DR. RENE ALVARO SEGURA-VELÁZQUEZ (Orcid ID : 0000-0001-8392-4919)

DR. IVAN SANCHEZ-BETANCOURT (Orcid ID : 0000-0001-8201-9669)

FIRST REPORT AND PHYLOGENETIC ANALYSIS OF PORCINE DELTACORONAVIRUS IN MEXICO

Claudia Pérez-Rivera [1], Humberto Ramírez-Mendoza[2], Susana Mendoza-Elvira[3], Rene Segura-Velázquez [4], José Ivan Sánchez-Betancourt [1*]

[1] Departamento de Medicina y Zootecnia de Cerdos FMVZ-UNAM. Av. Universidad No. 3000. Copilco, Del. Coyoacán, CP 04510 Mexico City, Mexico

[2] Departamento de Microbiología e Inmunología, Facultad de Medicina Veterinaria y Zootecnia (FMVZ), Universidad Nacional Autónoma de México (UNAM), CP 04510, Mexico City, Mexico.

[3] Laboratorio de Microbiología. Facultad de Estudios Superiores Cuautitlán, UNAM.

[4] Unidad de Investigación, Facultad de Medicina Veterinaria y Zootecnia (FMVZ), Universidad Nacional Autónoma de México (UNAM), CP 04510, México City, México.

*Correspondence: Dr. José Iván Sánchez Betancourt, Departamento de Medicina y Zootecnia de Cerdos FMVZ-UNAM. Av. Universidad No. 3000. Copilco, Del. Coyoacán, CP 04510 Mexico City, Mexico. Email: aisb_7@yahoo.com.mx

Summary

Porcine deltacoronavirus has caused great economic losses in the swine industry worldwide. In this study, we carried out the first detection, sequencing and characterization of this virus in Mexico. We analyzed 885 rectal samples by multiplex RT-PCR to determine coinfections. In addition, the Spike gene was amplified, sequenced and analyzed phylogenetically. We found 85 positive samples for porcine deltacoronavirus, representing 9.6% of the total samples, and we determined that the most frequent coinfection was with porcine epidemic diarrhea virus (54.1%). Four sequences of Mexican isolates were most closely related to those of the United States. The antigenic regions and the glycosylation site of the strains obtained coincide with those previously reported. This relationship is probably related to the commercial exchange of pigs between the US and Mexico and the geographical proximity of these two countries.

Keywords: Porcine deltacoronavirus, Mexico, phylogenetic analysis, prevalence, spike gene.

Introduction

Porcine deltacoronavirus (PDCoV) is an emerging virus that was recently described. This viral disease together with PED (Porcine epidemic diarrhea) have caused a significant economic impact due to the high mortality rate in piglets (Jung, Hu, & Saif, 2016; J. Zhang, 2016). PDCoV belongs to the recently classified subgenus *Buldecovirus* of the *Deltacoronavirus* genus (ICTV, 2018). The subfamily *Orthocoronavirinae* is divided into 4 genera, *Alphacoronavirus, Betacoronavirus, Gammacoronavirus* and *Deltacoronavirus*. The first two genera originate in bats, while the third and fourth genera originate in birds (Woo et al., 2012). These viruses can affect several species of mammals and birds and cause various clinical conditions (neurological, digestive, respiratory) (Wang, Byrum, & Zhang, 2014).

Coronaviruses are unique within RNA viruses because their genome is particularly large (~ 27.5 to ~ 31 kb). These enveloped viruses, with a single chain in the 5'-3' direction and a helical nucleocapsid, contain at least 7 open reading frames (ORF), which code for 4 structural proteins (Masters, 2006; Sawicki, 2009). Coronaviruses received their name due to one of their surface proteins, which is spicule-shaped, giving the appearance of a crown. This so-called S or "Spike" protein is glycosylated and plays a fundamental role in the binding and entry of the virus into the cell (Masters, 2006). The genome organization of PDCoV is in the following order: 5´ untranslated region (UTR), replicase (ORF 1ab), spike gene (S), envelope gene (E), membrane gene (M), nonstructural gene 6 (NS6), nucleocapsid gene (N), NS7 gene, and 3´UTR. Particularly, PDCoV lacks ORF 3 and non-structural protein 1 (Lee & Lee, 2014; Woo et al., 2012; Q. Zhang & Yoo, 2016).

In pigs, at least six coronaviruses are known to cause diseases. These belong to three of the four known genera. Within the *Alphacoronaviruses* are the transmissible gastroenteritis virus (TGEV), the porcine epidemic diarrhea virus (PEDV), the porcine respiratory coronavirus (PRCV) and the HKU2-related bat coronavirus, which were described in China in 2016 related to acute diarrhea in pigs (Zhou et al., 2018). In the *Betacoronavirus,* there is the porcine hemagglutinating encephalomyelitis virus (PHEV), and finally, in the genus of the *Deltacoronaviruses,* there is the newly described porcine deltacoronavirus (PDCoV) (Chen et al., 2015; ICTV, 2018).

Porcine deltacoronavirus (PDCoV) was detected for the first time in Hong Kong in samples collected in 2009. This study aimed to determine the variability of bat coronaviruses due to the importance that the *Coronaviridae* family has in Hong Kong since the presentation of acute respiratory diseases in humans. In this work, pigs did not present an apparent clinical sign; however, PDCoV was detected (Woo et al., 2012).

Subsequently, in early 2014, PDCoV was reported from the United States of America (USA) and Canada, and it caused heavy economic losses to the swine industry due to the presentation of a clinical enteric disease (Marthaler, Jiang, Collins, & Rossow, 2014; Song et al., 2015; Wang et al., 2014). The described infection was indistinguishable from that caused by PEDV or TGEV. However, the first reports in the USA mentioned that the mortality in piglets was lower (30-40%) than that normally observed in outbreaks of PED (Marthaler, Raymond, et al., 2014; Wang et al., 2014).

Experimental studies in piglets show that the most frequent signs related to PDCoV are watery diarrhea, vomiting and dehydration (Ma et al., 2015). The authors note that PDCoV infections are common in pigs and that coinfections are frequent, especially with the porcine epidemic diarrhea virus and rotavirus C (Hu et al., 2015; Marthaler, Raymond, et al., 2014; Song et al., 2015). Likewise, some authors speculated that coinfections are likely to cause higher mortality rates (Song et al., 2015).

Despite the impact that the disease has had in several pig-producing countries, until this year, we did not know about its presence or the genetic characteristics of the virus in Mexico. The objective of this work was to identify the presence of this virus in Mexico and to analyze the genomic sequence of the S gene (Spike).

Materials and Methods

Sample collection and RNA extraction

To monitor the prevalence and sequence properties of PDCoV in Mexico, 885 porcine rectal swabs were collected from 5 different regions of the country from 2014 to 2017. The sampling locations are shown in Figure 1. These samples were preserved at -70 ° C until use, later they were diluted 1:5 with 1X PBS (pH 7.4) and then centrifuged at 5000 rpm for 15 minutes. Once the sample was centrifuged, the supernatant was collected, and the total RNA was extracted using a commercial kit following the manufacturer's recommendations (kit QIAamp Viral RNA. Qiagen Cat. 52906). The resulting yield of RNA extracted was 60 µL at an average concentration of 7 ng / µL.

Nucleic acid detection

Real Time-PCR

RT-qPCR was carried out to address the frequency of PDCoV monoinfection and coinfection(s) with PEDV and TGEV. We used the VetMAX™ PEDV/TGEV/SDCoV kit (Applied Biosystems Cat. A33402) following the manufacture´s recommendations.

Phylogenetic analysis

Spike (S) protein gene amplification

To amplify the S gene, RNA from the samples that were positive by RT-qPCR, was reverse-transcribed (RT) using the reverse primer 5´CACTATGTCTGACGCAGAAG3´ and Superscript II (Invitrogen, San Diego, CA). The RT products were then used to perform PCR using primers specifically targeting the S gene of PDCoV (forward, 5´CAGAATGCAGAGAGCTCTAT3´ and reverse, 5´CGTCAGACATAGTGAGTGTT3´) based on the sequence PDCoV/USA/Illinois121/2014 (GenBank accession no. KJ481931) using the NCBI tool (National Center for Biotechnology Information) primer-BLAST. Amplification of fragments of 3734 bp was obtained.

A high-fidelity polymerase was used in the PCR (LongAmp Hot Start Taq 2X Master Mix, Biolabs® cat #M0533S) following the manufacture´s recommendations. The protocol was carried out in a LABNET Multigen Optimax thermocycler following denaturation at 94 °C for 30 s; 35 cycles of 94 °C x 15 s, 52 °C alignment temperature for 30 s, and 65 °C x 4 min; and a final extension of 65 °C for 10 min.

Sequencing

The cDNA obtained from the S gene amplification was subjected to sequencing using the Ion Personal Genome Machine platform (Ion Torrent Thermo Fisher Scientific) following the supplier's specifications for DNA sequencing. The obtained readings were filtered with the FastQC plug-in v3.4.1.1, selecting only those with a Q score ≥ 20 (6,074,916 reads); the obtained depth was greater than 300X.

Alignment and phylogenetic analysis of the S gene (tree construction)

The alignment of the sequences obtained from the amplification of the S gene was performed using the Clustal W program in MEGA7 software (Kumar, Stecher, & Tamura, 2016)**.** For the construction of the tree, the neighbor-joining distance-based method was used, and the bootstrap analysis consisted of 1000 repetitions using the same software.

Spike protein analysis

To predict the epitopes, the S protein was analyzed with Protean, DNAStar V.7.1 (Madison, WI, EE. UU.). The Jameson-Wolf method predicted the potential antigenicity sites, which were compared with those reported by Mai, K and collaborators based on the sequence HKU15-155 (GenBank accession no. AFD29194.1). Likewise, the surface properties, including hydrophobicity, accessibility and flexibility, were analyzed by the Kyte-Doolittle, Plot-Emini and Karplus-Schulz methods, respectively, also in Protean software. Furthermore, the positions of the epitopes were predicted using the online service http://www.cbs.dtu.dk/services/BepiPred/. The modeling of the amino acid sequence

of the protein was carried out in SWISS-MODEL (https://swissmodel.expasy.org/). The glycosylation sites were analyzed with NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc).

Results and Discussion

PDCoV infections in Mexico have probably not been adequately addressed due to the high prevalence PED in the country since 2013. Pig farms have experienced sporadic outbreaks of diarrhea in pigs less than 3 weeks old, often without determining the causative agent, while assuming that the causative agent is PEDV, possibly masking the presence of deltacoronavirus.

A total of 885 samples were analyzed from 5 different regions of Mexico, of which 85 (9.6%) were PDCoV positive (Table 1, Figure 1). The most common coinfection was PDCoV/PEDV, found in 54.1% of the total deltacoronavirus-positive cases (46/85), a result that coincides with that reported by other authors (Song et al., 2015; J. Zhang, 2016). Coinfections with TGEV are the least frequent (10.6%; 9/85), while 16.5% (14/85) of the samples presented PDCoV as the only infectious agent. The central region presented the highest percentage of positive samples (15.4%; 38/246), followed by the southeastern region (13.3%; 24/180), which are characterized by having highly productive pig farms (Table 1, Figure 1). In addition, it is interesting that a considerable number of samples (16/85) were positive for the three viral agents studied (PDCoV/PEDV/TGEV).

The S protein has the same structure in all coronaviruses. It is the most studied protein due to the role it plays in binding to the receptor and thus determining tropism, in addition to mediating the fusion of membranes for entry of the virus into the cell (Li et al., 2018; Shang et al., 2018). Within the coronaviruses, the genus *Deltacoronavirus* is the one with the smallest S protein; however, the structure is the same; it is divided into 2 subunits, S1 (1-573 aa) and S2 (574-1160 aa) (Thachil, Gerber, Xiao, Huang, & Opriessnig, 2015).

Four sequences of the complete S gene were obtained from 3 different regions, to which the following names were designated PDCoV Mex/Yuc 2015, PDCoV Mex/Oax 2017 (Southeast), PDCoV Mex/Edo Mex 2017 (Center) and PDCoV Mex/Qro 2017 (West). The S gene of the four Mexican sequences is composed of 3,480 nt and has a nucleotide identity between them of 99.4% on average. The highest identity was between the Edo Mex and Oax strains (99.8%) and the lowest identity value was between Yuc and Oax strains (99%). S gene sequences obtained from Mexican isolates shared 99.6% nucleotide identity with respect to North American strains. In addition, it presents a homology of 98 to 98.7% with Chinese sequences and 96.5 and 96.1% with sequences from Vietnam and Thailand, respectively.

A more detailed comparison analysis of Mexican strains was performed against the amino acid sequences of the S gene from the strains Ilinois121 / 2014 (GenBank accession no. AHL45007) and Ohio137 / 2014 (GenBank accession no. AIB07807), which have structures known by electron microscopy (Shang et al., 2018; Xiong et al., 2018). We observed the mutation of amino acids at 6 sites (110 E / D, 221 N / K, 510 NI, 534 KN, 550 IL, 624 AV).

We built a phylogenetic tree using the neighbor-joining method with 48 sequences from all countries that are available in GenBank. Three main groups were formed. In the first group are sequences from China (in green, Figure 2) and in the second, sequences from Thailand, Laos, and Vietnam are grouped (in orange). Finally, in the third group, we found sequences from the US, Japan and South Korea (in blue). This third group also comprises the sequences of the S gene from Mexico, which form 2 clades, one with sequences from 2017 and the second with the sequence from 2015. However, these clades are grouped closely with the strains from the United States.

Comparing the sequence of the 11 antigenic regions previously reported by Mai et al in 2017 against Mexican sequences, we found 6 mutations and 1 insertion in 4 of the 11 epitopes. We observed that at positions 38 to 53, a 1 amino acid insertion (N) increased the antigenic index, surface probability and hydrophilic level (Figure 3-C and S1) (Mai et al., 2017). Likewise, 17 glycosylations were located throughout the S1 region, similar those predicted by Xiaoli, X in 2018 to with the Illinois 2014 strain (Figure 3-A, B). Knowing the glycosylation sites is important in the study of the antigenicity of the virus because the sites are part of a viral strategy to evade the host immune system (Shang et al., 2018). We know that the immune response of B cells is against the spike protein (Mai et al., 2017); however, deeper studies would be necessary to evaluate whether these predictions coincide with the real antigenicity and pathogenicity of the protein.

In conclusion, we determined that porcine deltacoronavirus occurs in Mexico and that it is frequently associated with other pathogens, mainly PEDV and TGEV. To date, recombinations of PDCoV with other coronaviruses have not been reported, but the recombinant capacity of these viruses is known, as demonstrated by the event reported in Italy, where porcine enteric coronavirus (SeCoV) was found. The genome of the S gene of this virus has greater homology with the PED virus, while the rest of its genome has homology with the TGE virus (Boniotti et al., 2016).

Likewise, the Mexican strain is phylogenetically closer to those strains reported in the US. This supports the theory of the current global distribution of porcine deltacoronavirus. Further analysis of the structure and changes found in the S protein of the PDCoV from Mexico are necessary to determine its degree of pathogenicity and antigenicity.

Acknowledgments

Conflict of Interest

The authors declare no financial or commercial conflicts of interest.

References

Boniotti, M. B., Papetti, A., Lavazza, A., Alborali, G., Sozzi, E., Chiapponi, C., …
Marthaler, D. (2016). Porcine epidemic diarrhea virus and discovery of a
recombinant swine enteric coronavirus, Italy. *Emerging Infectious Diseases*,
22(1), 83–87. http://doi.org/10.3201/eid2201.150544

Chen, Q., Gauger, P., Stafne, M., Thomas, J., Arruda, P., Burrough, E., … 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. http://doi.org/http://dx.doi.org/10.1016/j.virol.2015.03.024

Hu, H., Jung, K., Vlasova, A. N., Chepngeno, J., Lu, Z., Wang, Q., & Saif, L. J.
(2015). Isolation and Characterization of Porcine Deltacoronavirus from Pigs
with Diarrhea in the United States. *Journal of Clinical Microbiology*, *53*(5),
JCM.00031-15. http://doi.org/10.1128/JCM.00031-15

ICTV. (2018). Virus Taxonomy: 2018 Release ICTV. Retrieved November 6, 2018,
from https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-
viruses-2011/w/posrna_viruses/222/coronaviridae

Jung, K., Hu, H., & Saif, L. J. (2016). Porcine deltacoronavirus infection: Etiology,
cell culture for virus isolation and propagation, molecular epidemiology and
pathogenesis. *Virus Research*, *226*, 50–59.
http://doi.org/10.1016/j.virusres.2016.04.009

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA 7. Molecular Biology and
Evolution. https://doi.org/10.1093/molbev/msw054

Lee, S., & Lee, C. (2014). Complete Genome Characterization of Korean Porcine
Deltacoronavirus Strain KOR/KNU14-04/2014. *Genome Announcements*, *2*(6),
4–5. http://doi.org/10.1128/genomeA.01191-14.Copyright

Li, W., Hulswit, R. J. G., Kenney, S. P., Widjaja, I., Jung, K., & Alhamo, M. A. (2018). Broad receptor engagement of an emerging global coronavirus may potentiate its diverse cross-species transmissibility. *Proceedings of the National Academy of Sciences*, *115*(20), 1–9. http://doi.org/10.1073/pnas.1802879115

Ma, Y., Zhang, Y., Liang, X., Lou, F., Oglesbee, M., Krakowka, S., & Li, J. (2015). Origin, Evolution, and Virulence of Porcine Deltacoronaviruses in the United States. *Microbiology*, *6*(2), 1–13. http://doi.org/10.1128/mBio.00064-15.Invited

Mai, K., Feng, J., Chen, G., Li, D., Zhou, L., Bai, Y., … Ma, J. (2017). The detection and phylogenetic analysis of porcine deltacoronavirus from Guangdong Province in Southern China. *Transboundary and Emerging Diseases*, (November 2016), 166–173. http://doi.org/10.1111/tbed.12644

Marthaler, D., Jiang, Y., Collins, J., & Rossow, K. (2014). Porcine Deltacoronavirus from the United States. *Genome Announcements*, *2*(2), 2–3. http://doi.org/10.1128/genomeA.00218-14.Copyright

Marthaler, D., Raymond, L., Jiang, Y., Collins, J., Rossow, K., & Rovira, A. (2014). Rapid detection, complete genome, and phylogenetic analysis of Porcine Deltacoronavirus (Technical Appendix). *Emerging Infectious Diseases*, *20*(8), 1347–1350. http://doi.org/http://dx.dpi.org/10.3201/eid2008.14-0526

Masters, P. S. (2006). The Molecular Biology of Coronaviruses. *Advances in Virus Research*, *66*, 193–292. http://doi.org/10.1016/S0065-3527(06)66005-3

Sawicki, S. G. (2009). Coronavirus Genome Replication. *Viral Genome Replication*. http://doi.org/10.1007/b135974

Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Tai, W., … Li, F. (2018). Cryo-electron microscopy structure of porcine deltacoronavirus spike protein in the prefusion state. *Journal of Virology*, *92*(4), 1–14. http://doi.org/10.1128/jvi.01556-17

Song, D., Zhou, X., Peng, Q., Chen, Y., Zhang, F., Huang, T., … 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. http://doi.org/10.1111/tbed.12399

Thachil, A., Gerber, P. F., Xiao, C. T., Huang, Y. W., & Opriessnig, T. (2015). Development and application of an ELISA for the detection of porcine deltacoronavirus IgG antibodies. *PLoS One*, *10*(4), e0124363. http://doi.org/10.1371/journal.pone.0124363

Wang, L., Byrum, B., & Zhang, Y. (2014). Detection and Genetic Characterization of Deltacoronavirus in Pigs, Ohio, USA, 2014. *Emerging Infectious Diseases*, *20*(7), 1227–30. http://doi.org/10.3201/eid2007.140296

Woo, P. C. Y., Lau, S. K. P., Lam, C. S. F., Lau, C. C. Y., Tsang, A. K. L., Lau, J. H. N., … 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 Deltacoronavi. *Journal of Virology*, *86*(7), 3995–4008. http://doi.org/10.1128/JVI.06540-11

Xiong, X., Tortorici, A., Snijder, J., Yoshioka, C., Walls, A., Li, W., … Veesier, D. (2018). Glycan shield and fusion activation of a deltacoronavirus spike glycoprotein fine-tuned for enteric infections. *J of Virology*, *92*(4), 1–16.

http://doi.org/https://doi.org/10.1128/JVI.01556-1

Zhang, J. (2016). Porcine deltacoronavirus: Overview of infection dynamics, diagnostic methods, prevalence and genetic evolution. *Virus Research*, *226*, 71–84. http://doi.org/10.1016/j.virusres.2016.05.028

Zhang, Q., & Yoo, D. (2016). Immune evasion of porcine enteric coronaviruses and viral modulation of antiviral innate signaling. *Virus Research*, *226*, 128–141. http://doi.org/10.1016/j.virusres.2016.05.015

Zhou, P., Fan, H., Lan, T., Yang, X. Lou, Shi, W. F., Zhang, W., … Ma, J. Y. (2018). Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. *Nature*, *556*(7700), 255–259. http://doi.org/10.1038/s41586-018-0010-9
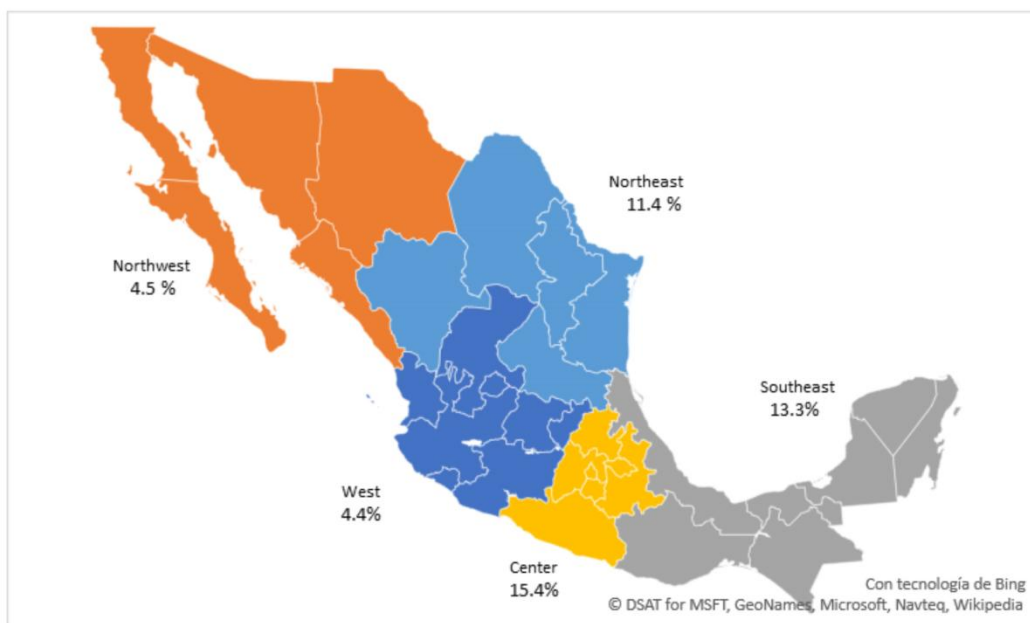
Table 1. Presence and prevalence of infections of different pathogens in samples by region

PDCoV, porcine deltacoronavirus; PEDV, porcine epidemic diarrhea virus; TGEV, transmissible gastroenteritis virus.

Region	n	No. of positive samples	% of positives	PDCoV (+ / %)	PDCoV/PEDv (+ / %)	PDCoV/TGEv (+ / %)	PDCoV/PEDv/TGEv (+ / %)
Center	246	38	15.4	**2** / 0.8	**24** /9.7	0	12 /4.9
Southeast	180	24	13.3	**6** / 3.3	10/ 5.5	6/ 3.3	2/ 1.1
Northeast	45	5	11.1	**2** / 4.4	2 /4.4	1 /2.2	0
Northwest	48	2	4.2	0	0	1/ 2.1	1/ 2.1
West	366	16	4.4	**4** / 1.1	10 / 2.7	1/ .3	1/ .3
Total	**885**	**85**	**9.6**	**14/ 1.6**	**46/ 5.2**	**9/ 1.0**	**16/ 1.8**

Figure 1. Regions of Mexico and percentage of PDCoV positives

Northwest
4.5 %

Northeast
11.4 %

Southeast
13.3%

West
4.4%

Center
15.4%

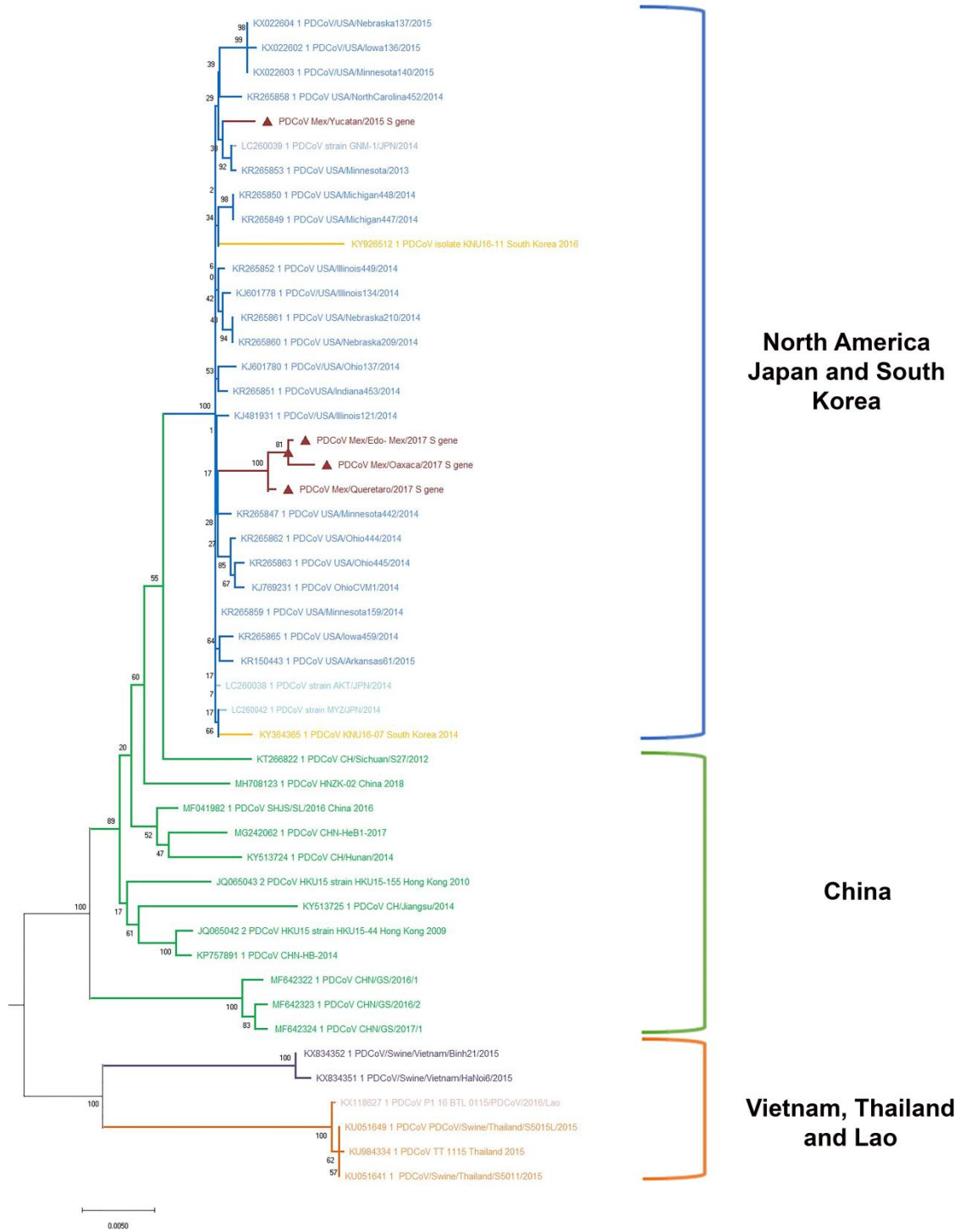Con tecnología de Bing
© DSAT for MSFT, GeoNames, Microsoft, Navteq, Wikipedia

Figure 3. Glycosylation sites on PDCoV spike.

A

B

C

Position	PDCoV MEX/YUC 2015 Aminoacid sequence
38-53	PTRNSSSLYSRANNNF
58-64	LPGYPTK
109-114	SDIQTY
147-150	GQHG
273-291	SSHEVEDGFYSDPKSAVRA
328-335	INGNTSYC
410-427	EGQTHTGTTSVHATDTSS
460-466	FTSPTGE
502-512	TSSNSTENIRF
549-561	ALVGTSTLQNTRP
570-574	EVEIP