (1-4) in 7 central provinces were: 52.7% for PPV1 ; 56.2% for PPV2; 5.5% for PPV3 and 8.2% for PPV4. Co-infection with PPV genotypes (1-4) was also detected in 4/7 studied provinces (Quang Tri, Thua Thien Hue, Quang Ngai and Binh Dinh); in which co-infection with 2 genotypes PPV1/PPV2 in pigs raised in Thua Thien Hue accounted for the highest rate with 67.9%.

2. We have sequenced the genes encoding structural proteins of 3 strains of PPV2 isolated from pigs raised in Quang Ngai, Thua Thien Hue and Quang Binh provinces with the size of 2,493 bp (GenBank code: OL913365, OL913366 and OL913367); The whole genome sequence of two PPV4 strains isolated in Quang Tri has the size of 5,367 bp (GenBank codes: MT434668 and MT434669), including 3 open reading frames: ORF1, ORF2 and ORF3 with corresponding sizes of 1,797 bp, 2,187 bp and 615 bp.

3. The results of analyzing the gene sequences encoding structural proteins of three strains of PPV2 have shown the detection of four nucleotide substitution mutations (444: T→A, 734: C→T, 820: A→C and 1794: T→A),

in which there are two mutations leading to inferred amino acid substitution at two positions: 245 (S→F) and 274 (K→Q) which are completely different from the reference strains. The results of genomic sequence analysis of two PPV4 strains revealed a nucleotide substitution mutation (124: G → A) leading to amino acid substitution (42: D → N) in the ORF1 open reading frame and eight sites of substitutions in the nucleotide sequence of genes encoding structural proteins (ORF2) (714: A→G, 906: C→T, 1365: A→T, 1405: C→G, 1593: A→G, 1563: A →G, 1977: G→A, and 2068: C→T) leading to four substitutions in the inferred amino acid sequence at positions 455 (E→D), 469 (I→V), 531 (H→ Q) and 690 (P→S), which were completely different from the reference strains. The inferred nucleotide/amino acid substitution mutations mainly concentrated in the structural protein (ORF2) gene sequences of PPV4 strains isolated.

4. The hylogenetic tree was built based on genes encoding structural protein of PPV2 (2,493 bp) and genome of PPV4 (5,387 bp). Three PPV2 strains isolated in Quang Binh, Thua Thien Hue and Quang Ngai are closely related to PPV2 strains originating from Europe and China. PPV4 strains isolated in Quang Tri are closely related to Romanian and Chinese PPV4 strains.

5. Diseases that have occurred in pigs recently are quite complicated, including a complex of reproductive and respiratory diseases, often with co-infection of many viruses: PCV, PRRSV, PPV... Therefore, updated information on genetic changes, especially in gene regions related to antigenic properties, immune response... of the virus, is very necessary,

especially for the purpose of developing methods detect pathogens that cause co-infection in pigs.

RECOMMENDATIONS

Although PPV circulates in pigs of all ages, this study was only performed on samples collected from commercially raised pigs, of slaughter age, collected at slaughterhouses. In addition, the project was carried out from September 2019 to August 2023, entirely during the COVID-19 pandemic, in which the sample collection period was affected, so there were certain limitations, especially It is not possible to collect samples with sows from breeding farms. Therefore, the author of the thesis would like to propose the following recommendations:

1. More PPV isolates in Vietnam need to be screened, to find out whether or not there is circulation of other PPV genotypes (PPV5-8), in order to provide data on PPV in pigs raised in Vietnam.

2. Conduct testing for the presence of PPV in pigs (sows, post-weaning, growing) and their sperm in order to control the sources of virus transmission.

3. Decode the entire genome/structural protein-encoding genes of circulating pig PPVs to compare with existing genotypes in different geographical regions to evaluate the molecular epidemiological characteristics of the pigs.

AUTHOR'S PUBLICATIONS RELATED TO THE THESIS

1. Nguyen Thi Dieu Thuy, **Nguyen Tran Trung**, Tran Quoc Dung, Do Vo Anh Khoa, Dinh Thi Ngoc Thuy, Tanja Opriessnig (2021). First investigation of the prevalence of parvoviruses in slaughterhouse pigs and genomic characterization of ungulate copiparvovirus 2 in Vietnam. *Archives of Virology*, 166: 779-788. (SCIE/Q2)

2. **Nguyen Tran Trung**, Tran Quoc Dung, Nguyen Thi Dieu Thuy (2023). Prevalence and structural protein encoding gene sequence (VP) of porcine parvovirus 2 (PPV2) in slaughtered pigs in cebpral provinces of Vietnam. *Hue University Journal of Science: Natural Science*, 132 (1D): 5-14.

3. **Nguyen Tran Trung**, Tran Quoc Dung, Dinh Thi Ngoc Thuy, and Nguyen Thi Dieu Thuy (2023). The nucleotide sequence of NS1/VP1 gene of Porcine Parvovirus 3 (PPV3) in finishing pigs raised in Quang Ngai and Quang Tri provinces. *Journal of Animal Husbandry Sciences and Technics*, 293: 59-65.