Phylogenetic analysis of porcine circovirus 3 (PCV3) circulating in Canadian pigs

Maodong Zhang[1], Chao Chun Liu[2], Yanyun Huang[1], Janet Hill[3], Maria Bravo Araya[1], Davor Ojkic[4], and C.A. Gagnon[5]

[1]Prairie Diagnostic Services Inc
[2]Simon Fraser University Department of Molecular Biology and Biochemistry
[3]University of Saskatchewan Western College of Veterinary Medicine
[4]University of Guelph
[5]Universite de Montreal Faculte de medecine veterinaire

April 16, 2024

Abstract

Porcine circovirus 3 (PCV3) has been detected in pigs worldwide and associated with several clinical signs. To investigate the genetic diversity of PCV3 strains circulating in Canada, 44 PCV3 positive samples from Saskatchewan (2/44), Manitoba (2/44), Quebec (4/44), Alberta (11/44) and Ontario (25/44) submitted to diagnostic laboratories in Canada between 2019 and 2021 were sequenced and analyzed. Phylogenetic analysis of capsid genes showed that all of the 44 Canadian strains classified into PCV3a and segregated into seven lineages with common amino acid changes observed at A24V, R27K, N56D, T77S, Q98R, L150I (F) and R168K positions. Future studies are required to determine whether the polymorphisms in capsid proteins, as revealed in this study, could be associated with differences in the pathogenicity or antigenicity of PCV3 strains. This is the first phylogenetic analysis of PCV3 strains among different provinces in Canada.

Introduction

Porcine circovirus 3 (PCV3), which belongs to the family *Circoviridae* and genus *Circovirus* , was first identified in 2015 in pigs in the United States displaying myocarditis, multisystemic inflammation, porcine dermatitis and nephropathy syndrome (PDNS) and reproductive failure (Palinski et al., 2017; Phan et al., 2016). Subsequently, PCV3 has been detected in many countries around the world raising concerns to the swine industry, although the virus has also been detected in asymptomatic pigs (Gu et al., 2020; Jiang et al., 2019; Klaumann et al., 2018). PCV3 is a small DNA virus with a 2000 nucleotide single-stranded circular genome containing three major open-reading frames (ORFs): ORF1 encoding replicase (rep) protein, ORF2 encoding the capsid (cap) protein and ORF3 encoding a protein with unknown functions (Klaumann et al., 2018; Palinski et al., 2017; Qi et al., 2019; Assao et al., 2021a). The cap protein is considered to be the major structural protein, which determines the antigenic properties of the virus, showing only approximately 37% amino acid sequence identity to that of PCV2 (Palinski et al., 2017; Nawagitgul et al., 2000). PCV3 strains have shown genetic stability and low mutation rates over time and across countries, with nucleotide similarity ranging from 97 to 100% on partial or complete sequences (Assao et al., 2021a; Assao et al., 2021b; Franzo et al., 2020; Plut et al., 2020; Qi et al., 2019). This work aimed to investigate the genetic diversity of PCV3 strains circulating in Canada and whether a main PCV3 strain circulating across different Canadian provinces.{Palinski, 2017 #400;Phan, 2016 #401}{Palinski, 2017 #400;Phan, 2016 #401}{Palinski, 2017 #400;Phan, 2016 #401}{Palinski, 2017 #400;Phan, 2016 #401}

The investigated sample set consisted of 44 samples including oral, processing and thoracic fluids (25/44),

feces (1/44) and tissues (18/44), which were submitted to diagnostic laboratories including the Centre de diagnostic vétérinaire de l'Université de Montréal (CDVUM) in Quebec, Animal Health Laboratory (AHL) in Ontario, and Prairie Diagnostic Services (PDS) in Saskatchewan (Supplementary Table 1). Available samples enrolled in this study were from different provinces including Saskatchewan (2/44), Manitoba (2/44), Quebec (4/44), Alberta (11/44) and Ontario (25/44) between 2019 and 2021. Nucleic acid extraction was performed using MagMAX^{TM} CORE Nucleic Acid Extraction Procedure (ThermoFisher, Texas, USA) as per the manufacturer's instructions. PCV3 detection was performed using a previously described real-time PCR assay (Palinski et al., 2017) and samples with quantification cycle (Cq) values at or below 30 were selected for PCR and sequencing of the full PCV3 ORF2 sequences. Two sets of primers with the first set (Forward 5'-GTG TAC AAT TAT TGC GTT GGG-3' and Reverse 5'- AAA ACA CAG CCG TTA CTT CAC C-3') and the second set of primers (Forward 5'-GCT TTG TCC TGG GTG AGC G-3' and Reverse 5'-CCT GCG GCA TCA AAA CAC G-3') were utilized separately to generate two overlapping fragments containing the full ORF2 sequence (Arruda et al., 2019). These two PCRs were carried out using Taq DNA polymerase master mix (Qiagen, Vilnius, Lithuania). The following cycling conditions were used: initial activation phase at 94 degC for 10 min; 45 two-step cycles of denaturation at 94 degC for 30 s; annealing at 55 degC for 30 s and extension at 68 degC for 45 s (Arruda et al., 2019). A 444 and 643 bp PCR products could be visualized using Qiagen QIAxcel instrument, respectively. Then PCR products were sent to Macrogen (Seoul, South Korea) for Sanger sequencing.

The Sanger sequencing chromatograms of PCR products were basecalled and assembled using Tracy v0.5.9 (Rausch et al., 2020) to obtain full ORF2 sequences. Multiple alignment of ORF2 sequences was performed using Mafft v7.475 (Katoh and Standley, 2013). To determine the best-fit nucleotide substitution model, jModelTest2 v2.1.10 (Darriba et al., 2012) was used to assess model fit to the sequence alignment across 88 nucleotide substitution models based on Bayesian Information Criterion estimates. The maximum likelihood tree was generated using RAxML-NG v1.0.3 (Kozlov et al., 2019). Maximum likelihood inference was performed using the best-fit HKY nucleotide model with invariant sites (+I) and Gamma distributed rates (+G). The robustness of the phylogenetic tree nodes was determined by bootstrap analysis of 1000 replicates. The phylogenetic tree was visualized using the ggtree package v2.4.1 in R (Yu, 2020). Pairwise sequence alignment and calculation of sequence identity were conducted using Jalview v2.11.1.4 (Waterhouse et al., 2009). All 44 sequences were deposited to the National Center for Biotechnology Information (NCBI) GenBank database (accession numbers provided in Supplementary Table 1).

Sequence analysis revealed that cap gene sequences of 44 Canadian strains had 96.43% to 100% nucleotide identity with each other and 88.28% to 100% nucleotide identity with those of 33 reference strains included in this study. According to the criteria of genetic distance of 6% at the ORF2 level (3% at the complete genome level) and over 90% bootstrap support, PCV3 is classified into two clades: clade 1 (PCV3a) and clade 2 (PCV3b) (Franzo et al., 2020). Clade 2 (PCV3b) contains only two sequences identified from Chinese farms (Franzo et al., 2020). After removing the reference sequence from clade 2 (PCV3b) (GenBank number MG372488) included in this study, the resulting pairwise nucleotide identity across all sequences ranged from 96.74% to 100%. This supports the statement of genetic stability and low mutation rates of PCV3 (Arruda et al., 2019; Assao et al., 2021a; Qi et al., 2019). All 44 Canadian sequences in this study were classified into clade 1 (PCV3a) along with included reference strains (Figure 1), which strengthens the hypothesis that clade 2 (PCV3b) including only two sequences is considered to be either recently emerged variants or the last descendant of previously circulating genotypes (Franzo et al., 2020). The phylogenetic tree of the cap gene (Figure 1) showed that Canadian strains were subdivided into seven lineages based on nucleotide patterns. A total of 12 Canadian strains (Lineage 1) clustered with strains detected in different countries. A cluster of four Canadian strains (Lineage 2) clustered with a strain from Malaysia and a cluster of 11 Canadian strains (Lineage 3) clustered with a strain in USA were found. Lineage 4 containing only one Canadian strain (12769/2021/MB-CAN) clustered with strains from China. Lineage 5 and 6 containing six Canadian strains respectively clustered separately. Lineage 7 containing four Canadian strains clustered with a strain from Germany. Assao et al., demostrated that at least six different PCV3 lineages circulating in North and South America based on amino acid patterns with only one available Canadian strains in that study (Assao

et al., 2021a). Further study is necessary to determine whether the nucleotide variations in cap gene, as revealed in this study, could be associated with differences in the pathogenicity of PCV3 strains.

Considering cap as the major structural protein and the main antigen of PCV3 (Klaumann et al., 2018), the amino acid sequences of the 44 Canadian PCV3 strains included in this study were compared to the reference strain NC_031753. Sequence analysis showed that the presence of 17 amino acid substitutions in the cap protein (Table 1). Common amino acid changes at A24V, R27K, N56D, T77S, Q98R, L150I (F) and R168K positions were observed (Table 1). It has been speculated that A24V and R27K might be related to immune escape and were used for lineage or subclade classification of PCV3 (Assao et al., 2021a; Gu et al., 2020; Fu et al., 2018). In this study, mutation of ^{24}ARRR^{27,} ^{24}ARRK^{27}and ^{24}VRRK^{27} was found without mutation of ^{24}VRRR^{27}, which was reported previously (Sun et al., 2018). Amino acid change at T77N position was observed on both sequences from AB and ON, along with other amino acids changes at R4K, A5P (T, S, G), F7Y, A75T, V134Y, K140T, S156G (T), L173V and G196E positions across different provinces (Table 1), suggesting the importance of these differences for understanding the evolutions of these strains.

Conclusion

Our results demonstrated that Canadian PCV3 strains submitted to diagnostic laboratories and sequenced in this study were not diverse. All of the Canadian strains classified into PCV3a and segregated into different clusters with strains from North America, South America, Asia and Europe. Amino acid polymorphisms found in this study indicated the importance of these differences for understanding the evolutions of these strains. This is the first phylogenetic analysis of PCV3 strains among different provinces in Canada. The new PCV3 sequences will allow further investigations to explore the evolutions of PCV3 and future studies are required to better determine the polymorphisms in cap proteins could be associated with differences in the pathogenicity or antigenicity of PCV3 strains.

Acknowledgements

We would like to thank Anatoliy Trokhymchuk, Dhinesh Periyasamy and Anju Tumber in Prairie Diagnostic Services, Inc. in addition to Veronique Bournival at CDVUM for technical assistance. C.A. Gagnon was financially supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) discovery grant and by a Canadian Swine Research and Development Cluster (CSRDC) grant.

References

Arruda, B., P. Pineyro, R. Derscheid, B. Hause, E. Byers, K. Dion, D. Long, C. Sievers, J. Tangen, T. Williams and K. Schwartz, 2019: PCV3-associated disease in the United States swine herd. *Emerg Microbes Infect,* 8, 684-698. https://doi.org/10.1080/22221751.2019.1613176

Assao, V. S., M. R. Santos, C. E. R. Pereira, F. Vannucci and A. Silva-Junior, 2021a: Porcine circovirus 3 in North and South America: Epidemiology and genetic diversity. *Transboundary and emerging diseases* . https://doi.org/10.1111/tbed.14238

Assao, V. S., M. R. Santos, N. C. L. Rosado, G. C. Bressan, J. L. R. Fietto, Y. F. Chang, P. M. P. Vidigal and A. Silva-Junior, 2021b: Genetic diversity of porcine circovirus 3 strains and the first detection of two different PCV3 strains coinfecting the same host in Minas Gerais, Brazil. *Arch Virol,* 166, 1463-1468. https://doi.org/10.1007/s00705-021-05032-y

Darriba, D., G. L. Taboada, R. Doallo and D. Posada, 2012: jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods,* 9, 772-772. https://doi.org/10.1038/nmeth.2109

Franzo, G., E. Delwart, R. Fux, B. Hause, S. Su, J. Zhou and J. Segales, 2020: Genotyping Porcine Circovirus 3 (PCV-3) Nowadays: Does It Make Sense? *Viruses,* 12. https://doi.org/10.3390/v12030265

Fu, X., B. Fang, J. Ma, Y. Liu, D. Bu, P. Zhou, H. Wang, K. Jia and G. Zhang, 2018: Insights into the epidemic characteristics and evolutionary history of the novel porcine circovirus type 3 in southern China.*Transboundary and emerging diseases,* 65, e296-e303. https://doi.org/10.1111/tbed.12752

Gu, W., Q. Shi, L. Wang, J. Zhang, G. Yuan, S. Chen, Y. Zuo and J. Fan, 2020: Detection and phylogenetic analysis of porcine circovirus 3 in part of northern China from 2016 to 2018. *Arch Virol,* 165, 2003-2011. https://doi.org/10.1007/s00705-020-04709-0

Jiang, H., D. Wang, J. Wang, S. Zhu, R. She, X. Ren, J. Tian, R. Quan, L. Hou, Z. Li, J. Chu, Y. Guo, Y. Xi, H. Song, F. Yuan, L. Wei and J. Liu, 2019: Induction of Porcine Dermatitis and Nephropathy Syndrome in Piglets by Infection with Porcine Circovirus Type 3. *Journal of Virology,* 93. https://doi.org/10.1128/jvi.02045-18

Katoh, K. and D. M. Standley, 2013: MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.*Molecular Biology and Evolution,* 30, 772-780. https://doi.org/10.1093/molbev/mst010

Klaumann, F., F. Correa-Fiz, G. Franzo, M. Sibila, J. I. Nunez and J. Segales, 2018: Current Knowledge on Porcine circovirus 3 (PCV-3): A Novel Virus With a Yet Unknown Impact on the Swine Industry.*Frontiers in Veterinary Science,* 5. https://doi.org/10.3389/fvets.2018.00315

Kozlov, A. M., D. Darriba, T. Flouri, B. Morel and A. Stamatakis, 2019: RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics,* 35, 4453-4455. https://doi.org/10.1093/bioinformatics/btz305

Nawagitgul, P., I. Morozov, S. R. Bolin, P. A. Harms, S. D. Sorden and P. S. Paul, 2000: Open reading frame 2 of porcine circovirus type 2 encodes a major capsid protein. *J Gen Virol,* 81, 2281-2287. https://doi.org/10.1099/0022-1317-81-9-2281

Palinski, R., P. Pineyro, P. Shang, F. Yuan, R. Guo, Y. Fang, E. Byers and B. M. Hause, 2017: A Novel Porcine Circovirus Distantly Related to Known Circoviruses Is Associated with Porcine Dermatitis and Nephropathy Syndrome and Reproductive Failure. *Journal of Virology,* 91. https://doi.org/10.1128/jvi.01879-16

Phan, T. G., F. Giannitti, S. Rossow, D. Marthaler, T. P. Knutson, L. Li, X. Deng, T. Resende, F. Vannucci and E. Delwart, 2016: Detection of a novel circovirus PCV3 in pigs with cardiac and multi-systemic inflammation. *Virology journal,* 13, 184. https://doi.org/10.1186/s12985-016-0642-z

Plut, J., U. Jamnikar-Ciglenecki, I. Golinar-Oven, T. Knific and M. Stukelj, 2020: A molecular survey and phylogenetic analysis of porcine circovirus type 3 using oral fluid, faeces and serum. *BMC Vet Res,* 16, 281. https://doi.org/10.1186/s12917-020-02489-y

Qi, S., M. Su, D. Guo, C. Li, S. Wei, L. Feng and D. Sun, 2019: Molecular detection and phylogenetic analysis of porcine circovirus type 3 in 21 Provinces of China during 2015-2017. *Transboundary and emerging diseases,* 66, 1004-1015. https://doi.org/10.1111/tbed.13125

Rausch, T., M. H.-Y. Fritz, A. Untergasser and V. Benes, 2020: Tracy: basecalling, alignment, assembly and deconvolution of sanger chromatogram trace files. *BMC Genomics,* 21, 230. https://doi.org/10.1186/s12864-020-6635-8

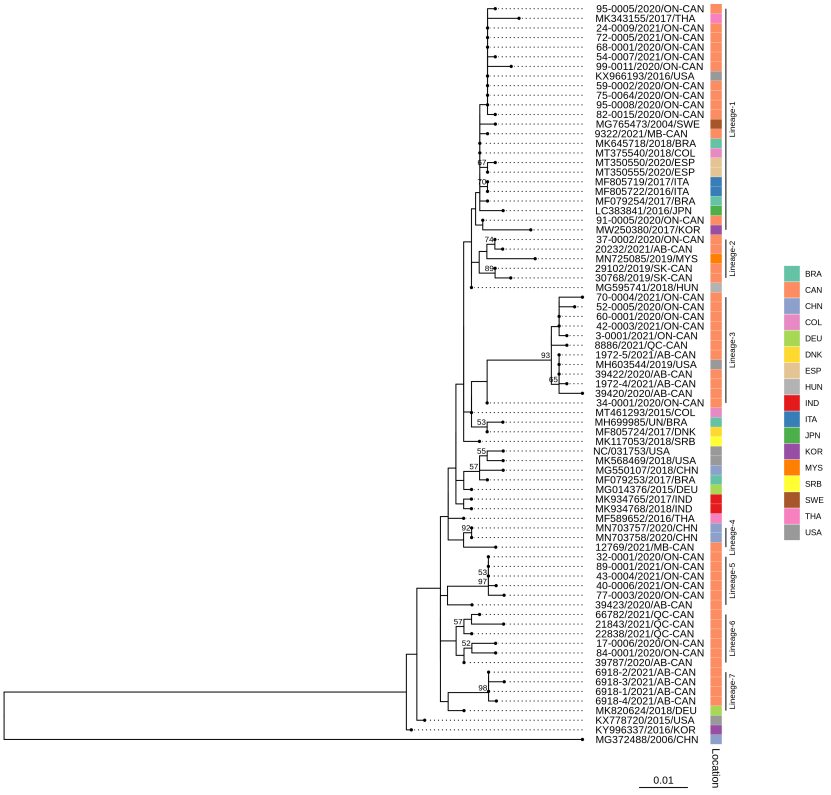
Sun, J., L. Wei, Z. Lu, S. Mi, F. Bao, H. Guo, C. Tu, Y. Zhu and W. Gong, 2018: Retrospective study of porcine circovirus 3 infection in China. *Transboundary and emerging diseases,* 65, 607-613. https://doi.org/10.1111/tbed.12853

Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton, 2009: Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25, 1189-1191. https://doi.org/10.1093/bioinformatics/btp033

Yu, G., 2020: Using ggtree to Visualize Data on Tree-Like Structures.*Curr Protoc Bioinformatics,* 69, e96. https://doi.org/10.1002/cpbi.96

Figure Legends

Figure 1. Phylogenetic analysis of PCV3 cap genes based on the nucleotides sequences. All reference strain sequences included in this study were downloaded from the NCBI GenBank database. The tree was constructed by the maximum-likelihood method using the HKY nucleotide model with invariant sites (+I) and Gamma distributed rates (+G) and 1000 bootstrap replicates. Bootstrap values are indicated for nodes with values >50%. Canadian PCV3 strains are indicated by light orange color.



Hosted file

`Table 1.docx` available at https://authorea.com/users/738133/articles/712617-phylogenetic-analysis-of-porcine-circovirus-3-pcv3-circulating-in-canadian-pigs