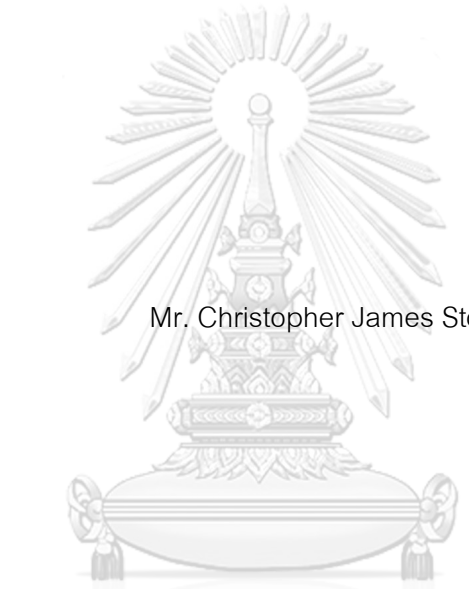
EVOLUTION AND PHYLOGENETIC STUDIES OF PORCINE EPIDEMIC DIARRHEA VIRUS AND PORCINE DELTACORONAVIRUS IN THAILAND FROM 2008 TO 2021

Mr. Christopher James Stott



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Veterinary Pathobiology
Department of Veterinary Pathology
FACULTY OF VETERINARY SCIENCE
Chulalongkorn University
Academic Year 2022

Copyright of Chulalongkorn University

วิวัฒนาการและการศึกษาทางวิวัฒนาการของเชื้อไวรัสพีอีดีและพอร์ซัยน์เดลตาโคโรนาไวรัสในประเทศไทยระหว่างปี พ.ศ. 2551 ถึง พ.ศ. 2564

นายคริสโตเฟอร์ เจมส์ สต๊อต

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาพยาธิชีววิทยาทางสัตวแพทย์ ภาควิชาพยาธิวิทยา
คณะสัตวแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title	EVOLUTION AND PHYLOGENETIC STUDIES OF PORCINE EPIDEMIC DIARRHEA VIRUS AND PORCINE DELTACORONAVIRUS IN THAILAND FROM 2008 TO 2021
By	Mr. Christopher James Stott
Field of Study	Veterinary Pathobiology
Thesis Advisor	Assistant Professor DACHRIT NILUBOL, Ph.D.

Accepted by the FACULTY OF VETERINARY SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

..... Dean of the FACULTY OF VETERINARY SCIENCE

(Professor SANIPA SURADHAT, Ph.D.)

DISSERTATION COMMITTEE

..... Chairman

(Associate Professor MORAKOT KAEWTHAMASORN, Ph.D.)

..... Thesis Advisor

(Assistant Professor DACHRIT NILUBOL, Ph.D.)

..... Examiner

(NAVAPON TECHAKRIENGKRAI, Ph.D.)

..... Examiner

(Associate Professor ANGKANA TANTITUVANONT, Ph.D.)

..... External Examiner

(Associate Professor KAMPON KAEOKET, Ph.D.)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

คริสโตเฟอร์ เจมส์ สต๊อต : วิวัฒนาการและการศึกษาทางวิวัฒนาการของเชื้อไวรัสพีอีดีและพอร์ซัยน์เดลตาโคโรนาไวรัสในประเทศไทยระหว่างปี พ.ศ. 2551 ถึง พ.ศ. 2564. (EVOLUTION AND PHYLOGENETIC STUDIES OF PORCINE EPIDEMIC DIARRHEA VIRUS AND PORCINE DELTACORONAVIRUS IN THAILAND FROM 2008 TO 2021) อ.ที่ปรึกษาหลัก : ผศ. ดร.เดชฤทธิ์ นิลอุบล

ไวรัสพีอีดี (porcine epidemic diarrhea virus) และพอร์ซัยน์เดลตาโคโรนาไวรัส (porcine deltacoronavirus) ได้ถูกจัดเป็นเชื้อก่อโรคซึ่งทำให้เกิดโรคทางลำไส้ที่สำคัญในอุตสาหกรรมการผลิตสุกรทั่วโลก ระหว่างปี พ.ศ. 2557 ถึง 2558 ได้มีการระบาดของเชื้อไวรัสพีอีดีในหลายจังหวัดในประเทศไทยและอาการวิทยามีแนวโน้มรุนแรงมากกว่าพีอีดีสายพันธุ์ดั้งเดิมใน พ.ศ. 2551 ถึง 2555 จึงเป็นที่น่าสงสัยว่าจะมีการรับเข้ามาซึ่งเชื้อสายพันธุ์ใหม่หรือการเกิดการกลายพันธุ์ของไวรัส ในปี พ.ศ. 2558 ได้มีการรายงานการพบไวรัสพอร์ซัยน์เดลตาโคโรนาไวรัสในประเทศไทยทำให้สถานการณ์ของโรคกลุ่มดังกล่าวในประเทศไทยแย่ลง จากการที่ไม่เคยมีการศึกษาเพื่อจำแนกสายพันธุ์ของไวรัสพีอีดีและพอร์ซัยน์เดลตาโคโรนาไวรัสในประเทศไทยมาก่อนการศึกษานี้จึงมีวัตถุประสงค์ที่จะศึกษาความหลากหลายทางพันธุกรรมของเชื้อไวรัสพีอีดีและพอร์ซัยน์เดลตาโคโรนาไวรัสในประเทศไทย จากการใช้ยีนสไปค์ (spike gene) ไวรัสพีอีดีในประเทศไทยสามารถจำแนกได้ออกเป็น 7 กลุ่มย่อย (subgroup) กลุ่มย่อยที่สอง (TH2) หนึ่ง (TH1) และสี่ (TH4) เป็นสายพันธุ์เด่นซึ่งก่อให้เกิดการระบาดในระหว่างปี พ.ศ. 2551 ถึง 2556 ปี พ.ศ. 2557 ถึง 2559 และ ปี พ.ศ. 2559 ถึง 2561 ตามลำดับ อัตราวิวัฒนาการของกลุ่มย่อยที่หนึ่งระหว่างปี พ.ศ. 2551 ถึง 2558 มีแนวโน้มที่สูงกว่ากลุ่มย่อยอื่นใน หลังจากปี พ.ศ. 2561 ไวรัสพีอีดีมีอุบัติการลดลงและการระบาดประปราย (sporadic case) ของลูกผสม (recombinant) สามารถพบได้เขตที่มีสุกรชุก สำหรับไวรัสพอร์ซัยน์เดลตาโคโรนาวัสซึ่งแหล่งที่มายังไม่ชัดเจนถึงแม้ว่าจากการศึกษาจะแสดงให้เห็นว่าไวรัสพอร์ซัยน์เดลตาโคโรนาไวรัสสายพันธุ์ไทยและเอเชียตะวันออกเฉยงใต้ (SEA) จะมาความใกล้เคียงกันมากกว่ากับสายพันธ์จีน จากการใช้ลำดับพันธุกรรมตลอดความยาวไวรัสพอร์ซัยน์เดลตาโคโรนาไวรัสสามารถจำแนกออกได้เป็น 3 กลุ่ม ได้แก่เอเชียตะวันออกเฉียงใต้ จีน (China) และอเมริกา (US) สายพันธ์เอเชียตะวันออกเฉียงใต้สามารถจำแนกออกเป็น 3 กลุ่มพันธุกรรมย่อย โดยกลุ่มที่หนึ่ง (SEA-1) เป็นสายพันธุ์ระบาดใหญ่ (pandemic variant) ในเอเชียตะวันออกเฉียงใต้ ซึ่งเป็นต้นตอของการระบาดในประเทศไทย ลาว และเวียดนามระหว่างปี พ.ศ. 2556 ถึง 2559 ผลการศึกษายังพบว่าการรวมกัน (recimbination) เป็นกลไกสำคัญในการวิวัฒนาการของไวรัสในกลุ่มเอเชียตะวันออกเฉียงใต้ การศึกษานี้ทำให้ได้รับข้อมูลเชิงลึกของสายพันธุ์ที่มีการระบาดของไวรัสพีอีดีและพอร์ซัยน์เดลตาโคโรนาไวรัสซึ่งสามารถทำไปใช้ในการพัฒนากลยุทธ์ในการป้องกันโรคที่มีประสิทธิภาพและข้อมูลทางพันธุกรรมของสายพันธุ์ที่พบในฟาร์มสามารถที่จะทำไปประยุกต์ใช้ในการผลิตวัคซีนได้

สาขาวิชา	พยาธิชีววิทยาทางสัตวแพทย์	ลายมือชื่อนิสิต .....
ปีการศึกษา	2565	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

5875502231 : MAJOR VETERINARY PATHOBIOLOGY

KEYWORD: Porcine epidemic diarrhea virus (PEDV), Porcine deltacoronavirus (PDCoV), Porcine enteric coronavirus (PEC), phylogenetic, phylogeography, phylodynamic, evolution, recombination, molecular epidemiology, Thailand

Christopher James Stott : EVOLUTION AND PHYLOGENETIC STUDIES OF PORCINE EPIDEMIC DIARRHEA VIRUS AND PORCINE DELTACORONAVIRUS IN THAILAND FROM 2008 TO 2021. Advisor: Asst. Prof. DACHRIT NILUBOL, Ph.D.

PEDV and PDCoV are recognized as important pathogens causing enteric disease in swine herds worldwide. During 2014-2015, several outbreaks of PEDV occurred in many provinces in Thailand, and the clinical signs were more severe than those of classical PEDV in Thailand during 2008-2012. Therefore, the new introduction of PEDV or mutated virus strains was suspected. In 2015, PDCoV was also reported in Thailand, making the situation of enteric disease in the Thai swine industry worse. Since genetic classification studies of PEDV and PDCoV had never been done in Thailand before, this study aimed to investigate the genetic diversity of PEDV and PDCoV in the swine population in Thailand. PEDV in Thailand can be classified into seven subgroups based on the spike gene. Subgroups TH2, TH1, and TH4 were the predominant strains during 2008-2013, 2014-2016, and 2016-2018, respectively. The substitution rate of TH1 was higher during 2008-2015 compared to other strains in Thailand. After 2018, the incidence of PEDV decreased, and sporadic cases of new recombinants could be observed. For PDCoV, the origin of the virus is still unclear, although the study demonstrated that Thai and SEA PDCoV are closer to those of the Chinese PDCoV. Using full-length sequences, PDCoV could be classified into three different genogroups, including SEA, China, and the US, while SEA PDCoV could be further classified into three different subgroups. Subgroup SEA-1, or the SEA pandemic variants, responded to the outbreak in Thailand, Lao PDR, and Vietnam, during 2013-2016. The results also suggested that recombination was an important mechanism in the PDCoV evolution of the SEA genogroups. This study provided insight into the circulating strains of PEDV and PDCoV that can be used to develop effective preventive strategies. The information on the evolutionary dynamics of the virus and genetic information of the present variants in the field can also be applied for vaccine development.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Field of Study:	Veterinary Pathobiology	Student's Signature .....
Academic Year:	2022	Advisor's Signature .....

ACKNOWLEDGEMENTS

This research was supported by the Thailand Research Fund (Grant Number MRG5080323), the National Research Council of Thailand (Grant Number 5080001), and the Agricultural Research Development Agency (Public organization; Grant number PRP5805021650). Partial funding was provided by the Special Task Force for Activating Research (STAR), Swine Viral Evolution, and Vaccine Research (SVEVR) at Chulalongkorn University (Grant number GSTAR 59-013-31-007). The first author is grateful for the Royal Golden Jubilee PhD Program (PHD/0204/2557) Scholarship of the Thailand Research Fund.

I would like to express my sincere gratitude to my principal advisor, Assistant Professor Dr. Dachrit Nilubol, for his guidance and advice throughout all stages of my study and thesis. I am also thankful to my committee members, Associate Professor Dr. Morakot Kaewthamasorn, Associate Professor Dr. Angkana Tantituvanont, and Dr. Navapon Techakriengkrai, and Associate Professor Dr. Kampon Kaeoket, the external examiner, for their valuable suggestions and for giving me the opportunity to defend my thesis.

I would like to acknowledge Dr. Gun Temeeyasen, Ms. Thitima Tripipat, and Dr. Dam Thi Vui for providing valuable information on PEDV, and Dr. Kepalee Saeng-chuto for sharing genetic sequences of PDCoV. I am grateful for the contributions of Ms. Kanokon Sawattrakool to the project, and to the members of the DN1214 laboratory and SVEVR research unit, including Dr. Adthakorn Madapong, Dr. Puwich Chaikhumwang, Dr. Anchalee Srijangwad, Dr. Sunit Meebumrung, Mr. Parin Watcharavongthip, and Ms. Pathumporn Jermsutjarit, as well as my seniors, Dr. Kittitat Lugsomya and Dr. Pattanapon Kayansamruaj, for their helpful advice and for the enjoyable moments we shared.

Lastly, I would like to express my heartfelt gratitude to my dear family members, Mrs. Pranee Nuchlamyong, my mother, Mrs. Pornthip Warren, my aunt, and Ms. Thitichaya Thanadechakul, my sister, for their unwavering support.

Christopher James Stott

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

TABLE OF CONTENTS

Page

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

LIST OF TABLES

Page

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

LIST OF FIGURES

Page

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER 1

IMPORTANCE AND RATIONALE

Porcine enteric coronaviruses (PECs) are a group of viruses causing gastroenteritis in pigs that belong to the family *Coronaviridae* in order *Nidovirales*. To date, there are four PECs discovered in pigs including transmissible gastroenteritis virus (TGEV), porcine epidemic diarrhea virus (PEDV), swine acute diarrhea syndrome coronavirus (SADS-CoV), and porcine deltacoronavirus (PDCoV).

In Thailand, there are three PECs have been recognized in swine farms including TGEV (the 1980s), PEDV (suspected cases in 1995 and later confirmed in 2007), and PDCoV (2015) (Srinuntapunt et al., 1995; Temeeyasen et al., 2014; Janetanakit et al., 2016). Since TGEV became the endemic disease in Thailand and there is evidence of porcine respiratory coronavirus (PRCV), the mutant lineage of TGEV, affecting swineherd, this phenomenon leads to the immunization against both TGEV and PRCV (Turlewicz-Podbielska and Pomorska-Mól, 2021), and only sporadic outbreak might have occurred in some unexposed area. Therefore, at the present, PEDV and PDCoV are the two main enteric coronaviruses that cause acute diarrhea and gastroenteritis in pigs.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Porcine epidemic diarrhea (PED) is a devastating enteric disease in pigs characterized by vomit and acute watery diarrhea leading to death due to severe dehydration (Chasey and Cartwright, 1978; Pensaert and De Bouck, 1978; Pospischil et al., 2002). Although pigs of all ages are susceptible to PED with high morbidity following infection, the disease causes high mortality in pigs less than a week of age (Pospischil et al., 2002) whereas the mortality in older pigs is lower compared to young pigs. Since its first recognition in the late 1970s, PED has continued to cause a severe economic

impact on the swine industry worldwide. At present, PED has become an endemic disease in several countries including Thailand (Temeeyasen et al., 2014).

At present, two genogroups of PEDV, genogroups 1 (G1) and 2 (G2), have been recognized (Sun et al., 2015). CV777, the primitive strain of PEDV was clustered in G1. The G2 variant, a higher virulent PEDV, was first reported in 2011 in China (Li et al., 2012a), although the presentation of similar isolates has been detected in Korea since 2007 (Park et al., 2007). At present, these G2 variants have been endemic, circulating globally and causing damage in many pig-producing countries worldwide. The differences between these two variants are based primarily on the spike gene region in which the G2 variant possesses two insertions of four (^{56}GENQ^{59}) and one (^{140}N) amino acid(s) at positions 55-60 and 140, respectively, and deletion of two amino acids (^{160}DG^{161}) at positions 160-161 (Li et al., 2012b).

At present, both genogroups of PEDV have been reported as emerging diseases worldwide (Chen et al., 2014b; Lee and Lee, 2014b; Lin et al., 2014; Hanke et al., 2015; Masuda et al., 2015; Ojkic et al., 2015; Theuns et al., 2015; Vui et al., 2015). Although many authors are believed that PEDV should be derived from somewhere close to the eastern China coastline and were distributed to other countries via bats, fomites, or animal food somehow, however, the origin of the virus is still doubtful. In Thailand, G2 is the most recognized as the source of the outbreak in several areas at the present although G1 was discovered in 2014.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

PDCoV is a recently emerged enteric coronavirus in pigs (Woo et al., 2012), characterized by watery diarrhea and villous atrophy (Chen et al., 2015) which currently is a threat to the swine industry worldwide, and has been recognized as the zoonotic pathogen (Lednicky et al., 2021; Kong et al., 2022), especially in human. Although there was no association with the clinical disease at the first time of detection (Woo et al., 2012), in 2014, there are pieces of evidence of PDCoV causing the disease in the USA (Marthaler et al., 2014). Soon after its emergence in the USA, PDCoV associated disease was subsequently reported in several countries, including China, South Korea, Thailand, Lao PDR in 2015 (Woo et al., 2012; Lee and Lee, 2014b; Marthaler et al., 2014; Song et al., 2015; Janetanakit et al., 2016; Lorsirigool et al., 2016; Saeng-Chuto et al., 2017), Vietnam (Saeng-Chuto et al., 2016), and Mexico (Pérez-Rivera et al., 2019).

PEDV and PDCoV are the enveloped, single-stranded, positive-sense RNA viruses in the genus *Alphacoronavirus* and *Deltacoronavirus*, respectively, in the family *Coronaviridae*. Its genome is approximately 28 kb for PEDV and 25 kb for PDCoV. The genome arrangement of PEDV is in the order of 5' untranslated region (UTR), open reading frame 1a/1b (ORF 1a/1b), spike (S), open reading frame 3 (ORF3) envelope (E), membrane (M), nucleocapsid (N) and 3' UTR, while the order of PDCoV genome is consist of 5' untranslated region (UTR), open reading frame 1a/1b (ORF 1a/1b), spike (S), envelope (E), membrane (M), nonstructural protein 6 (NS6), nucleocapsid (N), nonstructural protein 7 (NS7) and 3' UTR (Kocherhans et al., 2001; Woo et al., 2012).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

The phylodynamic framework is including the study of immunodynamics, epidemiology, and evolutionary biology of pathogens (Grenfell et al., 2004), especially viruses. Since computer performance has been increased and user-friendly software such as BEAST packages has been developed, the phylodynamic analysis based on the Bayesian framework become popular at the present for its reliability and flexibility (Drummond and Rambaut, 2007). The study aimed to investigate the evolutionary and epidemiology aspects of the phylodynamic of PEDV and PDCoV, using spike gene and full-length genome, respectively, and the underlying reason for the recurrent outbreak and the persistence of the virus in Thai swine herds should be conducted. The results would provide information on the evolutionary dynamics of the viruses in Thailand.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

]

CHAPTER 2

RESEARCH PROBLEM, OBJECTIVES OF STUDY, KEYWORDS, AND HYPOTHESIS

Research problem

1. How does the genetic diversity of PEDV and PDCoV in the swine population in Thailand?
2. How does introducing exotic virus variants affect the prevalence and distribution of PEDV and PDCoV in the piglet production process in Thai swine farms?

Objectives of Study

1. To study the genetic diversity of PEDV and PDCoV in the swine population in Thailand.
2. To assess the impact of external introduction on the prevalence and distribution of PEDV and PDCoV in the piglet production process in Thai swine farms.

จุฬาลงกรณ์มหาวิทยาลัย

Keywords (Thai):

โคโรนาไวรัส วิวัฒนาการ พอร์ซัยน์เดลตาโคโรนาไวรัส พีอีดีไวรัส ประเทศไทย

Keywords (English):

Coronavirus, evolution, Porcine deltacoronavirus, Porcine epidemic diarrhea virus, Thailand

Hypothesis

1. The genetic diversity of PEDV and PDCoV in the swine population in Thailand is influenced by the continuous evolution of the viruses through mutation and recombination, resulting in a greater variation in the viruses.
2. The external introduction of exotic virus variants significantly contributes to the increasing prevalence and distribution of PEDV and PDCoV in piglet production on Thai swine farms. The external introduction of the virus may also participate in the evolution mechanism of the virus, leading to the emergence of new variants and strains.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CHAPTER 3

LITERATURE REVIEW

PED is caused by the PED virus (PEDV), an enveloped, positive-sense, single-stranded RNA virus belonging to the genus *Alphacoronavirus*, family *Coronaviridae*, order *Nidovirales* (Kocherhans et al., 2001) (later assigned as subgenus Pedacovirus by International Committee on Taxonomy of Viruses together with BtCoV/512/2005). The PEDV genome is approximately 28 kb in length and comprises seven open reading frames (ORF), including ORF1a/1b, spike (S), ORF3, envelop, membrane, and nucleocapsid protein. The S protein is a glycosylated protein involving viral pathogenesis and is further divided into S1 and S2 domains. The S1 protein plays an important role in receptor recognition, whereas S2 plays an important role in facilitating receptor recognition and membrane fusion (Bosch et al., 2003) and contains domains that stimulate the production of neutralizing antibodies (Duarte and Laude, 1994; Chang et al., 2002; Cruz et al., 2006; Sun et al., 2007; Cruz et al., 2008; Sun et al., 2008). S gene is divergent and is important for understanding the genetic relatedness of PEDV field isolates, the epidemiological status of the virus, and vaccine development (Park et al., 2011; Li et al., 2012b; Sun et al., 2012). ORF3 encodes an accessory protein located between S and E. In addition, the ORF3 gene is the accessory gene, and it has been an important determinant of virulence in PEDV (Park et al., 2008). The vaccine-derived isolates have a unique deletion of 17 amino acids at positions 82 to 99 (Park et al., 2008). The other E and M genes are associated with viral envelope formation and release. The N protein is the important antigen produced in coronavirus-infected cells, making it a major viral target (Kocherhans et al., 2001).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

PEDV is like other Coronaviruses since the spike protein plays an important role in host entry, with S1 as the attachment and together with S2 as the fusion procedures. According to the multiplication of the virus, host-pathogen interaction signaling against the express protein of the virus induces affected and neighbored cell apoptosis, then leading to villous atrophy (Kim and Lee, 2014). PEDV mainly targets porcine small intestinal villous epithelial cells or enterocytes and believes that porcine aminopeptidase N (pAPN) is the predominant receptor for the virus because pAPN is abundant on the intestinal tissue surface of pigs (Delmas et al., 1992) . However, it was controversial in recent studies since the pAPN-knockout pigs showed the disease as the control, and pAPN lacks reaction in forming complexes with the S1 domain *in vitro* (Luo et al., 2019). Pigs less than a week of age are the most susceptible host since they lack turnover rate of villi, have immature immunity, and more susceptible to dehydration (Thomas et al., 2015). According to a recent study, PEDV shows different degrees of clinical signs and disease progression; for example, some mutated strains in European countries (Steinbach et al., 2016) or the US harbored deletion strains (Su et al., 2018).

The previous report has shown that PEDV can infect kidney cells such as human embryonic kidney cells (Zhang et al., 2017) or bat (*Eptesicus fuscus*) kidney cell lines *in vitro* (Banerjee et al., 2016), although there is no clinical case report in human, or, renal abnormality in other animals. Interestingly, previous reports suggest that pAPN as the receptor of virus adhesion is controversial, with the latest findings (Li et al., 2017), demonstrating that APN is not essential for PEDV cell entry. The APN was reported to be rich in the renal system (Chen et al., 2019) , which might be one reason for the study of kidney cell lines.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Although the origin of PEDV is still unknown, many authors agreed that it should be originated from the *Scotophilus* bat or somehow recombined in other hosts before being reservoir in this host because bats are the rich reservoir of many viruses, including coronaviruses (Huang et al., 2013; Chen et al., 2019). Furthermore, bats also play an important role as the carrier of mammalian viruses since these carriers are the only mammals that can fly (Zhang et al., 2013), whereas the previous study on phylogenetic relationships among coronaviruses also suggests that bats may be the natural reservoir of PEDV as other alphacoronavirus (Vijaykrishna et al., 2007). This theory was also supported by a recent study in which the novel strain of alphacoronavirus is still discoverable (Mendenhall et al., 2019). One study of another coronavirus, SARS, also suggests that the virus has to undergo multiple recombination events before affecting the new host (Zhang et al., 2005), and *Coronaviridae* was also the second predominantly genus associated with bats (Chen et al., 2014a). For PEDV, the recombination between two different types of the virus, TGEV, and PEDV, was first reported in Italy (Boniotti et al., 2016).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Porcine epidemic diarrhea was first described in 1971, and the first isolate, CV777, was reported in 1978 (Chasey and Cartwright, 1978; Pensaert and De Bouck, 1978). In North America, there was evidence of PEDV occurring in Canada in 1980 (Turgeon et al., 1980), although the virus was discovered later in 2014 (Ojkic et al., 2015). In Asia, there was evidence of the disease occurring in China from 1966 to 1973; however, the first isolate in China was found in 1984 (Sun et al., 2016), while the isolate in Japan was reported in 1983 (Takahashi et al., 1983). In 1988, Hofmann and Wyler (Hofmann and Wyler, 1988) were the first authors to report the propagation method of the virus in Vero cells; since then, many authors have studied the virus antigenicity,

immunogenicity, and serological responses (Carvajal et al., 1995; Utiger et al., 1995; Shibata et al., 2000; Liu et al., 2001; Shibata et al., 2001; Chang et al., 2002; de Arriba et al., 2002), including the development of the vaccine strains (Kweon et al., 1999; Song et al., 2007). The pathogen was becoming more concern in the pig industry since the disease was found in the USA (Marthaler et al., 2014), one of the largest pig industrial countries, resulting in raising more studies of the virus, and the situation was more in complexity.

PEDV in China seems to be the most variable compared to other regions. During the 1970s to 2000s, most of the outbreak cases were found sporadically due to the effect of the vaccine used in fields; however, new genotypes of the virus were found in the late 2000s (Chen et al., 2008; Chen et al., 2010) those were different from previous report genotype. In the 2010s, many studies have shown that PEDV was found widely spread through the China mainland with distinct variations, and genetic evolution was higher in degree resulting in differences between them from the primitive strain. (Tian et al., 2013; Hao et al., 2014; Sun et al., 2014). In Japan, the situation during the 1980s to 2000s was similar to that of China since the vaccine developed from an attenuated-83P strain capable of controlling the pathogen (Sato et al., 2011). However, since the outbreak of PEDV in the USA, the largest exporter of swine products to Japan, the pathogen seems to be introduced into Japan via these products, and almost all the strains found in Japan are close to those of US strains.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

From previous studies, the evolutionary rates of PEDV in Asian countries seem to be higher than those of other regions. Most studies aimed to study the evolutionary rate of the external structure elements such as spike and envelope (Sung et al., 2015). One study showed that the evolutionary rate of G2 and also China was higher than that of the worldwide dataset (Sung et al., 2015). Compared to the previous study of TGEV

(Sánchez et al., 1992), PEDV was much higher in substitution rate and even higher than HCoV-OC43 (Oong et al., 2017), one of the human alphacoronavirus, and, also higher than IBV (beta coronavirus) in avian was the virus which well-known in highly mutation rate and recombination events, the evolutionary rate of the hypervariable region of spike gene was 3-6 x 10^{-3} whereas the OC43 (alphacoronavirus) in humans was 6.41 x 10^{-4}. These suggest that PEDV has quite a high mutation rate among the alphacoronavirus.

Porcine deltacoronavirus (PDCoV) is a pathogen that shares a mimic syndrome as PEDV and is characterized by watery diarrhea and villous atrophy (Chen et al., 2015). PDCoV is a newly emerged pathogen that was initially detected during the molecular surveillance of coronavirus in Hong Kong in 2012 (Woo et al., 2012). At first, there was no associated clinical disease; however, the first evidence of PDCoV causing the disease was first described in 2014 (Marthaler et al., 2014). Later, PDCoV was subsequently reported in several countries, including China, South Korea, Thailand, Lao PDR, Vietnam, and Mexico (Woo et al., 2012; Lee and Lee, 2014a; Marthaler et al., 2014; Song et al., 2015; Janetanakit et al., 2016; Lorsirigool et al., 2016; Saeng-Chuto et al., 2016; Saeng-Chuto et al., 2017; Pérez-Rivera et al., 2019). Although the mortality rate caused by PDCoV is lower than PEDV, however, later studies suggest that PDCoV might become a highly pathogenic virus since researchers worldwide reported that the virus could solely trigger clinical signs, and the mortality rate is over 80% (Jung et al., 2015; Song et al., 2015; Zhao et al., 2019; Saeng-Chuto et al., 2020) those are in contrast with previous findings.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

PDCoV is an enveloped, single-stranded, positive-sense RNA virus in the genus *Deltacoronavirus*, family *Coronaviridae*. Its genome is approximately 25,000 kb in length (Woo et al., 2012). The genome arrangement is in the order of 5' untranslated region (UTR), open reading frame 1a/1b (ORF 1a/1b), spike (S), envelope (E), membrane (M), nonstructural protein 6 (NS6), nucleocapsid (N), nonstructural protein 7 (NS7) and 3' UTR (Woo et al., 2012). The S gene encodes the S protein comprising two domains, S1 and S2. The S1 domain functions in an important role involving attachment between the virus and host receptor. The S2 domain facilitates membrane fusion (Shang et al., 2018). E and M genes encode transmembrane proteins E and M, respectively, functioning in envelope formation and virus release (Woo et al., 2012).

Based on previous phylogenetic analyses, PDCoV has evolved into three separate groups, including China, the US, and Southeast Asia (SEA) (Saeng-Chuto et al., 2016). China PDCoV comprises PDCoV isolates detected in China and Hong Kong. US PDCoV is widespread, comprising isolates detected in the USA, South Korea, Japan, and Mexico. SEA PDCoV is unique, consisting of only PDCoV isolates detected in Southeast Asian countries, including Thailand, Lao PDR, and Vietnam. Although SEA PDCoV clustered separately from China and US PDCoV, viruses in these three clusters originated from the same common ancestor.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

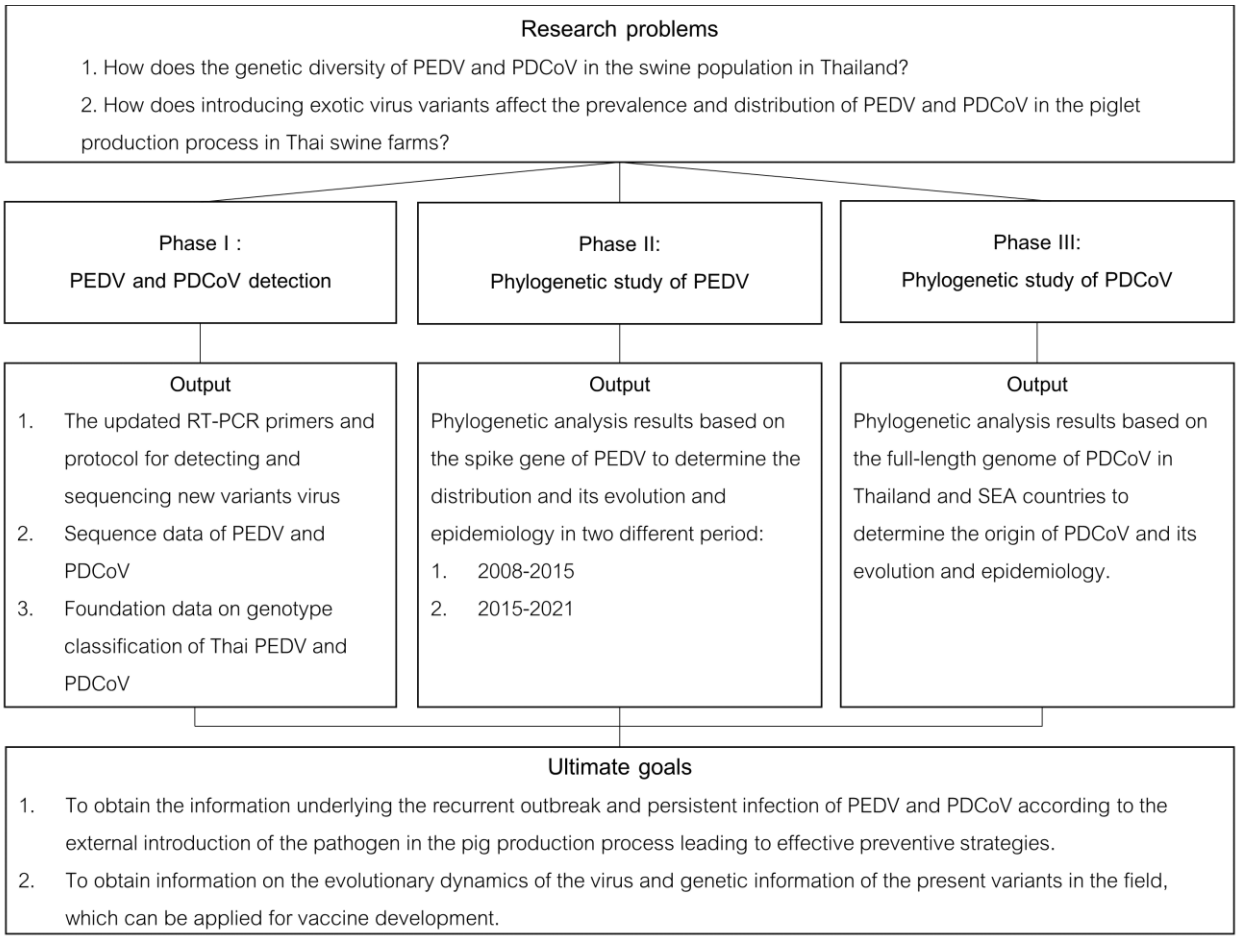
Although clinical diseases caused by PDCoV were first reported in the USA in 2014, retrospective investigations from several countries, including China and the US, reported the presence of PDCoV in intestinal samples of pigs affected with watery diarrhea as early as 2004 and 2013, respectively (Sinha et al., 2015; McCluskey et al., 2016; Saeng-Chuto et al., 2017; Pérez-Rivera et al., 2019; Saeng-Chuto et al., 2020). In Thailand, PDCoV was first reported in 2015 (Janetanakit et al., 2016). However, a subsequent retrospective study reported that PDCoV was present in Thailand as early

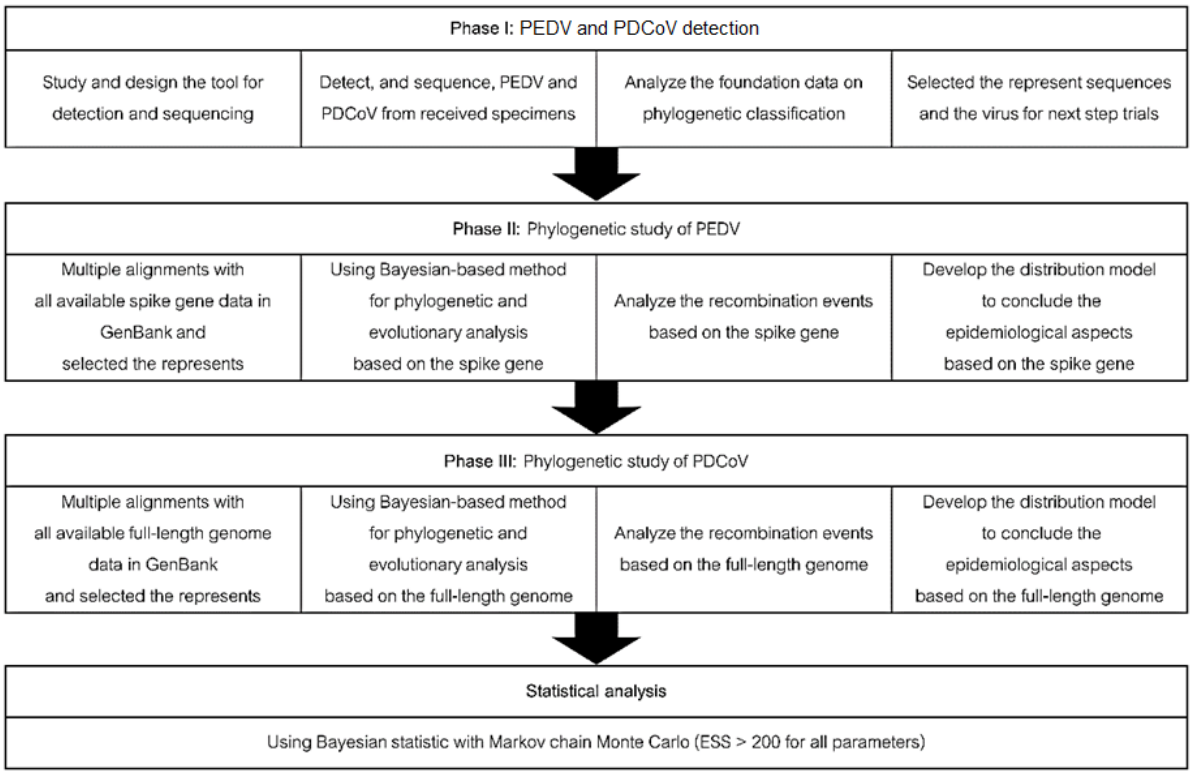
as 2013 (Saeng-Chuto et al., 2017). This suggests that PDCoV might circulate in swine farms, but clinical signs might confuse other highly pathogenic pathogens, including PEDV. In addition, recombinant viruses were detected in SEA PDCoV (Saeng-Chuto et al., 2020), in Vietnam in 2020, forming a new subgroup of SEA genogroup. Recombination between PDCoV isolates from Vietnam and the US was suspected of having given rise to this chimeric strain. (Saeng-Chuto et al., 2020). Based on these findings, SEA PDCoV appears to have more complexity in the evolution compared to that of the US and China PDCoV.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER 4

CONCEPTUAL FRAMEWORK





จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER 5

MATERIALS AND METHODS, AND ETHICAL STATEMENT

Ethical statement

The experiment was done under scientific ethics and regulations. The information about pathogen risk assessment and handling was submitted to the Institutional Biosafety Committee of the Faculty of Veterinary Science, Chulalongkorn University (CU-VET-IBC), protocol number 2231028.

Phase I: PEDV and PDCoV detection

Source of the specimen, RT-PCR, and sequencing

Intestinal samples were collected from 3- to 4-day-old piglets of the affected farms (n=36) from 2015 to 2021, those displaying the clinical features associated with swine enteric coronavirus, including vomiting and watery diarrhea. Intestinal samples were minced into small pieces and re-suspended in PBS in a 25 ml centrifuge tube and clarified by centrifugation for 15 minutes at 4500 rpm with a 4°C temperature maintained. The supernatant was transferred to the new tube and repeated in the centrifugation step for 5 minutes before being filtered through 0.22-micrometer filters and stored at - 80 °C until use. The clarified supernatants were subjected to RT-PCR.

The total RNA was extracted from the supernatant using the Nucleospin® RNA Virus kit (Macherey-Nagel Inc., PA, USA), and cDNA was synthesized from the extracted RNA using the M-MuLV Reverse Transcriptase (New England BioLabs Inc., MA, USA) following to the manufacturer's instructions.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

PCR amplification was performed on the cDNA by targeting the PEDV spike gene (Park et al., 2007) and the PDCoV membrane gene, 67F (5'-ATCCTCCAAGGAGGCTATGC-3'), 560R (5'-GCGAATTCTGGATCGTTGTT-3'), and nucleocapsid (N) gene 41F (5'-TTTCAGGTGCTCAAAGCTCA-3'), 735R (5'-GCGAAAAGCATTTCCTGAAC-3') (Wang et al., 2014). The condition of PCR consisted of 5 minutes of initial denature followed by 35 cycles of 95°C denaturation for 1 minute, 57°C annealing for 45 seconds, and 72°C extensions, followed by a final extension for 5 minutes after the last cycles. The PCR products were purified by Nucleospin® Gel and PCR Clean-up kit (Macherey-Nagel Inc., PA, USA) before being subject to Sanger's sequencing by a third-party service to confirm the pathogens.

Sequence analysis

The good quality signal nucleotide obtained from Sanger's sequencing were chosen to compare with the previously published sequences using nucleotide BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1990) to determine the species of the pathogen and its genotype.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Phase II: Phylogenetic study of PEDV

Samples and complete spike gene sequencing

PEDV-positive samples were retrieved from the previous phase for RNA extraction using the Nucleospin® RNA Virus kit (Macherey-Nagel Inc., PA, USA) and cDNA synthesis using M-MuLV Reverse Transcriptase (New England BioLabs Inc., MA, USA). The complete spike gene detection was performed by RT-PCR using four specific primers consisting of S1-F (5'-ACG TAA ACA AAT GAG GTC TTT-3'), S1-R (5'-ATA CAC CAA CAC AGG CTC TGT-3'), S2-F (5'-GGT TTC TAC CAT TCT AAT GAC G-3') and S2-R (5'-GTA TTG AAA AAG TCC AAG AAA CA-3') (Lee et al., 2010). The condition of PCR

consisted of 5 minutes of initial denature at 94°C followed by 35 cycles of 94°C denaturation for 30 seconds, 55°C annealing for S1 and 58°C for S2 primers for 30 seconds, and 72°C extensions for 2 minutes, followed by a final extension for 10 minutes after the last cycles (Temeeyasen et al., 2014). The amplified products were purified using Nucleospin® Gel and a PCR Clean-up kit (Macherey-Nagel Inc., PA, USA) and were sent to the third-party sequencing service.

Phylogenetic tree and analyses

The dataset used in the phylogenetic study was separated into two periods, including the period from 2008-2015 and 2014-2021. The first period aimed to study the introduction of Thai PEDV and its distribution during the widespread re-emerged of PEDV up to 2014, while the second period aimed to study the subsequent introduction of multiple variants and the causes of sporadic outbreaks in Thailand from the widespread outbreak in 2014 to 2021.

The nucleotide and deduced amino acid sequences were aligned using the software MAFFT (Multiple Alignment using Fast Fourier Transform) (Katoh et al., 2002). The phylogenetic analysis was performed based on the nucleotide sequences of complete spike genes of the sequences of this experiment. Previously published spike genes from 2008 to 2015 in Thailand, along with 94 other PEDV spike sequences available in GenBank® (Sayers et al., 2019) from the previous study (Lee, 2015), were used for the first-period study. A phylogenetic tree was constructed using the Bayesian Markov chain Monte Carlo (BMCMC) with a best-fit substitution model chosen by MEGA (Kumar et al., 2016). The base frequency was set as empirical, and all other settings were left as default. The analysis was performed in at least 200 million states with logged every 10,000 states until ESS is more than 200; the genogroup described in

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

another study (Lee, 2015) was used to classify and compare with the sequences in this study.

For the second period study, all available spike gene sequences in GenBank® (up to the end of 2021) were recruited from two different methods; the first method was using the represent sequences at 99% of the genome sequences and spike gene sequences separately, from each country dataset, whereas the second method was using the represent at 99% from all available sequences from each country separated by year of collection, to fulfill the classification clades. The representative sequences were chosen using CD-HIT (Li and Godzik, 2006) The phylogenetic tree of the second period was analyzed using neighbor-joining with the best-fit model available in MEGA for the further classification of Thai PEDV. This second-period study dataset was used for recombination analysis and molecular evolutionary analysis as well.

Recombination analysis

Recombination events among the sequences were analyzed using automated RDP, GENECONV, BOOTSCAN, MaxChi, CHIMAERA, and SISCAN provided by RDP4 software RDP (Martin and Rybicki, 2000), GENECOV (Padidam et al., 1999), BOOTSCAN (Salminen et al., 1995), MAXCHI (Smith, 1992), CHIMAERA (Posada and Crandall, 2001), SISCAN (Gibbs et al., 2000), and 3SEQ (Boni et al., 2007), with potential recombination signals.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Molecular evolutionary analysis

The substitution analysis of Thai PEDV spike sequences, as described in the previous section, was performed using a BMCMC method implemented in the program BEAST package (Drummond and Rambaut, 2007; Ferreira and Suchard, 2008; Lemey et al., 2010; Drummond et al., 2012; Pybus et al., 2012). The coalescent Bayesian skyline tree prior (Drummond et al., 2002; Drummond et al., 2005) and empirical base frequencies was applied under three different models for rate variation among branches: the strict molecular clock model (STR), the uncorrelated lognormal relaxed-clock model (LOG), and the uncorrelated exponential relaxed-clock model (EXP). Maximum clade credibility trees were annotated using TreeAnnotator, and a phylogenetic tree with timeline (chronogram), estimated divergences, posterior probability, and 95% HPD displays were generated using FigTree 1.4 (Rambaut and Drummond, 2014).

Phylogeographic analysis

Phylogeographic distribution was estimated using the BEAST package. The result was converted and displayed in GoogleEarth (Lozano-Fuentes et al., 2008) using SPREAD 1.0.7 (Bielejec et al., 2011). The sequences which have 100% identity from the same location and time of the outbreak were excluded to avoid statistical error. Collection dates were used as the tip dates; if there was any missing data, the mid-values of the month or year were compensated. The GPS-coordinated locations were applied using the location of states or provinces available from the information available in GenBank® and previous reports. The data will be accepted in case the effective sample size (ESS) of all parameters is more than 200.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Phase III: Phylogenetic study of PDCoV

Samples and full-length genome sequencing

PDCoV-positive samples were retrieved from the previous phase for RNA extraction using the Nucleospin® RNA Virus kit (Macherey-Nagel Inc., PA, USA) and cDNA synthesis using M-MuLV Reverse Transcriptase (New England BioLabs Inc., MA, USA). The full-length genomes detection was performed by RT-PCR using 26 specific primers adapted from previous publications (Wang et al., 2014; Lorsirigool et al., 2017). The condition of PCR consisted of 5 minutes of initial denature at 95°C followed by 35 cycles of 95°C denaturation for 30 seconds, 55°C annealing for 45 seconds, and 72°C extensions for 1 minute, followed by a final extension for 5 minutes after the last cycles. The amplified products were purified using Nucleospin® Gel and a PCR Clean-up kit (Macherey-Nagel Inc., PA, USA) and were applied to the third-party sequencing service.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 1 Primers used in the PDCoV experiment.

Pair number	Name	Sequence (5' to 3')
1	PDCoV_ORF1ab1_F	ACATGGGGACTAAAGATAAAAATTATAGC
	PDCoV_ORF1ab1_R	TCGTGTCAGCAAGAGATGTACCA
2	PDCoV_ORF1ab2_F	AGGACAAGGAACTTGCTGAGTTAGC
	PDCoV_ORF1ab2_R	GCCTTATTGAGTATCTTAGTAGTTGCC
3	PDCoV_ORF1ab3_F	CAATTTACACTGTTGGTAATCGCATGC
	PDCoV_ORF1ab3_R	GCAGGTGTTACTGGTACTATAGG
4	PDCoV_ORF1ab4_F	CAGGCTGTTGAAGACAAACCTTCTGATA
	PDCoV_ORF1ab4_R	GTCTTCAGGTAGACGCAGTATTGCTG
5	PDCoV_ORF1ab5_F	AGAACCAAGCCAAGGAACTTGACC
	PDCoV_ORF1ab5_R	GGCATTGTAATTACAAGGCCAGCC
6	PDCoV_ORF1ab6_F	GCTTAATTCAGGCTATACAATTGGCAC
	PDCoV_ORF1ab6_R	CCATTGGTCGCTTGCTTAATGTCTG
7	PDCoV_ORF1ab7_F	GCTTCTCGTATTGAGAAGTATTACCC
	PDCoV_ORF1ab7_R	ACAAGATCACCATTAATGTCAATAGCC
8	PDCoV_ORF1ab8_F	CAGGACACTAGACAACTCTATTGC

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Pair number	Name	Sequence (5’ to 3’)
8	PDCoV_ORF1ab8_R	GACACTTGACTATAAAGTCGCGGCA
9	PDCoV_ORF1ab9_F	ATTGGAAAATTCCGTGGTGACCAGTG
	PDCoV_ORF1ab9_R	CAGTGTAACCATACAGTCGCTGAGC
10	PDCoV_ORF1ab10_F	TGGATACCGAACACATTTCTGCGCAC
	PDCoV_ORF1ab10_R	CATTCCGCCTTTGCAATGTTAACAGC
11	PDCoV_ORF1ab11_F	ACGATCTGTTGAGATGAACGAATCTC
	PDCoV_ORF1ab11_R	CAAGAACTGCATGGCTCATGAGTC
12	PDCoV_ORF1ab12_F	ATCCTGGTGTTGATGGACGGTGC
	PDCoV_ORF1ab12_R	TCTGGACATTCAACAGTGTTAGGATG
13	PDCoV_ORF1ab13_F	CCCACTTGAGTATGATGGATTTCAGT
	PDCoV_ORF1ab13_R	GCATGTTTGGCATAGAACGATCACAC
14	PDCoV_ORF1ab14_F	CCAAATTCTATGGTGGTTGGGACAA
	PDCoV_ORF1ab14_R	TCGACGAATGCAATCACCACAGCG
15	PDCoV_ORF1ab15_F	GCTCAGATGTATGAGCAAAGTCCTAC
	PDCoV_ORF1ab15_R	TGATAGTCGAGAAGATATACTGTGCAG
16	PDCoV_ORF1ab16_F	TACCCACACGTACAACTGTTGAGTGC
	PDCoV_ORF1ab16_R	TGCCGATCCAACTCTGAACATAGGTG
17	PDCoV_ORF1ab17_F	ATCCACACTTGGGTTTCTTCCATCAC

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Pair number	Name	Sequence (5’ to 3’)
17	PDCoV_ORF1ab17_R	CAATAGGCTGTTCACTACTACAATCTG
18	PDCoV_ORF1ab18_F	TCAACTTGGCTCCACTCATCTACTAC
	PDCoV_ORF1ab18_R	GGCATGTGCTTGTCTTATGATGTCAAT
19	PDCoV_ORF1ab19_F	CACCTAACGCAAGCTCCAAGAACG
	PDCoV_ORF1ab19_R	GGTAGCCAGGAAGAACGCCAACA
20	PDCoV_S1_F	ATGCAGAGAGCTCTATTGATTATGACC
	PDCoV_S1_R	CTTCGCCAAAATCCATGTGTGCAG
21	PDCoV_S2_F	CAATAGCATGCCAGCGCTCTTCTCA
	PDCoV_S2_R	TGGTATTTCAACTTCGCCATCGTATAG
22	PDCoV_S3_F	CATCCACATTACAGAATACTCGAC CA
	PDCoV_S3_R	TGAGTAACATATGCATTAAGTGCAGC
23	PDCoV_S4_F	CATTATCACACCTGACTGCACAGCT
	PDCoV_S4_R	CTACCATTCCTTAAACTTAAAGGACG
24	PDCoV_EM_F	AAGGTATACCGACGACCAACCAACAC
	PDCoV_EM_R	CTGCAGATGGCAGTTGCACATTACAT
25	PDCoV_Nsp6N_F	AGAAATCCGCCACAGGATATGGT
	PDCoV_Nsp6N_R	GAAGGGGTCAACTCTGAAACCTTG
26	PDCoV_NNsp7_F	GGTTCGGGAGCTGACACTTCTATTAA
	PDCoV_NNsp7_R	GCTCCATCCCCCCTATAAGCCAA

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Phylogenetic tree and analyses

The alignment of nucleotide and amino acid sequences was utilized using MAFFT (Katoh et al., 2002) to determine the genetic relationship between Thai PDCoV isolates and other PDCoV strains. Additionally, 100 full-length genomes of PDCoV isolates available in GenBank® were included in the analysis. Nucleotide identity will be performed using the sequence identity matrix function implemented in BioEdit software (Hall, 1999). The phylogenetic tree based on the full-length genome and spike gene sequences was separately constructed using the Tamura-Nei model, with the gamma distribution and 1,000 bootstraps, using maximum likelihood in MEGA-X (Kumar et al., 2018). The model used for the maximum likelihood tree was chosen by the model selection function in MEGA-X. Due to the previous finding of Thai PDCoV as the SEA lineage (Saeng-Chuto et al., 2016), the SEA dataset, including previously published genomes of Viet Nam and Lao PDR, was used in order to compare with other lineages. The dataset of SEA lineage was applied for all analyses in the PDCoV study.

Recombination analysis

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Recombination events within Thai full-length genome sequences in the experiment were analyzed with the 100 references PDCoV isolates using RDP4. Recombination events which strongly supported by at least six of seven detection tools available in RDP4 automated analysis, including RDP (Martin and Rybicki, 2000), GENECOV (Padidam et al., 1999), BOOTSCAN (Salminen et al., 1995), MAXCHI (Smith, 1992), CHIMAERA (Posada and Crandall, 2001), SISCAN (Gibbs et al., 2000), and 3SEQ (Boni et al., 2007), with potential recombination signals, will be accepted. Major and minor parents and breakpoints would also be detected, and recombination fragments will then be generated. The recombination rate plot was used to compute the

recombination rate (Auton and McVean, 2007). Genes were separately analyzed, and genes with recombination incidences were exported as the partition according to the recombination breakpoints to apply for Bayesian analyses in the BEAST package.

Molecular evolutionary analysis

The nucleotide (nt) sequences of PDCoV were aligned using MAFFT and then modified using BioEdit. The datasets were prepared and generated in BEAUTI, including the full-length genome, spike, envelope, membrane, nucleocapsid, NS6, and NS7. The partition at the recombination breakpoints was employed to segregate the datasets containing recombination incidences (Martin et al., 2015), with the aim of reducing recombinational biases during analysis.

Molecular evolutionary rates and the estimated time to the most recent common ancestor (TMRCA) were estimated using BEAST (Drummond and Rambaut, 2007; Ferreira and Suchard, 2008; Lemey et al., 2010; Drummond et al., 2012; Pybus et al., 2012) and BEAGLE (Ayres et al., 2012). The general time-reversible model plus gamma distribution plus invariant sites (GTR+G+I) (Tavaré, 1986; Yang, 1994) was applied with the Bayesian skyline tree prior (Drummond et al., 2002; Drummond et al., 2005) and flexible local clock (FLC) (Fourment and Darling, 2018). Collection dates were used as tip dates. If there is any missing data, the mid-values of the month or year were compensated. The analysis was performed at least 500 million generations per run, with 10,000 states logged to qualify 200 ESS. The maximum clade credibility (MCC) tree was built by choosing common ancestor heights using TreeAnnotator in the BEAST package, 10% burn-in was applied, and the posterior probabilities and estimated divergences were estimated. he substitution rates were interpreted in Tracer 1.6 (Rambaut and Drummond, 2013) and compared with US and China datasets. The annotated

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

phylogenetic tree was generated and displayed using FigTree 1.4 (Rambaut and Drummond, 2014).

Positive selection, and InDels analysis

A genome scan for signals of natural selection was performed to determine the variability of selective pressure among different gene regions. Ratios of non-synonymous to synonymous nucleotide substitutions rates (dN/dS) were estimated for the entire alignment using the Synonymous Non-synonymous Analysis Program (SNAP) (Korber, 2000) to compare the codon selected among the global, SEA, US, and China codons. If there was an ambiguous nucleotide in the previously published sequences, the sequence was duplicated and replaced from the IUPAC nucleotide code to base nucleotides (R = A or G, Y = C or T, S = G or C, W = A or T, K = G or T, and M = A or C(Johnson, 2010)). A dN/dS ratio >1 indicated positive selection (Yang and Bielawski, 2000). Insertion deletions (InDels) were additionally estimated and displayed.

Phylogeographic analysis

For the phylogeographic analysis, the previously described dataset was used for analysis. The nucleotide sequences of the full-length genomes of PDCoV were aligned using MAFFT and then edited using BioEdit. In order to decrease biases in the analysis, RDP4 was employed to estimate the partition breakpoint. The phylogeography and distribution model was performed using BEAST with BEAGLE. Collection dates were used as the tip dates; if there was any missing data, the mid-values of the month or year were compensated. The GPS-coordinated locations were applied using the location of states or provinces from the information available in GenBank® and previous reports. The data will be accepted if the effective sample size (ESS) of all parameters exceeds 200.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The results, including this study's substitution rate and ESS, were interpreted using Tracer 1.6. MCC trees were annotated by choosing the common ancestor heights using TreeAnnotator in the BEAST package. The chronogram, estimated divergences, posterior probability, and 95% high population density (95% HPD) information were transformed into the represented tree in FigTree 1.4. The represented tree was then converted using SPREAD 1.0.7 (Bielejec et al., 2011), and displayed on Google Earth (Lozano-Fuentes et al., 2008).

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

CHAPTER 6

DURATION OF STUDY, PLACE OF STUDY, EQUIPMENT AND REAGENTS, BUDGET, AND ADVANTAGE OF THIS STUDY

Duration of Study

Month	1-4	5-8	9-12	13-16	17-20	21-24
Phase I: PEDV and PDCoV detection						
Month	25-28	29-32	33-36	37-40	41-44	45-48
Phase I: PEDV and PDCoV detection						
Phase II: Phylogenetic study of PEDV						
Phase III: Phylogenetic study of PDCoV						
Month	49-52	53-56	57-60	61-64	65-68	69-72
Phase I: PEDV and PDCoV detection						
Phase II: Phylogenetic study of PEDV						
Phase III: Phylogenetic study of PDCoV						
Data analysis						
Month	73-76	77-80	81-84	85-88	89-92	93-96
Data analysis						
Thesis writing						

Places of Study

- Department of Veterinary Microbiology, Faculty of Veterinary Science, Chulalongkorn University

Equipment and reagents

- Digital Fuzzy-control Autoclaves (Daihan MaXterile™)
- Microcentrifuge (Sorvall™ Legend™ Micro 17)
- Centrifuge (Sorvall™ Legend™ RT+)
- Thermal Cycler (BIO-RAD T100™)
- Protein quantification spectrometer (Colibri Titertek Berthold)
- Vortex Mixer (Sigma Aldrich Vortex-Genie® 2)
- Gel Documentation (Vilber lourmat CN-08)
- Downflow cabinet class 2 (Claus Damm DanLAF type VFRS 1206)
- Ultra-low temperature (ULT) freezers (Thermo Scientific™ - Revco™ RDE Series)
- Laboratory bottle 250/500/1000/2000 ml (Duran®)
- Measuring cylinder with hexagonal base 150/1000/2000 ml (Duran®)
- Glass beaker 50/250/500/1000/2000 ml (Duran®)
- Micropipette 0.2-2/2-20/20-200/100-1000 ml (Finpipette™ F1)
- Micropipette tips 10/200/1000 microliter (Axygen®)
- Microcentrifugetube 1.5 ml (Axygen®)
- PCR tube 200 microliter (Axygen®)
- Polypropylene centrifuge tube 15 ml (Corning®)
- Polypropylene Centrifuge tube 50 ml (Corning®)
- Sodium hypochlorite (Clorox®)
- Phosphate buffer solution (Bio Basic)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

- Tris-Borate-EDTA buffer (Bio Basic)
- Agarose A (Bio Basic)
- Nucleospin® RNA Virus kit (Macherey-Nagel Inc.)
- M-MuLV Reverse Transcriptase (New England BioLabs Inc.)
- Nucleospin® Gel and PCR Clean-up kit (Macherey-Nagel Inc.)
- GoTaq® Green Master Mix (Promega)

Budget

Categories	**Baht**
Facility fees	100,000
Sample collection wares and reagents	50,000
Template preparation wares and reagents	200,000
RT-PCR reagents	200,000
Product clean-up reagents	160.000
Sequencing services	320,000
Miscellaneous expenses	103,000
Total	1,133,000

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Advantages of this study

1. To obtain the information underlying the recurrent outbreak and persistent infection of PEDV and PDCoV according to the external introduction of the pathogen in the pig production process leading to effective preventive strategies.
2. To obtain information on the evolutionary dynamics of the virus and genetic information of the present variants in the field, which can be applied for vaccine development.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER 7

PHASE I: PEDV AND PDCoV DETECTION RESULTS

Source of the specimen, RT-PCR, and sequencing

Two hundred eighty-eight complete spike gene sequences were used for phylogenetic study, including 21 previous published sequences (KC764952-KC764960, KF724935-KF724938, KR610991-KR610994, KY000559, KX981899, KY828922, and LC496368) and 267 sequences retrieved from this study. All samples were collected from 71 different outbreak sources from 10 provinces, including Buriram (n=2), Chachoengsao (n=3), Chanthaburi (n=1), Chonburi (n=4), Lopburi (n=1), Nakhon Ratchasima (n=2), Nakhon Panom (n=2), Nakhon Pathom (n=2), Ratchaburi (n=44), and Saraburi (n=10) (Table 2).

All provinces in this study were affected by the widespread outbreaks of PEDV during 2014-2015 (Figure 1); however, from 2017, only the samples from Ratchaburi and Saraburi were PEDV positive. The incidences of PEDV were highest in 2015 and in Ratchaburi province (Table 2).

Chulalongkorn University

Twelve complete genome sequences were used for the phylogenetic study, including seven previously published sequences (KU051641, KU051649, KU984334, KX361343-KX361345, and MH700627) and five sequences retrieved from this study. All samples were collected from 9 different outbreak sources in 4 provinces, including Chachoengsao (n=1), Chonburi (n=4), Nakhon Panom (n=1), and Ratchaburi (n=3) (Table 2 and Figure 1).

Table 2 Sources of the specimens used in this study separated by year of collection and the province affected by the outbreaks by year (RBR = Ratchaburi; SRI = Saraburi; NPT = Nakhon Pathom; CBI = Chonburi; NMA = Nakhon Ratchasima; CCO = Chachoengsao; NPM = Nakhon Panom; LRI = Lopburi; CTI = Chanthaburi; BRM = Buriram)

PEDV positive samples (complete spike gene)														
Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Sources	2	-	1	3	4	1	18	14	11	1	8	4	1	3
Samples	4	-	1	5	8	2	48	61	38	2	63	37	6	13
Provinces	RBR	-	RBR	RBR	RBR	RBR	RBR	RBR	RBR	RBR	RBR	RBR	RBR	RBR
				SRI			SRI	SRI	SRI			SRI		SRI
							NPT	NPT	NMA					
							CBI	CBI	CCO					
							NMA	NPM	NPM					
							CCO	LRI						
								CTI						
								BRM						
PDCoV positive samples (complete genome)														
Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Sources	-	-	-	-	-	2	-	3	4	-	-	-	-	-
Samples	-	-	-	-	-	2	-	4	6	-	-	-	-	-
Provinces	-	-	-	-	-	RBR	-	CBI	RBR	-	-	-	-	-
								CCO	CBI					
									NKP					

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

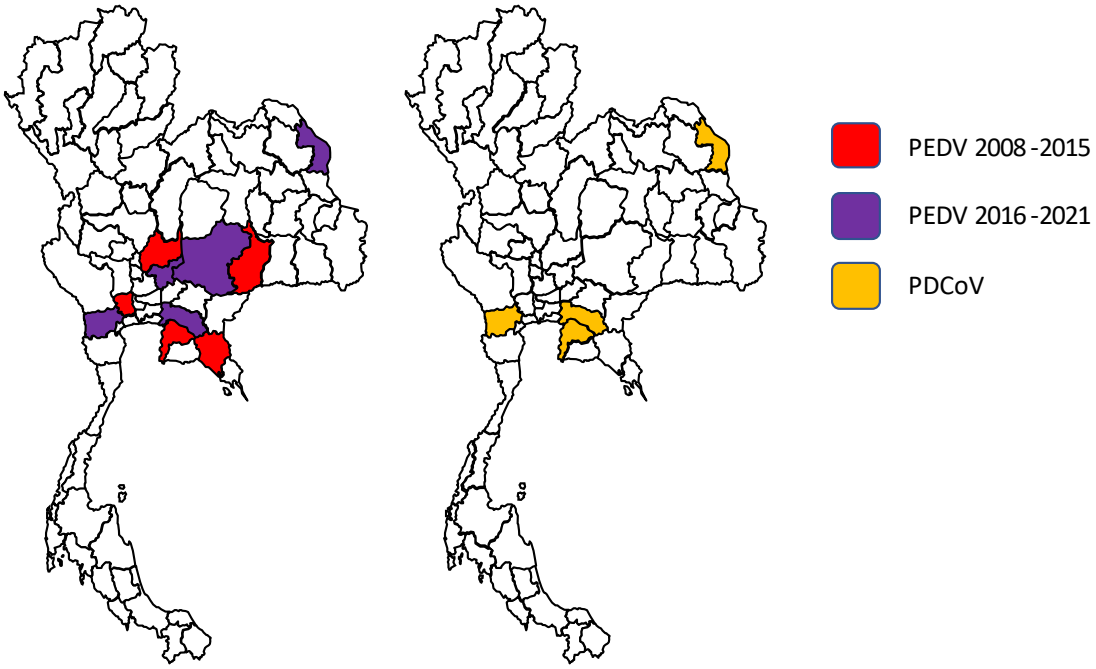


Figure 1 Sources of the specimens used in this study and positively detected by PEDV primers (left panel) and PDCoV primers (right panel). The red color denotes the provinces that only had widespread outbreaks during 2008-2015, while the purple color denotes the provinces that also had the outbreak since widespread outbreaks and after.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER 8

PHASE II: PHYLOGENETIC STUDY OF PEDV

Part I: The analysis of 2008-2015 PEDV

This work has been published in the topic of

Evolutionary and epidemiological analyses based on spike genes of porcine epidemic diarrhea virus circulating in Thailand in 2008-2015

Infection, Genetics and Evolution

Jun 2017, Volume 50, Pages: 70-76

(Appendix A)

Christopher James Stott, Gun Temeeyasen, Thitima Tripipat, Pavita Kaewprommal, Angkana Tantituvanont, Jittima Piriyapongsa, and Dachrit Nilubol

Phylogenetic tree and analyses

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

All PEDV samples were grouped based on spike gene sequence data using clades previously reported (Lee, 2015) in which PEDV evolved into two separated clades, including G1 and G2. Each clade was divided into two subclades, including G1a, G1b, G2a, and G2b. Subclade G1a included the first period of PEDV found in Belgium, the UK (CV777, CH/S, Br1/87), and attenuated vaccine strains from China and Korea. Subclade G1b, including the China, Korean, the US, and Europe strains reported during 2014-2015. G2 further evolved into G2a and G2b. G2a included strains from China, Korea, and Thailand. G2b included mainly PEDV strains from the US and countries that reported the detection of US-like PEDV.

Thai PEDVs are classified into six subgroups, including TH1-TH6 (Figure 2). Based on spike gene identity and previously described clades (Lee, 2015) , most Thai sequences belong to the G2a clade. The genetic differences among these six subgroups are greater than 3% between subgroups; therefore, we classified three subgroups of Thai PEDV to G2a, including TH1, TH2, and TH3. Subgroups TH1 (n=79) and TH2 (n=30) were the two dominant groups in Thailand. TH1 included the first Thai PEDV strain published in 2008 and the latest samples we found in 2015. In the endemic group 1 (TH1), there were 26 samples from 19 outbreaks with the 9-nucleotide insertion of CAA GGG AAT when aligned with the outbreak samples in Thailand and reference sequence (NC_003436; position 688th-689th), which caused the amino acid of reference sequence changed from N to T (229 of the reference) and insertion of REY at the site between 229th -230^{3rd}. TH3 only contained one sample in 2011, which might be a new introduction of the virus or the recombinant one. However, the sequence seems recombinant because it shares 97% identity with some sequences in both G1 and G2 groups when blasting with nucleotide BLAST.

Subgroups TH4 and TH6 were in genogroup G2b. TH4 was applied for CBR1 and CBR2 (KR610993 and KR610994) sequences in July 2014; these sequences belong to G2b and are very close to Vietnam sequences (99% identity) in 2013, so we can conclude them as a new introduction of the exotic strains. TH6 is the latest subgroup that contained the new introduction of the exotic strain (P1915-NPF-071511A) in 2015 that belongs to the G2b or US-like subgroup.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

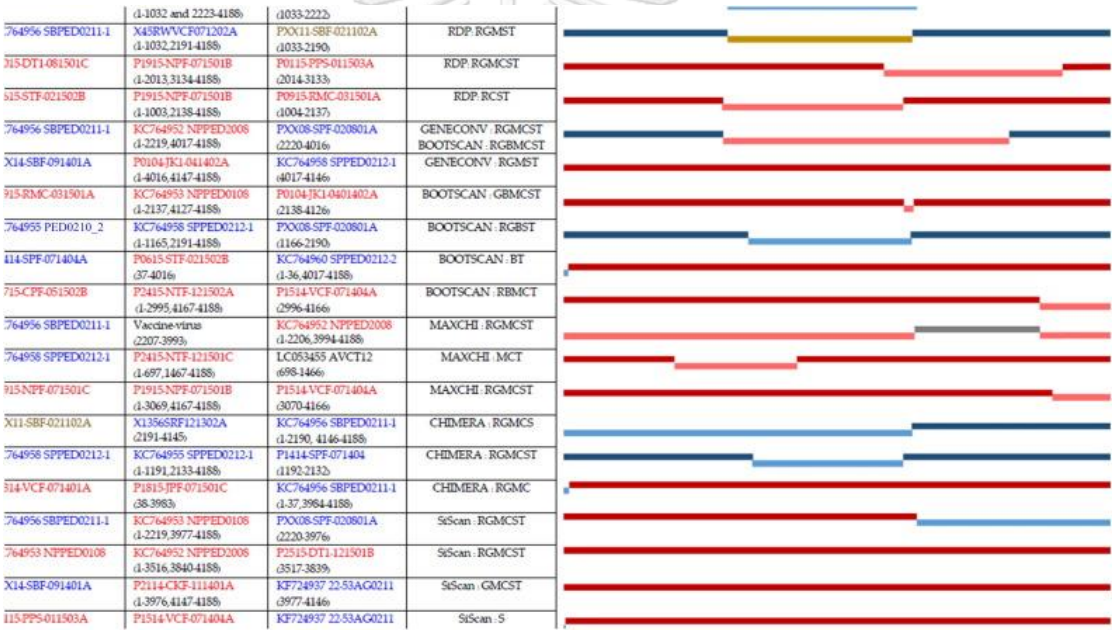
Subgroup TH5 was in genogroup G1a. TH5 belonged to the latter group of samples in October 2014, EAS1 and EAS2 (KR610991 and KR610992), the first group of a strain belonging to the G1 clade that we found in Thailand. The Thai PEDV in this genogroup were genetically distinct from PEDV strains that recently emerged in European countries. These strains were close to the vaccine strain, so we isolated the virus vaccine used in the farms that experienced the outbreak of the strain, including an attenuated strain (LC053455).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 2 The phylogenetic tree is based on the complete spike gene of Thai PEDV sequences and reference sequences (Lee, 2015). PEDVs have evolved into two genogroups, including G1 and G2. G1 and G2 are each further divided into two sub-groups, including G1a, G1b, G2a, and G2b. TH1-TH6 represent Thai PEDV subgroups of the samples collected from 2008 to 2015.

Recombination analysis

A total of 22 recombination events based on the spike, with known parents, were detected from RDP4 sequence analysis, as shown in Figure 3. Recombination rate plot located five recombination sites (cutoff with >0.01 Rho/bp), position 1st-117th (0.06131 Rho/bp), position 1048th-1206th (0.05912 Rho/bp), position 2224th-2266th (0.02536 Rho/bp), position 3070th-3116th (0.01947 Rho/bp), and position 4102nd-4142nd (0.06236) (Figure 4).



	(1-1032 and 2223-4188)	(1033-2222)	
764956 SBPED0211-1	X45RWVCF071202A (1-1032,2191-4188)	PXX11-SBF-021102A (1033-2190)	RDP.RGMST
015-DT1-081501C	P1915-NPF-071501B (1-2013,3134-4188)	P0115-PPS-011503A (2014-3133)	RDP.RGMCST
515-STF-021502B	P1915-NPF-071501B (1-1003,2138-4188)	P0915-RMC-031501A (1004-2137)	RDP.RCST
764956 SBPED0211-1	KC764952 NPPED2008 (1-2219,4017-4188)	PXX08-SPF-020801A (2220-4016)	GENECONV : RGMCST BOOTSCAN : RGBMCST
X14-SBF-091401A	P0104-JK1-041402A (1-4016,4147-4188)	KC764958 SPPED0212-1 (4017-4146)	GENECONV : RGMST
915-RMC-031501A	KC764953 NPPED0108 (1-2137,4127-4188)	P0104-JK1-0401402A (2138-4126)	BOOTSCAN : GBMCST
764955 PED0210_2	KC764958 SPPED0212-1 (1-1165,2191-4188)	PXX08-SPF-020801A (1166-2190)	BOOTSCAN : RGBST
414-SPF-071404A	P0615-STF-021502B (37-4016)	KC764960 SPPED0212-2 (1-36,4017-4188)	BOOTSCAN : BT
715-CPF-051502B	P2415-NTF-121502A (1-2995,4167-4188)	P1514-VCF-071404A (2996-4166)	BOOTSCAN : RBMCT
764956 SBPED0211-1	Vaccine virus (2207-3993)	KC764952 NPPED2008 (1-2206,3994-4188)	MAXCHI : RGMCST
764958 SPPED0212-1	P2415-NTF-121501C (1-697,1467-4188)	LC053455 AVCT12 (698-1466)	MAXCHI : MCT
915-NPF-071501C	P1915-NPF-071501B (1-3069,4167-4188)	P1514-VCF-071404A (3070-4166)	MAXCHI : RGMCST
X11-SBF-021102A	X1356SRF121302A (2191-4145)	KC764956 SBPED0211-1 (1-2190, 4146-4188)	CHIMERA : RGMCS
764958 SPPED0212-1	KC764955 SPPED0212-1 (1-1191,2133-4188)	P1414-SPF-071404 (1192-2132)	CHIMERA : RGMCST
514-VCF-071401A	P1815-JPF-071501C (38-3983)	KC764956 SBPED0211-1 (1-37,3984-4188)	CHIMERA : RGMC
764956 SBPED0211-1	KC764953 NPPED0108 (1-2219,3977-4188)	PXX08-SPF-020801A (2220-3976)	StScan : RGMCST
764953 NPPED0108	KC764952 NPPED2008 (1-3516,3840-4188)	P2515-DT1-121501B (3517-3839)	StScan : RGMCST
X14-SBF-091401A	P2114-CKF-111401A (1-3976,4147-4188)	KF724937 22-53AG0211 (3977-4146)	StScan : GMCST
115-PPS-011503A	P1514-VCF-071404A	KF724937 22-53AG0211	StScan : S

Figure 3 Recombination events with the name of isolation are parted in major and minor parents and their recombinants (left panel). And their recombination in this study provided the ladder between the 1st–4188th position (right panel).

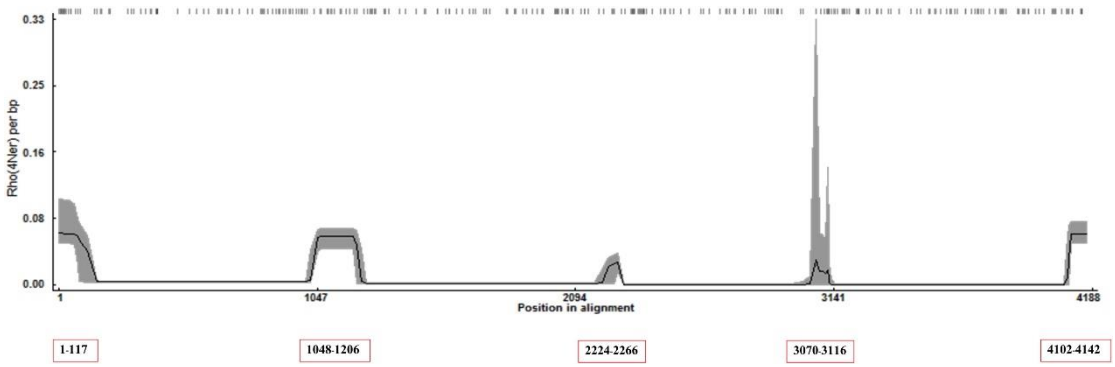


Figure 4 The recombination rate plot is provided with Rho/bp values; the number inside the red box presents the position in alignment related to the plot.

Recombination analysis results of PXX11-SBF-021102A obtained from RDP and CHIMERA are different. The RDP result showed that this sequence is the minor parent, and together with X45RWVCF071202A, they were recombined into SBPED0211-1, whereas the CHIMERA shows that it is the recombinant sequence of X1356SRF121302A and SBPED0211-1. On the other hand, SBPED0211-1 has the highest potential evidence of recombination since RDP, GENECONV and BOOTSCAN, and SiScan can detect it.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Molecular evolutionary analysis

The evolutionary rates were calculated for four datasets of PEDV groups, including overall (all field strains of Thailand; n=115), endemic (predominant TH1 and Th2 subgroups; n=109), TH1 (n=79), and TH2 (n=30). Generally, a higher evolutionary rate is present in Thai PEDV strains than in the US sequences. In comparison among Thai strains, subgroup TH1 (predominant endemic subgroup) had a higher substitution rate than the TH2 subgroup (Figure 5 and Table 3).

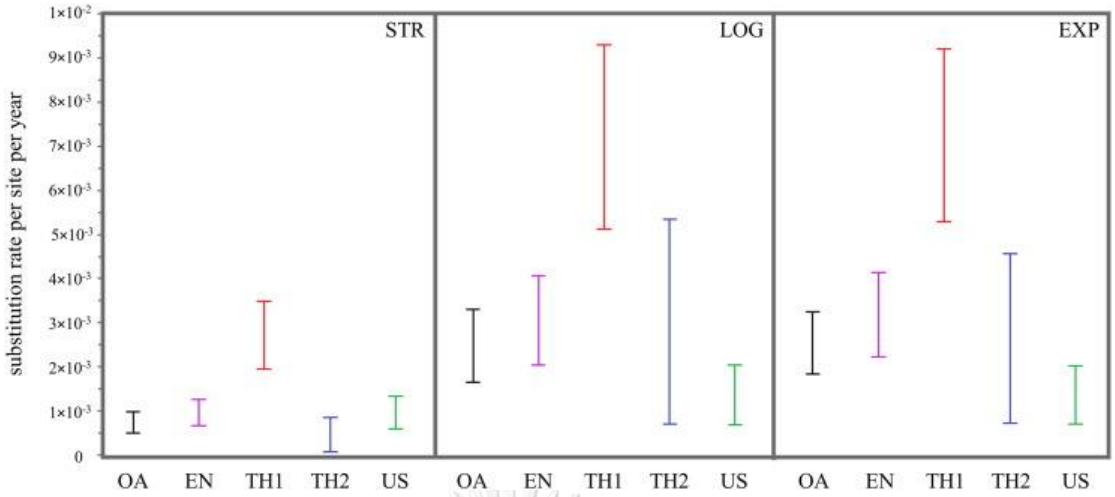


Figure 5 Substitution rate comparison between overall (OA), endemic (EN), TH1, and TH2 of the Thai and US datasets. The results were shown in substitution rate per site per year of strict (STR), uncorrelated relaxed logarithm (LOG), and exponential (EXP) prior clock.

The divergence times among Thai PEDV sequences were also estimated using the BEAST exponential relaxed-clock method. All Thai PEDV sequences share a common ancestor presented in 1999, with a 95% HPD interval of 1987-2004 (Figure 6). The TH2 subgroup evolved independently from all other samples. The emergence of the TH2 common ancestor was estimated to be around 2002, whereas the common ancestor of all other samples arose approximately in 2003. The TH1, TH3, and TH5 subgroups have a shared ancestor, which was presented around 2004. The TH1 subgroup emerged around 2005-2006, and the TH3/TH5 subgroups evolved later in 2007. The TH4 and TH6 subgroups descended from their common ancestor, which emerged recently in 2012.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 3 Substitution rate of each dataset and clock model performed in this study.

Prior clock	Value	Overall	Endemic clusters	TH1	TH2	USA
Strict	mean	7.365E-04	9.574E-04	2.681E-03	4.301E-04	9.501E-04
	SEM	6.868E-06	5.520E-06	6.988E-06	6.206E-06	4.880E-06
	SD	1.258E-04	1.512E-04	3.987E-04	2.206E-04	1.884E-04
	variance	1.583E-08	2.285E-08	1.590E-07	4.867E-08	3.549E-08
	95% HPD min	4.945E-04	6.684E-04	1.951E-03	6.335E-05	5.970E-04
	95% HPD max	9.812E-04	1.256E-03	3.485E-03	8.537E-04	1.336E-03
Exponential	mean	2.558E-03	3.215E-03	7.247E-03	2.425E-03	1.367E-03
	SEM	2.261E-05	2.636E-05	4.886E-05	5.819E-05	1.462E-05
	SD	3.628E-04	4.847E-04	9.900E-04	1.092E-03	3.441E-04
	variance	1.316E-07	2.349E-07	9.800E-07	1.192E-06	1.184E-07
	95% HPD min	1.860E-03	2.253E-03	5.317E-03	7.273E-04	7.141E-04
	95% HPD max	3.274E-03	4.150E-03	9.233E-03	4.572E-03	2.031E-03
Logarithm	mean	2.512E-03	2.242E-03	7.170E-03	3.125E-03	1.336E-03
	SEM	3.369E-05	2.303E-05	6.225E-05	9.146E-05	1.759E-05
	SD	4.235E-04	3.711E-04	1.053E-03	1.298E-03	3.519E-04
	variance	1.794E-07	1.377E-07	1.109E-06	1.685E-06	1.239E-07
	95% HPD Interval	1.667E-03	1.524E-03	5.153E-03	7.257E-04	6.999E-04

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Phylogeographic distribution

Due to the ESS values, the exponential molecular clock model was chosen to display the geographic distribution of PEDV in Thailand, as shown in Figure 7. Based on the phylogeographical analysis, there was a geographic separation between PEDV stains in Thailand. The TH2 subgroup was confined mainly to the western regions. In contrast, the TH1 subgroup was found across Thailand, especially the northeast region of Thailand, and has become the dominant subgroup since 2014.

Figure 6 Phylogenetic tree with time and their estimated divergences of PEDV in Thailand based on spike gene. Red, blue, yellow, aqua, purple, and green colors were represented for TH1, TH2, TH3, TH4, TH5, and TH6, respectively. TH-S-INS labeled parenthesis located the sequences with the unique insertion. Node labels denote the time of estimated divergences, whereas branch labels denote the posterior probability values.

Since its emergence in 2007, PED has been endemic in the western region of Thailand, including Ratchaburi and Nakhon Pathom provinces, and sporadic outbreaks have been reported in the eastern region, including Chonburi, Chanthaburi and Chachoengsao provinces. However, farms in the eastern regions of Thailand experienced PED outbreaks in 2012-2013, and the sequencing results demonstrated the emergence of the TH1 group, which could be responsible for the outbreak. The spread of the TH1 group to the northeast region could be due to the transportation of culled sows.

Figure 7 The geographic distribution of TH1 (left panel), TH2 (center panel), and TH3–6 (right panel) was estimated by the location and time of the outbreaks in Thailand during 2008–2015. Yellow (TH3), green (TH4 and TH5), and greenish blue (TH6) denoted the sporadic cases of the exotic strains.

Discussion and conclusion

In the first study of PEDV, the genetic diversity and evolutionary dynamics of PEDV were conducted using the sample from the first discovered strain in 2008 to the widespread outbreak samples during 2014-2015. Based on the phylogenetic analysis of the complete spike gene, several important findings involving the genetic evolution of PEDV in Thailand were unveiled. The findings included the multiple introductions of genetically distinct PEDV variants influencing the development of Thai PEDV strains. Thai circulating strains forming clades separated from other countries and undergoing high mutation rates, especially the TH1 subgroup during the widespread outbreak between 2014-2015. One mechanism that might involve the high mutation rate was recombination.

The multiple introductions included the presence of PEDV in both genogroup 1 and 2. The introduction of the G1a variant in Thailand was not that surprising since a previous study demonstrated that those PEDV stains in the G1a group were vaccine-like stains (Cheun-Arom et al., 2015). The emergence of PEDV variants in this group might be due to the heavy use of modified live PEDV vaccines, modified live (MLV), and killed virus (KV) vaccines during 2013–2014. MLV and KV in Thailand drastically decreased compared to 2013–2014 due to the efficacy of intramuscular vaccines, which provided a limited degree of success against PEDV. In addition, MLV was reported to cause PEDV outbreaks with mild clinical disease (Cheun-Arom et al., 2015).

The emergence of a US-like variant, as shown in the G2b group (TH6), was indeed intriguing, and this is the first report showing the emergence of a US-like PEDV in Thailand. TH6 is the latest subgroup, which contained the new introduction of the exotic PEDV strain (P1915-NPF-071511A) in 2015 that belongs to the G2b or US-like subgroup.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

This sequence shares 100% similarity with NIG-1/JPN/2014, GNM-2/JPN/2014, MIE-1/JPN/2014, PC21A, KUIDL-PED-2014-002, KUIDL-PED-2014-007, USA/Ohio75/2013, USA/Ohio69/2013, USA/Ohio60/2013, and USA/Indiana/17846/2013, suggesting the first US-like strain in Thailand. However, the ability of this strain to develop into the endemic strain in the Thai swine industry was still questionable and further investigation is required to monitor the development of these exotic strains.

Whether or not multiple introductions influence the strain development of Thai PEDV in the study was unknown. However, although multiple introductions were evident, those strains were not yet forming their clusters. Thai PEDV developed its cluster separated from other countries. PEDV variants that have been dominant in Thai swine farms are strains in the G2a group and can be further divided into two subgroups, including TH1 and TH2. We found that Thailand strains are unique and underwent mutation and recombination with some other strains, resulting in a pandemic between 2014 and 2015. We identified the sequence insertion, which is unique only in Thailand strains (TH-S-INS) and was mainly detected in the TH1 subgroup, particularly the samples in the outbreaks between 2014 and 2015. This insertion was also detected in 2 samples of the TH2 subgroup, which was sampled in 2014, and no sample that belonged to subgroup TH2 was detected in 2015. These suggested that the strains in TH1 might undergo positive selection pressure somehow, and some experienced an insertion mutation event as described in the result. The first sample containing this insertion of TH2 was P1414-SPF-071403A and was collected in July 2014, whereas another 27 samples from 8 outbreaks collected in the same month are TH1. P1314-VCF-071401A is the representative sequence of the samples that contain this insertion of TH1 at 99% identity.

According to the evolutionary rate analysis, the overall group of Thailand sequences (TH1–6; n = 115) evolved with a relatively greater substitution rate than the US sequences when estimating under lognormal and exponential relaxed clock models. The US sequences showed a higher substitution rate under the strict clock model; however, this rate was less than that of the Thailand endemic group (TH1–2; 109). The TH1 subgroup showed a distinctly higher evolutionary rate than other analyzed groups. Its substitution rate was over two-fold compared to the endemic or TH2 group and was over 5-fold compared to the US group as determined by uncorrelated relaxed-clock models. This event may suggest that there is any positive pressure on this subgroup.

Besides being influenced by the multiple introductions, the relatively greater evolutionary rate of Thai PEDV was partially due to recombination, as supported by the results of the recombination analysis. In comparison between TH1 and the TH2 groups, TH1 is more in population parted in the events, whereas TH2 is more in the number of recombination evidence. These results may explain why TH1 recombinants were chosen to be the subgroup suitable to cause an outbreak, whereas TH2 keeps undergoing recombination but does not usually cause an outbreak. TH2 might be the predominant group that could adapt to survive and persist in the host without triggering symptom development; therefore, some were not included in the study. On the other hand, TH2 is not preferred for proliferation in the host. However, we did not perform active surveillance to take the samples from subclinical herds, so we could not conclude. For TH3, we did further analysis with the dataset used in phylogenetic analysis, and there is possible that PXX11-SBF-021102A is the recombinant of NPPED2008 (KC764952) and Chinju99 (AV167585) of G2 and G1 genogroup as major and minor parents, respectively. This information supported why the TH3 subgroup is located between genogroup 1 and 2, and very close and shares the same estimated divergence with the TH5 subgroup.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The results of geographic distribution show that the western part of Thailand (mainly in Ratchaburi) is the endemic area of most strains; however, TH2 seems to be limited in these areas, whereas TH1 is spreading from the west-central-east of Thailand. The pinpoint denoted another new introduction (TH3–6) since they could not perform analysis. These results and previously described suggested that TH1 is becoming the predominant subgroup endemic in Thailand since 2014.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

PHASE II: PHYLOGENETIC STUDY OF PEDV

Part II: The analysis of 2014-2021 PEDV

Phylogenetic tree and analyses

A spike gene-based phylogenetic tree was conducted based on neighbor-joining (Figure 8) to determine Thai PEDV clusters. Up to 2021, the Global pandemic PEDV (G2) could be further classified into three different genogroups, including G2a, G2b, and G2c. According to the previous finding between 2008-2015, TH4 has become one of the predominant clades in Thailand, although the lineage that responded to the outbreak during 2016-2021 is not the Vietnamese strains (CBR1 and CBR2). Interestingly, a new cluster of Thai PEDV was discovered and called TH7, the recombinant strain between the Thai endemic strain and the exotic Chinese strain in genogroup G2b, which is the same as TH4 and located between TH2 and TH4 subgroups.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

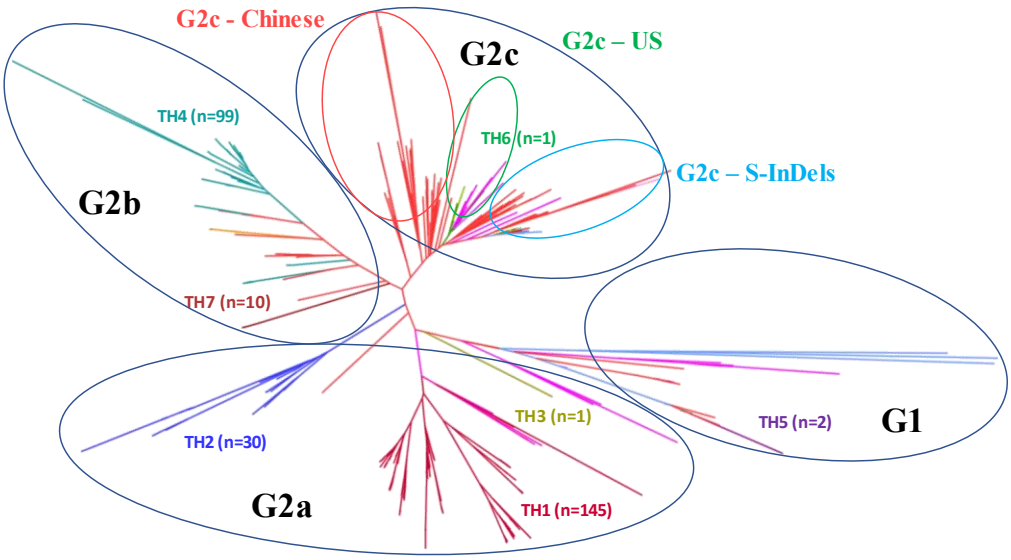


Figure 8 A radial phylogenetic tree was constructed using the Neighbor-joining method to determine the clusters of PEDV in Thailand. The different colors labeled the Thai PEDV clusters on the same color branches, including TH1 (Red), TH2 (blue), TH3 (brownish yellow), TH4 (greenish blue), TH5 (purple), TH6 (green), and TH7 (brownish red).

Recombination analysis

Five events were found to be related to Thai PEDV detected with all available detection methods in RDP, including the recombination between the primitive strain of TH1 and TH2 in 2008, which resulted in TH2 strains during 2011-2014 (Figure 9a). The second event was the recombination between TH1/2008 and TH1/2016, resulting in the TH1 strain responding to the outbreak in 2016 (Figure 9b). The third event was the recombination between TH1 (2016) and TH4 (2018, Chinese), resulting in TH1/TH4 hybrid strains in Thailand during 2016-2021 (Figure 9c), and one of these hybrid strains underwent recombination with a mutated endemic strain (TH1) resulting to the recombinant in TH4 subgroup (Figure 9d). The last event was the recombination between TH4 in 2018 and TH1 in 2015, resulting in the TH4 strain in 2020 (Figure 9e).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

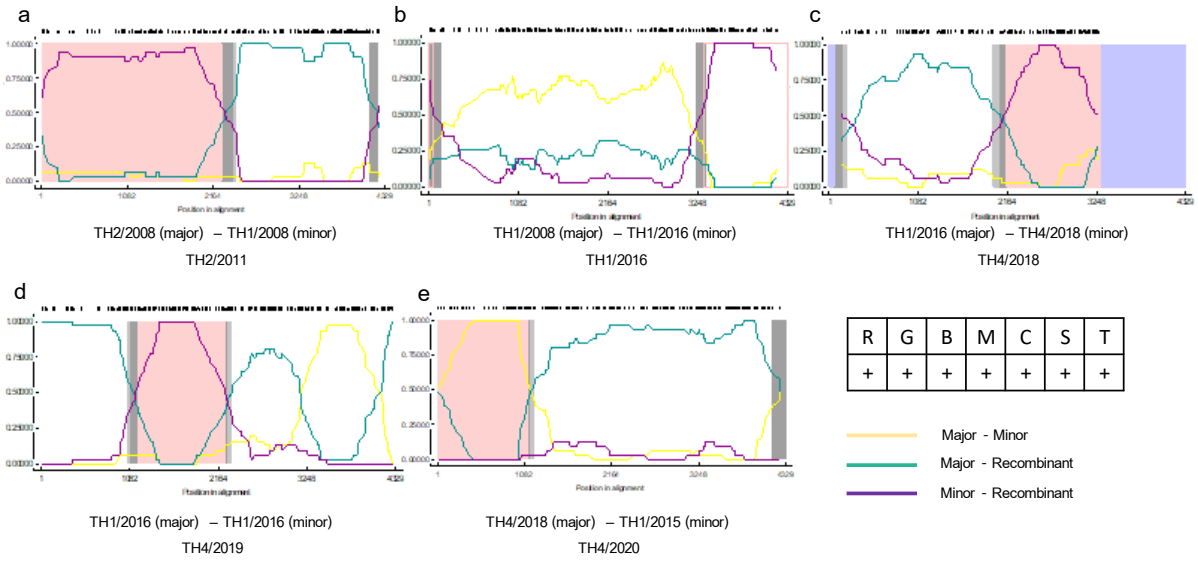
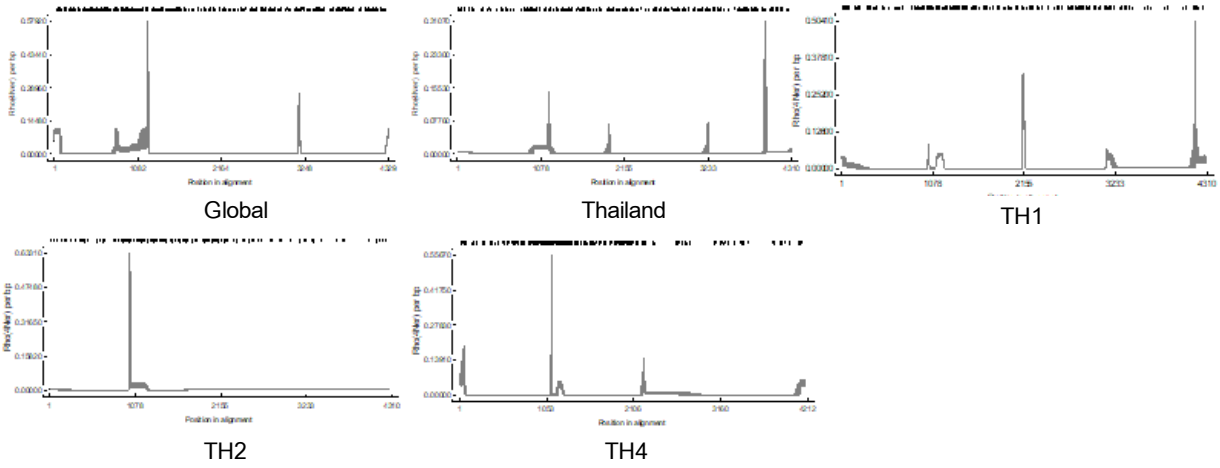


Figure 9 Recombination analysis of Thai PEDV spike gene using RDP4, detectable with all protocol available in the software.

Recombination rate plot analysis demonstrated that the recombination rate between the three predominant strains was similar. However, the global dataset has shown a 2-fold higher recombination rate when compared to those of Thailand subgroups (Figure 10).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



Spike	Global	Thailand	TH1	TH2	TH4
Watterson theta per site	4.684×10^{-2}	2.759×10^{-2}	2.018×10^{-2}	1.882×10^{-2}	2.063×10^{-2}
Average rho per site	1.095×10^{-2}	3.552×10^{-3}	6.479×10^{-3}	5.561×10^{-3}	6.39×10^{-3}
Watterson theta	207.441	121.104	87.882	81.474	90.56
Rho	48.478	15.592	28.66	24.073	28.052
Rho/theta	2.337×10^{-1}	1.287×10^{-1}	3.261×10^{-1}	$2.955 \times \times 10^{-1}$	3.098×10^{-1}

Figure 10 Recombination rate plot of PEDV spike gene examined by LDHat method available in RDP4.

Molecular evolutionary analysis

The chronogram demonstrated that some of the TH4 parted in the recombination event from the previous section is the TH1 lineage which strongly supports that the outbreaks in endemic areas such as Ratchaburi and some parts of Saraburi are affected by the recombinant between the Chinese strain and the endemic strain (TH1) in those areas.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

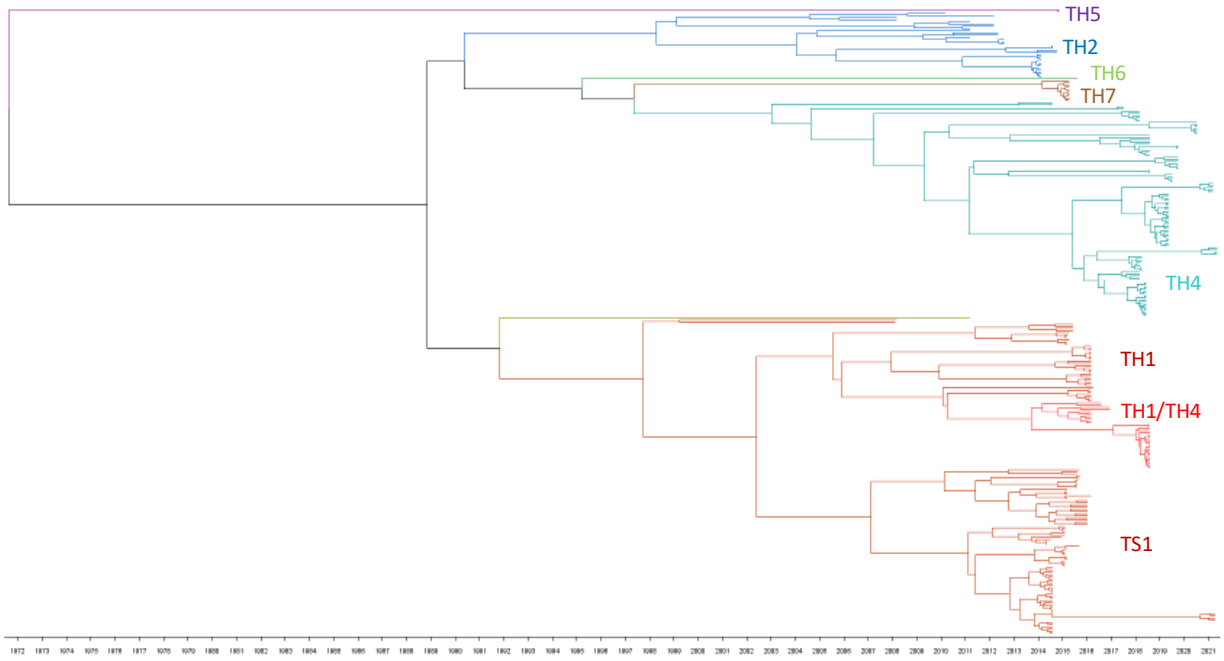


Figure 11 Chronogram of all Thai PEDV including TH1 (Thai endemic strain), TS1 (TH1 with insertion), TH1/TH4 (TH4 recombinant from TH1), TH2 (Thai classical strain), TH3 (recombinant between TH1 and TH2), TH4 (Chinese derived strain), TH5 (G1 strain), TH6, and TH7 (recombinant between TH2 and TH4).

Substitution rate was also conducted from the dataset in this study to compare with previous findings. The TH1 dataset from 2014 to 2021 was lower in substitution rate than the previous study dataset (2008 to 2015), and the TH4 dataset was relatively lower than TH1. Although substitution rates of TH1 and TH4 are quite low, the rates are still higher than that of TH2. Interestingly, the overall dataset is relatively higher than the previous finding (Table 4).

Table 4 Substitution rate of the strict clock model performed in this study, including TH1 (2014 to 2021), TH2 (from previous study), TH4 (2016 to 2020), and overall (all Thai PEDV from 2008 to 2021).

Prior clock	Value	TH1	TH2	TH4	Overall 2008-2021
Strict	mean	8.511×10^{4}	4.301×10^{4}	7.503×10^{4}	1.180×10^{3}
	SEM	3.977×10^{6}	6.206×10^{6}	3.355×10^{6}	6.1696E-6
	SD	1.3418×10^{4}	2.206×10^{4}	1.4277×10^{4}	1.034×10^{4}
	variance	1.800×10^{8}	4.867×10^{8}	2.038×10^{8}	1.070×10^{8}
	95% HPD min	5.930×10^{4}	6.335×10^{5}	4.728×10^{4}	9.845×10^{4}
	95% HPD max	1.114×10^{3}	8.537×10^{4}	1.027×10^{3}	1.395×10^{3}

Discussion and conclusion

Since the new introduction of Chinese PEDV, the PEDV situation has become more complex. The difference between the previous phylogenetic tree, which was established before a large portion of TH4 was discovered in Thailand, is the classification of genogroup two or G2, especially G2b. In this study, the reference sequences were prepared according to the result from the preliminary study using three different methods, including Bayesian analysis using BEAST package, maximum-likelihood (ML) using MEGA-X, and neighbor-joining (NJ) as a result in this study using MEGA-X. Interestingly, by using enough references, all methods showed similar results; conversely, G1 becomes more difficult to classify since there are several exotic strains. The second study here also suggested that a previous well-known phylogenetic tree of the spike gene from several previous studies, including the first-period study here, must be corrected because S-InDels was classified into the wrong clade, G1b. The reason for

this phylogenetic error is due to the missing data to contribute to a good phylogenetic clade since we needed more samples during the widespread outbreak. The pilot study used previously published sequences without the unpublished Thai PEDV retrieved from this study; however, providing enough references shows that both the complete genome and spike gene phylogenetic tree can classify these S-InDels into G2 (Figure 12). Thence, the references from the pilot study were applied to the second-period study to help classify Thai PEDV since using the common NJ and ML methods was difficult to classify Thai PEDV lineages.

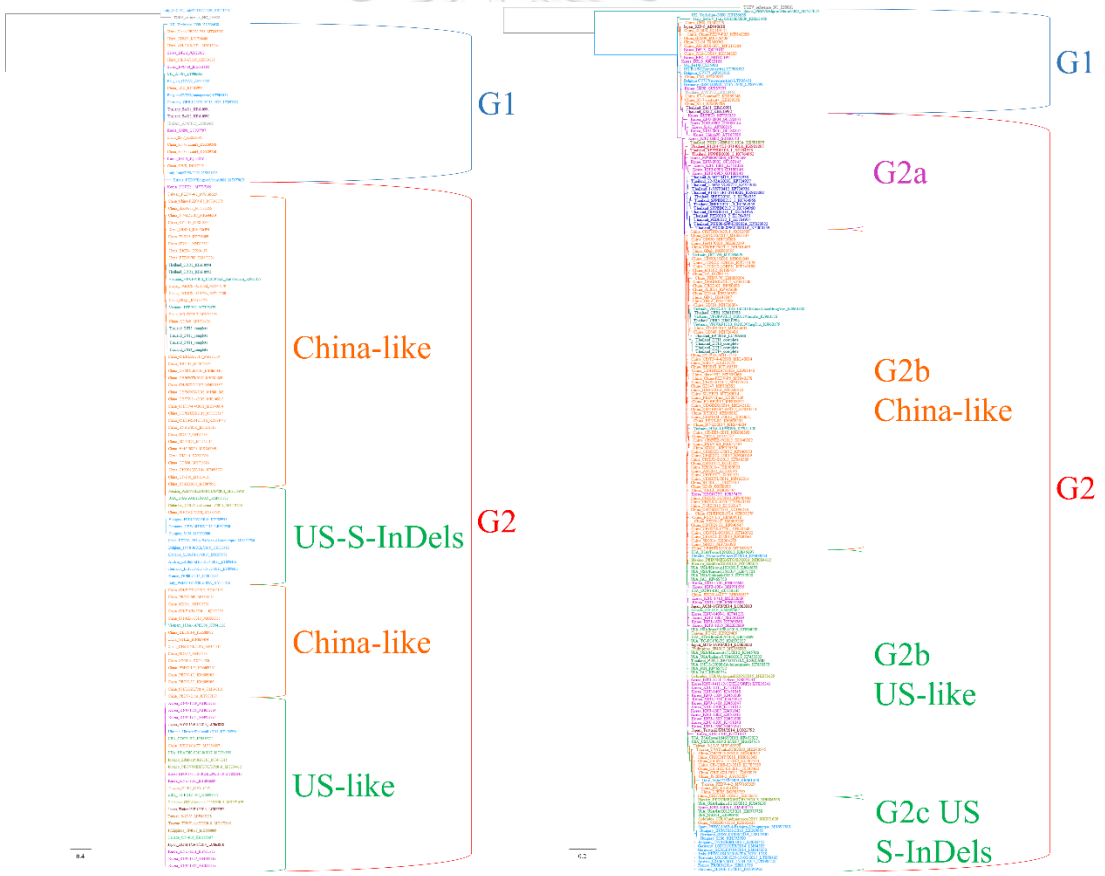


Figure 12 Phylogenetic tree of PEDV worldwide based on the complete genome (left panel) and spike gene (right panel) conducted using maximum-likelihood.

The recombination analysis suggested that several strains in subgroups TH1 and TH4 underwent recombination during 2016-2018, causing sporadic outbreaks during 2019-2021. This finding is unidirectional with the finding of TH2 during 2008-2013; those might be due to the adaptive ability of the virus to survive on the farm or might be due to some selective pressure in the endemic region. Some strains underwent multiple recombinations during 2016-2018 and caused severe damage to the affected herds. With the substitution rate, Thai PEDV tends to decrease in substitution of their genetics, especially in the predominant Thai strain but tend to evolve via recombination mechanism instead. However, the overall substitution rate is higher than before; this might cause the accumulation of different nucleotides in their genetic element, although the amino acids did not change.

From the chronogram phylogenetic tree result, some TH4 strains were classified to TH1 lineage; these TH4 strains were parted in the recombination between TH1 and TH4 in 2018 (Figure 9c). Hypothetically, these strains were also recombined with TH1, resulting in a new strain in TH4. These findings suggested that introducing the exotic strain, especially Chinese PEDV, can be the reason for re-emerging PEDV, particularly in the endemic area such as Ratchaburi.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

One of the new variants in TH4 with the insertion also demonstrated that Thai PEDV underwent multiple recombinations determined using SimPlot (Lole et al., 1999) (Figure 13). The result suggests that these TH4 strains underwent at least two recombination events related to the Chinese strains' GDS53, GDgh, and 7C.

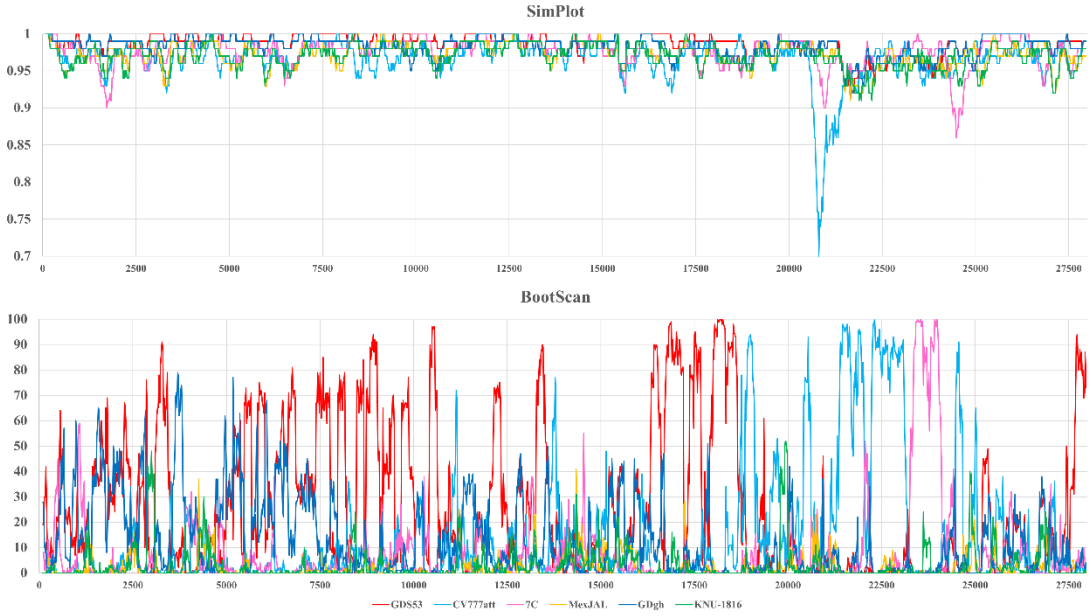


Figure 13 Simplot and BootScan recombination pattern analysis of DTI strain (TH4) complete genome using GDS53 (red), CV777att (light blue), 7C (pink), MexJAL (yellow), GDgh (dark blue), and KNU-1816 (green), as the representative genome of the possible parent.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

CHAPTER 9

PHASE III: PHYLOGENETIC STUDY OF PDCOV

This work has been published in the topic of

The phylodynamics of emerging porcine deltacoronavirus in Southeast Asia

Transboundary and Emerging Diseases

September 2022, Volume 69, Issue 5, Pages: 2816-2827

(Appendix B)

Christopher James Stott, Kanokon Sawattrakool, Kepalee Saeng-Chuto, Angkana Tantituvanont and Dachrit Nilubol

Samples and full-length genome sequencing

The full-length genome sequences of SEA, US, and China PDCoV were 35,411–35,417, 35,430–35,431, and 35,380–35,431 nucleotides (nt) in length, respectively. The genome organization is similar to all previously reported PDCoV genomes: 5'-ORF1a/1b-S-E-M-NS6-N-NS7–3' (Table 5). UTRs are at both ends (5´ UTR, nt 535, and 3´ UTR, nt 725). The lengths of ORF1a/1b, S, E, M, and N genes were 18,788, 3,480-3483, 252, 654, and 1,029 nt, respectively. S gene sequences of SEA, US, and China PDCoV were 3,480-3,483, 3,483, and 3,480-3,483 nucleotides (nt) long, respectively.

Phylogenetic tree and analyses

Phylogenetic tree analysis based on the full-length genome sequences demonstrated that PDCoV isolates were divided into two groups, G1 and G2 (Figure 14). G1 was divided into G1a (China) and G1b (US). G1a and G1b comprised isolates mainly reported in China and the US, respectively. In contrast, isolates mainly reported in SEA countries are clustered in G2 (SEA) group and divided into SEA-1, SEA-2, and SEA-2r. The genetic identities between G1 and G2, based on the full-length genome and spike gene, are shown in PDCoV Table 1. The intralineage nucleotide identities of G2 (SEA), G1a (US), and G1b (China) were 98.00-100.00%, 99.60-100.00%, and 97.30-100.00%, respectively.

Deletions/insertions (InDels) are present in spike and membrane genes. InDels found in the spike gene of each genogroup caused its difference from 3,480 to 3,483 nt, while the InDels found in the membrane (CHN/Sichuan/2017; MK211169) at the 280th-291st position, causing its difference from 642 to 654 nt.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Table 5 Complete genomes and spike genes length, GC contents, identity, and similarity of porcine deltacoronavirus sequences used in this study (Global), Southeast Asia (SEA), the US, and China.

Complete genomes				
	Global	SEA	US	China
Length (nt)	35380-35431	35411-35417	35430-35431	35380-35431
GC contents (%)	43.04-43.40	43.04-43.22	43.15-43.27	43.12-43.37
Intra-lineage identity	96.6-100	98-100	99.6-100	97.3-100
with SEA identity	-	-	97.0-98.1	96.6-98.2
with US identity	-	97.0-98.1	-	98.1-99.3
with China identity	-	96.6-98.2	98.1-99.3	-
Spike genes				
	Global	SEA	US	China
Length (nt)	3480-3483	3480-3483	3483	3480-3483
GC contents (%)	41.31-41.92	41.34-41.83	41.63-41.83	41.31-41.81
Intra-lineage identity (similarity)	94.9-100 (95-100)	96-100 (96.6-100)	99.5-100 (99.2-100)	95.9-100 (95-100)
with SEA identity (similarity)	-	-	95.7-99.4 (96.4-99.2)	95-99.2 (95-99.1)
with US identity (similarity)	-	95.7-99.4 (96.4-99.2)	-	97.3-98.9 (97-99.4)
with China identity (similarity)	-	95-99.2 (95-99.1)	97.3-98.9 (97-99.4)	-

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

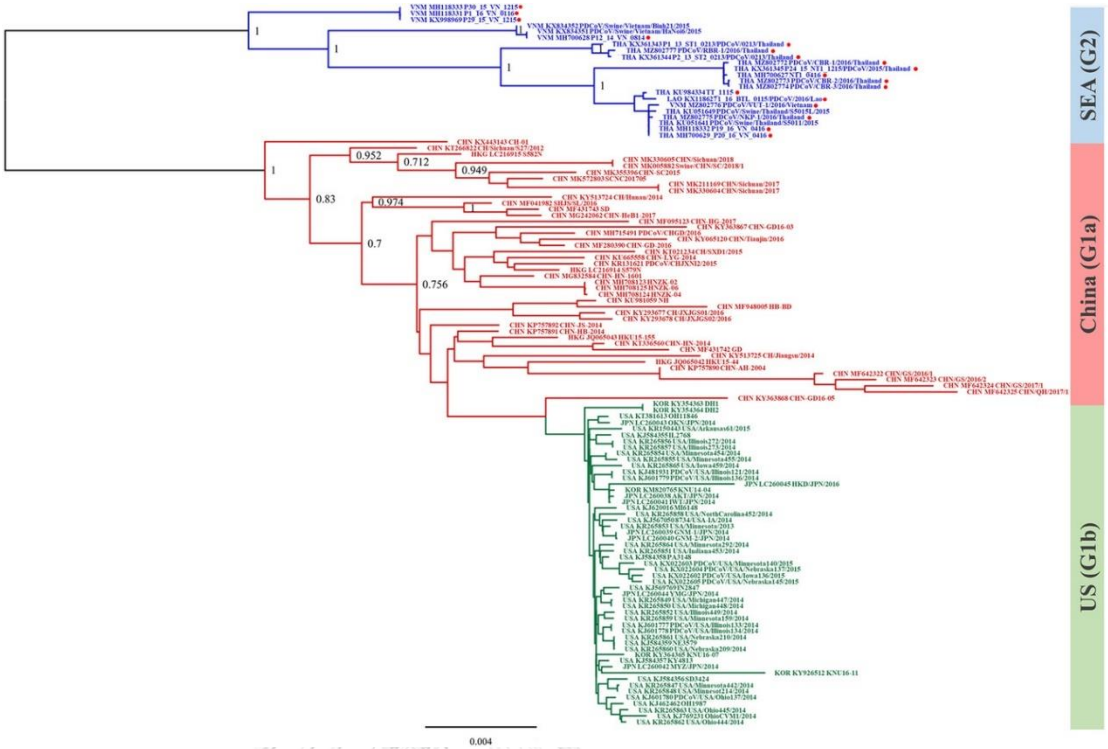


Figure 14 The phylogenetic tree of PDCoV conducted by maximum likelihood in MEGA-X shows that the virus could be classified into three genogroups: SEA (blue), China (red), and the US (green). The number at the nodes denotes the probability of performing by 1000 bootstraps.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Recombination analysis

Recombination analysis based on the full-length genome, as determined by at least six detection methods available in RDP4, demonstrated that five isolates in SEA PDCoV were recombinants. These isolates include PDCoV/Binh21/2015 (KX834352) and PDCoV/HaNoi6/2015 (KX834351) and P30_15_VN1215 (MH118332), P2_13_ST2_0313 (KX361344) and PDCoV/CBR-3/2016 from Vietnam and Thailand, respectively (Figure 15). High potential events from Vietnam are PDCoV/Binh21/2015 and P30_15_VN_1215,

which are recombinant between China PDCoV, CHN-GD16–03 (KY363867), and CHN/Tianjin/2016 (KY065120), and between Vietnam and US PDCoV, P12_14_VN_0814 (MH700628) and USA/Minnesota159/2014 (KR265859). In contrast, Thai PDCoV P2_13_ST2_0313 is the recombinant between Hong Kong HKU15–155 (JQ065043) and China PDCoV CH/Sichuan/S27/2012 (KT266822).

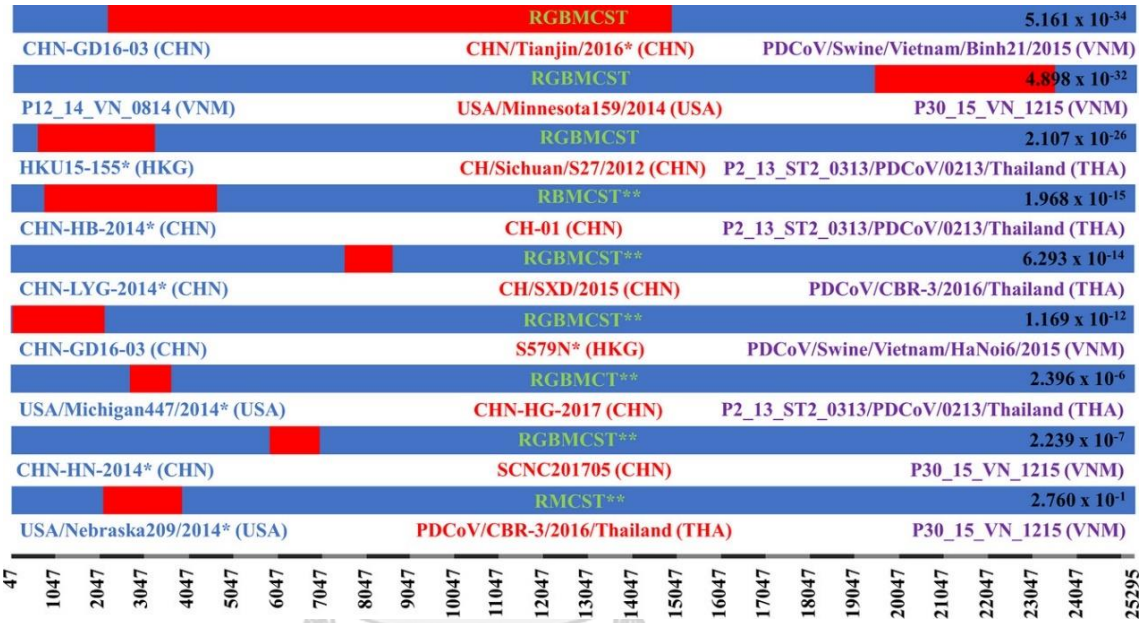
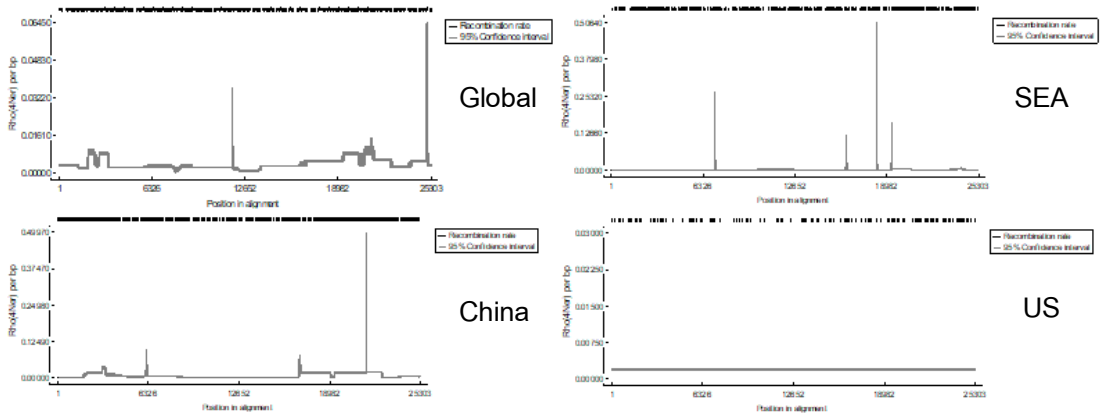


Figure 15 Recombination pattern of PDCoV full-length genome determination by RDP4 with at least 6 of 7 detection tools in the program. Blue shows the major parent, red shows the minor parent, purple shows the recombinant parent, and green shows the initial character of the detection tool available in RDP4 (RDP (R), GENECOV (G), BOOTSCAN (B), MAXCHI (M), CHIMAERA (C), SISCAN (S), and 3SEQ (T)). The gray/black ruler shows the length of the analyzed sequences. The average P-value from RDP is displayed in the right panel, and the asterisks denote the partial parent.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The intralineage recombination rate (theta) based on the full-length genomes of SEA, China, and the US was estimated at 8.784 x 10^{-3}, 1.993 x 10^{-2}, and 3.632 x 10^{-3}, respectively. The recombination rates based on spike genes of SEA and China were estimated at 2.065 x 10^{-2} and 2.715 x 10^{-2}, respectively. (Figures 16 and 17).



Complete genome	Global	SEA	China	US
Watterson theta per site	2.197 x 10^{-2}	8.784 x 10^{-3}	1.993 x 10^{-2}	3.632 x 10^{-3}
Average rho per site	3.647 x 10^{-3}	6.572 x 10^{-4}	7.841 x 10^{-3}	1.82 x 10^{-3}
Watterson theta	558.198	223.029	506.247	93.06
Rho	92.693	16.687	199.2	46.633
Rho/theta	1.661 x 10^{-1}	7.482 x 10^{-2}	3.9348 x 10^{-1}	5.011 x 10^{-1}

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

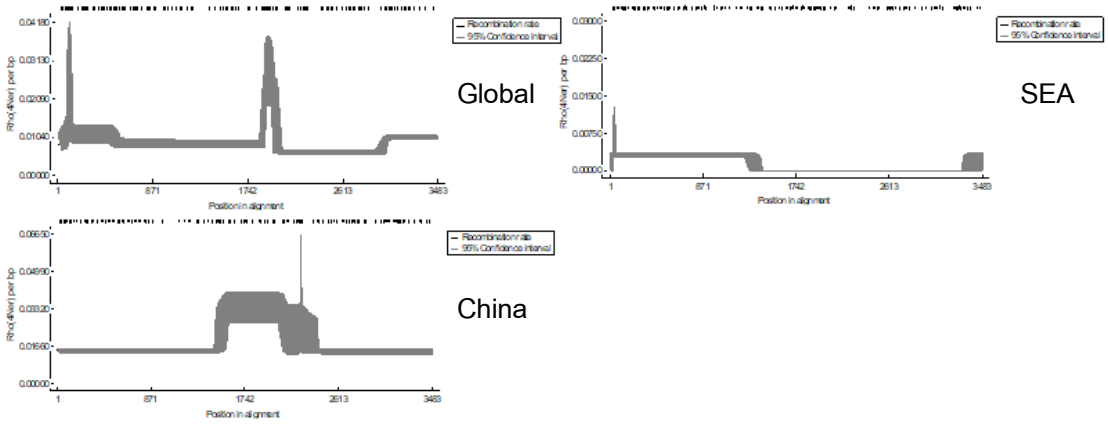
Figure 16 Recombination rate plot of complete genomes examined by LDHat method available in RDP4.

Molecular evolutionary analysis

Based on the full-length genome, the evolutionary analysis demonstrated that PDCoV could be grouped into two genogroups, G1 and G2 (Figure 18). G1 further evolved into two subgroups, including G1a (China) and G1b (US). In contrast, G2 further evolved into two subgroups, including G2a (SEA-1 or Thailand) and G2b (SEA-2 or Vietnam). The G2a (SEA-1 or Thailand) subgroup was further divided into eastern and western groups. Notably, recombinants from Vietnam clustered separately in a new subgroup, SEA-2r.

TMRCAs of G1 and G2 were estimated at 1989.17 with a 95% HPD range from 1986.81-1991.50, whereas the divergence of G2 was estimated at 1999.62 (1997.86-2001.32; 95% HPD) and diverged to SEA-1 and SEA-2 and SEA-2r clades. The divergence of the SEA-2 and SEA-2r clades was estimated at 2002.88 (2001.40–2004.56; 95% HPD), and the divergence of the Thailand subgroup was estimated at 2006.87 (2005.78–2007.91; 95% HPD).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



Spike	Global	SEA	China	US
Watterson theta per site	2.965 x 10^{-2}	1.803 x 10^{-2}	2.715 x 10^{-2}	NA
Average rho per site	9.065 x 10^{-3}	1.354 x 10^{-3}	1.827 x 10^{-2}	NA
Watterson theta	107.794	64.191	96.762	NA
Rho	32.952	4.82	65.111	NA
Rho/theta	3.057 x 10^{-1}	7.509 x 10^{-2}	6.729 x 10^{-1}	NA

Figure 17 Recombination rate plot of spike genes examined by LDHat method available in RDP4.

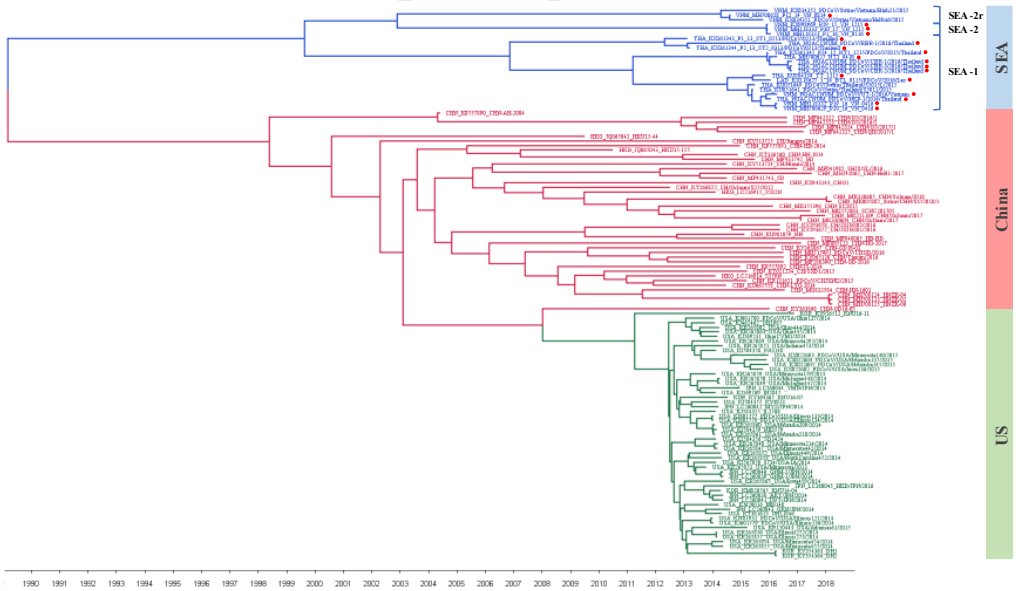


Figure 18 Phylogenetic trees of the full-length genome of PDCoV generated using BEAST. PDCoV could be clustered into three clades, the SEA (blue), China (red), and the US (green). Nodes denote the posterior probability. Red dots indicate the sequences in this study.

Interestingly, Vietnam strains existed in all subgroups. The first strain was in SEA-1, sharing a monophyletic group with Thailand and Lao PDR strains. The second subgroup was SEA-2, which shared the same ancestor with SEA-1, whereas SEA-2r formed a paraphyletic group from SEA-1 and SEA-2.

Evolutionary rates of PDCoV are shown in Figure 19. Based on the full-length genome, the substitution rate was 6.13 x 10^{-4} substitutions/site/year (s/s/y). In the comparison of the encoding genes, the envelope was the highest (1.28 x 10^{-3} s/s/y), followed by NS7 (9.93 x 10^{-4} s/s/y), nucleocapsid (8.26 x 10^{-4} s/s/y), NS6 (8.19 x 10^{-4} s/s/y), spike (7.59 x 10^{-4} s/s/y), and membrane (4.94 x 10^{-4} s/s/y).

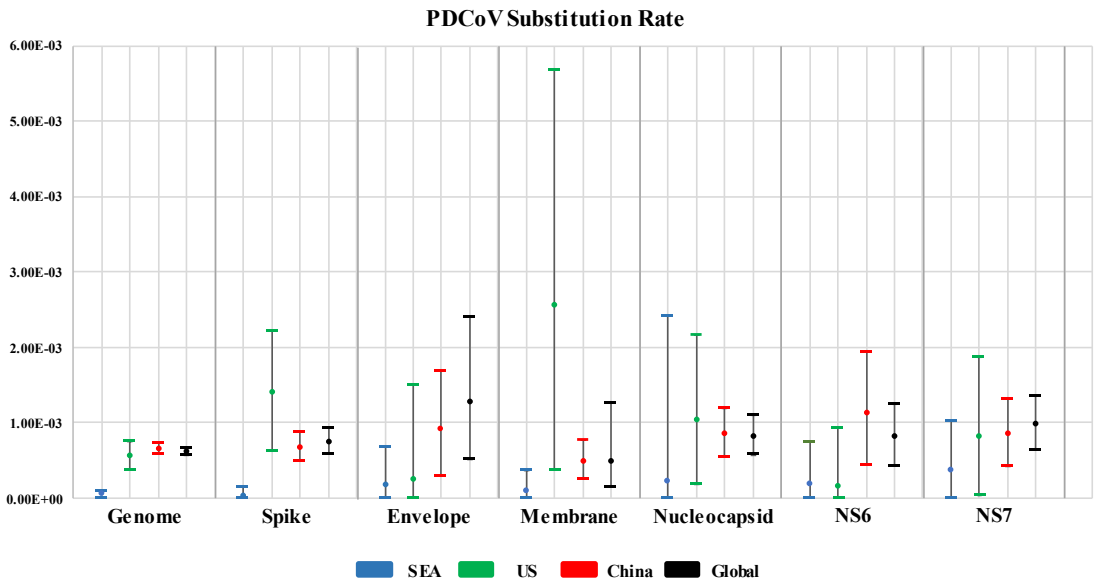


Figure 19 Substitution rate of global PDCoV in this study of a full-length genome (genome) in comparison with the structural genes (spike, envelope, membrane, and nucleocapsid) and the accessory genes (NS6 and NS7) determined using BEAST and interpreted by using TRACER 1.6. Dot located the mean; upper and lower lines located the upper and lower borders of the dataset.

For the individual datasets, the China dataset had the highest substitution rate of the full-length genome (6.57 x 10-4 s/s/y), envelope (9.23 x 10-4 s/s/y), NS6 (1.13 x 10-3 s/s/y), and NS7 (8.55 x 10-4 s/s/y), while the US dataset had the highest substitution rates of spikes (1.41 x 10-3 s/s/y), membranes (2.57 x 10-3 s/s/y), and nucleocapsids (1.04 x 10-4 s/s/y). Interestingly, the NS7 substitution rate of the SEA dataset was the highest compared to its other genes.

Positive selection, and InDels analysis

Structural protein genes including S, E, M, and N, together with NS6 and NS7, were separately analyzed. Positive selection pressure (dN/dS > 1) was found mainly in NS6 and NS7 genes (APPENDIX B, Supplementary material 7). In spike gene, 28 (2.41%), 28 (2.41%), and 14 (1.21%) positively selected codons were detected in the global, SEA, and China datasets, respectively. The cumulative dN/dS of the global, SEA, US, and China codons were 0.38, 0.38, 0.62 and 0.43, respectively, whereas 329 (28.34%), 328 (28.25%), 39 (3.36%), and 256 (22.05%) codons were under negative pressure, respectively (APPENDIX B, Supplementary material 7a). It is noteworthy that although the S gene of SEA PDCoV is under negative pressure, spike codon sites are under highly positive selection pressure, especially at position 642nd. In the E gene, one (1.19%) positively selected codon was detected in the global dataset.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

The cumulative dN/dS of the global, SEA, US, and China codons were 0.21, 0, 0 and 0.36, respectively, whereas 20 (23.81%), 2 (2.38%), 4 (4.76%), and 18 (21.43%) codons were under negative pressure (APPENDIX B, Supplementary material 7b). In the M gene, two (0.92%) and one (0.46%) codon were detected to be positively selected in the global and Chinese datasets, respectively. The cumulative dN/dS of the global, SEA, US, and China codons were 0.11, 0.19, 0.08 and 0.23, respectively, whereas 49

(22.48%), 14 (6.42%), 8 (3.67%), and 32 (14.68%) codons were under negative pressure, respectively (APPENDIX B, Supplementary material 7c). In the N gene, 7 (2.04%) positively selected codons were found in the global dataset. The cumulative dN/dS of the global, SEA, US, and China codons were 0.25, 0.27, 0.41, and 0.41, respectively, whereas 82 (23.91%), 36 (10.50%), 14 (10.08%), and 62 (18.08%) codons were under negative pressure, respectively (APPENDIX B, Supplementary material 7d). In contrast to structural protein genes, NS6 and NS7 genes were under positive pressure. In NS6, one (1.05%) positively selected codon was found in the global and Chinese datasets. The cumulative dN/dS of the global, SEA, US, and China had values of 0.97, 0.83, 7.8, and 1.05, respectively. whereas 19, 4, 1, and 12 codons were under negative pressure, respectively (APPENDIX B, Supplementary material 7e). In NS7, four (1.99%), two (1%), and one (0.49%) positively selected codon were detected in the global, China, and SEA datasets, respectively. The cumulative dN/dS were higher than one in all datasets.

Although the US dataset has no poorly selected codon at any site, the cumulative dN/dS of the global, SEA, US, and China datasets were 6.28, 5.55, 2.35, and 8.26, respectively. The negatively selected codon of NS7 tended to be the lowest when compared to other genes, which were detected in 11 (5.47%), 5 (2.49%), 4 (1.99%), and 5 (2.49%) sites of codons in the global, SEA, US, and China datasets, respectively (APPENDIX B, Supplementary material 7f).

InDels were found in the spike gene of the global, SEA, and China datasets and the membrane gene in the global and China datasets (APPENDIX B, Supplementary material 7g). The cumulative InDels of the spike genes of the global, SEA, and China were 1.12, 1.12, and 1.93, respectively, while the cumulative InDels of the membrane genes globally and in China were 0.08 and 0.25, respectively. Interestingly, although the cumulative dN/dS of the present study demonstrated that only NS6 and NS7 gene of the

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

US and China datasets, and the global, SEA, US, and China datasets, respectively, were under positive pressure, several sites of structural protein genes were positively selected especially in the global dataset.

Phylogeographic analysis

The results estimated that TMRCA of PDCoV circulated in the southern part of China close to the border of Vietnam during 1990, and its descendant approached SEA in 2000 and became divergent from the SEA strains. The virus was then introduced to Thailand during 2007–2011 prior to the divergences of eastern Thai (SEA pandemic) and western Thai in 2011 and 2012, respectively; therefore, the eastern subcluster became the hotspot of PDCoV in Thailand, which distributed the virus to Lao PDR and Vietnam in 2015 (Figure 20 and 21).

The Vietnam strains were part of 3 clades, and the cluster in SEA-1 shared the monophyletic group as those of the Thailand cluster. The cluster in SEA-2 was found only in Vietnam and was the most primitive cluster of Vietnam, and the cluster in SEA-2r found in Vung Tau was in a paraphyletic group that separated from the SEA-1 and SEA-2 clusters.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

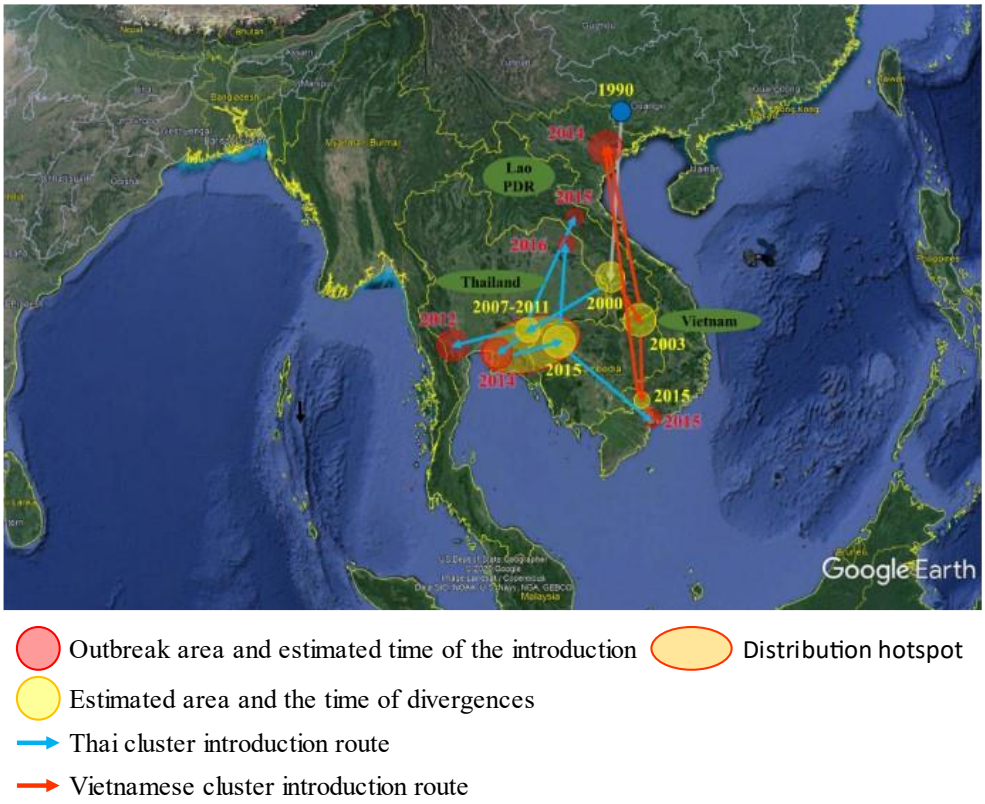


Figure 20 The phylogeographic distribution of PDCoV in this study shows the estimated times of the divergences that occurred in Southeast Asia (SEA).

The Vietnam strains were part of 3 clades, and the cluster in SEA-1 shared the monophyletic group as those of the Thailand cluster. The cluster in SEA-2 was found only in Vietnam and was the most primitive cluster of Vietnam, and the cluster in SEA-2r found in Vung Tau was in a paraphyletic group that separated from the SEA-1 and SEA-2 clusters.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Figure 21 Phylogeographic trees of PDCoV were generated using BEAST. PDCoV could be geographically clustered into three lineages, the SEA (blue), China (red), and the US (green). Nodes denote the year of estimated divergences, node bars denote the 95% high population density, and branches denote the posterior probability. Red dots indicate the sequences in this study.

Discussion and conclusion

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

The more complex evolution of SEA PDCoV compared to that of the US and China PDCoV has been interesting. Questions have been raised about the origin of the virus and what factors influence the evolution of SEA PDCoV. Therefore, phylodynamic analyses were conducted to hypothetically fulfill the missing information. Based on phylogenetic and evolutionary analyses, the results demonstrated that PDCoV isolates evolved into two separated genogroups, G1 (China and US) and G2 (SEA), sharing a common ancestor. G1 is further evolved into 2 clusters, including G1a (China) and G1b (US). G2 (SEA) is further divided into Thailand (G2a or SEA-1) and Vietnam (G2b or SEA-2) clusters. From the study, SEA PDCoV hypothetically descended from the same

ancestor as China strains, and the common ancestor of SEA PDCoV was hypothetically introduced to Vietnam, although the most primitive strain of SEA was found in Thailand. It is noteworthy that E and N genes had higher substitution rates than the S gene.

Furthermore, compared to China and SEA PDCoV, the substitution rate of US PDCoV was highest, whereas the recombination rate was lowest. In contrast with the US PDCoV, the recombination rate of SEA PDCoV and China was higher than that of the US PDCoV, whereas the substitution rate of SEA PDCoV was lower. The recombination rate of SEA PDCoV compared to US PDCoV, which lacked external introduction, was supported by recombination events in both Thailand and Vietnam clusters of the SEA genogroup and Vietnam recombinants formed a new subcluster (SEA-2r). Interestingly, the positively selected sites were detected in the NS7 gene; this could suggest that the external introduction of genetically distinct PDCoV could influence the genetic development of SEA PDCoV.

Based on the results of the present study, it is speculated that PDCoV originally circulated in China, in the southern region close to the Vietnam border, a decade before its first introduction into the Southeast Asia region via Vietnam. The results also suggest multiple introductions of genetically distinct PDCoV to Vietnam. The first and second introductions were estimated in 2003 and 2015, respectively. The second introduction simultaneously occurred in both Vietnam and Thailand, which was estimated around 2007-2011 in Thailand before the introduction to Vietnam in 2015. This introduction results in 3 and 2 different subclusters in Vietnam and Thailand. Thailand PDCoV could be geographically clustered into two subclusters: eastern and western. Isolates of the western subcluster found in 2013 were the most primitive strain. However, the recently emerged eastern subcluster (SEA-1) found since 2015 has been the predominant subcluster in Thailand and SEA PDCoV, including Lao PDR and Vietnam strains.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Interestingly, the emergence of the SEA-1 PDCoV, estimated during 2007-2015, created a hotspot of viral distribution throughout the SEA.

The substitution rates of the spike gene in this study were higher when compared to the antigenic site of a partial gene in a previous study of other coronaviruses, such as TGEV (Enjuanes et al., 1992) and IBV (Cavanagh et al., 1998), and the S gene of BCoV (Vijgen et al., 2005) but were lower than that of the PEDV S gene (Stott et al., 2017). Compared to SARS (Zhao et al., 2004), the genomic sequence substitution rate of PDCoV in this study was lower. This information suggests that the PDCoV genome and spike gene did not evolve rapidly compared to those of coronaviruses such as SARS and PEDV. Agreeing with the substitution rates of the envelope gene and nucleocapsid, those the highest in this study were reminded of SARS that the envelope, membrane, and nucleocapsid genes were the highest compared to other genes. These findings oppose the nature of coronavirus, in which these genes are usually conserved (Woo et al., 2010). Envelope and membrane proteins are needed to assemble the virus and can induce the activity of alpha interferon (Baudoux et al., 1998; de Haan et al., 1998). Mutation of the envelope gene in other coronaviruses attenuates virulence (Regla-Nava et al., 2015), and the envelope gene of most coronaviruses also plays an important role in host interactions, especially pathogenic inflammatory responses (DeDiego et al., 2014).

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

In this study, the substitution rates of the full-length genome, envelope, NS6, and NS7 of China were higher than those of SEA and US, whereas most of the structural encoding genes of the US, except the envelope, were higher than others. These findings suggest that US PDCoV was under higher pressure through structural genes than PDCoV in SEA and China. Interestingly, China PDCoV tends to have a higher substitution rate of nonstructural encoding genes, suggesting that external introduction

might influence structural protein changes in China PDCoV, whereas the substitution mechanism seems important for hybrid viruses in fields, which results in these strains lacking a substitution rate. Compared with the US and China datasets, although SEA PDCoV lacked a substitution rate, the NS6 substitution rate tended to be high. From a previous study, NS6 of PDCoV was found to be an IFN-β antagonist via interaction with RIG-I/MDA5, resulting in a reduction in RLR-mediated IFN-β production (Fang et al., 2018), while NS7 showed the ability to downregulate α-actinin-4 (ACTN4) and mitochondrial carbamoyl phosphate synthase 1 (CPS1) (Choi and Lee, 2019). In our study, the phenomenon of the accessory gene mutation appeared to be important in China, and SEA PDCoV.

Although the substitution rate of PDCoV genomic sequences was lower than that of many other coronaviruses, PDCoV was undergoing their mutation through these genes (envelope and nucleocapsid). Recent studies in calves and chickens have shown that the virus is not yet well adapted to the host (Jung et al., 2017; Liang et al., 2019). These findings relate to the phenomenon of the high mutation rate of the envelope gene and nucleocapsid.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

The recombination result shows that the SEA parental strains descended from the ancestor of the strains from China, such as Guang Dong or Si Chuan, and the strains from Hong Kong. The strains found in SEA before 2014 lacked information; however, the spike genes of several isolates in Vietnam were close to the Guang Dong strain (APPENDIX B, Supplementary material 8). The recombination results also correlated with a previous report (Saeng-Chuto et al., 2020) suggesting P29_15_VN_1215 (KX998969), P30_15_VN_1215 (MH118331), and P1_16_VN_0116 (MH118331) of Vietnam were the recombinant between the primitive strain of Vietnam with the external introduction of the

US strain and these strains were virulence compared to the primitive strain which not triggering severe clinical signs.

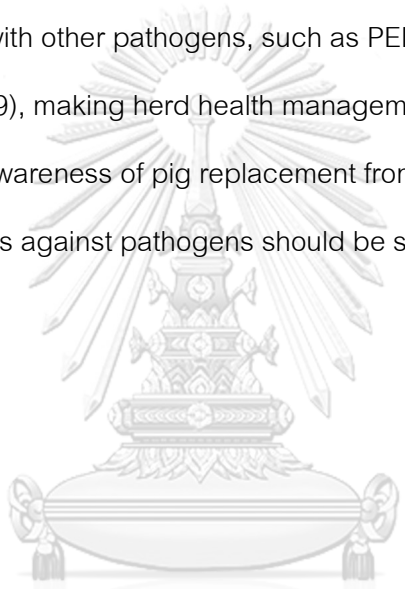
The recombination rate of the genome and spike gene showed the high potential of recombination in China and SEA; however, there is a lack of recombination rates in the US. Although the recombination rate of the US was very low, the substitution of the US was the highest, in contrast to that of the SEA, which shows a higher recombination rate but lacks the substitution rate, while China has a high potential of mutation via two mechanisms. The SEA results suggest that the influence of external introduction from other countries was the primary mechanism of the virus mutation. Unfortunately, we did not study the potential of the substitution hotspot caused by the recombination event that occurred in the China strain.

From a recent study (He et al., 2020), the substitution rate of the PDCoV spike gene was estimated at 1.7 x 10–3 s/s/y, while the substitution rates of Thailand, China (early strains excluded), and the US were estimated at 2.71, 1.58, and 1.21 (x 10–3 s/s/y), respectively. The estimated value of the US dataset was similar, although the substitution rate of others tended to be higher than that in our study. The reason for this difference might be affected by the difference in dataset preparation; we aimed to compare the entire SEA dataset with the China and US subgroups and the different sequences and substitution models used in this study.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

According to the results of this study, the SEA PDCoV shall be derived from China, and the SEA strains were close to that of the Guang Dong strain. The SEA strains were geographically distinct in several parts of Thailand and Vietnam. There was also influence by the multiple introductions from the US and China PDCoV throughout the SEA, especially in Vietnam. However, we still lack information on the nearest divergence

of PDCoV in SEA, and these strains are undergoing recombination events and positive selection pressure making the virus geographically distinctly genetically driven and causing the situation of the outbreak complicated since several studies raising concern about the high pathogenic PDCoV, determined by triggering the severe clinical signs by the virus itself (Jung et al., 2015; Saeng-Chuto et al., 2017; Xu et al., 2018). In this study, the SEA strains from both Thailand and Vietnam showed many properties related to a previous report of highly pathogenic PDCoV from Tianjin and Guangdong (Xu et al., 2018), and coinfection with other pathogens, such as PEDV, was also reported in recent years (Zhang et al., 2019), making herd health management difficult to eliminate the pathogen from herds. Awareness of pig replacement from external sources, quarantine, and preventive strategies against pathogens should be successfully achieved.



จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ 1990. Basic local alignment search tool. J Mol Biol. 215(3): 403-410.

Auton A and McVean G 2007. Recombination rate estimation in the presence of hotspots. Genome research. 17(8): 1219-1227.

Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, Huelsenbeck JP, Ronquist F, Swofford DL, Cummings MP, Rambaut A and Suchard MA 2012. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. Systematic biology. 61(1): 170-173.

Banerjee A, Rapin N, Miller M, Griebel P, Zhou Y, Munster V and Misra V 2016. Generation and Characterization of Eptesicus fuscus (Big brown bat) kidney cell lines immortalized using the Myotis polyomavirus large T-antigen. Journal of Virological Methods. 237: 166-173.

Baudoux P, Carrat C, Besnardeau L, Charley B and Laude H 1998. Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes. J Virol. 72(11): 8636-8643.

Bielejec F, Rambaut A, Suchard MA and Lemey P 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics. 27(20): 2910-2912.

Boni MF, Posada D and Feldman MW 2007. An Exact Nonparametric Method for Inferring Mosaic Structure in Sequence Triplets. Genetics. 176(2): 1035-1047.

Boniotti MB, Papetti A, Lavazza A, Alborali G, Sozzi E, Chiapponi C, Faccini S, Bonilauri P, Cordioli P and Marthaler D 2016. Porcine Epidemic Diarrhea Virus and Discovery of a Recombinant Swine Enteric Coronavirus, Italy. Emerg Infect Dis. 22(1): 83-87.

Bosch BJ, van der Zee R, de Haan CA and Rottier PJ 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J Virol. 77(16): 8801-8811.

Carvajal A, Lanza I, Diego R, Rubio P and Carmenes P 1995. Evaluation of a blocking ELISA using monoclonal antibodies for the detection of porcine epidemic diarrhea virus and its antibodies. J Vet Diagn Invest. 7(1): 60-64.

Cavanagh D, Mawditt K, Adzhar A, Gough RE, Picault JP, Naylor CJ, Haydon D, Shaw K and

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Britton P 1998. Does IBV change slowly despite the capacity of the spike protein to vary greatly? Adv Exp Med Biol. 440: 729-734.

Chang SH, Bae JL, Kang TJ, Kim J, Chung GH, Lim CW, Laude H, Yang MS and Jang YS 2002. Identification of the epitope region capable of inducing neutralizing antibodies against the porcine epidemic diarrhea virus. Mol Cells. 14(2): 295-299.

Chasey D and Cartwright S 1978. Virus-like particles associated with porcine epidemic diarrhoea. Research in veterinary science. 25(2): 255.

Chen J, Wang C, Shi H, Qiu H, Liu S, Chen X, Zhang Z and Feng L 2010. Molecular epidemiology of porcine epidemic diarrhea virus in China. Arch Virol. 155(9): 1471-1476.

Chen JF, Sun DB, Wang CB, Shi HY, Cui XC, Liu SW, Qiu HJ and Feng L 2008. Molecular characterization and phylogenetic analysis of membrane protein genes of porcine epidemic diarrhea virus isolates in China. Virus Genes. 36(2): 355-364.

Chen L, Liu B, Yang J and Jin Q 2014a. DBatVir: the database of bat-associated viruses. Database : the journal of biological databases and curation. 2014: bau021-bau021.

Chen Q, Gauger P, Stafne M, Thomas J, Arruda P, Burrough E, Madson D, Brodie J, Magstadt D, Derscheid R, Welch M and Zhang J 2015. Pathogenicity and pathogenesis of a United States porcine deltacoronavirus cell culture isolate in 5-day-old neonatal piglets. Virology. 482: 51-59.

Chen Q, Li G, Stasko J, Thomas JT, Stensland WR, Pillatzki AE, Gauger PC, Schwartz KJ, Madson D, Yoon KJ, Stevenson GW, Burrough ER, Harmon KM, Main RG and Zhang J 2014b. Isolation and characterization of porcine epidemic diarrhea viruses associated with the 2013 disease outbreak among swine in the United States. J Clin Microbiol. 52(1): 234-243.

Chen Y-N, Hsu H-C, Wang S-W, Lien H-C, Lu H-T and Peng S-K 2019. Entry of Scotophilus Bat Coronavirus-512 and Severe Acute Respiratory Syndrome Coronavirus in Human and Multiple Animal Cells. Pathogens (Basel, Switzerland). 8(4): 259.

Cheun-Arom T, Temeeyasen G, Srijangwad A, Tripipat T, Sangmalee S, Vui DT, Chuanasa T, Tantituvanont A and Nilubol D 2015. Complete Genome Sequences of Two Genetically Distinct Variants of Porcine Epidemic Diarrhea Virus in the Eastern Region of Thailand. Genome Announc. 3(3).

Choi S and Lee C 2019. Functional Characterization and Proteomic Analysis of Porcine

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Deltacoronavirus Accessory Protein NS7. J Microbiol Biotechnol. 29(11): 1817-1829.

Cruz DJ, Kim CJ and Shin HJ 2006. Phage-displayed peptides having antigenic similarities with porcine epidemic diarrhea virus (PEDV) neutralizing epitopes. Virology. 354(1): 28-34.

Cruz DJ, Kim CJ and Shin HJ 2008. The GPRLQPY motif located at the carboxy-terminal of the spike protein induces antibodies that neutralize Porcine epidemic diarrhea virus. Virus Res. 132(1-2): 192-196.

de Arriba ML, Carvajal A, Pozo J and Rubio P 2002. Isotype-specific antibody-secreting cells in systemic and mucosal associated lymphoid tissues and antibody responses in serum of conventional pigs inoculated with PEDV. Vet Immunol Immunopathol. 84(1-2): 1-16.

de Haan CA, Kuo L, Masters PS, Vennema H and Rottier PJ 1998. Coronavirus particle assembly: primary structure requirements of the membrane protein. J Virol. 72(8): 6838-6850.

DeDiego ML, Nieto-Torres JL, Jimenez-Guardeño JM, Regla-Nava JA, Castaño-Rodriguez C, Fernandez-Delgado R, Usera F and Enjuanes L 2014. Coronavirus virulence genes with main focus on SARS-CoV envelope gene. Virus Res. 194: 124-137.

Delmas B, Gelfi J, L'Haridon R, Vogel, Sjöström H, Norén and Laude H 1992. Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. Nature. 357(6377): 417-420.

Drummond AJ, Nicholls G, Rodrigo A and Solomon W 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics. 161.

Drummond AJ and Rambaut A 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology. 7(1): 214.

Drummond AJ, Rambaut A, Shapiro B and Pybus OG 2005. Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. Molecular Biology and Evolution. 22(5): 1185-1192.

Drummond AJ, Suchard MA, Xie D and Rambaut A 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 29(8): 1969-1973.

Duarte M and Laude H 1994. Sequence of the spike protein of the porcine epidemic diarrhoea virus. J Gen Virol. 75 (Pt 5): 1195-1200.

Enjuanes L, Suñé C, Gebauer F, Smerdou C, Camacho A, Antón IM, González S, Talamillo A, Méndez A, Ballesteros ML and et al. 1992. Antigen selection and presentation to protect

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

against transmissible gastroenteritis coronavirus. Vet Microbiol. 33(1-4): 249-262.

Fang P, Fang L, Ren J, Hong Y, Liu X, Zhao Y, Wang D, Peng G and Xiao S 2018. Porcine Deltacoronavirus Accessory Protein NS6 Antagonizes Interferon Beta Production by Interfering with the Binding of RIG-I/MDA5 to Double-Stranded RNA. J Virol. 92(15).

Ferreira MAR and Suchard MA 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. The Canadian Journal of Statistics / La Revue Canadienne de Statistique. 36(3): 355-368.

Fourment M and Darling AE 2018. Local and relaxed clocks: the best of both worlds. PeerJ. 6: e5140-e5140.

Gibbs MJ, Armstrong JS and Gibbs AJ 2000. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics. 16(7): 573-582.

Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA and Holmes EC 2004. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. Science. 303(5656): 327-332.

Hall TA 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic acids symposium series:95-98.

Hanke D, Jenckel M, Petrov A, Ritzmann M, Stadler J, Akimkin V, Blome S, Pohlmann A, Schirrmeier H, Beer M and Hoper D 2015. Comparison of porcine epidemic diarrhea viruses from Germany and the United States, 2014. Emerg Infect Dis. 21(3): 493-496.

Hao J, Xue C, He L, Wang Y and Cao Y 2014. Bioinformatics insight into the spike glycoprotein gene of field porcine epidemic diarrhea strains during 2011-2013 in Guangdong, China. Virus Genes. 49(1): 58-67.

He WT, Ji X, He W, Dellicour S, Wang S, Li G, Zhang L, Gilbert M, Zhu H, Xing G, Veit M, Huang Z, Han GZ, Huang Y, Suchard MA, Baele G, Lemey P and Su S 2020. Genomic Epidemiology, Evolution, and Transmission Dynamics of Porcine Deltacoronavirus. Mol Biol Evol. 37(9): 2641-2654.

Hofmann M and Wyler R 1988. Propagation of the virus of porcine epidemic diarrhea in cell culture. J Clin Microbiol. 26(11): 2235-2239.

Huang Y-W, Dickerman AW, Piñeyro P, Li L, Fang L, Kiehne R, Opriessnig T, Meng X-J and Griffin DE 2013. Origin, Evolution, and Genotyping of Emergent Porcine Epidemic

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Diarrhea Virus Strains in the United States. mBio. 4(5): e00737-00713.

Janetanakit T, Lumyai M, Bunpapong N, Boonyapisitsopa S, Chaiyawong S, Nonthabenjawan N, Kesdaengsakonwut S and Amonsin A 2016. Porcine Deltacoronavirus, Thailand, 2015. Emerging infectious diseases. 22(4): 757-759.

Johnson AD 2010. An extended IUPAC nomenclature code for polymorphic nucleic acids. Bioinformatics. 26(10): 1386-1389.

Jung K, Hu H, Eyerly B, Lu Z, Chepngeno J and Saif LJ 2015. Pathogenicity of 2 porcine deltacoronavirus strains in gnotobiotic pigs. Emerging infectious diseases. 21(4): 650-654.

Jung K, Hu H and Saif LJ 2017. Calves are susceptible to infection with the newly emerged porcine deltacoronavirus, but not with the swine enteric alphacoronavirus, porcine epidemic diarrhea virus. Archives of Virology. 162(8): 2357-2362.

Katoh K, Misawa K, Kuma K-i and Miyata T 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research. 30(14): 3059-3066.

Kim Y and Lee C 2014. Porcine epidemic diarrhea virus induces caspase-independent apoptosis through activation of mitochondrial apoptosis-inducing factor. Virology. 460-461: 180-193.

Kocherhans R, Bridgen A, Ackermann M and Tobler K 2001. Completion of the porcine epidemic diarrhoea coronavirus (PEDV) genome sequence. Virus genes. 23(2): 137-144.

Kong F, Wang Q, Kenney SP, Jung K, Vlasova AN and Saif LJ 2022. Porcine Deltacoronaviruses: Origin, Evolution, Cross-Species Transmission and Zoonotic Potential. Pathogens. 11(1).

Korber B 2000. HIV signature and sequence variation analysis. Computational analysis of HIV molecular sequences. 4: 55-72.

Kumar S, Stecher G, Li M, Knyaz C and Tamura K 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular biology and evolution. 35(6): 1547-1549.

Kumar S, Stecher G and Tamura K 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 33(7): 1870-1874.

Kweon CH, Kwon BJ, Lee JG, Kwon GO and Kang YB 1999. Derivation of attenuated porcine epidemic diarrhea virus (PEDV) as vaccine candidate. Vaccine. 17(20-21): 2546-2553.

Lednicky JA, Tagliamonte MS, White SK, Elbadry MA, Alam MM, Stephenson CJ, Bonny TS,

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Loeb JC, Telisma T, Chavannes S, Ostrov DA, Mavian C, Beau De Rochars VM, Salemi M and Morris JG 2021. Independent infections of porcine deltacoronavirus among Haitian children. Nature. 600(7887): 133-137.

Lee C 2015. Porcine epidemic diarrhea virus: an emerging and re-emerging epizootic swine virus. Virology journal. 12(1): 1-16.

Lee D-K, Park C-K, Kim S-H and Lee C 2010. Heterogeneity in spike protein genes of porcine epidemic diarrhea viruses isolated in Korea. Virus Research. 149(2): 175-182.

Lee S and Lee C 2014a. Complete Genome Characterization of Korean Porcine Deltacoronavirus Strain KOR/KNU14-04/2014. Genome Announc. 2(6).

Lee S and Lee C 2014b. Outbreak-related porcine epidemic diarrhea virus strains similar to US strains, South Korea, 2013. Emerg Infect Dis. 20(7): 1223-1226.

Lemey P, Rambaut A, Welch JJ and Suchard MA 2010. Phylogeography takes a relaxed random walk in continuous space and time. Mol Biol Evol. 27(8): 1877-1885.

Li W and Godzik A 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 22(13): 1658-1659.

Li W, Li H, Liu Y, Pan Y, Deng F and Song Y 2012a. New variants of porcine epidemic diarrhea virus, China, 2011. Emerg Infect Dis. 18.

Li W, Li H, Liu Y, Pan Y, Deng F, Song Y, Tang X and He Q 2012b. New variants of porcine epidemic diarrhea virus, China, 2011. Emerg Infect Dis. 18(8): 1350-1353.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Li W, Luo R, He Q, van Kuppeveld FJM, Rottier PJM and Bosch B-J 2017. Aminopeptidase N is not required for porcine epidemic diarrhea virus cell entry. Virus Research. 235: 6-13.

Liang Q, Zhang H, Li B, Ding Q, Wang Y, Gao W, Guo D, Wei Z and Hu H 2019. Susceptibility of Chickens to Porcine Deltacoronavirus Infection. Viruses. 11(6).

Lin CN, Chung WB, Chang SW, Wen CC, Liu H, Chien CH and Chiou MT 2014. US-like strain of porcine epidemic diarrhea virus outbreaks in Taiwan, 2013-2014. J Vet Med Sci. 76(9): 1297-1299.

Liu C, Kokuho T, Kubota T, Watanabe S, Inumaru S, Yokomizo Y and Onodera T 2001. A serodiagnostic ELISA using recombinant antigen of swine transmissible gastroenteritis virus nucleoprotein. J Vet Med Sci. 63(11): 1253-1256.

Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard

HW and Ray SC 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol. 73(1): 152-160.

Lorsirigool A, Saeng-chuto K, Madapong A, Temeeyasen G, Tripipat T, Kaewprommal P, Tantituvanont A, Piriyapongsa J and Nilubol D 2017. The genetic diversity and complete genome analysis of two novel porcine deltacoronavirus isolates in Thailand in 2015. Virus Genes. 53(2): 240-248.

Lorsirigool A, Saeng-Chuto K, Temeeyasen G, Madapong A, Tripipat T, Wegner M, Tuntituvanont A, Intrakamhaeng M and Nilubol D 2016. The first detection and full-length genome sequence of porcine deltacoronavirus isolated in Lao PDR. Arch Virol. 161(10): 2909-2911.

Lozano-Fuentes S, Elizondo-Quiroga D, Farfan-Ale JA, Loroño-Pino MA, Garcia-Rejon J, Gomez-Carro S, Lira-Zumbardo V, Najera-Vazquez R, Fernandez-Salas I, Calderon-Martinez J, Dominguez-Galera M, Mis-Avila P, Morris N, Coleman M, Moore CG, Beaty BJ and Eisen L 2008. Use of Google Earth™ to strengthen public health capacity and facilitate management of vector-borne diseases in resource-poor environments. Bulletin of the World Health Organization. 86(9): 718-725.

Luo L, Wang S, Zhu L, Fan B, Liu T, Wang L, Zhao P, Dang Y, Sun P, Chen J, Zhang Y, Chang X, Yu Z, Wang H, Guo R, Li B and Zhang K 2019. Aminopeptidase N-null neonatal piglets are protected from transmissible gastroenteritis virus but not porcine epidemic diarrhea virus. Scientific Reports. 9(1): 13186.

Marthaler D, Jiang Y, Collins J and Rossow K 2014. Complete Genome Sequence of Strain SDCV/USA/Illinois121/2014, a Porcine Deltacoronavirus from the United States. Genome announcements. 2(2): e00218-00214.

Martin D and Rybicki E 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics. 16(6): 562-563.

Martin DP, Murrell B, Golden M, Khoosal A and Muhire B 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution. 1(1): vev003.

Masuda T, Murakami S, Takahashi O, Miyazaki A, Ohashi S, Yamasato H and Suzuki T 2015. New porcine epidemic diarrhoea virus variant with a large deletion in the spike gene identified in domestic pigs. Arch Virol. 160(10): 2560-2568.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

McCluskey BJ, Haley C, Rovira A, Main R, Zhang Y and Barder S 2016. Retrospective testing and case series study of porcine delta coronavirus in U.S. swine herds. Preventive Veterinary Medicine. 123: 185-191.

Mendenhall IH, Kerimbayev AA, Strochkov VM, Sultankulova KT, Kopeyev SK, Su YCF, Smith GJD and Orynbayev MB 2019. Discovery and Characterization of Novel Bat Coronavirus Lineages from Kazakhstan. Viruses. 11(4).

Ojkic D, Hazlett M, Fairles J, Marom A, Slavic D, Maxie G, Alexandersen S, Pasick J, Alsop J and Burlatschenko S 2015. The first case of porcine epidemic diarrhea in Canada. Can Vet J. 56(2): 149-152.

Oong XY, Ng KT, Takebe Y, Ng LJ, Chan KG, Chook JB, Kamarulzaman A and Tee KK 2017. Identification and evolutionary dynamics of two novel human coronavirus OC43 genotypes associated with acute respiratory infections: phylogenetic, spatiotemporal and transmission network analyses. Emerging microbes & infections. 6(1): e3-e3.

Padidam M, Sawyer S and Fauquet CM 1999. Possible Emergence of New Geminiviruses by Frequent Recombination. Virology. 265(2): 218-225.

Park SJ, Kim HK, Song DS, Moon HJ and Park BK 2011. Molecular characterization and phylogenetic analysis of porcine epidemic diarrhea virus (PEDV) field isolates in Korea. Arch Virol. 156(4): 577-585.

Park SJ, Moon HJ, Luo Y, Kim HK, Kim EM, Yang JS, Song DS, Kang BK, Lee CS and Park BK 2008. Cloning and further sequence analysis of the ORF3 gene of wild- and attenuated-type porcine epidemic diarrhea viruses. Virus Genes. 36(1): 95-104.

Park SJ, Moon HJ, Yang JS, Lee CS, Song DS, Kang BK and Park BK 2007. Sequence analysis of the partial spike glycoprotein gene of porcine epidemic diarrhea viruses isolated in Korea. Virus Genes. 35(2): 321-332.

Pensaert M and De Bouck P 1978. A new coronavirus-like particle associated with diarrhea in swine. Archives of virology. 58(3): 243-247.

Pérez-Rivera C, Ramírez-Mendoza H, Mendoza-Elvira S, Segura-Velázquez R and Sánchez-Betancourt JI 2019. First report and phylogenetic analysis of porcine deltacoronavirus in Mexico. Transboundary and Emerging Diseases. 66(4): 1436-1441.

Posada D and Crandall KA 2001. Evaluation of methods for detecting recombination from DNA

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

sequences: Computer simulations. Proceedings of the National Academy of Sciences. 98(24): 13757-13762.

Pospischil A, Stuedli A and Kiupel M 2002. Update on porcine epidemic diarrhea. JOURNAL OF SWINE HEALTH AND PRODUCTION. 10(2): 81-85.

Pybus OG, Suchard MA, Lemey P, Bernardin FJ, Rambaut A, Crawford FW, Gray RR, Arinaminpathy N, Stramer SL, Busch MP and Delwart EL 2012. Unifying the spatial epidemiology and molecular evolution of emerging epidemics. Proceedings of the National Academy of Sciences. 109(37): 15066-15071.

Rambaut A and Drummond AJ 2013. Tracer 1.6 [computer program]. In.

Rambaut A and Drummond AJ 2014. FigTree 1.4.2 [computer program]. In.

Regla-Nava JA, Nieto-Torres JL, Jimenez-Guardeño JM, Fernandez-Delgado R, Fett C, Castaño-Rodríguez C, Perlman S, Enjuanes L and DeDiego ML 2015. Severe acute respiratory syndrome coronaviruses with mutations in the E protein are attenuated and promising vaccine candidates. J Virol. 89(7): 3870-3887.

Saeng-Chuto K, Jermsutjarit P, Stott CJ, Vui DT, Tantituvanont A and Nilubol D 2020. Retrospective study, full-length genome characterization and evaluation of viral infectivity and pathogenicity of chimeric porcine deltacoronavirus detected in Vietnam. Transbound Emerg Dis. 67(1): 183-198.

Saeng-Chuto K, Lorsirigool A, Temeeyasen G, Vui DT, Stott CJ, Madapong A, Tripipat T, Wegner M, Intrakamhaeng M, Chongcharoen W, Tantituvanont A, Kaewprommal P, Piriyapongsa J and Nilubol D 2016. Different Lineage of Porcine Deltacoronavirus in Thailand, Vietnam and Lao PDR in 2015. Transbound Emerg Dis.

Saeng-Chuto K, Stott CJ, Wegner M, Senasuthum R, Tantituvanont A and Nilubol D 2017. Retrospective investigation and evolutionary analysis of a novel porcine deltacoronavirus strain detected in Thailand from 2008 to 2015. Arch Virol. 162(7): 2103-2108.

Salminen MO, Carr JK, Burke DS and McCutchan FE 1995. Identification of Breakpoints in Intergenotypic Recombinants of HIV Type 1 by Bootscanning. AIDS Research and Human Retroviruses. 11(11): 1423-1425.

Sánchez CM, Gebauer F, Suñé C, Mendez A, Dopazo J and Enjuanes L 1992. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. Virology. 190(1): 92-105.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Sato T, Takeyama N, Katsumata A, Tuchiya K, Kodama T and Kusanagi K-i 2011. Mutations in the spike gene of porcine epidemic diarrhea virus associated with growth adaptation in vitro and attenuation of virulence in vivo. Virus genes. 43(1): 72-78.

Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD and Karsch-Mizrachi I 2019. GenBank. Nucleic Acids Res. 47(D1): D94-d99.

Shang J, Zheng Y, Yang Y, Liu C, Geng Q, Tai W, Du L, Zhou Y, Zhang W and Li F 2018. Cryo-electron microscopy structure of porcine deltacoronavirus spike protein in the prefusion state. Journal of virology. 92(4): e01556-01517.

Shibata I, Ono M and Mori M 2001. Passive protection against porcine epidemic diarrhea (PED) virus in piglets by colostrum from immunized cows. J Vet Med Sci. 63(6): 655-658.

Shibata I, Tsuda T, Mori M, Ono M, Sueyoshi M and Uruno K 2000. Isolation of porcine epidemic diarrhea virus in porcine cell cultures and experimental infection of pigs of different ages. Vet Microbiol. 72(3-4): 173-182.

Sinha A, Gauger P, Zhang J, Yoon K-J and Harmon K 2015. PCR-based retrospective evaluation of diagnostic samples for emergence of porcine deltacoronavirus in US swine. Veterinary Microbiology. 179(3): 296-298.

Smith JM 1992. Analyzing the mosaic structure of genes. J Mol Evol. 34(2): 126-129.

Song D, Zhou X, Peng Q, Chen Y, Zhang F, Huang T, Zhang T, Li A, Huang D, Wu Q, He H and Tang Y 2015. Newly Emerged Porcine Deltacoronavirus Associated With Diarrhoea in Swine in China: Identification, Prevalence and Full-Length Genome Sequence Analysis. Transbound Emerg Dis. 62(6): 575-580.

Song DS, Oh JS, Kang BK, Yang JS, Moon HJ, Yoo HS, Jang YS and Park BK 2007. Oral efficacy of Vero cell attenuated porcine epidemic diarrhea virus DR13 strain. Res Vet Sci. 82(1): 134-140.

Srinuntapunt S, Trongwongsa L, Antarasena C, Sangsuwan W and Prommuang P 1995. Porcine epidemic diarrhea in Trang Province. J Thai Vet Med Assoc. 46: 11-19.

Steinbach F, Dastjerdi A, Peake J, La Rocca SA, Tobin FP, Frossard J-P and Williamson S 2016. A retrospective study detects a novel variant of porcine epidemic diarrhea virus in England in archived material from the year 2000. PeerJ. 4: e2564-e2564.

Stott CJ, Sawattrakool K, Saeng-Chuto K, Tantituvanont A and Nilubol D 2022. The

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

phylodynamics of emerging porcine deltacoronavirus in Southeast Asia. Transbound Emerg Dis. 69(5): 2816-2827.

Stott CJ, Temeeyasen G, Tripipat T, Kaewprommal P, Tantituvanont A, Piriyapongsa J and Nilubol D 2017. Evolutionary and epidemiological analyses based on spike genes of porcine epidemic diarrhea virus circulating in Thailand in 2008-2015. Infect Genet Evol. 50: 70-76.

Su Y, Hou Y, Prarat M, Zhang Y and Wang Q 2018. New variants of porcine epidemic diarrhea virus with large deletions in the spike protein, identified in the United States, 2016-2017. Archives of virology. 163(9): 2485-2489.

Sun D, Feng L, Shi H, Chen J, Cui X, Chen H, Liu S, Tong Y, Wang Y and Tong G 2008. Identification of two novel B cell epitopes on porcine epidemic diarrhea virus spike protein. Vet Microbiol. 131(1-2): 73-81.

Sun D, Wang X, Wei S, Chen J and Feng L 2016. Epidemiology and vaccine of porcine epidemic diarrhea virus in China: a mini-review. The Journal of veterinary medical science. 78(3): 355-363.

Sun DB, Feng L, Shi HY, Chen JF, Liu SW, Chen HY and Wang YF 2007. Spike protein region (aa 636789) of porcine epidemic diarrhea virus is essential for induction of neutralizing antibodies. Acta Virol. 51(3): 149-156.

Sun M, Ma J, Wang Y, Wang M, Song W, Zhang W, Lu C and Yao H 2015. Genomic and epidemiological characteristics provide new insights into the phylogeographical and spatiotemporal spread of porcine epidemic diarrhea virus in Asia. Journal of clinical microbiology. 53(5): 1484-1492.

Sun R, Leng Z, Zhai SL, Chen D and Song C 2014. Genetic variability and phylogeny of current Chinese porcine epidemic diarrhea virus strains based on spike, ORF3, and membrane genes. ScientificWorldJournal. 2014: 208439.

Sun RQ, Cai RJ, Chen YQ, Liang PS, Chen DK and Song CX 2012. Outbreak of porcine epidemic diarrhea in suckling piglets, China. Emerg Infect Dis. 18(1): 161-163.

Sung MH, Deng MC, Chung YH, Huang YL, Chang CY, Lan YC, Chou HL and Chao DY 2015. Evolutionary characterization of the emerging porcine epidemic diarrhea virus worldwide and 2014 epidemic in Taiwan. Infect Genet Evol. 36: 108-115.

Takahashi K, Okada K and Ohshima K-i 1983. An Outbreak of Swine Diarrhea of a New-Type

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Associated with Coronavirus-Like Particles in Japan. The Japanese Journal of Veterinary Science. 45(6): 829-832.

Tavaré S 1986. Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. In: American Mathematical Society: Lectures on Mathematics in the Life Sciences. Amer Mathematical Society. 57-86.

Temeeyasen G, Srijangwad A, Tripipat T, Tipsombatboon P, Piriyapongsa J, Phoolcharoen W, Chuanasa T, Tantituvanont A and Nilubol D 2014. Genetic diversity of ORF3 and spike genes of porcine epidemic diarrhea virus in Thailand. Infect Genet Evol. 21: 205-213.

Theuns S, Conceicao-Neto N, Christiaens I, Zeller M, Desmarets LM, Roukaerts ID, Acar DD, Heylen E, Matthijnssens J and Nauwynck HJ 2015. Complete genome sequence of a porcine epidemic diarrhea virus from a novel outbreak in belgium, january 2015. Genome Announc. 3(3).

Thomas JT, Chen Q, Gauger PC, Giménez-Lirola LG, Sinha A, Harmon KM, Madson DM, Burrough ER, Magstadt DR, Salzbrenner HM, Welch MW, Yoon K-J, Zimmerman JJ and Zhang J 2015. Effect of Porcine Epidemic Diarrhea Virus Infectious Doses on Infection Outcomes in Naïve Conventional Neonatal and Weaned Pigs. PloS one. 10(10): e0139266-e0139266.

Tian Y, Yu Z, Cheng K, Liu Y, Huang J, Xin Y, Li Y, Fan S, Wang T, Huang G, Feng N, Yang Z, Yang S, Gao Y and Xia X 2013. Molecular characterization and phylogenetic analysis of new variants of the porcine epidemic diarrhea virus in Gansu, China in 2012. Viruses. 5(8): 1991-2004.

Turgeon DC, Morin M, Jolette J, Higgins R, Marsolais G and DiFranco E 1980. Coronavirus-like particles associated with diarrhea in baby pigs in Quebec. Can Vet J. 21(3): 100-xxiii.

Turlewicz-Podbielska H and Pomorska-Mól M 2021. Porcine Coronaviruses: Overview of the State of the Art. Virologica Sinica. 36(5): 833-851.

Utiger A, Tobler K, Bridgen A, Suter M, Singh M and Ackermann M 1995. Identification of proteins specified by porcine epidemic diarrhoea virus. Adv Exp Med Biol. 380: 287-290.

Vijaykrishna D, Smith GJD, Zhang JX, Peiris JSM, Chen H and Guan Y 2007. Evolutionary Insights into the Ecology of Coronaviruses. Journal of Virology. 81(8): 4012.

Vijgen L, Keyaerts E, Moës E, Thoelen I, Wollants E, Lemey P, Vandamme AM and Van Ranst M

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

2005. Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. J Virol. 79(3): 1595-1604.

Vui DT, Thanh TL, Tung N, Srijangwad A, Tripipat T, Chuanasa T and Nilubol D 2015. Complete genome characterization of porcine epidemic diarrhea virus in Vietnam. Arch Virol. 160(8): 1931-1938.

Wang L, Byrum B and Zhang Y 2014. Detection and genetic characterization of deltacoronavirus in pigs, Ohio, USA, 2014. Emerg Infect Dis. 20(7): 1227-1230.

Woo PC, Lau SK, Lam CS, Lau CC, Tsang AK, Lau JH, Bai R, Teng JL, Tsang CC, Wang M, Zheng BJ, Chan KH and Yuen KY 2012. Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. J Virol. 86(7): 3995-4008.

Woo PCY, Huang Y, Lau SKP and Yuen KY 2010. Coronavirus genomics and bioinformatics analysis. Viruses. 2(8): 1804-1820.

Xu Z, Zhong H, Zhou Q, Du Y, Chen L, Zhang Y, Xue C and Cao Y 2018. A Highly Pathogenic Strain of Porcine Deltacoronavirus Caused Watery Diarrhea in Newborn Piglets. Virologica Sinica. 33(2): 131-141.

Yang Z 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. Journal of Molecular Evolution. 39(3): 306-314.

Yang Z and Bielawski JP 2000. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 15(12): 496-503.

Zhang F, Luo S, Gu J, Li Z, Li K, Yuan W, Ye Y, Li H, Ding Z, Song D and Tang Y 2019. Prevalence and phylogenetic analysis of porcine diarrhea associated viruses in southern China from 2012 to 2018. BMC Vet Res. 15(1): 470.

Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, Tachedjian M, Zhu Y, Zhou P, Jiang X, Ng J, Yang L, Wu L, Xiao J, Feng Y, Chen Y, Sun X, Zhang Y, Marsh GA, Crameri G, Broder CC, Frey KG, Wang L-F and Wang J 2013. Comparative Analysis of Bat Genomes Provides Insight into the Evolution of Flight and Immunity. Science. 339(6118): 456-460.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Zhang J, Guo L, Xu Y, Yang L, Shi H, Feng L and Wang Y 2017. Characterization of porcine epidemic diarrhea virus infectivity in human embryonic kidney cells. Arch Virol. 162(8): 2415-2419.

Zhang XW, Yap YL and Danchin A 2005. Testing the hypothesis of a recombinant origin of the SARS-associated coronavirus. Arch Virol. 150(1): 1-20.

Zhao Y, Qu H, Hu J, Fu J, Chen R, Li C, Cao S, Wen Y, Wu R, Zhao Q, Yan Q, Wen X and Huang X 2019. Characterization and Pathogenicity of the Porcine Deltacoronavirus Isolated in Southwest China. Viruses. 11(11): 1074.

Zhao Z, Li H, Wu X, Zhong Y, Zhang K, Zhang YP, Boerwinkle E and Fu YX 2004. Moderate mutation rate in the SARS coronavirus genome and its implications. BMC Evol Biol. 4: 21.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

APPENDIX A

Evolutionary and epidemiological analyses based on spike genes of porcine epidemic diarrhea virus circulating in Thailand in 2008-2015

Infection, Genetics and Evolution

Jun 2017, Volume 50, Pages: 70-76

Christopher James Stott, Gun Temeeyasen, Thitima Tripipat, Pavita Kaewprommal, Angkana Tantituvanont, Jittima Piriyapongsa, and Dachrit Nilubol

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Infection, Genetics and Evolution 50 (2017) 70–76

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Evolutionary and epidemiological analyses based on spike genes of porcine epidemic diarrhea virus circulating in Thailand in 2008–2015

Christopher J. Stott [a], Gun Temeeyasen [a], Thitima Tripipat [a], Pavita Kaewprommal [c], Angkana Tantituvanont [b], Jittima Piriyapongsa [c], Dachrit Nilubol [a,*]

[a] Department of Veterinary Microbiology, Faculty of Veterinary Science, Chulalongkorn University, Thailand
[b] Department of Pharmaceutics and Industrial Pharmacy, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Thailand
[c] Genome Technology Research Unit, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathumthani, Thailand

ARTICLE INFO

Article history:
Received 7 November 2016
Received in revised form 17 February 2017
Accepted 20 February 2017
Available online 21 February 2017

Keywords:
Porcine epidemic diarrhea virus
Epidemiology
Evolutionary rate
Thailand

ABSTRACT

Porcine epidemic diarrhea (PED) has been endemic causing sporadic outbreaks in Thailand since 2007. In 2014–2015, several herds had experienced severe PED outbreaks and the reason of the re-current outbreaks was unknown. Whether or not the introduction of exotic strains or continual evolution of existing PEDV, genetic analyses would provide a more understanding in its evolutionary pattern. In the study, 117 complete spike gene sequences of Thai PED virus (PEDV) collected from 2008 to 2015 were clustered along with 95 references of PEDV spike sequences, and analyzed with the US sequences dataset (n = 99).

The phylogenetic analysis demonstrated that Thai PEDV spike sequences were genetically diverse and had been influenced by multiple introduction of exotic strains. Although Thai PEDV have been evolved into 6 subgroups (TH1–6), Subgroup TH1 strains with the unique 9 nucleotides (CAA GGG AAT) insertion between 688th–689th position of spike (changing amino acid from N to TREY) insertion has become the dominant subgroup since 2014. Thai PEDV spike gene have higher evolutionary rate compare to that of the US sequences. One contributing factor would be the intra-recombination between subgroups. Thailand endemic strain should be assigned into new subclade of G2 (Thai pandemic variant).

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Porcine epidemic diarrhea (PED) is a devastating enteric disease in pigs, characterized by vomiting and acute watery diarrhea leading to death due to severe dehydration. Pigs at all ages are susceptible to PED infection, but mortality is higher in younger pigs, especially in piglets at younger than a week of age. Since its first recognition in the late 1970s (Chasey and Cartwright 1978; Pensaert and De Bouck 1978), PED has continued to cause a severe economic impact in swine industry worldwide. At present, PED has been reported as re-emerging disease worldwide (Chen et al. 2014; Dennis et al. 2015; Lee et al. 2014; Lin et al. 2014; Masuda et al. 2015; Ojkic et al. 2015; Theuns et al. 2015; Vui et al. 2015).

PED virus (PEDV), a causative agent of PED, is an enveloped single-stranded positive-sense RNA virus belonging to the genus *Alphacoronavirus*, family *Coronaviridae*, order *Nidovirales*. At present, two PEDV variants including genogroups 1 and 2 are currently recognized (Lee 2015; Sun et al. 2015). The PEDV genome is consisting of seven open reading frames (ORFs) organizing in order as following; ORF1a and ORF1b, spike (S), ORF3, envelope (E), membrane (M), and nucleoprotein (N) (Kocherhans et al. 2001). Two-thirds of the genome is occupied by ORF1a and ORF1b, which encode nonstructural proteins. The other 4 structural proteins including S, E, M, and N are located downstream of the ORF1a/1b gene. The S protein is a glycosylated protein involving with viral pathogenesis and be further divided into S1 and S2 domains. The S protein attaches to the host cellular receptors resulting to virus entry by membrane fusion (Bosch et al. 2003) and contains domain that stimulates the production of neutralizing antibodies (Chang et al. 2002; Cruz et al. 2006, 2008; Duarte and Laude 1994; Sun et al. 2008; Sun et al. 2007). The S gene is divergent and important for understanding the genetic relatedness of PEDV field strains, the epidemiological status of the virus and vaccine development (Li et al. 2012; Park et al. 2011; Sun et al. 2012). ORF3 encodes an accessory protein located between S and E. In addition, the ORF3 gene is the only accessory gene, and has been an important determinant of virulence in PEDV (Park et al. 2008). The vaccine-derived strains have unique deletion of 17 amino acids at position 82 to 99 (Park et al. 2008). The other E and M genes are associated with viral envelope formation and release. The N protein is the predominant antigen produced in coronavirus-infected cells, making it a major viral target (Kocherhans et al. 2001).

In Thailand, PED was first emerged in 2007 and the variant responsible for the outbreaks was of genogroup 2 (Temeeyasen et al. 2014). At present, PED has developed into an endemic stage in which many

* Corresponding author.
E-mail address: ndachrit@chula.ac.th (D. Nilubol).

http://dx.doi.org/10.1016/j.meegid.2017.02.014
1567-1348/© 2017 Elsevier B.V. All rights reserved.

herds experience recurrent outbreaks and both PEDV genogroups 1 and 2 are currently existing in the population (Cheun-Arom et al. 2015). In 2014–2015, several herds reported the severe diarrhea outbreaks. Whether or not the novel introduction or continual development of PEDV in Thai swine farms causing the outbreak is not known and the information on PEDV evolutionary dynamics in Thailand has been deficiency.

Since increasing of computer performance and development of graphic user interface of BEAST format support software, BEAST packages become fast, flexible and reliable tools for molecular sequences analysis based on Bayesian framework (Drummond and Rambaut 2007). Our study was aimed to investigate the genetic diversity of PEDV clinical samples and determine the evolutionary rate of PEDV in Thailand. The genetic analysis was based on the complete spike gene of Thai PEDV collected from 2008 to 2015 along with that of reference sequences from other countries provided in GenBank®. The combination of the sequences with time of isolation analyzing with BEAST can provide us the estimated time of divergences and the evolutionary rate of suspected group of population. The results would provide us the information on the evolutionary dynamics of the viruses in Thailand.

2. Materials and methods

2.1. Source of specimens and PEDV isolation

A total of 117 Thai PEDV spike gene sequences including 99 PEDV clinical samples that were genetically characterized during 2014–2015 together with 18 Thai sequences previously reported (Temeeyasen et al. 2014) (Table S1) were used in the genetic analyses. The 115 PEDV sequences were collected from 24 pig farms located in Ratchaburi, Nakhon Pathom, Saraburi, Lopburi, Nakhon Ratchasima, Buriram, Chonburi, Chanthaburi and ChaCherngsao from 2008 to 2015 that experienced PED outbreaks (55 outbreaks) whereas another 2 sequences are not the field strains, the AVCT12 (LC053455) and a vaccine strain sequence from the vaccine using in the farm that experienced the outbreak of EAS (KR610991). In each outbreak, intestinal samples were collected from 3- to 4-day-old piglets that displayed the clinical features associated with PED, including vomiting and watery diarrhea. Intestinal samples were minced into small pieces and suspended in PBS and clarified by centrifugation. The supernatant was filtered through 0.22 μm filters and stored at −80 °C until use. The clarified supernatants were subjected to RT-PCR.

2.2. Cloning, plasmid purification and sequence determination

Sequencing was performed as previously reported (Temeeyasen et al. 2014). Briefly, total RNA was extracted from the supernatant using the Nucleospin®RNA Virus kit (Macherey-Nagel Inc., PA, USA) in accordance with the manufacturer's instructions. cDNA was synthesized from the extracted RNA using M-MuLV Reverse Transcriptase (New England BioLabs Inc., MA, USA).

PCR amplification was performed on the cDNA. To amplify the complete S genes, PCR amplification was performed using previously reported primers (Lee et al. 2010) and using Platinum® *Taq* DNA polymerase High Fidelity (Invitrogen, CA, USA). The PCR products were cloned into plasmid vectors for the subsequent transformation of *Escherichia coli* cells by using a commercial kit (pGEM-T® Easy Vector System I (Promega, WI, USA), and the controls were included at all stages of cloning and transformation. Bacterial transformant colonies were growth in Luria Broth (LB) agar for 18 h and randomly selected from each sample for plasmid purification using the Nucleospin® Plasmid kit (Macherey-Nagel Inc., PA, USA), and the 10 selected colonies were grown in LB broth for 24 h and subjected to plasmid isolation and sequencing. Sequencing reactions were performed by third party (1st BASE DNA Sequencing Services, Singapore).

2.3. Sequence analyses

2.3.1. Phylogenetic analysis

The nucleotide and deduced amino acid sequences were aligned using the software, MAFFT (Multiple Alignment using Fast Fourier Transform) (Katoh et al. 2002). A phylogenetic analysis was performed based on the nucleotide sequences of complete spike genes of 117 Thai PEDV spike sequences collected from 2008 to 2015 along with 94 other PEDV spike sequences available in GenBank® (Table S1) from previous study (Lee 2015) using the MEGA6.0 program (Tamura et al. 2013). A phylogenetic tree was constructed using Bayesian Markov chain Monte Carlo (BMCMC) with best-fit substitution model, TN93 + G + I and base frequency was set as empirical, all other settings were left as default. The analysis was performed until ESS > 200, the genogroup described in other study (Lee 2015) was used to classify and compare with the sequences in this study.

2.3.2. Evolutionary analysis

The evolutionary analysis of 115 Thai PEDV spike sequences except of those two which not the field strains as described in previous part were performed analysis in comparison 99 US PEDV spike sequences (Table S2) using a BMCMC method implemented in the program BEAST v1.8.3 (Drummond and Rambaut 2007; Drummond et al. 2012). The best fitted evolutionary model GTR + G + I was determined by maximum likelihood model selection function of MEGA 6.0 (Tamura et al. 2013) and was used in all BEAST analyses. The coalescent Bayesian skyline tree prior (Drummond et al. 2002; Drummond et al. 2005) and empirical base frequencies were applied under three different models for rate variation among branches: the strict molecular clock model (STR), the uncorrelated lognormal relaxed-clock model (LOG), and the uncorrelated exponential relaxed-clock model (EXP) (Drummond et al. 2006). For each analysis, 200 million states were applied at least with logged every 10,000 states, the states must be increased in order to gain ESS > 200. If there was another independent run, the log and tree data will be combined with the first 10% discarded as burn-in using LogCombinerv1.8.3, and, the resulting files were displayed using Tracer 1.6.0 (Rambaut and Drummond 2013). Maximum clade credibility trees were annotated using TreeAnnotator 1.8.3 and phylogenetic tree with timeline, estimated divergences, posterior probability and 95% HPD displays were generated using FigTree 1.4.2 (Rambaut and Drummond 2014).

2.3.3. Phylogeographic distribution

Phylogeographic distribution of the endemic subgroup was also estimated using BEAST 1.8.4 under TN93 + G + I substitute model. 600 million with logged 10,000 states was performed at least for subgroup TH1 and TH2. All other settings were performed as described in previous part. The results were converted and displayed in GoogleEarth (Lozano-Fuentes et al. 2008) using SPREAD 1.0.7 (Bielejec et al. 2011). The sequences which have 100% identity from the same location and time of the outbreak were excluded.

2.3.4. Recombinant analysis

Recombination events among the sequences (n = 117) were analyzed using automated RDP, GENECONV, BOOTSCAN, MaxChi, CHIMAERA, and SISCAN provided by RDP4 software (Martin et al. 2015). The evidence events with known potential parents and recombinants were displayed.

3. Results

3.1. Phylogenetic and genetic analyses

All PEDV samples were grouped based on spike gene sequence data using clades previously reported (Lee 2015) in which PEDV evolved into 2 separated clades including G1 and G2. Each clades was further divided

Fig. 1. The phylogenetic tree based on complete spike gene of Thai PEDV sequences along with reference sequences (Lee 2015). PEDV are evolved into two genogroups including G1 and G2. G1 and G2 are each further divided into two sub-groups including G1a, G1b, G2a and G2b. TH1–TH6 represent Thai PEDV subgroups of the samples collected during 2008 to 2015. Numbers above branches denote posterior probability values.

into 2 subclades including subclades G1a, G1b, G2a and G2b. Subclade G1a included mainly the first period of PEDV found in Belgium, UK, (CV777, CH/S, Br1/87) and attenuated strains of vaccine from China and Korea. Subclade G1b included mostly the strains of China together with Korea, USA and Europe strains that were reported during 2014–2015. G2 was further evolved into G2a and G2b. G2a included strains from China, Korea and Thailand. G2b included mainly PEDV strains from US, and from countries which reported the detection of US-like PEDV.

Thai PEDV are classified into 6 subgroups including TH1–TH6 (Fig. 1). Based on spike gene identity and the clades previously described (Lee 2015), most of Thai sequences belong to G2a clade. Among these 6 subgroups, the genetic differences are >3% between subgroups, we therefore classified 3 subgroups of Thai PEDV to G2a including TH1, TH2 and TH3. Subgroups TH1 (n = 79) and TH2 (n = 30) were the two dominant groups in Thailand. TH1 included the first Thai PEDV strain that was published in 2008 along with the latest samples we found in 2015. In the endemic group 1 (TH1), there were 26 samples from 19 outbreaks with the 9 nucleotides insertion of CAA GGG AAT (TH-S-INS) when aligned with the reference sequence (NC_003436; position 688th–689th) which caused the amino acid changed from N to T (229 of the reference) and insertion of REY at the site between 229th–230th. TH3 only contained one sample in 2011 which might be a new introduction of the virus or the recombinant one. However, the sequence seems to be a recombinant one because it shares 97% identity with some sequences in both G1 and G2 group when blasting with nucleotide BLAST.

Subgroups TH4 and TH6 were in genogroup G2b. TH4 was applied for CBR1 and CBR2 (KR610993 and KR610994) sequences in July 2014, these sequences belong to G2b and very close to Vietnam sequences (99% identity) in 2013, so we can conclude them as a new introduction of the exotic strains. TH6 is the latest subgroup which contained the new introduction of the exotic strain (P1915-NPF-071511A) in 2015 that belong to G2b or US-like subgroup.

Subgroup TH5 was in genogroup G1a. TH5 belongs to the latter group of samples in October 2014, EAS1 and EAS2 (KR610991 and 610992) which the first group of strain that belong to G1 clade that we found in Thailand. The Thai PEDV in this genogroup were genetically distinct from PEDV strains recently emerged in European countries. These strains were close to the vaccine strain so we did isolation the virus vaccine using in the farms experienced to the outbreak of the strain, and including of an AVCT12, both the attenuated strains are assigned into THX subgroup in this study.

Nucleotide and amino acid sequences of TH1–6 represent sequences and nearest identity to each subgroup are provided in Figs. S3 and S4 in Supplementary materials.

3.2. Evolutionary analysis

The evolutionary rates were calculated for 4 datasets of PEDV groups including overall (all field strains of Thailand; n = 115), endemic (predominant TH1 and Th2 subgroups; n = 109), TH1 (n = 79), and TH2 (n = 30). In general, the higher evolutionary rate is present in Thai PEDV strains than in the US sequences represent. In comparison among Thai strains, subgroup TH1 (predominant endemic subgroup) had higher rate of substitution than TH2 subgroup (Fig. 2; Table S3).

The divergence times among Thai PEDV sequences were also estimated using BEAST exponential relaxed-clock method. All Thai PEDV sequences share a common ancestor that presented in year 1999, with 95% HPD interval of 1987–2004 (Fig. 3, Table S3). The TH2 subgroup seems to evolve independently from all other samples. The emergence of TH2 common ancestor was estimated to be around year 2002 whereas the common ancestor of all other samples arose approximately in 2003. The TH1, TH3, and TH5 subgroups have a shared ancestor which presented around 2004. The TH1 subgroup emerged around 2005–2006 and the TH3/TH5 subgroups evolved later in 2007. The TH4 and TH6 subgroups descended from their common ancestor which emerged recently in 2012.

3.3. Phylogeographic distribution

Due to the ESS values, exponential molecular clock model was chosen to display the geographic distribution of PEDV in Thailand as shown in Fig. 4. Based on the phylogeographical analysis, there was geographic separation between PEDV stains in Thailand. TH2 subgroup was confined mainly in the western regions. In contrast, TH1 subgroup was found across Thailand, especially the northeast region of Thailand, and has become the dominant subgroup since 2014.

Since its emergence in 2007, PED has been endemic in the western region of Thailand including Ratchaburi and Nakhon Pathom provinces and sporadic outbreaks have been reported in the eastern region including Chonburi, Chanthaburi and ChaCherngsao provinces. However, farms in the eastern regions of Thailand experienced PED outbreaks in 2012–2013 and the sequencing results demonstrated the emergence

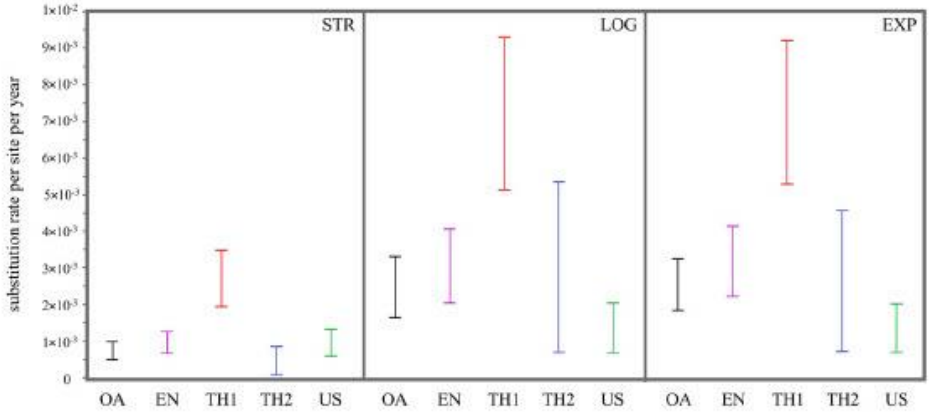


Fig. 2. Evolutionary rate comparison between overall (OA), endemic (EN), TH1 and TH2 of Thai dataset, and, US dataset. The results were show in substitution rate per site per year of strict (STR), uncorrelated relaxed logarithm (LOG) and exponential (EXP) prior clock.

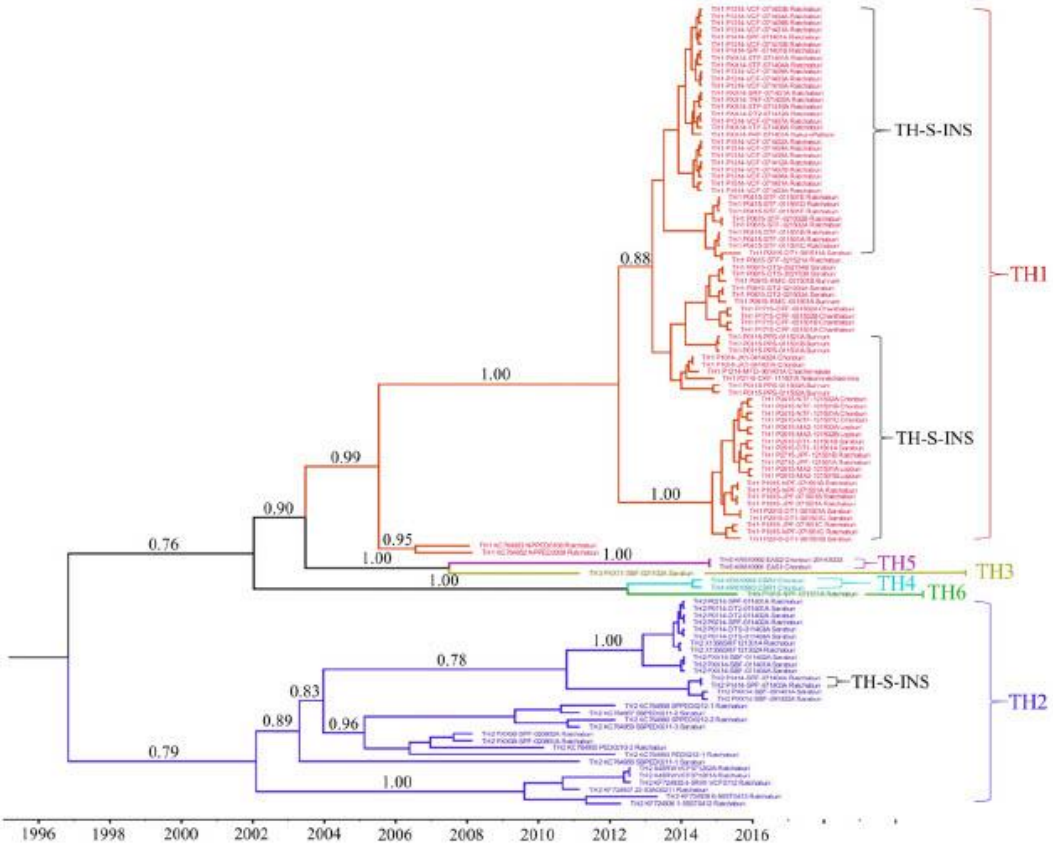


Fig. 3. Phylogenetic tree with time and their estimated divergences of PEDV in Thailand based on spike gene. Red, blue, yellow, aqua, purple and green colors were represented for TH1, TH2, TH3, TH4, TH5 and TH6, respectively. TH-S-INS labeled parenthesis located the sequences with the unique insertion. Node labels denote the time of estimated divergences whereas branch labels denote the posterior probability values.

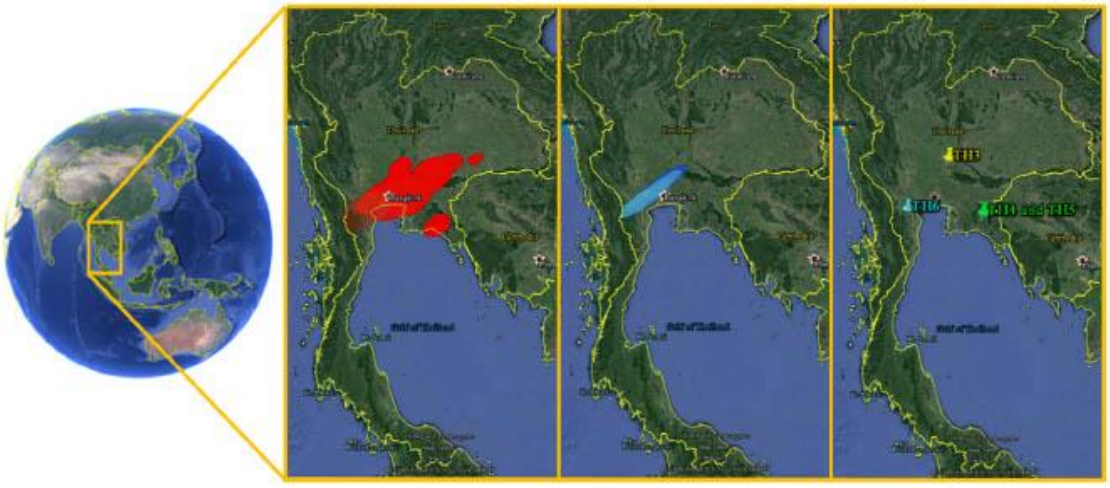


Fig. 4. Geographic distribution of TH1 (left panel), TH2 (center panel) and TH3–6 (right panel) estimated by the location and time of the outbreaks in Thailand during 2008–2015.

of the TH1 group which could be responsible for the outbreak. The spread of the TH1 group to the northeast region could be due to the transportation of culled sows.

3.4. Recombinant analysis

A total of 22 recombination events based on spike, with known parents were detected from RDP4 sequence analysis as shown in Fig. S1 (left panel). Recombination rate plot (Fig. S2) located 5 recombination sites (cutoff with > 0.01 Rho/bp), position 1–117 (0.06131 Rho/bp), position 1048–1206 (0.05912 Rho/bp), position 2224–2266 (0.02536 Rho/bp), position 3070–3116 (0.01947 Rho/bp), and position 4102–4142 (0.06236). The recombination events and their patterns are show in Fig. S1 (right panel).

Recombination analysis results of PXX11-SBF-021102A obtained from RDP and CHIMERA are different. The RDP result showed that this sequence is the minor parent, and together with X45RWVCF071202A, they were recombined into SBPED0211-1 whereas the CHIMERA show that it is the recombinant sequence of X1356SRF121302A and SBPED0211-1. In the other hand, SBPED0211-1 has a highest potential evidence of recombination since it can be detected by RDP, GENECONV and BOOTSCAN, and SiScan. Recombination event 16 (Fig. S1) were occurred in the same farm experienced with several outbreaks.

4. Discussion

In this study, the genetic diversity and evolutionary dynamics of PEDV samples collected between 2008 and 2015 in Thailand were investigated. Based on the phylogenetic analysis of complete spike gene, several important findings involving the genetic evolution of PEDV in Thailand were unveiled. The findings included the multiple introductions of genetically distinct PEDV variants influencing the development of Thai PEDV strains. Thai endemic PEDV strains forming clades separated from other countries and undergoing high mutation rate. One mechanism involving in the high mutation rate was the recombination.

The multiple introduction included the presence of PEDV in G1a and G2b subclades. The introduction of G1a variant in Thailand was not surprised. A previous study demonstrated that those PEDV stains in G1a group were vaccine-like stains (Cheun-Arom et al. 2015). The emergence of PEDV variants in this group was due to the heavy use of modified live PEDV vaccines, both modified live (MLV) and killed virus (KV) vaccines during 2013–2014. At present, the use of MLV and KV in Thailand drastically decreased compared to 2013–2014. This was due to the efficacy of intramuscular vaccines in which provided a limited degree of success against PED. In addition, MLV was reported to cause PEDV outbreaks with mild clinical disease (Cheun-Arom et al. 2015). The emergence of a US-like variant as shown in G2b group (TH6), was indeed intriguing and this is the first report showing the emergence of US-like PEDV in Thailand. TH6 is the latest subgroup in which contained the new introduction of the exotic PEDV strain (P1915-NPF-071511A) in 2015 that belong to G2b or US-like subgroup. This sequence share 100% similarity with NIG-1/JPN/2014, GNM-2/JPN/2014, MIE-1/JPN/2014, PC21A, KUIDL-PED-2014-002, KUIDL-PED-2014-007, USA/Ohio75/2013, USA/Ohio69/2013, USA/Ohio60/2013, and USA/Indiana/17846/2013, suggesting the first US-like strain in Thailand. However, the abilities of this strain to develop into the endemic strain in the Thai swine industry was still questionable and further investigation is required to monitor the development of these exotic strains.

Whether or not multiple introduction influences the strain development of Thai PEDV in the study was not known. However, although multiple introduction was evident, those strains were yet to form their own clusters. In contrast, Thai PEDV developed its own cluster separated from other countries. PEDV variants that has been dominant in Thai swine farms are strains in G2a group and can be further divided into 2 subgroups including TH1 and TH2. We found that Thailand strains are unique and underwent mutation and recombination with some other strains which resulted in pandemic outbreak between 2014 and 2015. We identified the sequence insertion which is unique only in Thailand strains (TH-S-INS) and was mainly detected in TH1 subgroup particularly the samples in the outbreaks between 2014 and 2015. This insertion was also detected in 2 samples of TH2 subgroup which was sampled in 2014 and no sample that belong to subgroup TH2 was detected in 2015. These suggested that the strains in TH1 might underwent positive selection pressure somehow and some of them experienced an insertion mutation event as described in the result. The first sample which containing this insertion of TH2 was P1414-SPF-071403A, and, was collected in July 2014 whereas another 27 samples from 8 outbreaks which collected at the same month are TH1. P1314-VCF-071401A is the represent sequence of the samples that containing this insertion of TH1 at 99% identity.

According to the evolutionary rate analysis, we found that the overall group of Thailand sequences (TH1–6; n = 115) evolved with relatively greater substitution rate compared to the US sequences when estimating under lognormal and exponential relaxed clock models. The US sequences showed higher substitution rate under the strict clock model; however, this rate was less than that of Thailand endemic group (TH1–2; 109). TH1 subgroup showed distinctly higher evolutionary rate than other analyzed groups. Its substitution rate was over than two fold compared to the endemic or TH2 group and was over than 5 fold compared to the US group as determined by uncorrelated relaxed-clock models. This event may suggests that there is any positive pressure on this subgroup.

Besides influenced by the multiple introduction, the relatively greater evolutionary rate of Thai PEDV was partially due to recombination as supported by the results of the recombination analysis. In comparison between TH1 and the TH2 groups, TH1 is more in population parted in the events whereas TH2 is more in number of recombination evidences. These results may explain that TH1 recombinants were chosen to be the subgroup that suitable to cause an outbreaks whereas TH2 is keep undergoing recombination but not usually cause an outbreaks. TH2 might be the predominant group which could adapt to survive and persist in the host without triggering symptom development, therefore some were not included in the study. On the other hand, TH2 might not be preferred for proliferation in the host. However, we did not perform active surveillance to take the samples from subclinical herds so we could not make a final conclusion. For TH3, we did further analysis with the dataset that using in phylogenetic analysis, and, there is possible that PXX11-SBF-021102A is the recombinant of NPPED2008 (KC764952) and Chinju99 (AV167585) of G2 and G1 genogroup as major and minor parents, respectively. This information supported why TH3 subgroup is locate between genogroup 1 and 2 (Fig. 1), and, very close and share the same estimated divergence to TH5 subgroup (Fig. 3). The results of geographic distribution show that the western part of Thailand (mainly in Ratchaburi) is the endemic area of most strains, however, TH2 seem to be limited in these area whereas TH1 is spreading from west-central-east of Thailand. Another new introductions (TH3–6) were denoted by the pinpoint since they could not performed analysis. These results together with previous described suggested that TH1 is becoming the predominant subgroup which endemic in Thailand since 2014.

5. Conclusion

Thai PEDV was genetically diverse and the genetic development has been influenced by the multiple introduction of exotic strains. Such introductions were not developed into epidemic stage. In contrast, Thai PEDV strains developed their own subgroups separated from other countries and underwent relatively higher evolutionary than the US sequences. There was an evidence of intra-recombination between sub-groups that can be one factor to accelerate the evolutionary rate of Thai PEDV. Samples in the sub-group with high

evolutionary rate become dominant causing not only geographic separation, but also influence the development of other subgroup as the unique insertion could be found in both TH1 and TH2 whereas the identity between subgroup is >97%. We did provided the information of P1314-VCF-071401A, P1414-SPF-071403A, PXX11-SBF-021102A, P1915-NPF-071511A and PXX08-SPF-020801A on GenBank® as following accession number: KX981897, KX981898, KX981899, KX981900 and KY000559, respectively. PXX08-SPF-020801A was the first strain in TH2 genogroup found in 2008.

Author contributions

C.J.S. and D.N. conceived and designed the experiments; C.J.S., G.T., P.K., T.T. and J.P. performed the experiments; C.J.S. and D.N. analyzed the data; C.J.S., T.T. and A.T. contributed reagents/materials/analysis tools; C.J.S. and D.N. composed the manuscript.

Conflicts of interest

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Acknowledgments

This research was financially supported by the National Research Council of Thailand and Agricultural Research Development Agency (Public organization; PRP5805021650). Partial funding was provided by Special Task Force for Activating Research (STAR), Swine Viral Evolution and Vaccine Research (SVEVR), Chulalongkorn University (GSTAR 59-013-31-007). The first author is supported by the Royal Golden Jubilee PhD program (PHD/0204/2557) Scholarship of the Thailand Research Fund (TRF) and the Ratchadaphiseksomphot Endowment Fund 2013 of Chulalongkorn University (CU-56-527-HR).

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi: 10.1016/j.meegid.2017.02.014. These data include the Google maps of the most important areas described in this article.

References

Bielejec, F., Rambaut, A., Suchard, M.A., Lemey, P., 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 27, 2910–2912.
Bosch, B.J., van der Zee, R., de Haan, C.A., Rottier, P.J., 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J. Virol. 77, 8801–8811.
Chang, S.H., Bae, J.L., Kang, T.J., Kim, J., Chung, G.H., Lim, C.W., Laude, H., Yang, M.S., Jang, Y.S., 2002. Identification of the epitope region capable of inducing neutralizing antibodies against the porcine epidemic diarrhea virus. Mol. Cells 14, 295–299.
Chasey, D., Cartwright, S., 1978. Virus-like particles associated with porcine epidemic diarrhoea. Res. Vet. Sci. 25, 255.
Chen, Q., Li, G., Stasko, J., Thomas, J.T., Stensland, W.R., Pillatzki, A.E., Gauger, P.C., Schwartz, K.J., Madson, D., Yoon, K.J., Stevenson, G.W., Burrough, E.R., Harmon, K.M., Main, R.G., Zhang, J., 2014. Isolation and characterization of porcine epidemic diarrhea viruses associated with the 2013 disease outbreak among swine in the United States. J. Clin. Microbiol. 52, 234–243.
Cheun-Arom, T., Temeeyasen, G., Srijangwad, A., Tripipat, T., Sangmalee, S., Vui, D.T., Chuanasa, T., Tantituvanont, A., Nilubol, D., 2015. Complete genome sequences of two genetically distinct variants of porcine epidemic diarrhea virus in the eastern region of Thailand. Genome Announc. 3, e00634-00615.
Cruz, D.J., Kim, C.J., Shin, H.J., 2006. Phage-displayed peptides having antigenic similarities with porcine epidemic diarrhea virus (PEDV) neutralizing epitopes. Virology 354, 28–34.
Cruz, D.J., Kim, C.J., Shin, H.J., 2008. The GPRLQPY motif located at the carboxy-terminal of the spike protein induces antibodies that neutralize porcine epidemic diarrhea virus. Virus Res. 132, 192–196.
Dennis, H., Maria, J., Anja, P., Mathias, R., Julia, S., Valerij, A., Sandra, B., Anne, P., Horst, S., Martin, B., Dirk, H., 2015. Comparison of porcine epidemic diarrhea viruses from Germany and the United States, 2014. Emerg. Infect. Dis. 21, 493.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 1–8.
Drummond, A.J., Nicholls, G., Rodrigo, A., Solomon, W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161.
Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185–1192.
Drummond, A.J., Ho, S., Phillips, M., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol.
Duarte, M., Laude, H., 1994. Sequence of the spike protein of the porcine epidemic diarrhoea virus. J. Gen. Virol. 75 (Pt 5), 1195–1200.
Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.
Kocherhans, R., Bridgen, A., Ackermann, M., Tobler, K., 2001. Completion of the porcine epidemic diarrhoea coronavirus (PEDV) genome sequence. Virus Genes 23, 137–144.
Lee, C., 2015. Porcine epidemic diarrhea virus: an emerging and re-emerging epizootic swine virus. Virol. J. 12, 193.
Lee, D.-K., Park, C.-K., Kim, S.-H., Lee, C., 2010. Heterogeneity in spike protein genes of porcine epidemic diarrhea viruses isolated in Korea. Virus Res. 149, 175–182.
Lee, S., Park, G.-S., Shin, J.-H., Lee, C., 2014. Full-genome sequence analysis of a variant strain of porcine epidemic diarrhea virus in South Korea. Genome Announc. 2.
Li, W., Li, H., Liu, Y., Pan, Y., Deng, F., Song, Y., Tang, X., He, Q., 2012. New variants of porcine epidemic diarrhea virus, China, 2011. Emerg. Infect. Dis. 18, 1350–1353.
Lin, C.N., Chung, W.B., Chang, S.W., Wen, C.C., Liu, H., Chien, C.H., Chiou, M.T., 2014. US-like strain of porcine epidemic diarrhea virus outbreaks in Taiwan, 2013–2014. J. Vet. Med. Sci. 76, 1297–1299.
Lozano-Fuentes, S., Elizondo-Quiroga, D., Farfan-Ale, J.A., Loroño-Pino, M.A., Garcia-Rejon, J., Gomez-Carro, S., Lira-Zumbardo, V., Najera-Vazquez, R., Fernandez-Salas, I., Calderon-Martinez, J., Dominguez-Galera, M., Mis-Avila, P., Morris, N., Coleman, M., Moore, C.G., Beaty, B.J., Eisen, L., 2008. Use of Google Earth™ to strengthen public health capacity and facilitate management of vector-borne diseases in resource-poor environments. Bull. World Health Organ. 86, 718–725.
Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 1.
Masuda, T., Murakami, S., Takahashi, O., Miyazaki, A., Ohashi, S., Yamasato, H., Suzuki, T., 2015. New porcine epidemic diarrhoea virus variant with a large deletion in the spike gene identified in domestic pigs. Arch. Virol. 160, 2565–2568.
Ojkic, D., Hazlett, M., Fairles, J., Marom, A., Slavic, D., Maxie, G., Alexandersen, S., Pasick, J., Alsop, J., Burlatschenko, S., 2015. The first case of porcine epidemic diarrhea in Canada. Can. Vet. J.: La revue veterinaire canadienne 56, 149–152.
Park, S.J., Moon, H.J., Luo, Y., Kim, H.K., Kim, E.M., Yang, J.S., Song, D.S., Kang, B.K., Lee, C.S., Park, B.K., 2008. Cloning and further sequence analysis of the ORF3 gene of wild- and attenuated-type porcine epidemic diarrhea viruses. Virus Genes 36, 95–104.
Park, S.J., Kim, H.K., Song, D.S., Moon, H.J., Park, B.K., 2011. Molecular characterization and phylogenetic analysis of porcine epidemic diarrhea virus (PEDV) field isolates in Korea. Arch. Virol. 156, 577–585.
Pensaert, M., De Bouck, P., 1978. A new coronavirus-like particle associated with diarrhea in swine. Arch. Virol. 58, 243–247.
Rambaut, A., Drummond, A.J., 2013. Tracer 1.6 [Computer Program].
Rambaut, A., Drummond, A.J., 2014. FigTree 1.4.2 [Computer Program].
Sun, D.B., Feng, L., Shi, H.Y., Chen, J.F., Liu, S.W., Chen, H.Y., Wang, Y.F., 2007. Spike protein region (aa 636789) of porcine epidemic diarrhea virus is essential for induction of neutralizing antibodies. Acta Virol. 51, 149–156.
Sun, D., Feng, L., Shi, H., Chen, J., Cui, X., Chen, H., Liu, S., Tong, Y., Wang, Y., Tong, G., 2008. Identification of two novel B cell epitopes on porcine epidemic diarrhea virus spike protein. Vet. Microbiol. 131, 73–81.
Sun, R.Q., Cai, R.J., Chen, Y.Q., Liang, P.S., Chen, D.K., Song, C.X., 2012. Outbreak of porcine epidemic diarrhea in suckling piglets, China. Emerg. Infect. Dis. 18, 161–163.
Sun, M., Ma, J., Wang, Y., Wang, M., Song, W., Zhang, W., Lu, C., Yao, H., 2015. Genomic and epidemiological characteristics provide new insights into the phylogeographical and spatiotemporal spread of porcine epidemic diarrhea virus in Asia. J. Clin. Microbiol. 53, 1484–1492.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol.
Temeeyasen, G., Srijangwad, A., Tripipat, T., Tipsombatboon, P., Piriyapongsa, J., Phoolcharoen, W., Chuanasa, T., Tantituvanont, A., Nilubol, D., 2014. Genetic diversity of ORF3 and spike genes of porcine epidemic diarrhea virus in Thailand. Infect. Genet. Evol. 21, 205–213.
Theuns, S., Conceição-Neto, N., Christiaens, I., Zeller, M., Desmarets, L.M.B., Roukaerts, I.D.M., Acar, D.D., Heylen, E., Matthijnssens, J., Nauwynck, H.J., 2015. Complete genome sequence of a porcine epidemic diarrhea virus from a novel outbreak in Belgium, January 2015. Genome Announc. 3, e00506–e00515.
Vui, D.T., Thanh, T.L., Tung, N., Srijangwad, A., Tripipat, T., Chuanasa, T., Nilubol, D., 2015. Complete genome characterization of porcine epidemic diarrhea virus in Vietnam. Arch. Virol. 160, 1931–1938.

APPENDIX B

The phylodynamics of emerging porcine deltacoronavirus in Southeast Asia

Transboundary and Emerging Diseases

September 2022, Volume 69, Issue 5, Pages: 2816-2827

Christopher James Stott, Kanokon Sawattrakool, Kepalee Saeng-Chuto, Angkana Tantituvanont and Dachrit Nilubol

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Check for updates

Received: 29 March 2021 | Revised: 17 November 2021 | Accepted: 5 December 2021

DOI: 10.1111/tbed.14434

Transboundary and Emerging Diseases WILEY

ORIGINAL ARTICLE

The phylodynamics of emerging porcine deltacoronavirus in Southeast Asia

Christopher James Stott[1,2] | Kanokon Sawattrakool[1] | Kepalee Saeng-Chuto[1] | Angkana Tantituvanont[3] | Dachrit Nilubol[1]

[1] Department of Veterinary Microbiology, Faculty of Veterinary Science, Chulalongkorn University, Bangkok, Thailand

[2] Akkhraratchakumari Veterinary College, Walailak University, Nakhon Si Thammarat, Thailand

[3] Department of Pharmaceutics and Industrial Pharmacy, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok, Thailand

Correspondence
Dachrit Nilubol, Department of Veterinary Microbiology, Faculty of Veterinary Science, Chulalongkorn University, Bangkok 10330, Thailand.
Email: dachrit@gmail.com

Funding information
Chulalongkorn University (Fundamental fund), Grant/Award Numbers: 65413100200013, CUFRB65_food(26)190_31_09); Special Task Force for Activating Research (STAR); Swine Viral Evolution, and Vaccine Research (SVEVR); Chulalongkorn University, Grant/Award Number: GSTAR 59-013-31-007; Royal Golden Jubilee PhD program, Grant/Award Number: PHD/0204/2557; Scholarship of the Thailand Research Fund

Abstract

Porcine deltacoronavirus (PDCoV), a recently emerging pathogen, causes diarrhoea in pigs. A previous phylogenetic analysis based on spike genes suggested that PDCoV was divided into three different groups, including China, the United States, and Southeast Asia (SEA). SEA PDCoV, however, is genetically separated from China and the United States but shares a common ancestor. Its origin and evolution have yet been identified. Herein, phylodynamic analyses based on the full-length genome were performed to investigate the origin and evolution of SEA PDCoV. In the study, 18 full-length genome sequences of SEA PDCoV identified in 2013–2016 together with PDCoV from other regions were used in analyses. The results demonstrated that PDCoV was classified into two genogroups including G1 and G2. G1 is further evolved into G1a (China) and G1b (US). G2 (SEA) group is further evolved into three clades, including SEA-1 (Thailand), SEA-2 (Vietnam) and SEA-2r (Vietnam recombinant) clades. The time to the most recent common ancestor (MRCA) of global PDCoV was estimated to be approximately 1989–1990 and possibly have been circulated in SEA more than a decade. SEA PDCoV is genetically diverse compared to China and U.S. PDCoV. The substitution rate of SEA PDCoV was lower than those of China and the United States, but the recombination rate of SEA was higher. Recombination analyses revealed four potential recombinant events in SEA PDCoV, suggesting that they were derived from the same ancestor of China PDCoV. The SEA-2r subgroup was potentially recombinant between SEA-2 and U.S. strains. In conclusion, the major mechanisms driving the complex evolution and genetic diversity of SEA PDCoV were multiple introductions of exotic PDCoV strains followed by recombination.

KEYWORDS
coronavirus, epidemiology, evolution, genetic, Lao PDR, mutation, orcine deltacoronavirus, SEA, Thailand, Vietnam

1 | INTRODUCTION

Porcine deltacoronavirus (PDCoV) causes a disease characterized by watery diarrhoea and villous atrophy (Chen et al., 2015), which currently is a threat to the swine industry worldwide. PDCoV is a newly emerged pathogen which was initially detected during the molecular surveillance of coronavirus in Hong Kong in 2012 (P. C. Woo et al., 2012). Although it was first detected, there was no association with clinical disease. The first evidence of PDCoV causing the disease was first evident in Ohio, USA in 2014 (D. Marthaler et al., 2014). Soon

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

after its emergence in the United States, PDCoV-associated disease was subsequently reported in several countries, including China, South Korea, Thailand, Lao PDR in 2015 (A. Lorsirigool et al., 2016; Janetanakit et al., 2016; Lee & Lee, 2014; Marthaler et al., 2014; P. C. Woo et al., 2012; Saeng-Chuto et al., 2017; Song et al., 2015), Vietnam (Saeng-Chuto et al., 2017) and Mexico (Pérez-Rivera et al., 2019). Although clinical diseases caused by PDCoV are similar to that of porcine epidemic diarrhoea (PED), mortality is lower. In addition, the virus is detected concurrently with PED virus. However, later studies suggest that PDCoV might become a highly pathogenic virus since researchers worldwide reported that the virus can solely trigger clinical signs, and the mortality rate is over 80% (Kwonil Jung et al., 2015; Saeng-Chuto et al., 2020; Song et al., 2015; Y. Zhao et al., 2019). This is in contrast with previous findings.

PDCoV is an enveloped, single-stranded, positive-sense RNA virus in the genus *Deltacoronavirus*, family *Coronaviridae*. Its genome is approximately 25 kb in length (P. C. Woo et al., 2012). [Correction added on 23 February 2021, after first online publication: 25, 000 kb was changed to 25 kb in the preceding sentence.] The genome arrangement is in the order of 5′ untranslated region (UTR), open reading frame 1a/1b (ORF 1a/1b), spike (S), envelope (E), membrane (M), non-structural protein 6 (NS6), nucleocapsid (N), non-structural protein 7 (NS7) and 3′ UTR (P. C. Woo et al., 2012). S gene encodes S protein comprising two domains, S1 and S2. S1 domain functions in an important role involving attachment between the virus and host receptor. S2 domain facilitates membrane fusion (Shang et al., 2018). E and M genes encode transmembrane proteins E and M, respectively, functioning in envelope formation and virus release (P. C. Y. Woo et al., 2010).

Based on previous phylogenetic analyses, PDCoV has evolved into three separated groups, including China, the United States, and Southeast Asia (SEA) (Saeng-Chuto et al., 2020). China PDCoV comprises PDCoV isolates detected in China and Hong Kong. U.S. PDCoV is widespread comprising isolates detected in the United States, South Korea, Japan and Mexico. SEA PDCoV is unique consisting of only PDCoV isolates detected in Southeast Asian countries including Thailand, Lao PDR and Vietnam (Pérez-Rivera et al., 2019; Saeng-Chuto et al., 2020). Although SEA PDCoV clustered separately from China and U.S. PDCoV, viruses in these three clusters originated from a common ancestor.

Although clinical diseases caused by PDCoV were first reported in the United States in 2014, retrospective investigations from several countries including China and the United States reported the presence of PDCoV in intestinal samples of pigs affected with watery diarrhoea as early as 2004 and 2013, respectively (McCluskey et al., 2016; Pérez-Rivera et al., 2019; Saeng-Chuto et al., 2017, 2020; Sinha et al., 2015). In Thailand, PDCoV was first emerged in 2015 (Janetanakit et al., 2016). However, a subsequent retrospective study reported that PDCoV was present in Thailand as early as 2013 (Saeng-Chuto et al., 2020). This suggests that PDCoV might circulate in swine farms, but clinical signs might confuse other highly pathogenic pathogens, including PEDV.

In addition, recombinant viruses were detected in SEA PDCoV (Saeng-Chuto et al., 2020). In Vietnam, recombinant PDCoV detected in 2020, formed a new subgroup of SEA genogroup. These chimeric PDCoVs were considered as the results of the recombination between Vietnam and U.S. isolates (Saeng-Chuto et al., 2020). Based on these findings, SEA PDCoV appears to have a complex evolution compared to that of U.S. and China PDCoV. Questions have been raised as to how PDCoV was introduced into this SEA region and where the origin of the virus was. Phylogeography and molecular evolutionary analyses of SEA PDCoV are needed to fulfil the missing information. This information would provide a link to the global PDCoV through available information. Therefore, in this study, phylodynamic analyses were performed using Bayesian probability-based software and BEAST packages (Drummond & Rambaut, 2007).

2 | MATERIALS AND METHODS

2.1 | PDCoV detection and sequences analysis

2.1.1 | Samples

Eighteen intestinal samples were collected from pig farms in Thailand, Vietnam, and Lao PDR experiencing PDCoV outbreaks during 2013–2016. Following collection, intestinal samples were minced and suspended in PBS to generate a 10% (w/v) homogenate. These samples were then clarified by centrifugation. The supernatant was filtered through 0.22 μm filters and stored at −80°C until use. The clarified supernatants were subjected to PDCoV detection using previously described PCR methods with PDCoV specific primers for M and N genes (Wang et al., 2014).

2.1.2 | The full-length genome characterization

Positive PDCoV samples by PCR were selected and further characterized for the full-length genomes using previously described methods (Lorsirigool et al., 2017; Saeng-Chuto et al., 2017). Briefly, total RNA was extracted from supernatants using Nucleospin RNA Virus kit (Macherey-Nagel Inc., PA, USA). cDNA was synthesized from extracted RNA using M-MuLV Reverse Transcriptase (New England BioLabs Inc., MA, USA). PCR was performed using 26 specific primers (Supporting Information 1). Specific PCR bands were purified by a Nucleospin® Gel and PCR Clean-up kit (Macherey-Nagel Inc.). The purified PCR products were subjected for sequencing at First BASE Laboratories Inc. (Selangor, Malaysia) using an ABI Prism 3730XL DNA sequencer (Applied Biosystems Inc., Carlsbad, CA, USA).

2.1.3 | Phylogenetic tree and genetic analyses

Eighteen PDCoV were characterized for their full-length genome sequences in the present study. Sequencing data are deposited in GenBank under accession numbers listed in Supporting Information 2. To determine the genetic relationship between 18 SEA PDCoV isolates and other PDCoV isolates, nucleotide (nt) and amino acid (aa)

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sequences were aligned using MAFFT (Katoh et al., 2002), together with 100 other PDCoV isolates available in GenBank (Supporting Information 2). Nucleotide identity was performed using the sequence identity matrix function implemented in BioEdit software (Hall, 1999). The phylogenetic tree based on the full-length genome and spike gene sequences were separately constructed using the Tamura-Nei model, with the gamma-distribution and 1000 bootstraps, using maximum likelihood in MEGA-X (Kumar et al., 2018). The model used for the maximum likelihood tree was chosen by the model selection function in MEGA-X.

3 | RECOMBINATION ANALYSIS

Recombination events within 18 full-length genome sequences in this study together with the 100 reference PDCoV isolates were analyzed using RDP4. Recombination events are strongly supported by at least six of seven detection tools available in RDP4 automated analysis, including RDP (D. Martin & Rybicki, 2000), GENECOV (Padidam et al., 1999), BOOTSCAN (Salminen et al., 1995), MAXCHI (Smith, 1992), CHIMAERA (Posada & Crandall, 2001), SISCAN (Gibbs et al., 2000) and 3SEQ (Boni et al., 2007), with potential recombination signals accepted. Major and minor parents and breakpoints were also detected, and recombination fragments were generated. The recombination rate plot was used to compute the recombination rate (Auton & McVean, 2007). Genes were separately analyzed and genes with recombination incidences were saved as the partition according to the recombination breakpoints to apply for further analyses.

4 | MOLECULAR EVOLUTIONARY ANALYSIS

Molecular evolutionary analyses were conducted using 18 full-length genome sequences of SEA PDCoV from the present study together with 1000 full-length genomes of other PDCoV available in GenBank. The nucleotide (nt) sequences of PDCoV were aligned using MAFFT (Katoh et al., 2002) and then modified using BioEdit (Hall, 1999). The datasets were prepared and generated in BEAUTI, including the full-length genome, spike, envelope, membrane, nucleocapsid, NS6 and NS7. To reduce the recombinatorial biases in the analysis, the datasets with the recombination incidences (Martin et al., 2015), the partition were separated by the recombination breakpoints.

Molecular evolutionary rates and the estimated time to the most recent common ancestor (TMRCA) were estimated using BEAST (Drummond & Rambaut, 2007; Drummond et al., 2012; Ferreira & Suchard, 2008; Lemey et al., 2010; Pybus et al., 2012) and BEAGLE (Ayres et al., 2012). The general time-reversible model plus gamma distribution plus invariant sites (GTR+G+I) (Tavaré, 1986; Yang, 1994) was applied with Bayesian skyline tree prior (Drummond et al., 2002; Drummond, Rambaut, Shapiro & Pybus, 2005) and flexible local clock (FLC; Fourment & Darling, 2018). Collection dates were used for tip dates. If there were any missing data, the mid-values of the month or year were filled. The analysis was performed at least 500 million generations per run, with 10,000 states logged to qualified 200 ESS. The maximum clade credibility (MCC) tree was built by choosing common ancestor heights using TreeAnnotator in the BEAST package, 10% burn-in was applied, and the posterior probabilities and estimated divergences were estimated. The substitution rates were interpreted in Tracer 1.7.1 (Rambaut & Drummond, 2013) and compared with U.S. and China datasets, and the annotated phylogenetic tree was generated and displayed using FigTree 1.4.3 (Rambaut & Drummond, 2014).

5 | PHYLOGEOGRAPHIC ANALYSIS

For the phylogeographic analysis, the previously described dataset was used for analysis. The nucleotide sequences of the full-length genomes of PDCoV were aligned using MAFFT and then edited using BioEdit. The partition breakpoint was estimated by RDP4 to reduce the biases in the analysis. The phylogeography and distribution model was performed using BEAST 1.10.4 with BEAGLE. The GTR+G+I model was applied, and a thousand million states with logging every 10,000 states were performed using the FLC and the Bayesian skyline tree prior. Collection dates were used for tip dates, if there were any missing data, the mid-values of the month or year were filled. The GPS coordinated locations were applied using the location of states or provinces available from the information available in GenBank and previous reports. The data would be accepted in case of the effective sample size (ESS) of all parameters greater than 200.

The results, including the substitution rate and ESS of this study, were interpreted using Tracer 1.7.1. MCC trees were annotated by choosing the common ancestor heights using TreeAnnotator in the BEAST package. The chronogram, estimated divergences, posterior probability and 95% high population density (95% HPD) information were transformed into the phylogeographic tree in FigTree 1.4.3. The phylogeographic tree was then converted and displayed in Google Earth (Lozano-Fuentes et al., 2008) by using SPREAD 1.0.7 (Bielejec et al., 2011).

6 | POSITIVE SELECTION ANALYSIS

A genome scan for signals of natural selection was performed to determine the variability of selective pressure among different gene regions. Ratios of non-synonymous to synonymous nucleotide substitutions rates (dN/dS) were estimated for the entire alignment using the Synonymous Non-synonymous Analysis Program (SNAP) (Korber, 2000; Nei & Gojobori, 1986; Ota & Nei, 1994) to compare the codon selected among the global, SEA, U.S. and China codons. If there was an ambiguous nucleotide in the previously published sequences, the sequence was duplicated and replaced from the IUPAC nucleotide code to base nucleotides (R=A or G, Y = C or T, S = G or C, W = A or T, K = G or T, and M = A or C; Johnson, 2010). A dN/dS ratio > 1 indicated positive selection (Z. Yang & Bielawski, 2000). Insertion deletions (InDels) were additionally estimated and displayed.

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Full-length genome and spike gene length, GC contents, identity and similarity of porcine deltacoronavirus sequences used in this study (Global), Southeast Asia (SEA), the United States and China

Full-length genomes				
	Global	SEA	USA	China
Length (nt)	25,380–25,431	25,411–25,417	25,430–25,431	25,380–25,431
GC contents (%)	43.04–43.40	43.04–43.22	43.15–43.27	43.12–43.37
Intra-lineage identity	96.6–100	98–100	99.6–100	97.3–100
With SEA identity	–	–	97.0–98.1	96.6–98.2
With U.S. identity	–	97.0–98.1	–	98.1–99.3
With China identity	–	96.6–98.2	98.1–99.3	–
Spike genes				
	Global	SEA	USA	China
Length (nt)	3480–3483	3480–3483	3483	3480–3483
GC contents (%)	41.31–41.92	41.34–41.83	41.63–41.83	41.31–41.81
Intra-lineage identity (similarity)	94.9–100 (95–100)	96–100 (96.6–100)	99.5–100 (99.2–100)	95.9–100 (95–100)
With SEA identity (similarity)	–	–	95.7–99.4 (96.4–99.2)	95–99.2 (95–99.1)
With U.S. identity (similarity)	–	95.7–99.4 (96.4–99.2)	–	97.3–98.9 (97–99.4)
With China identity (similarity)	–	95–99.2 (95–99.1)	97.3–98.9 (97–99.4)	–

[Correction added on 23 February 2021, after first online publication: The full-length genome sequences of Global, SEA, US and China PDCoV were changed to "25,380–25,431, 25,411–25,417, 25,430–25,431 and 25,380–25,431 nucleotides (nt)".]

7 | RESULTS

7.1 | PDCoV detection and sequences analysis

The full-length genome sequences of SEA, U.S. and China PDCoV were 25,411—25,417, 25,430—25,431 and 25,380—25,431 nucleotides (nt) in length, respectively. [Correction added on 23 February 2021, after first online publication: The full-length genome sequences of SEA, US and China PDCoV were changed to "25,411—25,417, 25,430—25,431 and 25,380—25,431 nucleotides (nt)" in the preceding sentence.] The genome organization is similar to that of all previously reported PDCoV genomes: 5′-ORF1a/1b-S-E-M-NS6-N-NS7—3′ (Table 1). UTRs are present at both ends (5t′ UTR, nt 535 and 3t′ UTR, nt 725). The lengths of ORF1a/1b, S, E, M and N genes were 18,788, 3480-3483, 252, 654 and 1,029 nt, respectively. S gene sequences of SEA, U.S. and China PDCoV were 3480–3483, 3483 and 3480-3483 nucleotides (nt) in length, respectively. Phylogenetic tree analysis based on the full-length genome sequences demonstrated that PDCoV isolates were divided into two groups, G1 and G2 (Figure 1). G1 is further divided into G1a (China) and G1b (U.S.). G1a and G1b comprised isolates mainly reported in China and the United States, respectively. In contrast, isolates mainly reported in SEA countries are clustered in G2 (SEA) group and further divided into SEA-1, SEA-2 and SEA-2r. The genetic identities between G1 and G2, based on the full-length genome and spike gene, are shown in Table 1. The intralineage nucleotide identities of G2 (SEA), G1a (US) and G1b (China) were 98.00–100.00%, 99.60–100.00% and 97.30–100.00%, respectively.

Deletions/insertions (InDels) are present in spike and membrane genes. InDels found in spike gene of each genogroup caused its difference from 3480 to 3483 nt, while the InDels found in the membrane (CHN/Sichuan/2017; MK211169) at the 280th–291st position, causing its difference from 642 to 654 nt.

7.2 | Recombination analysis

Recombination analysis based on the full-length genome, as determined by at least six detection methods available in RDP4, demonstrated that five isolates in SEA PDCoV were recombinants. These isolates include PDCoV/Binh21/2015 (KX834352) and PDCoV/HaNoi6/2015 (KX834351) and P30_15_VN1215 (MH118332), P2_13_ST2_0313 (KX361344) and PDCoV/CBR-3/2016 from Vietnam and Thailand, respectively (Figure 2). High potential events from Vietnam are PDCoV/Binh21/2015 and P30_15_VN_1215, which are recombinant between China PDCoV, CHN-GD16—03 (KY363867) and CHN/Tianjin/2016 (KY065120), and between Vietnam and US PDCoV, P12_14_VN_0814 (MH700628) and USA/Minnesota159/2014 (KR265859). In contrast, Thai PDCoV P2_13_ST2_0313 is the recombinant between Hong Kong HKU15—155 (JQ065043) and China PDCoV CH/Sichuan/S27/2012 (KT266822).

The intralineage recombination rate (theta) based on the full-length genomes of SEA, China and the United States was estimated at 8.784×10^{-3}, 1.993×10^{-2} and 3.632×10^{-3}, respectively. The recombination rates based on spike genes of SEA and China were estimated at 2.065×10^{-2} and 2.715×10^{-2}, respectively (Supporting Information 3).

The recombination analysis suggests that the global dataset of full-length genomes comprises 57 breakpoints, including three partitions in the 5′UTR, one partition between the 5′UTR end and ORF1a, 29

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

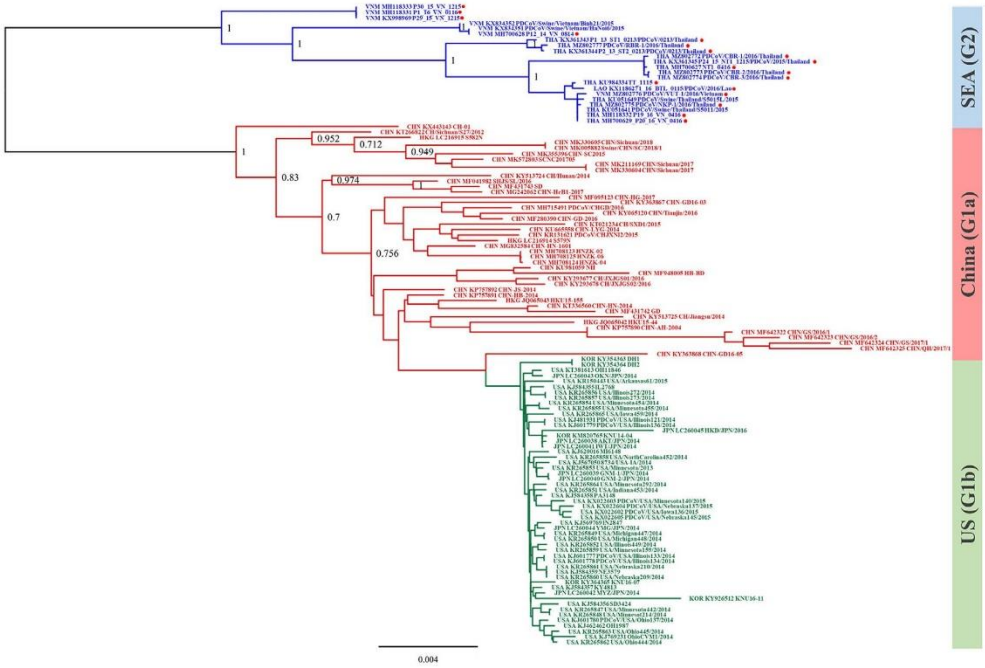


FIGURE 1 Phylogenetic tree of porcine deltacoronavirus (PDCoV) conducted by maximum likelihood in MEGA-X showing that the virus could be classified into three genogroups: Southeast Asia (SEA) (blue), China (red), and the United States (green). The number at the nodes denotes the probability of performing by 1000 bootstraps

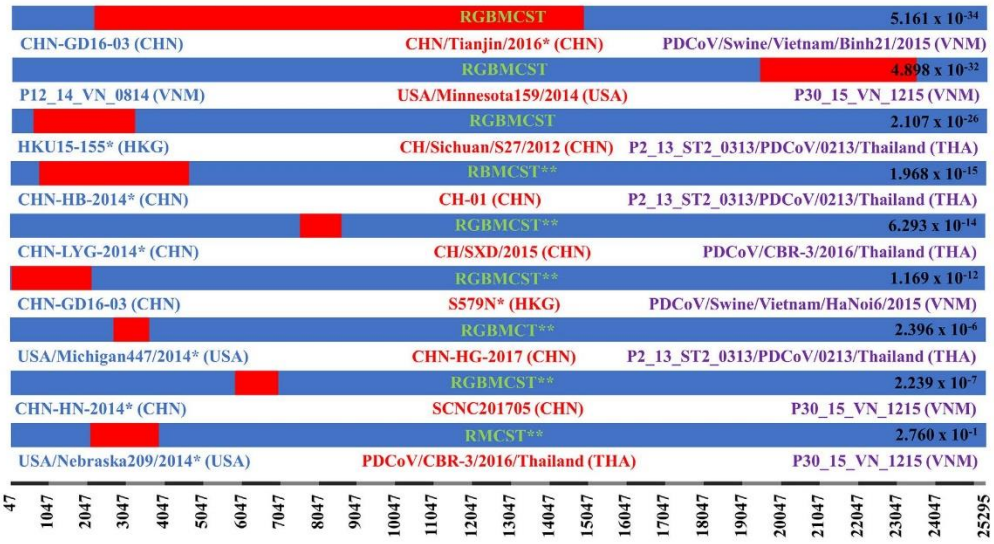


FIGURE 2 Recombination pattern of porcine deltacoronavirus (PDCoV) full-length genome determination by RDP4 with at least six of seven detection tools in the program. Blue shows the major parent, red shows the minor parent, purple shows the recombinant parent and green shows the initial character of the detection tool available in RDP4 (RDP (R), GENECOV (G), BOOTSCAN (B), MAXCHI (M), CHIMAERA (C), SISCAN (S) and 3SEQ (T)). The grey/black ruler shows the length of the analyzed sequences. The average *p*-value from RDP is displayed in the right panel, and the asterisks denote the partial parent

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

partitions at ORF1a, one partition between ORF1a and ORF1b, 14 partitions in ORF1b, one partition between ORF1b and the spike, six partitions in the spike, one partition among the spike, envelope and membrane, one partition among the membrane, NS6, nucleocapsid, and NS7, and one partition among the nucleocapsid, NS7, and 3′UTR (Supporting Information 4a). The breakpoints were also found at the spike gene, including 11 partitions (Supporting Information 4b).

7.3 | Molecular evolutionary analysis

The evolutionary analysis demonstrated that PDCoV could be grouped into two genogroups, G1 and G2, based on the full-length genome (Figure 3a). G1 further evolved into two subgroups, including G1a (China) and G1b (U.S.). In contrast, G2 further evolved into two subgroups, including G2a (SEA-1 or Thailand) and G2b (SEA-2 or Vietnam). The G2a (SEA-1 or Thailand) subgroup was further divided into two groups: eastern and western groups. Notably, recombinants from Vietnam clustered separately in a new subgroup, SEA-2r.

TMRCAs of G1 and G2 were estimated at 1989.17 with a 95% HPD range from 1986.81 to 1991.50, whereas the divergence of G2 was estimated at 1999.62 (1997.86–2001.32; 95% HPD) and diverged to SEA-1 and SEA-2 and SEA-2r clades. The divergence of the SEA-2 and SEA-2r clades was estimated at 2002.88 (2001.40–2004.56; 95% HPD), and the divergence of the Thailand subgroup was estimated at 2006.87 (2005.78–2007.91; 95% HPD).

Interestingly, Vietnam strains existed in all subgroups. The first strain was in SEA-1, sharing a monophyletic group with Thailand and Lao PDR strains. The second subgroup was SEA-2, which shared the same ancestor with SEA-1, whereas SEA-2r formed a paraphyletic group from SEA-1 and SEA-2.

Evolutionary rates of PDCoV are shown in Figure 3b. Based on the full-length genome, the substitution rate was 6.13×10^{-4} substitutions/site/year (s/s/y). In the comparison of the encoding genes, the envelope was the highest (1.28×10^{-3} s/s/y), followed by NS7 (9.93×10^{-4} s/s/y), nucleocapsid (8.26×10^{-4} s/s/y), NS6 (8.19×10^{-4} s/s/y), spike (7.59×10^{-4} s/s/y), and membrane (4.94×10^{-4} s/s/y).

For the individual datasets, the China dataset had the highest substitution rate of full-length genome (6.57×10^{-4} s/s/y), envelope (9.23×10^{-4} s/s/y), NS6 (1.13×10^{-3} s/s/y) and NS7 (8.55×10^{-4} s/s/y), while the U.S. dataset had the highest substitution rates of spikes (1.41×10^{-3} s/s/y), membranes (2.57×10^{-3} s/s/y) and nucleocapsids (1.04×10^{-4} s/s/y). Interestingly, the NS7 substitution rate of the SEA dataset was the highest when compared to its other genes.

7.4 | Phylogeographic analysis

The results estimated that TMRCA of PDCoV circulated in the southern part of China close to the border of Vietnam during 1990, and its descendant approached SEA in 2000 and became divergent from the SEA strains. The virus was then introduced to Thailand during 2007–2011 prior to the divergences of eastern Thai (SEA pandemic) and western Thai in 2011 and 2012, respectively; therefore, the eastern subcluster became the hotspot of PDCoV in Thailand, which distributed the virus to Lao PDR and Vietnam in 2015 (Figure 4 and Supporting Information 5).

The Vietnam strains were part of three clades, and the cluster in SEA-1 shared the monophyletic group as those of the Thailand cluster. The cluster in SEA-2 was found only in Vietnam and was the most primitive cluster of Vietnam, and the cluster in SEA-2r found in Vung Tau was in a paraphyletic group that separated from the SEA-1 and SEA-2 clusters (Supporting Information 5 and 6).

7.5 | Positive selection analysis

Structural protein genes including S, E, M and N, together with NS6 and NS7, were separately analyzed. Positive selection pressure (dN/dS > 1) was found mainly in NS6 and NS7 genes (Supporting Information 7). In spike gene, 28 (2.41%), 28 (2.41%), and 14 (1.21%) positively selected codons were detected in the global, SEA and China datasets, respectively. The cumulative dN/dS of the global, SEA, U.S. and China codons were 0.38, 0.38, 0.62 and 0.43, respectively, whereas 329 (28.34%), 328 (28.25%), 39 (3.36%) and 256 (22.05%) codons were under negative pressure, respectively (Supporting Information 7a). It is noteworthy that although the S gene of SEA PDCoV is under negative pressure, spike codon sites are under highly positive selection pressure, especially at 642 position. In the E gene, one (1.19%) positively selected codon was detected in the global dataset. The cumulative dN/dS of the global, SEA, U.S. and China codons were 0.21, 0, 0 and 0.36, respectively, whereas 20 (23.81%), 2 (2.38%), 4 (4.76%) and 18 (21.43%) codons were under negative pressure (Supporting Information 7b). In the M gene, two (0.92%) and one (0.46%) codon were detected to be positively selected in the global and China datasets, respectively. The cumulative dN/dS of the global, SEA, U.S. and China codons were 0.11, 0.19, 0.08 and 0.23, respectively, whereas 49 (22.48%), 14 (6.42%), 8 (3.67%) and 32 (14.68%) codons were under negative pressure, respectively (Supporting Information 7c). In the N gene, seven (2.04%) positively selected codons were found in the global dataset. The cumulative dN/dS of the global, SEA, U.S., and China codons were 0.25, 0.27, 0.41 and 0.41, respectively, whereas 82 (23.91%), 36 (10.50%), 14 (10.08%) and 62 (18.08%) codons were under negative pressure, respectively (Supporting Information 7d). In contrast to structural protein genes, NS6 and NS7 genes were under positive pressure. In NS6, one (1.05%) positively selected codon was found in the global and China datasets. The cumulative dN/dS of the global, SEA, the United States and China had values of 0.97, 0.83, 7.8, and 1.05, respectively, whereas 19, 4, 1 and 12 codons were under negative pressure, respectively (Supporting Information 7e). In NS7, four (1.99%), two (1%) and one (0.49%) positively selected codon were detected in the global, China and SEA datasets, respectively. The cumulative dN/dS was higher than one in all datasets. Although the U.S. dataset has no poorly selected codon at any site, the cumulative dN/dS of the global, SEA, U.S. and China datasets were 6.28, 5.55, 2.35 and 8.26, respectively. The negatively selected codon of NS7 tended to be the lowest when compared to other genes, which were detected in 11 (5.47%), 5 (2.49%), 4 (1.99%) and 5 (2.49%)

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

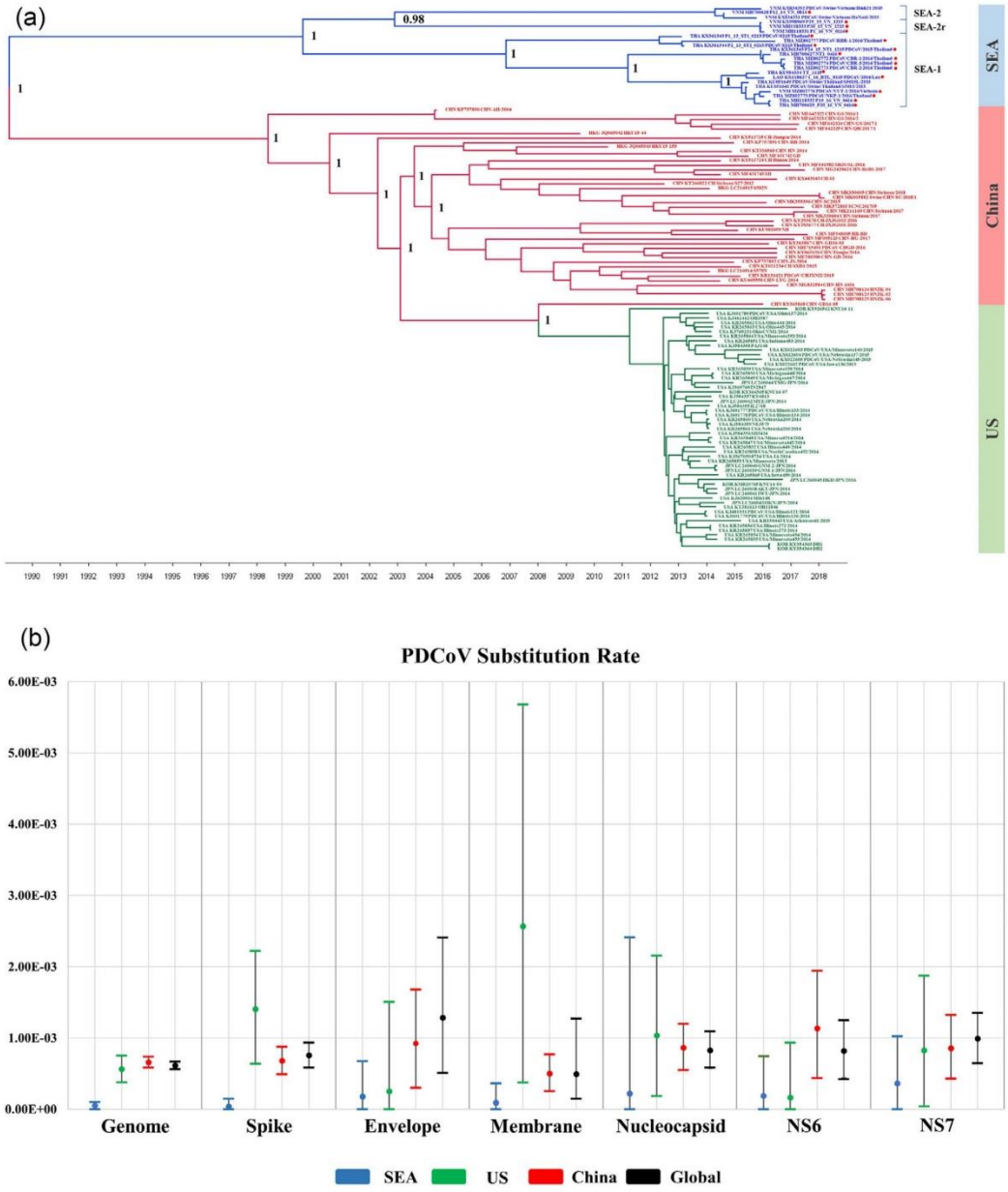


FIGURE 3 (a) Phylogenetic trees of the full-length genome of porcine deltacoronavirus (PDCoV) generated using BEAST. PDCoV could be clustered into three clades, the Southeast Asia (SEA) (blue), China (red), and the United States (green). Nodes denote the posterior probability. Red dots indicate the sequences in this study. (b) Substitution rate of global PDCoV in this study of a full-length genome (genome) in comparison with the structural genes (spike, envelope, membrane and nucleocapsid) and the accessory genes (NS6 and NS7) determined using BEAST and interpreted by using TRACER. Dots represent the mean, and upper and lower lines represent the upper and lower borders of the dataset

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

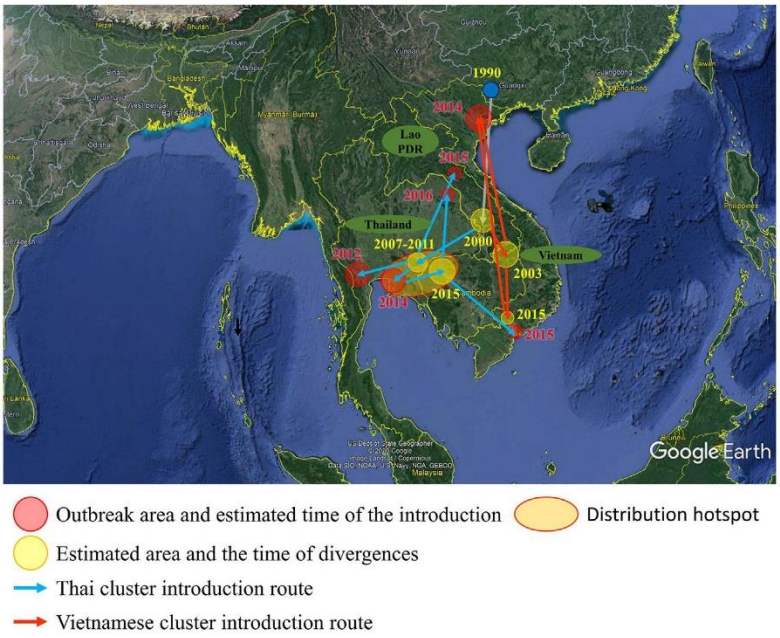


FIGURE 4 Phylogeographic distribution of porcine deltacoronavirus (PDCoV) in this study shows the estimated times of the divergences that occurred in Southeast Asia (SEA)

sites of codons in the global, SEA, the U.S. and China datasets, respectively (Supporting Information 7f).

InDels were found in the spike gene of the global, SEA and China datasets and the membrane gene in the global and China datasets (Supporting Information 7g). The cumulative InDels of the spike genes of the global, SEA and China were 1.12, 1.12 and 1.93, respectively, while the cumulative InDels of the membrane genes globally and in China were 0.08 and 0.25, respectively. Interestingly, although the cumulative dN/dS of the present study demonstrated that only NS6 and NS7 gene of the U.S. and China datasets, and the global, SEA, U.S. and China datasets, respectively, were under positive pressure, several sites of structural protein genes were positively selected especially in the global dataset.

8 | DISCUSSION

The complex evolution of SEA PDCoV compared to that of U.S. and China PDCoV has been interesting. Questions have been raised as to where the origin of the virus is and what factors influence the evolution of SEA PDCoV. Therefore, phylodynamic analyses were conducted to hypothetically fulfil the missing information. Based on phylogenetic and evolutionary analyses, the results demonstrated that PDCoV isolates are evolved into two separated genogroups, G1 (China and U.S.) and G2 (SEA), sharing a common ancestor. G1 is further evolved into two clusters including G1a (China) and G1b (US). G2 (SEA) is further divided into Thailand (G2a or SEA-1) and Vietnam (G2b or SEA-2) clusters. From the study, SEA PDCoV hypothetically descended from the same ancestor as China strains, and the common ancestor of SEA PDCoV was hypothetically introduced to Vietnam, although the most primitive strain of SEA was found in Thailand. It is noteworthy that E and N genes had substitution rates higher than that of the S gene. Furthermore, compared to China and SEA PDCoV, the substitution rate of US PDCoV was highest, whereas the recombination rate was lowest. In contrast with the US PDCoV, the recombination rate of SEA PDCoV and China was higher than that of the US PDCoV, whereas the substitution rate of SEA PDCoV was lower. The higher recombination rate of SEA PDCoV compared to US PDCoV lacking external introduction was supported by recombination events that occurred in both Thailand and Vietnam clusters of the SEA genogroup and Vietnam recombinants formed a new subcluster (SEA-2r). Interestingly, the positively selected sites were detected in the NS7 gene. This could suggest that the external introduction of genetically distinct PDCoV could influence the genetic development of SEA PDCoV.

It is speculated, based on the results of the present study, that PDCoV originally circulated in China, in the southern region close to the Vietnam border, a decade before its first introduction into the Southeast Asia region via Vietnam. The results also suggest that multiple introductions of genetically distinct PDCoV to Vietnam. The first and second introductions were estimated in 2003 and 2015, respectively. The second introduction was simultaneously occurred in both Vietnam and Thailand, which was estimated around 2007–2011 in

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Thailand prior to the introduction in Vietnam in 2015. This introduction results in three and two different subclusters in Vietnam and Thailand. Thailand PDCoV could be geographically clustered into two subclusters, including eastern and western subclusters. Isolates of the western subcluster found in 2013 were the most primitive strain. However, recently emerged eastern subcluster (SEA-1) found since 2015 has been the predominant subcluster in Thailand and SEA PDCoV, including Lao PDR and Vietnam strains. Interestingly, the emergence of the SEA-1 PDCoV, which was estimated during 2007–2015, was creating the hotspot of the viral distribution throughout the SEA.

The substitution rates of the spike gene in this study were higher when compared to the antigenic site of a partial gene in a previous study of other coronaviruses, such as TGEV (Enjuanes et al., 1992) and IBV (Cavanagh et al., 1998), and the S gene of BCoV (Vijgen et al., 2005) but were lower than that of the PEDV S gene (Stott et al., 2017). Compared to SARS (Z. Zhao et al., 2004), the genomic sequence substitution rate of PDCoV in this study was lower. This information suggests that the PDCoV genome and spike gene did not evolve rapidly when compared to those of coronaviruses such as SARS and PEDV. The highest substitution rates of the envelope and nucleocapsid genes of PDCoV in this study are in agreement with that of SARS in which the envelope, membrane, and nucleocapsid genes were the highest when compared to other genes. These findings are opposed to the nature of coronavirus, in which these genes are usually conserved (P. C. Y. Woo et al., 2010). Envelope and membrane proteins are needed to assemble the virus and can induce the activity of alpha interferon (Baudoux et al., 1998; de Haan et al., 1998). Mutation of the envelope gene in other coronaviruses attenuates virulence (Regla-Nava et al., 2015), and the envelope gene of most coronaviruses also plays an important role in host interactions, especially pathogenic inflammatory responses (DeDiego et al., 2014).

In this study, the substitution rates of the full-length genome, envelope, NS6 and NS7 of China were higher than those of SEA and U.S. PDCoV, whereas most of the structural encoding genes of U.S. PDCoV, except the envelope, were higher than others. These findings suggest that U.S. PDCoV was under higher pressure through structural genes than PDCoV in SEA and China. Interestingly, China PDCoV tends to have a higher substitution rate of non-structural encoding genes, suggesting that external introduction might influence structural protein changes in China PDCoV, whereas the substitution mechanism seems to be important for hybrid viruses in fields, which results in these strains lacking a substitution rate. Compared with the U.S. and China datasets, although SEA PDCoV lacked a substitution rate, the NS6 substitution rate tended to be high. From a previous study, NS6 of PDCoV was found to be an IFN-β antagonist via interaction with RIG-I/MDA5, resulting in a reduction in RLR-mediated IFN-β production (Fang et al., 2018), while NS7 showed the ability to downregulate α-actinin-4 (ACTN4) and mitochondrial carbamoyl phosphate synthase 1 (CPS1; Choi & Lee, 2019). In our study, the phenomenon of the accessory gene mutation was appeared to be important in China and SEA PDCoV.

Although the substitution rate of PDCoV genomic sequences was lower than that of many other coronaviruses, PDCoV was undergoing their mutation through these genes (envelope and nucleocapsid). Recent studies in calves and chickens have shown that the virus is not yet well adapted in the host (K. Jung et al., 2017; Liang et al., 2019). These findings might relate to the phenomenon of the high mutation rate of the envelope gene and nucleocapsid.

The recombination result shows that the SEA parental strains descended from the ancestor of the strains from China, such as Guang Dong or Si Chuan, and the strains from Hong Kong. The strains found in SEA before 2014 lacked information; however, the spike genes of several isolates in Vietnam were close to the Guang Dong strain (Supporting Information 8). The recombination results also correlated with a previous report (Saeng-Chuto et al., 2020) suggesting P29_15_VN_1215 (KX998969), P30_15_VN_1215 (MH118331), and P1_16_VN_0116 (MH118331) of Vietnam were the recombinant between the primitive strain of Vietnam with the external introduction of the U.S. strain and these strains were virulent compared to the primitive strain, not triggering severe clinical signs.

The recombination rate of the genome and spike gene showed the high potential of recombination in China and SEA; however, there is a lack of recombination rates in the United States. Although the recombination rate of the U.S. strain was very low, the substitution of the U.S. strain was the highest, in contrast to that of the SEA strain, which shows a higher recombination rate but lacks the substitution rate, while the China strain has a high potential of mutation via two mechanisms. The SEA results suggest that the influence of external introduction from other countries was the primary mechanism of the virus mutation. Unfortunately, we did not study the potential of the substitution hotspot caused by the recombination event that occurred in the China strain.

From a recent study (He et al., 2020), the substitution rate of the PDCoV spike gene was estimated at 1.7×10^{-3} s/s/y, while the substitution rates of Thailand, China (early strains excluded) and the United States were estimated at 2.71, 1.58 and 1.21 (×10−3 s/s/y), respectively. The estimated value of the U.S. dataset was similar, although the substitution rate of others tended to be higher than that in our study. The reason for this difference might be affected by the difference in dataset preparation; we aimed to compare the entire SEA dataset with the China and U.S. subgroups and the different sequences and substitution models used in this study.

According to the results in this study, we could strongly consider that the SEA PDCoV shall be derived from China, and the SEA strains were close to that of the Guang Dong strain. The SEA strains were geographically distinct in several parts of Thailand and Vietnam, and there was also influence with the multiple introductions from both the U.S. and China PDCoV throughout the SEA especially in Vietnam; although we still lack the information of the nearest divergence of PDCoV in SEA, these strains are undergoing the recombination events and positive selection pressure making the virus geographically distinctly genetically drive and causing the situation of the outbreak complicated since the several studies raising concern on the high pathogenic PDCoVs, determined by triggering the severe clinical signs by the virus itself (Jung et al., 2015; Saeng-Chuto et al., 2017; Xu et al., 2018). In this study, the SEA strains from both Thailand and Vietnam showed

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

many properties related to a previous report of highly pathogenic PDCoVs from Tianjin and Guangdong (Xu et al., 2018), and coinfection with other pathogens, such as PEDV, was also reported in recent years (Zhang et al., 2019), making herd health management difficult to eliminate the pathogen from herds. Awareness of pig replacement from external sources, quarantine and preventive strategies against pathogens should be successfully achieved.

ACKNOWLEDGEMENTS

This study was supported by Chulalongkorn University (Fundamental fund, grant number 65413100200013, CUFRB65_food(26)190_31_09). Partial funding was provided by Special Task Force for Activating Research (STAR), Swine Viral Evolution, and Vaccine Research (SVEVR), Chulalongkorn University (Grant number GSTAR 59-013-31-007). The first author is supported by the Royal Golden Jubilee PhD program (PHD/0204/2557) Scholarship of the Thailand Research Fund.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICAL APPROVAL

All utilizable international, national and/or institutional guidelines for concern and usage of animals were followed. The animal challenge study was approved by the Institutional Animal Care and Use Committee of Chulalongkorn University (CU-IACUC) (1931018).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the Supporting Information of this article. The full-length sequences of the 18 PDCoV isolates have been deposited in GenBank under the accession numbers KU984334, KX118627, KX361343 to KX361345, KX998969, MH118331 to MH118333, MH700627 to MH700629 and MZ802772 to MZ802777.

ORCID

Dachrit Nilubol https://orcid.org/0000-0002-7691-6790

REFERENCES

Auton, A., & Mcvean, G. (2007). Recombination rate estimation in the presence of hotspots. *Genome Research*, *17*(8), 1219–1227. https://doi.org/10.1101/gr.6386707

Ayres, D. L., Darling, A., Zwickl, D. J., Beerli, P., Holder, M. T., Lewis, P. O., Huelsenbeck, J. P., Ronquist, F., Swofford, D. L., Cummings, M. P., Rambaut, A., & Suchard, M. A. (2012). Beagle: An application programming interface and high-performance computing library for statistical phylogenetics. *Systematic Biology*, *61*(1), 170–173. https://doi.org/10.1093/sysbio/syr100

Baudoux, P., Carrat, C., Besnardeau, L., Charley, B., & Laude, H. (1998). Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes. *Journal of Virology*, *72*(11), 8636–8643.

Bielejec, F., Rambaut, A., Suchard, M. A., & Lemey, P. (2011). Spread: Spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics*, *27*(20), 2910–2912. https://doi.org/10.1093/bioinformatics/btr481

Boni, M. F., Posada, D., & Feldman, M. W. (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics*, *176*(2), 1035–1047. https://doi.org/10.1534/genetics.106.068874

Cavanagh, D., Mawditt, K., Adzhar, A., Gough, R. E., Picault, J. P., Naylor, C. J., Haydon, D., Shaw, K., & Britton, P. (1998). Does IBV change slowly despite the capacity of the spike protein to vary greatly? *Advances in Experimental Medicine and Biology*, *440*, 729–734.

Chen, Q.i, Gauger, P., Stafne, M., Thomas, J., Arruda, P., Burrough, E., Madson, D., Brodie, J., Magstadt, D., Derscheid, R., Welch, M., & Zhang, J. (2015). Pathogenicity and pathogenesis of a United States porcine deltacoronavirus cell culture isolate in 5-day-old neonatal piglets. *Virology*, *482*, 51–59. doi:https://doi.org/10.1016/j.virol.2015.03.024

Choi, S., & Lee, C. (2019). Functional characterization and proteomic analysis of porcine deltacoronavirus accessory protein NS7. *Journal of Microbiology and Biotechnology*, *29*(11), 1817–1829. https://doi.org/10.4014/jmb.1908.08013

De Haan, C. A. M., Kuo, L., Masters, P. S., Vennema, H., & Rottier, P. J. M. (1998). Coronavirus particle assembly: Primary structure requirements of the membrane protein. *Journal of Virology*, *72*(8), 6838–6850.

Dediego, M. L., Nieto-Torres, J. L., Jimenez-Guardeño, J. M., Regla-Nava, J. A., Castaño-Rodriguez, C., Fernandez-Delgado, R., Usera, F., & Enjuanes, L. (2014). Coronavirus virulence genes with main focus on SARS-CoV envelope gene. *Virus Research*, *194*, 124–137. https://doi.org/10.1016/j.virusres.2014.07.024

Drummond, A. J., Nicholls, G., Rodrigo, A., & Solomon, W. (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, 161(3), 1307–1320.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*(1), 1–8. https://doi.org/10.1186/1471-2148-7-214

Drummond, A. J., Rambaut, A., Shapiro, B.,& O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, *22*(5), 1185–1192. https://doi.org/10.1093/molbev/msi103

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. https://doi.org/10.1093/molbev/mss075

Enjuanes, L., Suñé, C., Gebauer, F., Smerdou, C., Camacho, A., Antón, I. M., González, S., Talamillo, A., Méndez, A., Ballesteros, M. L., & Sánchez, C., (1992). Antigen selection and presentation to protect against transmissible gastroenteritis coronavirus. *Veterinary Microbiology*, *33*(1-4), 249–262.

Fang, P., Fang, L., Ren, J., Hong, Y., Liu, X., Zhao, Y., Wang, D., Peng, G., & Xiao, S. (2018). Porcine deltacoronavirus accessory protein NS6 antagonizes interferon beta production by interfering with the binding of RIG-I/MDA5 to double-stranded RNA. *Journal of Virology*, 92(15), e00712-18. https://doi.org/10.1128/jvi.00712-18

Ferreira, M. A. R., & Suchard, M. A. (2008). Bayesian analysis of elapsed times in continuous-time Markov chains. *The Canadian Journal of Statistics /La Revue Canadienne de Statistique*, *36*(3), 355–368.

Fourment, M., & Darling, A. E. (2018). Local and relaxed clocks: The best of both worlds. *PeerJ*, *6*, e5140-e5140. https://doi.org/10.7717/peerj.5140

Gibbs, M. J., Armstrong, J. S., & Gibbs, A. J. (2000). Sister-scanning: A Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*, *16*(7), 573–582. https://doi.org/10.1093/bioinformatics/16.7.573

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95–98.

He, W.-T., Ji, X., He, W., Dellicour, S., Wang, S., Li, G., Zhang, L., Gilbert, M., Zhu, H., Xing, G., Veit, M., Huang, Z., Han, G.-Z., Huang, Y., Suchard, M. A., Baele, G., Lemey, P., & Su, S. (2020). Genomic epidemiology, evolution, and transmission dynamics of porcine deltacoronavirus. *Molecular Biology and Evolution*, *37*(9), 2641–2654. https://doi.org/10.1093/molbev/msaa117

Janetanakit, T., Lumyai, M., Bunpapong, N., Boonyapisitsopa, S., Chaiyawong, S., Nonthabenjawan, N., Kesdaengsakonwut, S., & Amonsin, A. (2016). Porcine deltacoronavirus, Thailand, 2015. *Emerging Infectious Diseases*, *22*(4), 757–759. https://doi.org/10.3201/eid2204.151852

Johnson, A. D. (2010). An extended IUPAC nomenclature code for polymorphic nucleic acids. *Bioinformatics (Oxford, England)*, *26*(10), 1386–1389. https://doi.org/10.1093/bioinformatics/btq098

Jung, K., Hu, H., Eyerly, B., Lu, Z., Chepngeno, J., & Saif, L. J. (2015). Pathogenicity of 2 porcine deltacoronavirus strains in gnotobiotic pigs. *Emerging Infectious Diseases*, *21*(4), 650–654. https://doi.org/10.3201/eid2104.141859

Jung, K., Hu, H., & Saif, L. J. (2017). Calves are susceptible to infection with the newly emerged porcine deltacoronavirus, but not with the swine enteric alphacoronavirus, porcine epidemic diarrhea virus. *Archives of Virology*, *162*, 2357–2362, https://doi.org/10.1007/s00705-017-3351-z

Katoh, K. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066.

Korber, B. (2000). HIV sequence signatures and similarities. In G. A. Rodrigo & H. G. Learn (Eds.), *Computational and evolutionary analysis of HIV molecular sequences*. Kluwer Academic Publishers.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lee, S., & Lee, C. (2014). Complete genome characterization of Korean porcine deltacoronavirus strain KOR/KNU14-04/2014. *Genome Announcements*, *2*(6), e01191–14. https://doi.org/10.1128/genomeA.01191-14

Lemey, P., Rambaut, A., Welch, J. J., & Suchard, M. A. (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution*, *27*(8), 1877–1885. https://doi.org/10.1093/molbev/msq067

Liang, Zhang, Li, Ding, Wang, Gao, Guo, Wei, & Hu, (2019). Susceptibility of chickens to porcine deltacoronavirus infection. *Viruses*, *11*(6), 573. https://doi.org/10.3390/v11060573

Lorsirigool, A., Saeng-Chuto, K., Madapong, A., Temeeyasen, G., Tripipat, T., Kaewprommal, P., Tantituvanont, A., Piriyapongsa, J., & Nilubol, D. (2017). The genetic diversity and complete genome analysis of two novel porcine deltacoronavirus isolates in Thailand in 2015. *Virus Genes*, *53*(2), 240–248. https://doi.org/10.1007/s11262-016-1413-z

Lorsirigool, A., Saeng-Chuto, K., Temeeyasen, G., Madapong, A., Tripipat, T., Wegner, M., Tuntituvanont, A., Intrakamhaeng, M., & Nilubol, D. (2016). The first detection and full-length genome sequence of porcine deltacoronavirus isolated in Lao PDR. *Archives of Virology*, *161*(10), 2909–2911. https://doi.org/10.1007/s00705-016-2983-8

Eisen, L. (2008). Use of Google Earth™ to strengthen public health capacity and facilitate management of vector-borne diseases in resource-poor environments. *Bulletin of the World Health Organization*, *86*(9), 718–725. https://doi.org/10.2471/BLT.07.045880

Marthaler, D., Jiang, Y., Collins, J., & Rossow, K. (2014). Complete genome sequence of strain SDCV/USA/Illinois121/2014, a porcine deltacoronavirus from the United States. *Genome Announcements*, *2*(2), e00218-00214. https://doi.org/10.1128/genomeA.00218-14

Marthaler, D., Raymond, L., Jiang, Y., Collins, J., Rossow, K., & Rovira, A. (2014). Rapid detection, complete genome sequencing, and phylogenetic analysis of porcine deltacoronavirus. *Emerging Infectious Diseases*, *20*(8), 1347–1350. https://doi.org/10.3201/eid2008.140526

Martin, D., & Rybicki, E. (2000). RDP: Detection of recombination amongst aligned sequences. *Bioinformatics*, *16*(6), 562–563. https://doi.org/10.1093/bioinformatics/16.6.562

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, *1*(1), vev003. https://doi.org/10.1093/ve/vev003

Mccluskey, B. J., Haley, C., Rovira, A., Main, R., Zhang, Y., & Barder, S. (2016). Retrospective testing and case series study of porcine delta coronavirus in U.S. swine herds. *Preventive Veterinary Medicine*, *123*, 185–191. doi:https://doi.org/10.1016/j.prevetmed.2015.10.018

Nei, M., & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, *3*(5), 418–426. https://doi.org/10.1093/oxfordjournals.molbev.a040410

Ota, T., & Nei, M. (1994). Variance and covariances of the numbers of synonymous and nonsynonymous substitutions per site. *Molecular Biology and Evolution*, *11*(4), 613–619. https://doi.org/10.1093/oxfordjournals.molbev.a040140

Padidam, M., Sawyer, S., & Fauquet, C. M. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology*, *265*(2), 218–225. doi:https://doi.org/10.1006/viro.1999.0056

Pérez-Rivera, C., Ramírez-Mendoza, H., Mendoza-Elvira, S., Segura-Velázquez, R., & Sánchez-Betancourt, J. I. (2019). First report and phylogenetic analysis of porcine deltacoronavirus in Mexico. *Transboundary and Emerging Diseases*, *66*(4), 1436–1441. https://doi.org/10.1111/tbed.13193

Posada, D., & Crandall, K. A. (2001). Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proceedings of the National Academy of Sciences*, *98*(24), 13757–13762. https://doi.org/10.1073/pnas.241370698

Pybus, O. G., Suchard, M. A., Lemey, P., Bernardin, F. J., Rambaut, A., Crawford, F. W., Gray, R. R., Arinaminpathy, N., Stramer, S. L., Busch, M. P., & Delwart, E. L. (2012). Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *Proceedings of the National Academy of Sciences*, *109*(37), 15066–15071. https://doi.org/10.1073/pnas.1206598109

Rambaut, A., & Drummond, A. J. (2013). Tracer 1.6 [computer program].

Rambaut, A., & Drummond, A. J. (2014). FigTree 1.4.2 [computer program].

Regla-Nava, J. A., Nieto-Torres, J. L., Jimenez-Guardeño, J. M., Fernandez-Delgado, R., Fett, C., Castaño-Rodríguez, C., Perlman, S., Enjuanes, L., & Dediego, M. L. (2015). Severe acute respiratory syndrome coronaviruses with mutations in the E protein are attenuated and promising vaccine candidates. *Journal of Virology*, *89*(7), 3870–3887. https://doi.org/10.1128/jvi.03566-14

Saeng-Chuto, K., Jermsutjarit, P., Stott, C. J., Vui, D. T., Tantituvanont, A., & Nilubol, D. (2020). Retrospective study, full-length genome characterization and evaluation of viral infectivity and pathogenicity of chimeric porcine deltacoronavirus detected in Vietnam. *Transboundary and Emerging Diseases*, *67*(1), 183–198. https://doi.org/10.1111/tbed.13339

Saeng-Chuto, K., Lorsirigool, A., Temeeyasen, G., Vui, D. T., Stott, C. J., Madapong, A., Tripipat, T., Wegner, M., Intrakamhaeng, M., Chongcharoen, W., Tantituvanont, A., Kaewprommal, P., Piriyapongsa, J., & Nilubol, D. (2017). Different lineage of porcine deltacoronavirus in Thailand, Vietnam and Lao PDR in 2015. *Transboundary and Emerging Diseases*, *64*(1), 3–10. https://doi.org/10.1111/tbed.12585

Saeng-Chuto, K., Stott, C. J., Wegner, M., Senasuthum, R., Tantituvanont, A., & Nilubol, D. (2017). Retrospective investigation and evolutionary analysis of a novel porcine deltacoronavirus strain detected in Thailand from 2008 to 2015. *Archives of Virology*, *162*(7), 2103–2108. https://doi.org/10.1007/s00705-017-3331-3

Salminen, M. O., Carr, J. K., Burke, D. S., & Mccutchan, F. E. (1995). Identification of breakpoints in intergenotypic recombinants of HIV Type 1 by bootscanning. *AIDS Research and Human Retroviruses*, *11*(11), 1423–1425. https://doi.org/10.1089/aid.1995.11.1423

Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Tai, W., Du, L., Zhou, Y., Zhang, W., & Li, F. (2018). Cryo-electron microscopy structure of porcine deltacoronavirus spike protein in the prefusion state. *Journal of Virology*, *92*(4), e01556-01517.

Sinha, A., Gauger, P., Zhang, J., Yoon, K. -. J., & Harmon, K. (2015). PCR-based retrospective evaluation of diagnostic samples for emergence of porcine

18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

deltacoronavirus in US swine. *Veterinary Microbiology*, *179*(3), 296–298. https://doi.org/10.1016/j.vetmic.2015.06.005

Smith, J. (1992). Analyzing the mosaic structure of genes. *Journal of Molecular Evolution*, *34*(2), 126–129. https://doi.org/10.1007/bf00182389

Song, D., Zhou, X., Peng, Q., Chen, Y., Zhang, F., Huang, T., Zhang, T., Li, A., Huang, D., Wu, Q., He, H., & Tang, Y. (2015). Newly emerged porcine deltacoronavirus associated with diarrhoea in swine in China: Identification, prevalence and full-length genome sequence analysis. *Transboundary and Emerging Diseases*, *62*(6), 575–580. https://doi.org/10.1111/tbed.12399

Stott, C. J., Temeeyasen, G., Tripipat, T., Kaewprommal, P., Tantituvanont, A., Piriyapongsa, J., & Nilubol, D. (2017). Evolutionary and epidemiological analyses based on spike genes of porcine epidemic diarrhea virus circulating in Thailand in 2008–2015. *Infection, Genetics and Evolution*, *50*, 70–76. https://doi.org/10.1016/j.meegid.2017.02.014

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. In R. M. Miura (Ed.). *Some Mathematical Questions in Biology: DNA Sequence Analysis. Lectures on Mathematics in the Life Sciences*, Vol. 17, (pp. 57–86). American Mathematical Society.

Vijgen, L., Keyaerts, E., Moës, E., Thoelen, I., Wollants, E., Lemey, P., Vandamme, A. -. M., & Van Ranst, M. (2005). Complete genomic sequence of human coronavirus OC43: Molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *Journal of Virology*, *79*(3), 1595–1604. https://doi.org/10.1128/jvi.79.3.1595-1604.2005

Wang, L., Byrum, B., & Zhang, Y. (2014). Porcine coronavirus HKU15 detected in 9 US states, 2014. *Emerging Infectious Diseases*, *20*(9), 1594–1595. https://doi.org/10.3201/eid2009.140756

Woo, P. C., Lau, S. K., Lam, C. S., Lau, C. C., Tsang, A. K., Lau, J. H., Bai, R., Teng, J. L. L., Tsang, C. C. C., Wang, M., Zheng, B.-J., Chan, K.-H., & Yuen, K. Y. (2012). Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus. *Journal of Virology*, *86*(7), 3995–4008. https://doi.org/10.1128/jvi.06540-11

Woo, P. C. Y., Huang, Y.i, Lau, S. K. P., & Yuen, K. -. Y. (2010). Coronavirus genomics and bioinformatics analysis. *Viruses*, *2*(8), 1804–1820. https://doi.org/10.3390/v2081803

Xu, Z., Zhong, H., Zhou, Q., Du, Y., Chen, L.i, Zhang, Y., Xue, C., & Cao, Y. (2018). A highly pathogenic strain of porcine deltacoronavirus caused watery diarrhea in newborn piglets. *Virologica Sinica*, *33*(2), 131–141. https://doi.org/10.1007/s12250-018-0003-8

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *Journal of Molecular Evolution*, *39*(3), 306–314. https://doi.org/10.1007/BF00160154

Yang, Z., & Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends in Ecology & Evolution*, *15*(12), 496–503. https://doi.org/10.1016/s0169-5347(00)01994-7

Zhang, F., Luo, S., Gu, J., Li, Z., Li, K., Yuan, W., Ye, Y.u, Li, H., Ding, Z., Song, D., & Tang, Y. (2019). Prevalence and phylogenetic analysis of porcine diarrhea associated viruses in southern China from 2012 to 2018. *BMC Veterinary Research*, *15*(1), 470. https://doi.org/10.1186/s12917-019-2212-2

Zhao, Y., Qu, H., Hu, J., Fu, J., Chen, R., Li, C., Cao, S., Wen, Y., Wu, R., Zhao, Q., Yan, Q., Wen, X., & Huang, X. (2019). Characterization and pathogenicity of the porcine deltacoronavirus isolated in Southwest China. *Viruses*, *11*(11), 1074. https://doi.org/10.3390/v11111074

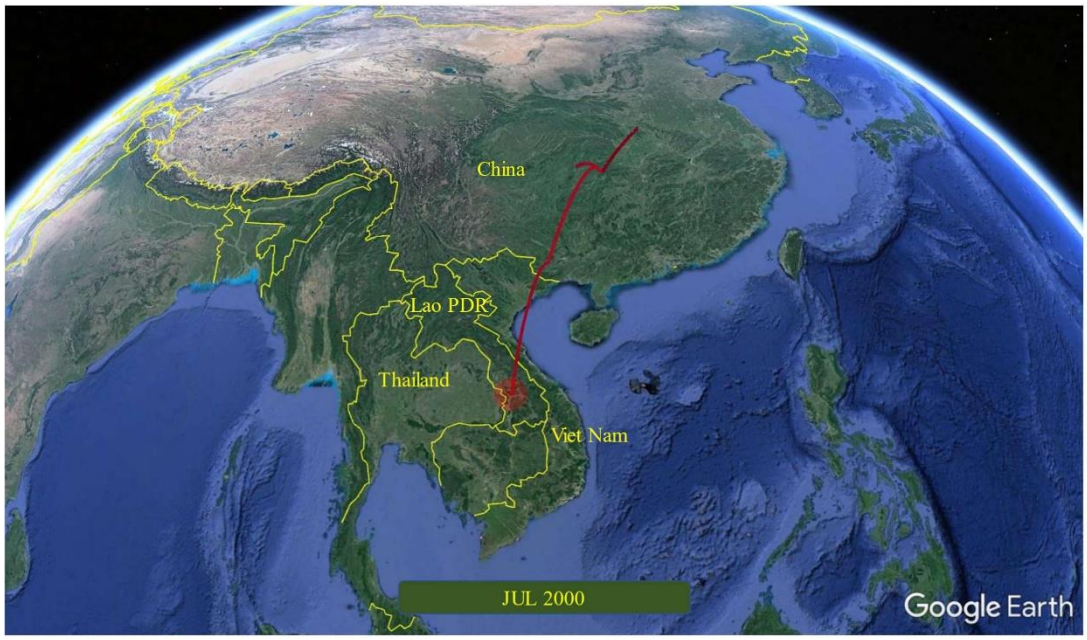
Zhao, Z., Li, H., Wu, X., Zhong, Y., Zhang, K., Zhang, Y.a-P., Boerwinkle, E., & Fu, Y. -. X. (2004). Moderate mutation rate in the SARS coronavirus genome and its implications. *BMC Evolutionary Biology*, *4*(1), 21. https://doi.org/10.1186/1471-2148-4-21

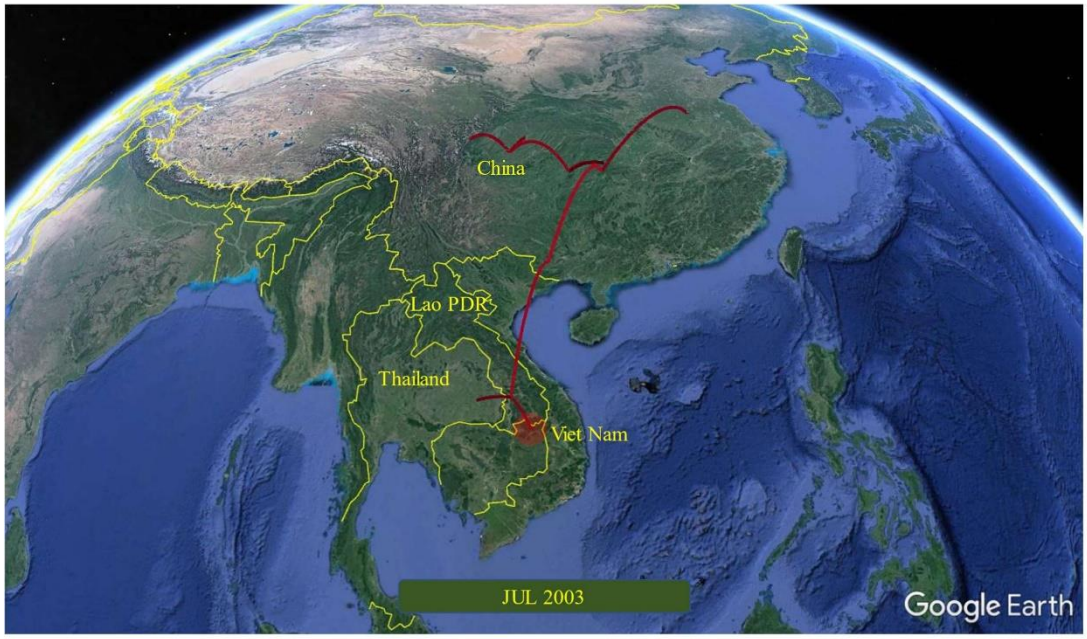
SUPPORTING INFORMATION

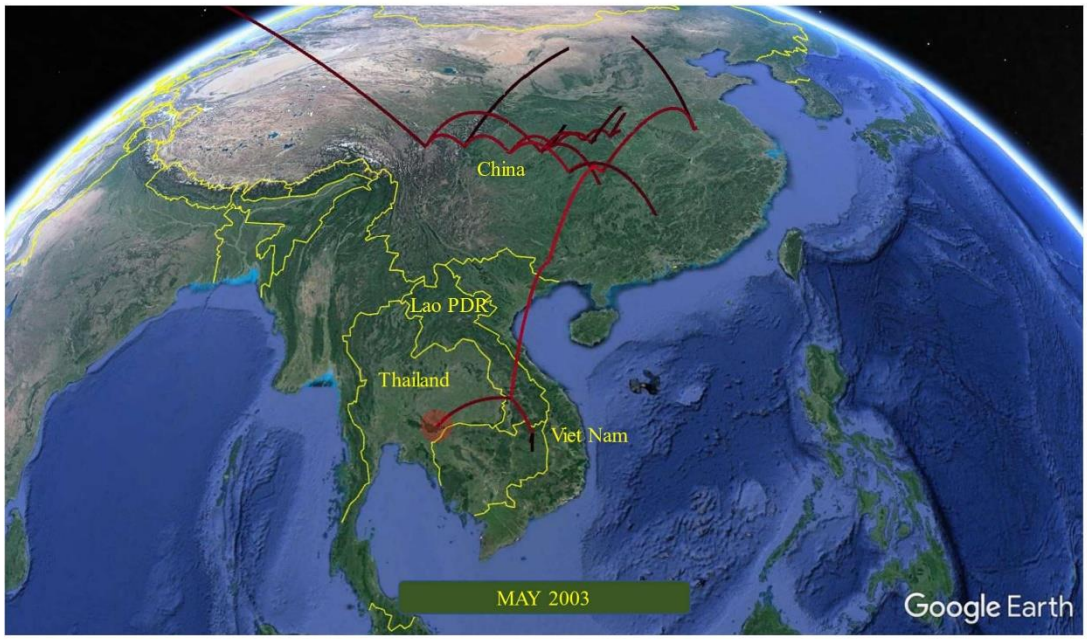
Additional supporting information may be found in the online version of the article at the publisher's website.

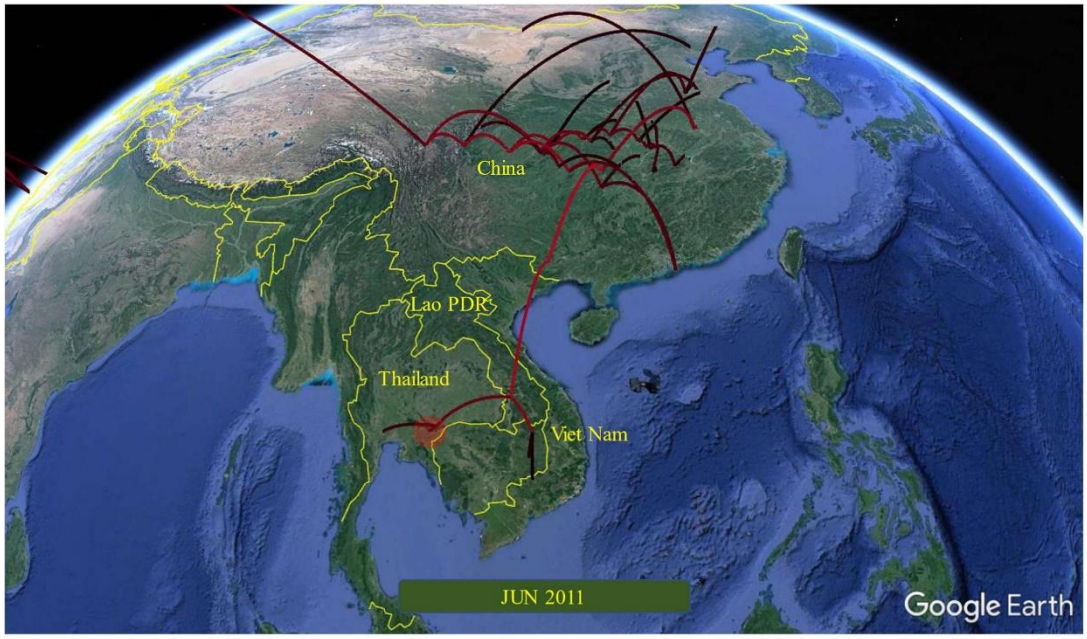
How to cite this article: Stott, C. J., Sawattrakool, K., Saeng-Chuto, K., Tantituvanont, A., & Nilubol, D. (2022). The phylodynamics of emerging porcine deltacoronavirus in Southeast Asia. *Transboundary and Emerging Diseases*, *69*, 2816–2827. https://doi.org/10.1111/tbed.14434

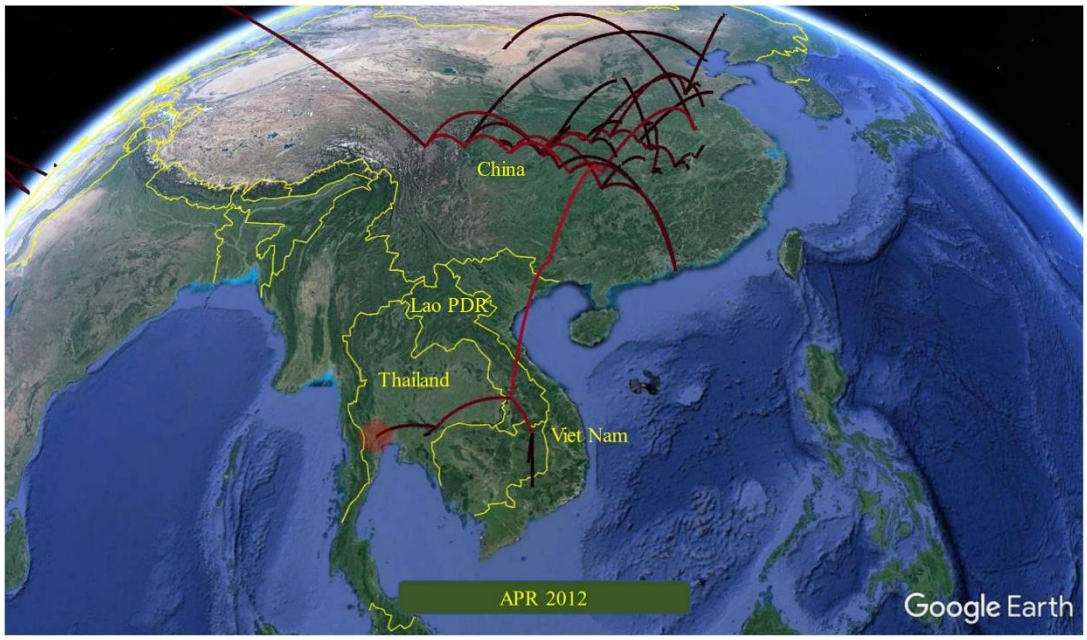
18651682, 2022, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/tbed.14434 by Office Of Academic Resources, Wiley Online Library on [20/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

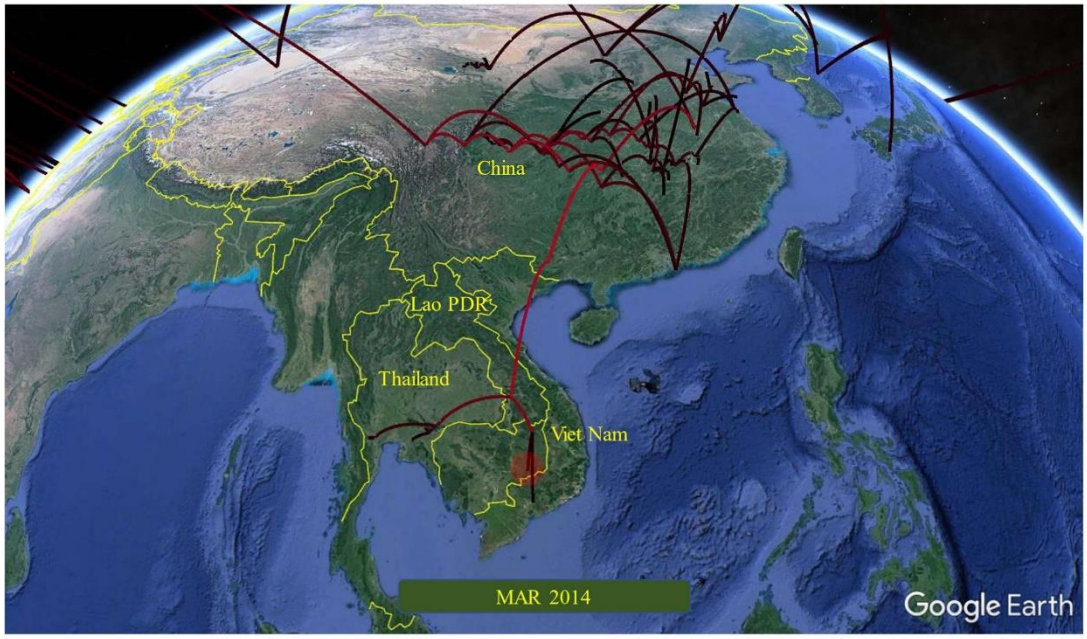

China
Lao PDR
Thailand
Viet Nam
JUL 2000
Google Earth

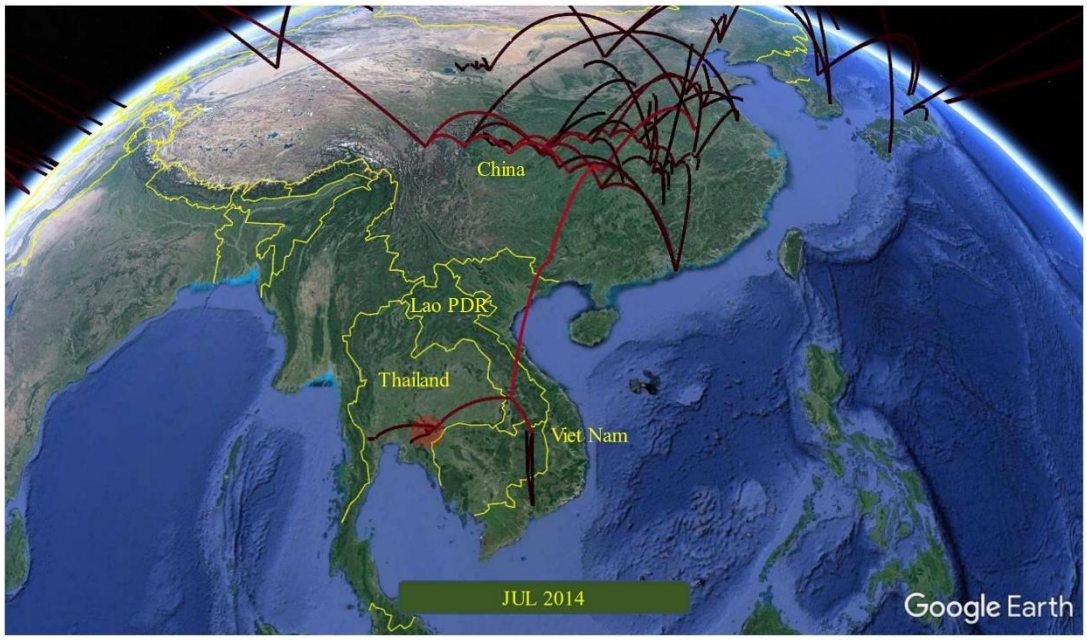

China
Lao PDR
Thailand
Viet Nam
JUL 2003
Google Earth

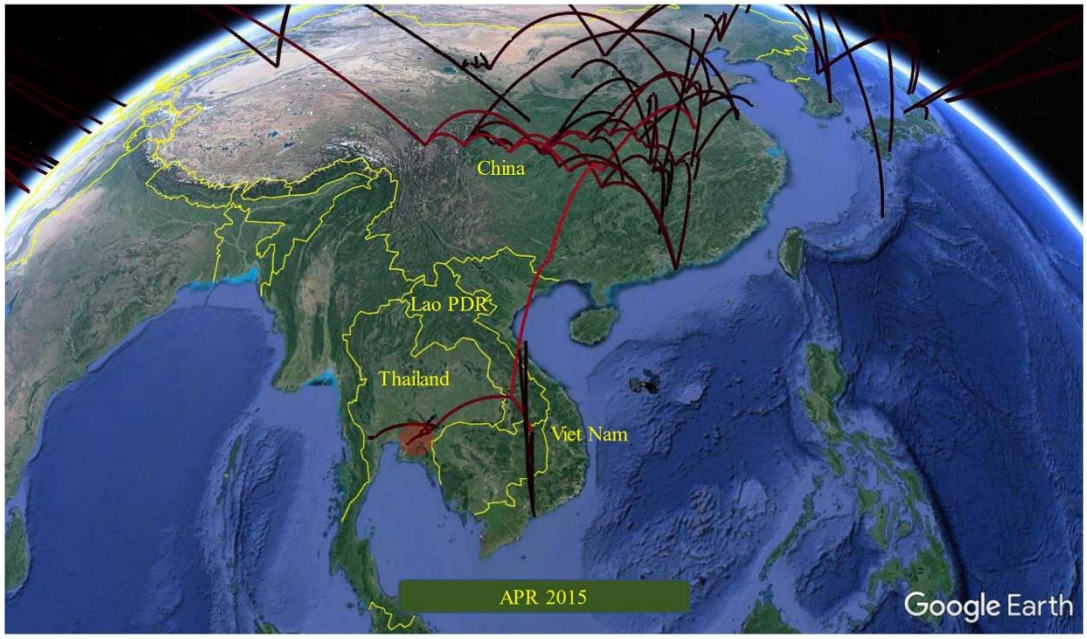

China
Lao PDR
Thailand
Viet Nam
MAY 2003
Google Earth

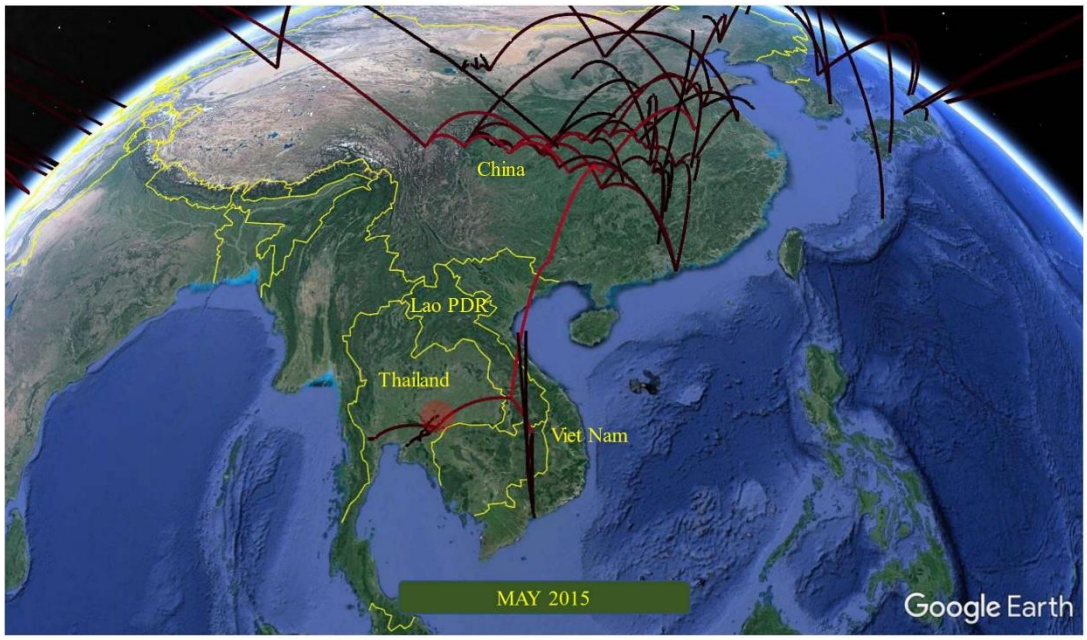

China
Lao PDR
Thailand
Viet Nam
JUN 2011
Google Earth

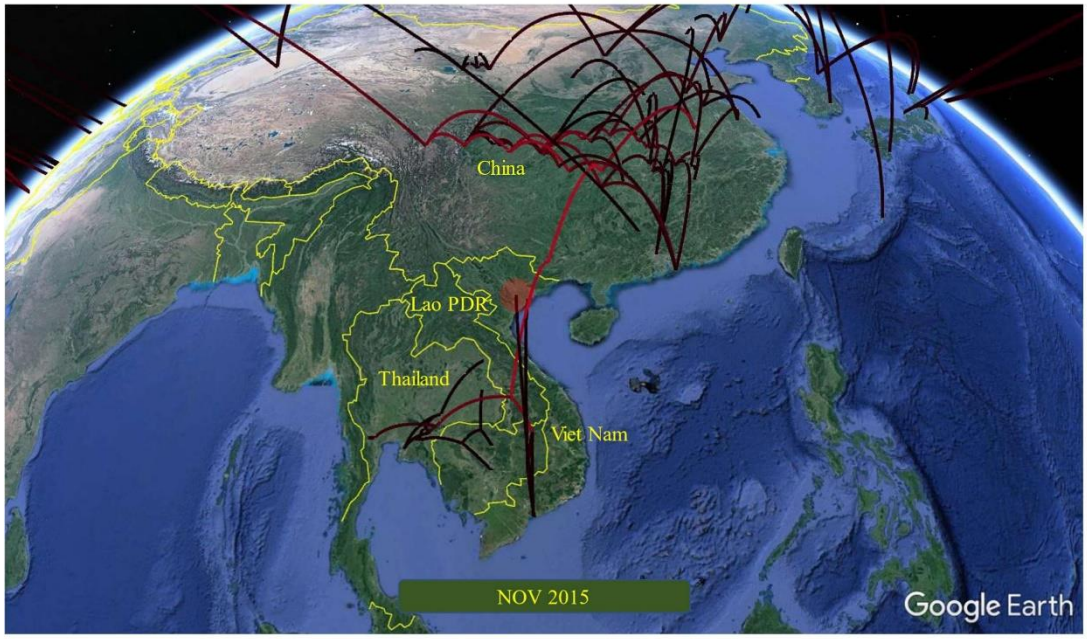

China
Lao PDR
Thailand
Viet Nam
APR 2012
Google Earth

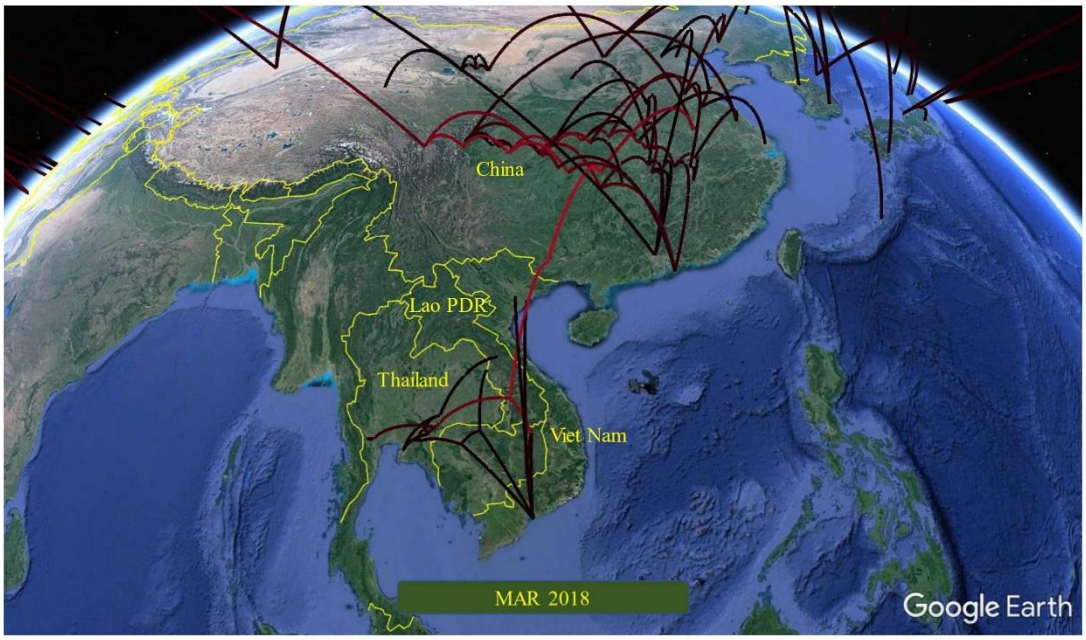

China
Lao PDR
Thailand
Viet Nam
MAR 2014
Google Earth


China
Lao PDR
Thailand
Viet Nam
JUL 2014
Google Earth

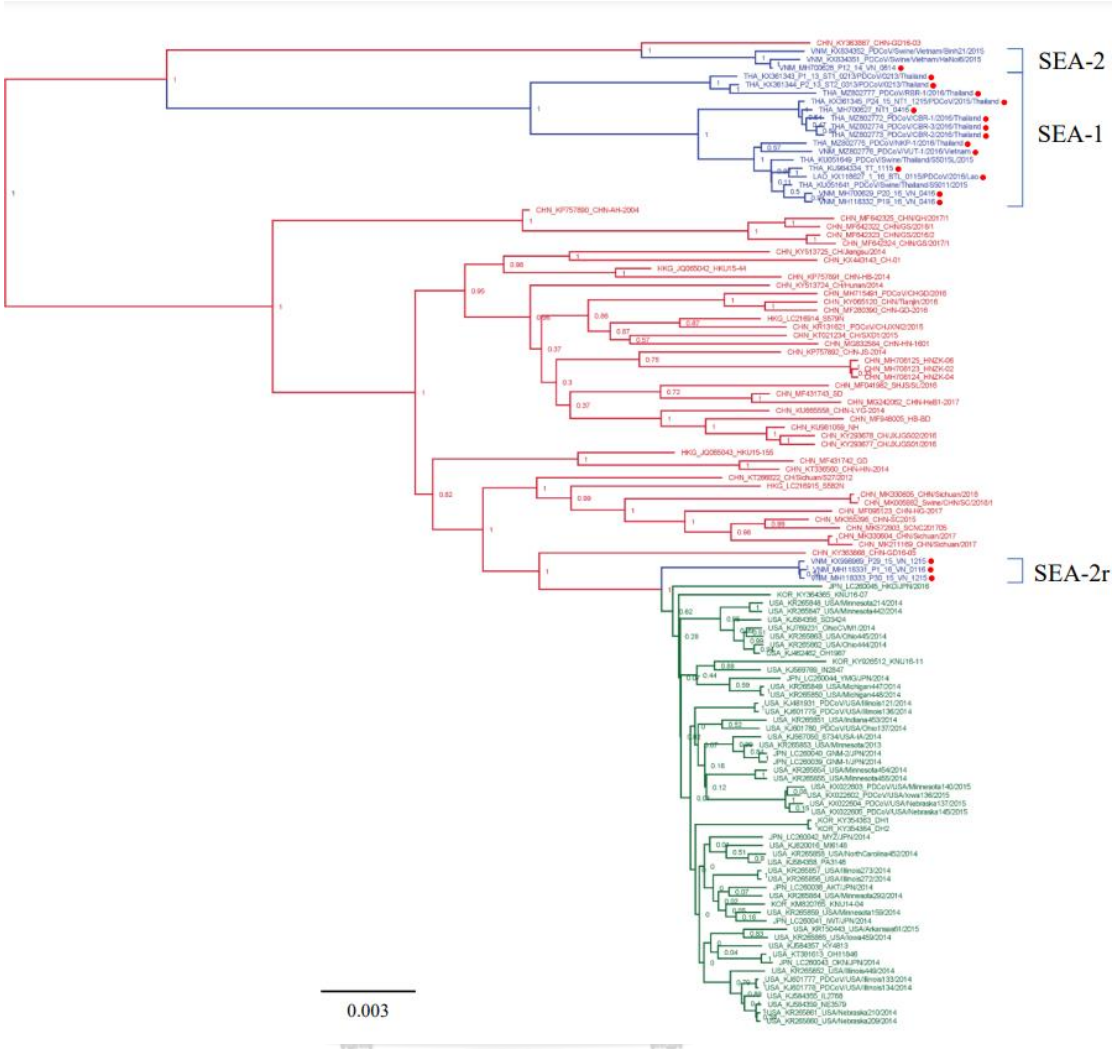

China
Lao PDR
Thailand
Viet Nam
APR 2015
Google Earth


China
Lao PDR
Thailand
Viet Nam
MAY 2015
Google Earth


China
Lao PDR
Thailand
Viet Nam
NOV 2015
Google Earth



Supplementary material 6 Phylogeographic snapshots when the important estimated divergences of SEA PDCoV occurred, analyzed by BEAST package, generated by SPREAD, and displayed using Google Earth, the center point of the red circle shows the estimated location of the divergence.



Supplementary material 8 Figure 22 Phylogeographic trees of PDCoV spike gene generated using BEAST. Nodes denote the posterior probability. Red dots indicate the sequences in this study.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

VITA

NAME	Christopher James Stott
DATE OF BIRTH	28 August 1986
PLACE OF BIRTH	Lopburi, Thailand
INSTITUTIONS ATTENDED	2009-2015 Doctor of Veterinary Medicine (first-class honors) Faculty of Veterinary Medicine, Mahasarakham University, Mahasarakham, Thailand
HOME ADDRESS	1/352 Popular Rd. 5, Bang Pood Sub-district, Pak Kret District, Nonthaburi Province, 11120, Thailand
PUBLICATION	Stott, CJ., Sawattrakool, K., Saeng-Chuto, K., Tantituvanont, A., and Nilubol, D. 2021. The phylodynamics of emerging porcine deltacoronavirus (PDCoV) in Southeast Asia. Transboundary and Emerging Diseases. First published: 20 December 2021, https://doi.org/10.1111/tbed.14434. Stott, CJ., Temeeyasen, G., Tripipat, T., Kaewprommal, P., Tantituvanont, A., Piriyapongsa, J. & Nilubol D. 2017. Evolutionary and epidemiological analyses based on spike genes of porcine epidemic diarrhea virus circulating in Thailand in 2008-2015. Infection, Genetics and Evolution. Jun;50:70-76.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University