The detection and phylogenetic analysis of porcine deltacoronavirus from Guangdong Province in Southern China

K. Mai | J. Feng | G. Chen | D. Li | L. Zhou | Y. Bai | Q. Wu | J. Ma

College of Animal Science, South China Agricultural University, Guangzhou, China

Correspondence
J. Ma, College of Animal Science, South China Agricultural University, Guangzhou, China.
Email: majy2400@scau.edu.cn

Funding Information
National Key Research and Development Program, Grant/Award Number: 2016YFD0501300

Summary
Porcine deltacoronavirus (PDCoV) is a newly discovered coronavirus that causes diarrhoea, vomiting and dehydration in sucking and nursing piglets. It was first reported in Hong Kong in 2012 and has since been discovered in the United States, Canada, South Korea, mainland China, Thailand and Laos. PDCoV has been experimentally proved to lead to diarrhoea in swine and it was detected positive in pigs in Guangdong, southern China. In our study, 252 faecal and intestinal samples from sucking piglets and sows with diarrhoea were surveyed for common enteric viruses. We found a prevalence of PDCoV (21.8%), porcine epidemic diarrhoea virus (65.5%), transmissible gastroenteritis virus (0%), rotavirus group A (25.0%) and porcine kobuvirus (68.7%). We isolated 13 PDCoV strains and discovered that PDCoV infections were often co-infections with kobuvirus rather than the commonly linked porcine epidemic diarrhoea virus. Phylogenetic analysis of S gene and N gene revealed that 11 of 13 PDCoV strains belonged to Chinese lineage. As for the left two strains, one single strain (CHN-GD16-05) belonged to American and Korean lineages while another strain (CHN-GD16-03) was similar to a Thai strain, but only in the S gene. This suggested a possible recombination event between the Thai and the newly described Chinese strain.

KEYWORDS
Porcine deltacoronavirus, prevalence, spike gene, nucleocapsid gene, phylogenetic analysis, Southern China

1 | INTRODUCTION

Coronaviruses belong to Coronavirinae subfamily within Coronaviridae family and the Nidovirales order. They are spherical, enveloped RNA viruses that cause enteric, respiratory and other diseases in humans and animals. In 2012 after the discovery of the Deltacoronavirus in animal species in Hong Kong (Woo et al., 2012), the classification of Coronavirinae subfamily was expanded. There are presently four genera in this subfamily: Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus. Five coronavirus strains have been isolated from pigs. These include the Alphacoronavirus species porcine epidemic diarrhoea virus (PEDV), transmissible gastroenteritis virus (TGEV) and porcine respiratory coronavirus (PRCV). Porcine haemagglutinating encephalomyelitis virus (PHEV) is a Betacoronavirus and porcine deltacoronavirus (PDCoV), a Deltacoronavirus.

Porcine deltacoronavirus was first identified in swine specimens collected in 2009 in Hong Kong (Woo et al., 2012). It was identified later in several states of USA (Li et al., 2014; Marthaler, Jiang, Collins, & Rossow, 2014; Wang, Byrum, & Zhang, 2014) and has now spread to South Korea, mainland China, Thailand and Laos (Lee & Lee, 2014; Lee et al., 2016; Lorsirigool et al., 2016; Madapong et al., 2016; Song et al., 2015). The emergence of PDCoV was most often associated with clinical infection symptoms including diarrhoea, vomiting and dehydration in pigs, especially in sucking and nursing piglets (Janetanakit et al., 2016; Ma et al., 2015; Marthaler, Raymond et al., 2014; Song et al., 2015). These worldwide outbreaks of PDCoV have caused economic losses to the global commercial swine.
industry (Chen, Gauger et al., 2015; Chen, Zhu et al., 2015; Hoomong et al., 2016; Song et al., 2015; Wang et al., 2014).

RT-PCR detection methods have provided data for the prevalence of PDCoV in China (Dong et al., 2015; Song et al., 2015; Zhai et al., 2016). A previous study (Zhai et al., 2016) reported that the current prevalence in Southern China was 1.54% (5/390), and three of the five detected strains were from Guangdong Province, revealing that the information regarding the molecular epidemiology of PDCoVs in Guangdong was still limited. The current study further investigates the prevalence, epidemiology and genomic properties of PDCoV in Guangdong.

## MATERIALS AND METHODS

### 2.1 Sample collection and detection methods

In order to monitor the prevalence and sequence properties of PDCoV in Guangdong Province, 252 faecal and intestinal samples were collected from 11 commercial swine farms between December 2015 and June 2016. These samples were all collected from sucking piglets and sows. We screened for PDCoV, PEDV, TGEV, Rotavirus A and porcine kobuvirus (PKV).

Thirteen PDCoV cell culture isolates were propagated in swine testis (ST) cells and plaque purified twice. These isolates were obtained from case submissions from the 252 swine samples we collected previously. Total RNA was extracted from samples using a commercial kit following the manufacturer’s instructions (Axygen Scientific). All RNA samples were stored at −80°C. RNA was reverse-transcribed using a PrimeScript RT Reagent Kit with gDNA Eraser (Takara, Biotechnology, Dalian, China) following the manufacturer’s recommendations (Axygen Scientific). Nucleotide sequences were submitted to GenBank with accession numbers (KY078891–KY078903 for nucleocapsid gene and KY078904–KY078916 for spike gene of CHN-GD16-01 to CHN-GD16-13, respectively). Reference sequences included 33 strains for the S gene and 33 strains of N gene from different farms of global isolates obtained from GenBank (Table S2 and S3). DNA sequences

### 2.2 Amplification, cloning and sequencing the Spike (S) protein and Nucleocapsid protein (N) genes

The complete S and N genes were amplified using three pairs of primers: PDCoV-S1, 5'- ATGCCAGAGGCTTATG-3' and 5'-TATTT CAACCTCGCCATC-3'; PDCoV-S2, 5'-CGACCACTCATTTAC-3' and PDCoV-N, 5'-CTACCACTCTTAACT-3'; and PDCoV-N, 5'- ATGGCCGACCAGTAGTC-3' and 5'-CTACGCTGCTGATTCCCTG-3' (based on CHJXN12/2015/China). PCR amplification was carried out using the LA Taq polymerase kit (Takara, Biotechnology, Dalian, China) using directions supplied by the manufacturer. PCR products were purified using the Gel Band Purification Kit (Omega Bio-Tek, USA) and then cloned into the PMD-19T vector (Takara) using an In-Fusion PCR Cloning Kit (Takara), the recombinant plasmids were sequenced by the Beijing Genomics Institute (Shenzhen, Guangdong).

### 2.3 Sequence alignment and phylogenetic analysis

Nucleotide sequences were submitted to GenBank with accession numbers (KY078891–KY078903 for nucleocapsid gene and KY078904–KY078916 for spike gene of CHN-GD16-01 to CHN-GD16-13, respectively). Reference sequences included 33 strains for the S gene and 33 strains of N gene from different farms of global isolates obtained from GenBank (Table S2 and S3). DNA sequences

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>RT-PCR detection of porcine enteric viruses PEDV, PDCoV, Rotavirus group A, PKV and TGEV in diarrhoeal faecal or intestinal samples from December 2015 to June 2016 in Guangdong Province</th>
</tr>
</thead>
<tbody>
<tr>
<td>Months of year</td>
<td>Diarrhoeal samples no. (%)</td>
</tr>
<tr>
<td></td>
<td>SP/S</td>
</tr>
<tr>
<td>December, 2015</td>
<td>36/17 (67.9/32.1)</td>
</tr>
<tr>
<td>January, 2016</td>
<td>26/18 (59.1/40.1)</td>
</tr>
<tr>
<td>February, 2016</td>
<td>31/5 (86.1/13.9)</td>
</tr>
<tr>
<td>March, 2016</td>
<td>15/1 (93.8/6.2)</td>
</tr>
<tr>
<td>April, 2016</td>
<td>34/2 (94.4/5.6)</td>
</tr>
<tr>
<td>May, 2016</td>
<td>18/5 (78.3/21.7)</td>
</tr>
<tr>
<td>June, 2016</td>
<td>43/1 (97.7/2.3)</td>
</tr>
<tr>
<td>Total</td>
<td>203/49 (80.6/19.4)</td>
</tr>
</tbody>
</table>

PEDV, porcine epidemic diarrhoea virus; PDCoV, porcine deltacoronavirus; PKV, porcine kobuvirus; TGEV, transmissible gastroenteritis virus; SP, Suckling piglets (<20 days); S, Sow.
were assembled using the Clustal W program of DNAStar V7.1 (Madison, WI, USA). Phylogenetic trees were constructed using the neighbour-joining method in MEGA 6.0 software with bootstrap analysis of 1,000 replicates. The percentage of replicate trees in which the associated taxa clustered are shown as nearby branches (Tamura, Nei, & Kumar, 2004; Tamura, Stecher, Peterson, Filipski, & Kumar, 2013).

2.4 | Prediction of epitopes for spike protein antigen sequence and N-glycosylation of PDCoV

The secondary structure and surface properties of S1 protein of PDCoV were analysed with Protein, DNAStar V.7.1 (Madison, WI, USA) to predict its antigenic epitopes. The secondary structures of these S proteins were analysed by Garnier–Robson and Chou–Fasman methods, and surface properties, including hydrophilicity, flexible regions, accessibility and antigenicity were analysed by Kyte–Doolittle, Karplus–Schulz, Emini and Jameson–Wolf methods, respectively. Also, further prediction of epitope positions were confirmed on online server at http://www.cbs.dtu.dk/services/BepiPred/. Determination of potential N-glycosylation sites of the S protein was predicted with software available at http://www.cbs.dtu.dk/services/NetNGlyc.

3 | RESULTS

3.1 | PDCoV detection

Samples (N = 252) from 11 swine farms in Guangdong Province were collected during a 7-month period. Approximately 22% were PDCoV positive while 65%, 25% and 69% contained PEDV, Rota A and PKV, respectively. Transmissible gastroenteritis virus was not detected in any samples. Sucking piglets showed slight lower PDCoV infection rate (21.7%, 44/203) than sows (22.4%, 11/49), while other types of pathogens are on the contrary (Table 1). In addition, co-infections with PEDV and PKV were the most common (44.4%, 79/178) (Table 2). PDCoV occurrence was examined further by an additional RT-PCR analysis using 13 PDCoV isolates we successfully purified in cell culture (Table 3). All 13 samples were collected from faeces and intestinal contents of sucking piglets. PDCoV-PEDV-PKV and PDCoV-PKV co-infection occurred with high frequency; Rota A was detected in three samples and PBV in only one. As for other porcine enteric pathogens, such as TGEV, PSV and PAstV, they were detected in any samples. Sucking piglets showed slight lower PDCoV infection rate than sows; TGEV, transmissible gastroenteritis virus; PEDV, porcine epidemic diarrhoea virus; Rota A, porcine rotavirus A; PKV, porcine kobuvirus.

3.2 | Homology and phylogenetic analysis of S gene and N gene of PDCoV isolates

The 13 PDCoV genomes were designated CHN-GD16-01 to CHN-GD16-13. For S gene, sequencing results revealed that CHN-GD16-03 and CHN-GD16-05 possessed 3,483 nucleotides, CHN-GD16-06 was included 3,477 nucleotides and other 10 strains composed of 3,480 nucleotides in size. Interestingly, the viral strain CHN-GD16-06 contained unique AAT and GTT deletions compared to the US and individual Asian strains. Their nucleotide identities were represented in dispersed clusters within the reference strains (Table S2). From this point, we used only the last two numbers of the names of our cultivated strains to eliminate confusion with strains from other studies, that is, CHN-GD16-01 = strain 01. Most of the viral strains (01, 02, 09, 12 and 13) had a high nucleotide similarity (0.985–0.990) with CHJXN12/China/2015 (Song et al., 2015). On the other hand, 04, 06, 07, 08, 10, 11 shared 0.988–0.992 similarity with CHN/JS/2014 (Dong et al., 2015). Unlike the other strains from this work, 03 had the highest nucleotide homology (0.971) with HKU15-44, a strain first identified from swine in Hong Kong. Interestingly, strain 05 had a maximum nucleotide identity (0.985) with KNU14-04 (Lee & Lee, 2014) from South Korea. Moreover, the neighbour-joining phylogenetic tree reconstructed from the PDCoV S genes in isolate 03 was located at the Thai branch. The others were clustered within the Chinese branch. Strain 05 was located at a single branch next to the US and Korean branches (Figure 1).

Within the N gene, our 13 purified samples contained 1,029 nt indicating a lack of insertions and deletions. These 13

<table>
<thead>
<tr>
<th>Virus</th>
<th>PDCoV-positive samples no.</th>
<th>% (n = 246)</th>
<th>PDCoV-positive samples no.</th>
<th>% (n = 55)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One pathogen only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDCoV</td>
<td>14</td>
<td>20.6</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>PEDV</td>
<td>32</td>
<td>47.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rota A</td>
<td>2</td>
<td>2.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKV</td>
<td>20</td>
<td>29.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>68</td>
<td>27.6</td>
<td>14</td>
<td>25.5</td>
</tr>
<tr>
<td>2 pathogens</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDCoV+PEDV</td>
<td>2</td>
<td>1.6</td>
<td>2</td>
<td>7.4</td>
</tr>
<tr>
<td>PDCoV+Rota A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PDCoV+PKV</td>
<td>25</td>
<td>19.3</td>
<td>25</td>
<td>92.6</td>
</tr>
<tr>
<td>PEDV+Rota A</td>
<td>3</td>
<td>2.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEDV+PKV</td>
<td>79</td>
<td>61.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rota A+PKV</td>
<td>20</td>
<td>15.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>129</td>
<td>52.4</td>
<td>27</td>
<td>49.1</td>
</tr>
<tr>
<td>3 pathogens</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDCoV+PEDV+Rota A</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PDCoV+PEDV+PKV</td>
<td>11</td>
<td>23.9</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>PDCoV+Rota A+PKV</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PEDV+Rota A+PKV</td>
<td>35</td>
<td>76.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td>18.7</td>
<td>11</td>
<td>83.6</td>
</tr>
<tr>
<td>4 pathogens</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDCoV+PEDV+Rota A+PKV</td>
<td>3</td>
<td>1.2</td>
<td>3</td>
<td>5.5</td>
</tr>
</tbody>
</table>

PDCoV, porcine deltacoronavirus; PEDV, porcine epidemic diarrhoea virus; TGEV, transmissible gastroenteritis virus; PKV, porcine kobuvirus.
samples were also had the highest nucleotide identities with different PDCoV strains. For example, strain 05 was most closely to KNU14-04 (S. Korea), OH11846 and Minnesota292 (0.993) (USA). The other strains were most closely to CH/SXD1/2015, Sichuan/S27 and CHN/GX01/2015 (Chen, Gauger et al., 2015; Chen, Zhu et al., 2015; Dong et al., 2015) originating in different regions in China (Table S3). The phylogenetic tree of the N gene indicated that strain 05 was located at the US branch, while other positive samples from this work were all clustered in a group within the Chinese branch (Figure 2).

Overall, combining the results of homology and phylogenetic analysis of the S and N genes, we concluded that most of the current PDCoV strains in Guangdong Province are closely related to other Chinese PDCoV strains. Only a few of PDCoV identities in strains 03 and 05 might have evolved from different ancestors.

### 3.3 | Spike protein B-cell antigenic epitope and N-glycosylation site prediction

We identified 11 potential and distinct B-cell antigenic epitopes in the S1 protein (Figure S1). Peptides less than 4 amino acid homology were not computed. These predicted epitopes all contained mostly β structure and high hydrophilicity, strong accessibility, good flexibility and a high, predicted antigenicity. We also identified 17 potential N-glycosylation sites in most of the isolates as well as the reference strains. Strain 02. USA/IL/2014 and Korea/SL5/2015 each contained 16 glycosylation sites with the lack of an N-glycosylation site at position 162. An additional deletion at position 41 was observed in Thailand strains PDCoV/S5011/2015, PDCoV/S5015L/2015 and PDCoV/S5024/2015.

### 4 | DISCUSSION

Until now, numerous reports had revealed that PEDV was the most common porcine enteric pathogen in pigs due to its high prevalence in diarrhoeal samples. PDCoV-PEDV co-infection occurred at a rate greater than PDCoV-Rota A and PDCoV-TGEV. But in our study, co-infection of PDCoV and PKV was greatly higher than PDCoV-PEDV co-occurrence (45.4%, 25/55 vs. 3.6%, 2/55, Table 2); additionally, PDCoV-PEDV-PKV co-infection also processed certain proportion on PDCoV-positive samples (20%, 11/55). PKV was also common pathogens causing diarrhoea and viral gastroenteritis in pigs, especially in piglets (Di, Angeloni, Tofani, Monini, & Ruggeri, 2015; Pfankuche et al., 2016) and it showed a high co-infection rate with PEDV (Wang, Lan, & Yang, 2016; Zhao et al., 2016). PKV detection has not been included with most other studies of PDCoV; therefore, evidence is lacking to prove that PKV co-infects more frequently than PEDV with PDCoV. This warrants further investigation.
FIGURE 1  Phylogenetic analysis using the PDCoV spike protein (S) gene from different species. The S gene strains identified in this work are indicated with "black triangles". GenBank accession numbers, countries and years of isolation are listed at the right side of each reference strain. The tree was constructed using MEGA 6.0 software with neighbour-joining (NJ) methods and 1,000 replicate sets on bootstrap analysis.
FIGURE 2  Phylogenetic analysis of the PDCoV N gene isolated from different species. The N genes from the strains of this work are indicated with “black diamonds”. GenBank accession numbers, countries and years of isolation are listed at the right side of each reference strains. The tree was constructed as per Figure 2, above.
The remaining 11 isolates all contained an AAT deletion resulting in an asparagine deletion. Furthermore, strain 06 possessed a GTT deletion, causing a valine deletion. More importantly, a C3447T substitution of its nucleotide sequence changed a glutamine codon into a stop codon, resulting in an 11 amino acid (QPTPSFKFKEW) deletion. Although amino acid deletions and insertions have been reported, it remains unknown whether these have any biological significance to PDCoV pathogenesis.

Two pairs of strains 07 and 08 from farm G and 10 and 11 from farm I were isolated from the same commercial swine farms, but at different times. Unlike 07 and 08 that were identical in nucleotide sequence, 10 and 11 showed nucleotide identities of 99.0% and 98.3%, respectively, resulting in 10 amino acid changes. Interestingly, these differences were in the S protein and concentrated within the S1 region. According to the two phylogenetic trees established for the S and N genes, 11 of 13 strains clustered within the Chinese category and had the highest nucleotide identities with five published Chinese strains. This indicated that they were more closely related to other strains from China.

Strain 05 isolated from farm E belonged to the US and Korean branch with the greatest identity with Korean strain KNU14-04 (Lee & Lee, 2014). As all reported PDCoV isolates from China are clustered into a single Chinese group, this is the first case of a US and Korean strain detected in diarrheal samples from pigs in China. Strain 03 is more related on a genomic scale to the Thai strains and had the lowest nucleotide identity with CHN/GD02/2015 and CHN/GX01/2015 (95.8%, Table S2) (Zhai et al., 2016), which were collected from Guangdong and nearby Guangxi provinces. However, the N gene of this strain revealed a completely reverse conclusion; that it belongs to the Chinese group with a low nt identity (96.7%) with the Thai strains S5011 and S015L (Janetanakit et al., 2016) (Fig. 2 and Table S3). Therefore, we inferred that a recombination event might occur in PDCoV strain 03.

Recombination events are often reported in PEDV studies and the majority are intrarecombinants with different strain lineages of the same kinds of enteric coronaviruses (Jarvis et al., 2015; Tian et al., 2014). A recent study (Boniotti et al., 2016; Valerij et al., 2016) identified a swine enteric coronavirus (SeCoV) from Italy as more closely related to PEDV in the S gene, while the remainder of the genome shared highest identity with TGEV. Recombination events have not been reported in PDCoV strains. Since we were lack of the full-length genome of this PDCoV strain, the mechanism of the construction of this chimera as well as with the recombination in strain 03 is unknown, something suppose to be further work and explain with it.

The spike protein is an important surface glycoprotein of coronavirus and plays a significant role in receptor binding and fusion of the viral and cellular membranes (Chakrabarti, Prabakaran, Xiao, & Dimitrov, 2005; Schwegmann-Weber et al., 2009). And also, it mediates interspecies transmission (Bosch, Van, de Haan, & Rottier, 2003). All coronavirus spike proteins share the same two functional components: an N-terminal subunit and a membrane-anchored subunit (C-terminal) that are covalently bound. The S protein of PDCoV can be similarly subdivided into S1 (1–573 aa) and S2 (574–1161 aa) regions (Thachil, Gerber, Xiao, Huang, & Opriessnig, 2015). S1 is a dominant viral antigen and an ELISA is available for its detection (Thachil et al., 2015). Three different groups of S1 proteins from coronavirus infections shared less than 12% amino acid sequence identities with each other (Wang, Deng et al., 2016). The B-cell response is directed against the spike protein of coronaviruses (Cao et al., 2015) and it plays an important role in pathogenesis of virus infection.

In our study, 11 potential B-cell antigenic epitopes of PDCoV (Figure S1) and 16 or 17 N-glycosylation sites were analysed. Interestingly, in amino acids 549–561 in strain 05, a T559I mutation greatly reduced the predicted level of its antigenicity, hydrophilicity, surface probability and flexibility (Figure S2 & Table S4). Whether this mutation has altered its antigenicity will be interesting to test in the laboratory. Overall, further studies are needed to confirm whether these alterations of B-cell antigenic epitopes and N-glycosylation sites affect the pathogenicity and antigenicity of each PDCoV strain.

In conclusion, diarrheal samples collected from pigs in Guangdong Province were screened to detect the prevalence of PDCoV. Phylogenetic analysis suggested that nearly all of the strains in mainland China were clustered into the Chinese lineage except one newly discovered PDCoV strain that had a close relationship with US and Korea strains. This complements the geographical lineage theory of global PDCoV distribution. The presence of another suspected recombinant strain will provide additional data to examine the diversity of the PDCoV genome.

ACKNOWLEDGEMENTS
The study was supported by The National Key Research and Development Program (2016YFD0501300) of China.

CONFLICT OF INTEREST
The authors declare no conflicts of interest.

REFERENCES


**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the supporting information tab for this article.