

Emergence of porcine delta-coronavirus pathogenic infections among children in Haiti through independent zoonoses and convergent evolution

John A. Lednicky,^{1,2†} Massimiliano S. Tagliamonte,^{1,3†} Sarah K. White,^{1,2} Maha A. Elbadry,^{1,2} Md. Mahbubul Alam,^{1,2} Caroline J. Stephenson,^{1,2} Tania S. Bonny,^{1,2} Julia C. Loeb,^{1,2} Taina Telisma,⁴ Sonese Chavannes,⁴ David A. Ostrov,^{1,3} Carla Mavian,^{1,3} Valerie Madsen Beau De Rochars,^{1,5} Marco Salemi,^{1,3*} J. Glenn Morris, Jr.^{1,6*}

1. Emerging Pathogens Institute, University of Florida, Gainesville, FL
2. Department of Environmental and Global Health, College of Public Health and Health Professions, University of Florida, Gainesville, FL
3. Department of Pathology, Immunology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL
4. Christianville Foundation, Gressier, Haiti
5. Department of Health Services Research, Management and Policy, College of Public Health and Health Professions, University of Florida, Gainesville, FL
6. Department of Medicine, College of Medicine, University of Florida, Gainesville, FL

† These authors contributed equally

*Corresponding authors:

Dr. Glenn Morris
Emerging Pathogens Institute
University of Florida
2055 Mowry Rd., Gainesville, FF
326100009
jgmmorris@epi.ufl.edu

Dr. Marco Salemi
Emerging Pathogens Institute
University of Florida
2055 Mowry Rd., Gainesville, FF
326100009
salemi@pathology.ufl.edu

Coronaviruses have caused three major epidemics since 2003, including the ongoing SARS-CoV-2 pandemic. In each case, coronavirus emergence in our species has been associated with zoonotic transmissions from animal reservoirs^{1,2}, underscoring how prone such pathogens are to spill over and adapt to new species. Among the four recognized genera of the family *Coronaviridae* – *Alphacoronavirus*, *Betacoronavirus*, *Deltacoronavirus*, *Gammacoronavirus*, – human infections reported to date have been limited to alpha- and betacoronaviruses³. We identify, for the first time, porcine deltacoronavirus (PDCoV) strains in plasma samples of three Haitian children with acute undifferentiated febrile illness. Genomic and evolutionary analyses reveal that human infections were the result of at least two independent zoonoses of distinct viral lineages that acquired the same mutational signature in the *nsp15* and the *spike* glycoprotein genes by convergent evolution. In particular, structural analysis predicts that one of the changes in the Spike S1 subunit, which contains the receptor-binding domain, may affect protein’s flexibility and binding to the host cell receptor. Our findings not only underscore the ability of deltacoronaviruses to adapt and potentially lead to human-to-human transmission, but also raise questions about the role of such transmissions in development of pre-existing immunity to other coronaviruses, such as SARS-CoV-2.

Coronaviruses are enveloped, positive-sense single-stranded RNA viruses that belong to the family *Coronaviridae*, subfamily *Coronavirinae*. They are responsible for respiratory syndromes and hepatic and gastroenteric diseases in birds and mammals⁴. Their host cell binding protein, the Spike (S) glycoprotein, binds to receptors relatively conserved across species^{5,6}; this characteristic can facilitate cross-species transmission, and lead to outbreaks and endemicity in the human population, as has already happened multiple times^{1,2}. Their propensity for recombination^{1,5,7,8} can also facilitate the escape from local fitness optima and broaden their species tropism^{1,5,8-10}.

Porcine deltacoronavirus (PDCoV) is a member of the genus *Deltacoronavirus*, and the virus was reported for the first time in Hong Kong, China, in 2012¹¹, and later in the Americas, in 2014¹². It causes gastrointestinal symptoms in piglets, with transient viremia, which may lead to dehydration and death^{13,14}. The virus infects villous epithelial cells of the entire large intestine, although the jejunum and ileum are the primary sites of infection^{13,15}. Studies^{6,16} have

demonstrated that PDCoV employs the host aminopeptidase N (APN) protein as an entry receptor through an interaction with it *via* domain B of its spike (S) glycoprotein⁶. Further cause of concern are reports of symptomatic infection, in experimental settings, of chickens and turkeys¹⁷. Human cells were also reported to be permissive to PDCoV infection⁶. Such evidence suggests that the binding of the PDCoV S glycoprotein to an interspecies conserved site (APN protein) may facilitate direct transmission to non-reservoir species, possibly including humans.

Haiti is part of the island of Hispaniola, and one of the poorest countries in the world¹⁸. Efforts were made to wipe out the local pig population in the 1980s to eliminate African swine fever from the area¹⁹. Pigs were then reintroduced from North American populations, mainly the USA and in small part, Canada^{19,20}. Pig farming in the country is at a subsistence level, and, to our knowledge, no PDCoV cases have been reported to date in pigs. Our group monitored occurrence of illness among children attending school within a school system operated by the Christianville Foundation in the Gressier region of Haiti from 2012 to 2020²¹⁻²⁴. During that time period, free medical care was provided to children at the school clinic. Among children seen at the clinic, respiratory and diarrheal illnesses were most common, followed by acute undifferentiated febrile illnesses (fever with no clear localizing symptoms) that were reported by approximately 16% of clinic attendees²¹.

Detection of human-PDCoV (Hu-PDCoV) strains in Haitian pediatric patients

This study was approved by the Institutional Review Board (IRB) at the University of Florida and the Haitian National IRB; written informed consent for sample collection was obtained from parents of participants, with assent from participants. Beginning in May 2014 plasma samples were collected from children presenting to the school clinic with acute undifferentiated febrile illness, with a total of 369 samples collected between May 2014 and December 2015. As described below, three of these children (aged 5-10 years) were subsequently found to be infected with coronavirus strains which clustered with PDCoV. Cases 1 and 3 were in patients from the main campus of the school (School A) that is attended by students from semi-urban areas, with students coming from a range of socioeconomic backgrounds. Case 2 was from a different campus (School B), which is located in the mountains approximately an hour's drive from School A, in a completely rural area, with students from very low socioeconomic backgrounds. All three children presented with a history of fever but

recovered uneventfully: Child 2 was febrile (40°C) when seen in the clinic; Child 1 and 2 reported cough and abdominal pain. Child 3, while reporting a fever, did not have acute symptoms when seen in clinic.

TABLE 1 | Sample data, children with PDCoV infection

Child case #	Collection date	School	PDCoV strain	GenBank No.	GenBank No. of closest PDCoV strain by BLAST analysis
1	Dec. 2014	School A	Haiti/Human/0081-4/2014	MW685622	KY065120 (Pig/China/2016)
2	March 2015	School B	Haiti/Human/0256-1/2015	MW685623	KR150443 (Pig/USA/2015)
3	April 2015	School A	Haiti/Human/0329-4/2015	MW685624	KY065120 (Pig/China/2016)

Nucleic acids purified as previously described^{25,26} from the plasma samples of the three children of Table 1 tested negative for alpha- and flavivirus RNAs^{25,26} by RT-PCR. To screen for viruses not detected by the RT-PCR, virus isolation was also attempted after inoculation of aliquots of the plasma onto Vero E6 cells^{25,26}. Nucleic acids, purified from the cell culture media 7, 14, 21, and 30 days post-inoculation (dpi) of the cells, again tested negative for alpha- and flavivirus RNAs. Moreover, they tested negative for the DNA and RNA of common human respiratory viruses using a GenMark Respiratory Panel²⁷. However, subtle cytopathic effects (CPEs) were observed in Vero E6 cell monolayers starting at about 11 dpi, suggesting that a virus had been isolated. The non-specific CPEs included granulation of the cells (Fig. 1). Since none of the tests produced evidence that could be used for a preliminary identification of a viral agent, an unbiased amplification and sequencing approach²⁸ was attempted for cells inoculated with plasma from sample 0081-4, which displayed more CPEs than cells inoculated with the

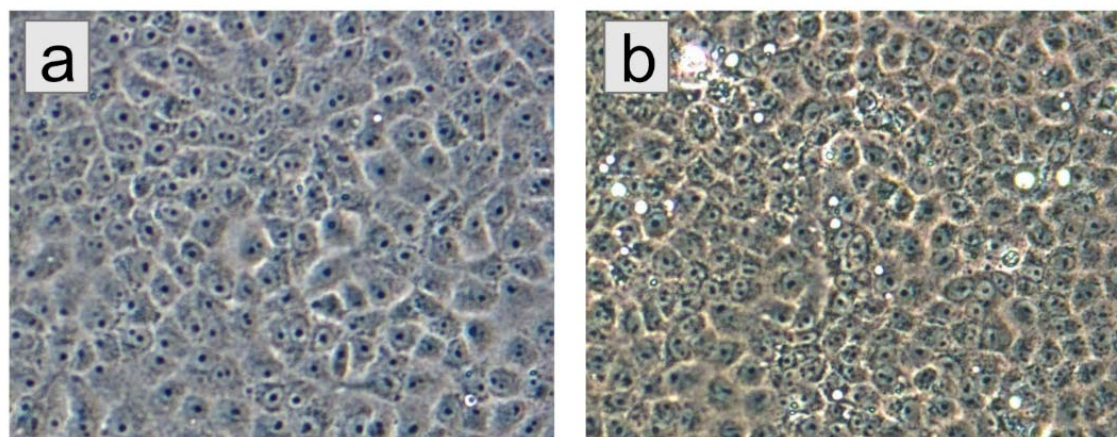


Figure 1: Non-specific CPEs formed by plasma from patient 0081-4 in Vero E6 cells. a) Mock-infected Vero E6 cells, 11 days post-inoculation with phosphate-buffered saline. b) Vero E6 cells 11 days post-inoculation with plasma from patient 0081-4. Original magnification at 200X

other two plasma samples. PCR amplification yielded seven amplicons. Sequence analyses indicated that six were African green monkey sequences from the Vero E6 cells, whereas one 401 bp amplicon had 100% identity with the corresponding genome sequence of various PDCoV strains, such as CHN-SC2015 (GenBank MK355396). Therefore, RNA purified from Vero E6 culture samples was re-tested using a pancoronavirus RT-PCR test that amplifies a 668 bp-region within the RNA-dependent RNA polymerase gene that encodes the most conserved protein domain of alpha-, beta-, gamma-, and delta CoVs²⁹, generating positive results. A follow-up test using RNA directly purified from plasma generated the same 668-bp amplicons, providing further indications that a PDCoV was present in the plasma and cell culture samples. In contrast, 20 other plasma samples and mock-inoculated Vero E6 cultures tested negative for PDCoV RNAs. At the time, we had no PDCoV strains in our laboratory. Following this preliminary identification, whole genome sequences for the three isolates were obtained by Sanger sequencing. The GenBank accession numbers corresponding to the sequenced genomes are given in Table 1.

Genomic characterization and evolution of Hu-PDCoV strains

Assessment of potential recombinants^{30,31} in a multiple sequence alignment including currently available full genome PDCoV sequences did not detect any signal for recombination in the viral strains from human patients (Supplemental Figures S1a, S1b), and displayed robust signal (Supplemental Figure S1c) for phylogeny inference. Strains from case 1 (designated as strain 0081-4) and case 3 (strain 0329-4), identified four months apart among children attending School A (Table 1), were highly similar (99.97%) and closely related (99.8%) to a pig strain detected one year later in Tianjin, China. Case 3 (strain 0246-1), who attended School B in the mountains, was infected with a variant strain similar to the other two Hu-PDCoV strains (98.9%), but most closely related to a pig strain detected in Arkansas, USA, in 2015. Despite such a high genomic similarity, the maximum likelihood (ML) tree inferred from full genome sequences clearly shows that strains from cases 1 and 3 and the strain from case 2 belong, respectively, to two distinct and well supported monophyletic clades: the first one clustering strains from Chinese pigs, and the second one clustering strains from USA pigs (Fig. 2). In other words, both epidemiology data and evolutionary relationships indicate that the

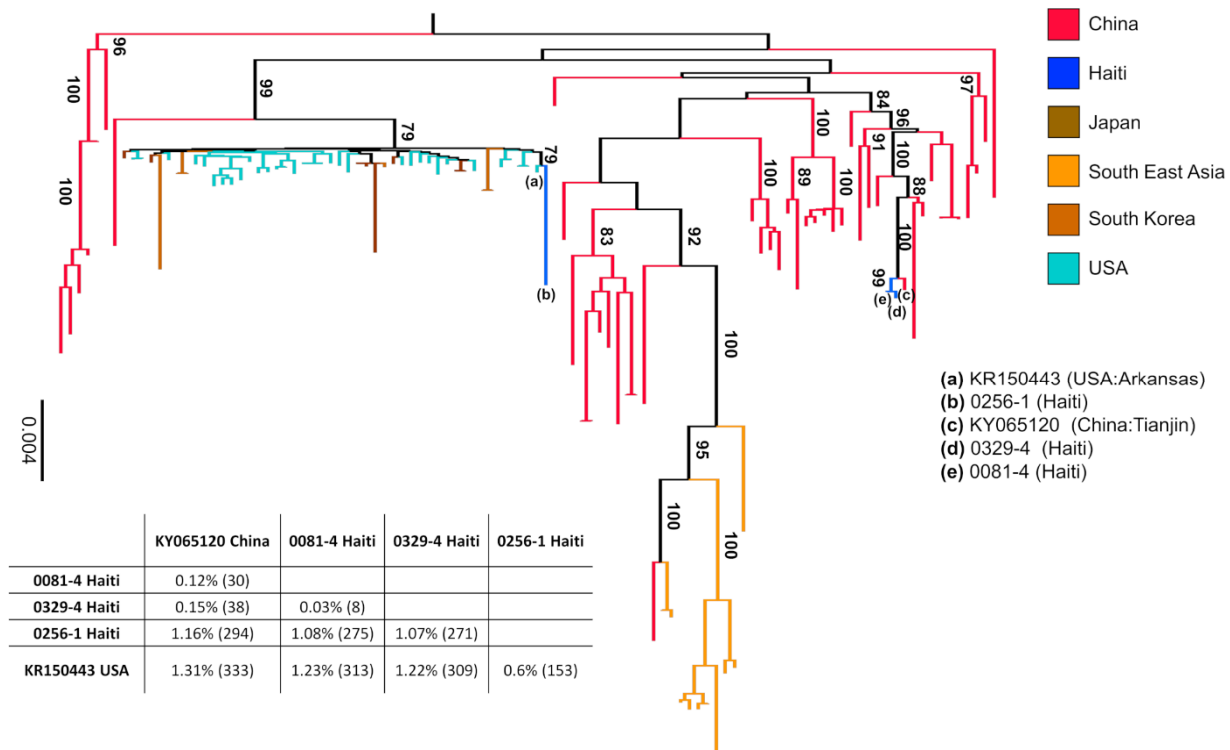


Figure 2 | Maximum likelihood tree of PDCoV strains. The tree was inferred from 109 PDCoV full genome sequences, including four sparrow CoV genomes for outgroup rooting (accession numbers MG812375, MG812376, MG812377, and MG812378, from Chen *et al.*, 2018) using the best fitting nucleotide substitution model (see Methods). For display purposes, outgroup sequences were removed from the figure. Vertical branches are

scaled in number of nucleotide substitutions per site according to the bar on the left of the tree. Bootstrap values > 75% (1000 replicates) are indicated along supported branches. The table at the bottom of the tree shows nucleotide % dissimilarity (and total number of nucleotide differences in parenthesis) between Haitian strains and their closest non-Haitian relatives, which are labelled in tree by progressive letters according to the legend on the right.

three PDCoV strains identified in the Haitian children were the result of at least two separate zoonotic transmissions that likely occurred within a similar time frame. The close evolutionary relationships of strain 0256-1 with a USA PDCoV sequence sampled in 2014-2015 is not surprising given the re-introduction of pigs from USA to Haiti after the local population was mostly wiped out following the African Swine Fever epidemic^{19,20}. On the other hand, the close phylogenetic relationship between Haitian strains 0081-4 and 0329-4 and a viral sequence isolated from a Chinese pig one year later (Table 1) is less obvious. These three sequences clearly share a most recent common ancestor (MRCA) with high bootstrap support (99%). The Haitian PDCoV strains may have been transmitted to the children through two independent zoonoses from the same animal, or different animals infected with a highly similar virus, or be the result of an initial zoonosis followed by human-to-human transmission. Since samples from pigs in the areas surrounding the two schools were not available, it is impossible at this time to

discern which scenario is the most likely. Regardless, the phylogeny demonstrates the circulation of two distinct PDCoV lineages in School A and School B, highlighting deltacoronaviruses' ability to spill over successfully in the human population.

Our next step was the calibration of a molecular clock to infer the time of the most recent ancestor (TMRCA) of Hu-PDCoV and their most closely related porcine strains. We tested for the presence of temporal signal in the sequence data set by calculating the linear regression between root-to-tip distances and sampling time in the ML tree³². After removal of sparrow outgroup sequences and the South East Asian clade outliers, the tree inferred from the remaining ($n = 94$) sequences showed sufficient signal to calibrate a molecular clock (see Supplemental Figure S2 for details). The topology of the Bayesian maximum clade credibility (MCC) tree obtained by enforcing a strict molecular clock confirmed the findings of the ML phylogeny (Fig. 3). Identical results were obtained by enforcing a relaxed clock model. According to the clock calibration, 0081-4 and 329-4 TMRCA dates back to October 2014, with 95% high posterior density intervals (95%HPD) essentially overlapping (October 2014 – January 2015) with the sampling dates (see Table 1) of the strains themselves. In turn, the Haitian strains diverged from their MRCA with the pig Chinese strain in July 2014 (95%HPD: April – August 2014). These results point to a recent introduction in the country before the virus infected the children and are compatible both with zoonotic events and/or limited human-to-human transmission. On the other hand, Haitian isolate 0256-1 TMRCA with USA isolate KR150443 dates back to 2011 (95% HPD: Feb-2011 - March-2012). Therefore, the strain had likely been circulating in pigs for a few years, as also suggested by its relatively long terminal branch in the ML tree (Fig. 2), before infecting the human patient. It is possible, of course, that we are missing several intermediate links along the 0256-1 branch, either from pig or other human strains, and the introduction in the country might have been more recent than suggested by the molecular clock analysis.

While phylogeny inference demonstrates that the Hu-PDCoV strains isolated from Haitian children belong to independent evolutionary lineages, introduced in humans through at least two separate zoonotic transmissions, a more in-depth analysis of the genomic changes occurred along the human lineages revealed an interesting pattern. Signature pattern analysis³³ shows that the three Haitian strains share a signature of five conserved amino acid residues in the ORF1a/b polyprotein and two in the Spike glycoprotein, unique among other PDCoV sequences

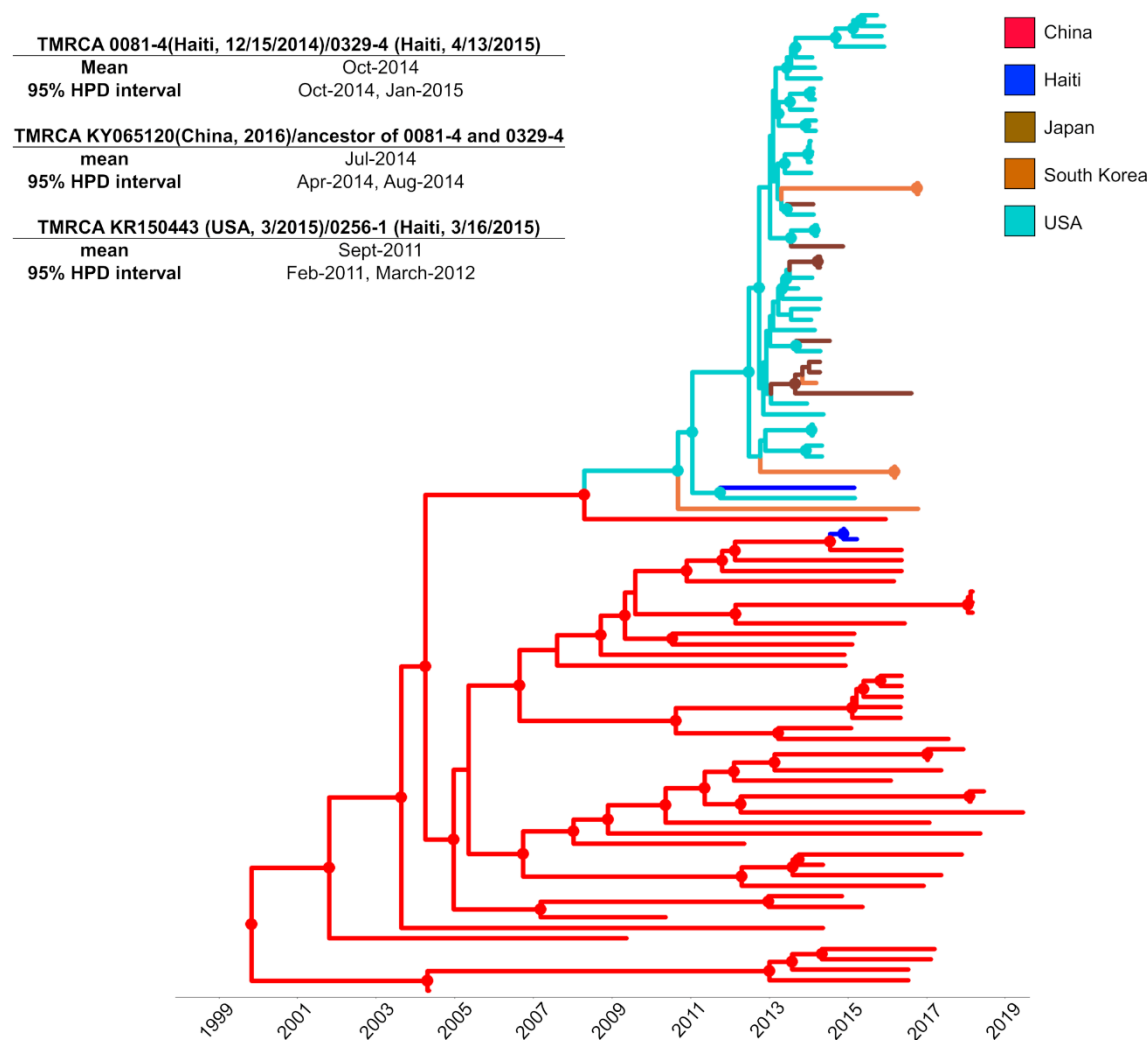


Figure 3 | Bayesian maximum clade credibility (MCC) tree of PDCoV strains. The MCC tree was inferred from a subset of 94 full genome strains that displayed sufficient temporal signal for molecular clock calibration. Branch lengths were scaled in time, according to the bar at the bottom, by enforcing a strict molecular clock and using

sampling dates to estimate PDCoV evolutionary rate. Circles at internal nodes indicate high posterior probability (PP) support >0.9. The table on the left shows the inferred time of the most recent common ancestor (TMRCAs) between Haitian strains and their closest phylogenetic relative, with 95% high posterior density intervals (95%HPD).

from pigs currently known (Fig. 4a). The sole exception is Chinese strain KY065120, which displays the same amino acid signature and is the one most closely related to Haitian strains from School A (Fig. 2 and Table 1) and may represent a porcine strain pre-adapted for effective transmission to humans. Indeed, the convergent evolution of identical amino acid changes along distinct phylogenetic lineages is highly suggestive of an adaptive response. Mutations in the first five ORF1/ab amino acids part of the Hu-PDCoV-specific signature (Fig. 4a) are located at sites that do not correspond to solved crystal structures. The other ORF1/ab mutation maps in the

nonstructural protein 15 (Nsp15): A30V (amino acid position numbered according to reference sequence JQ065043). The C-terminal domain of the Nsp15 protein possesses endoribonuclease

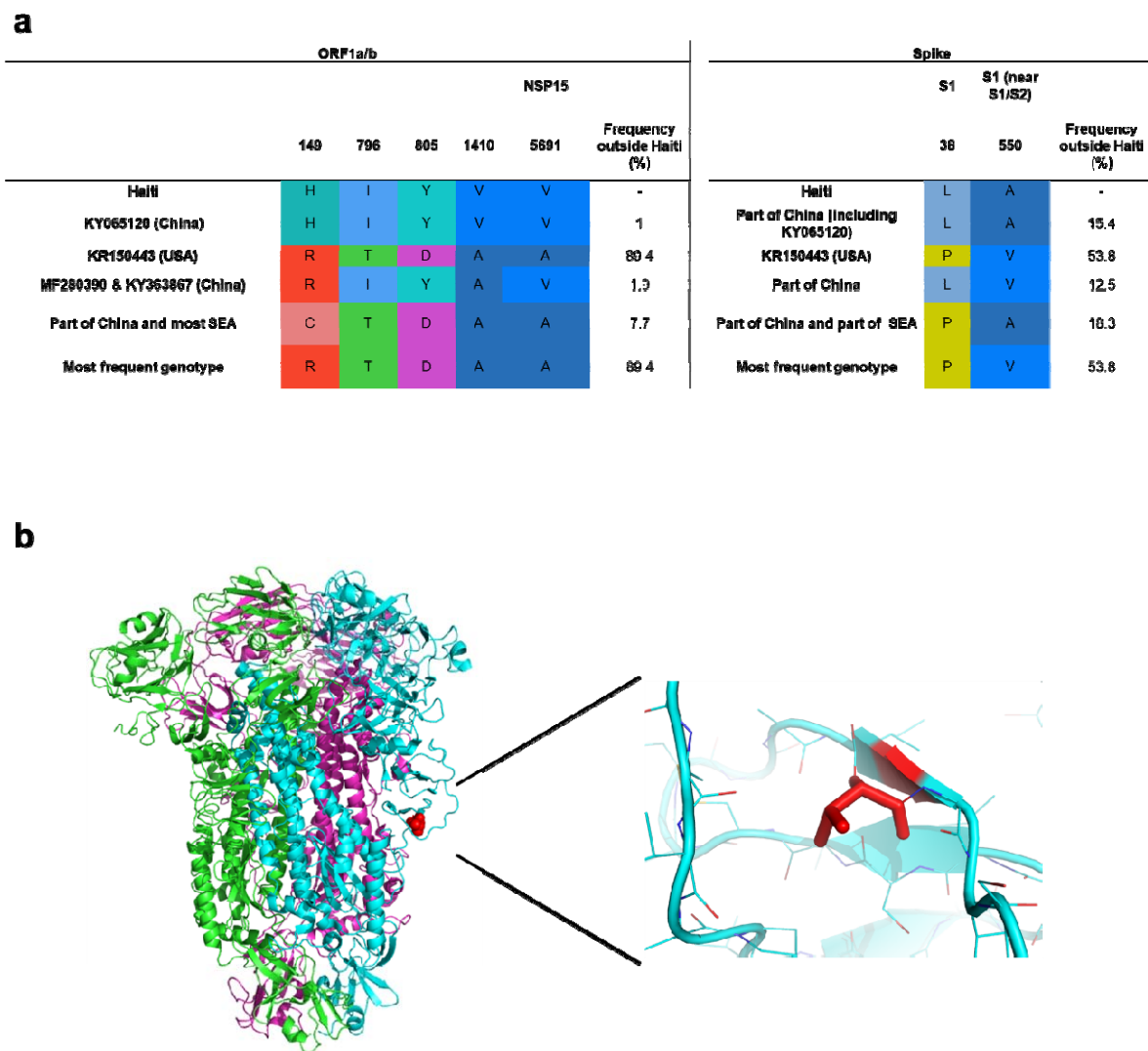


Figure 4 | Analysis of conserved amino acids in Hu-PDCoV strains. a) Amino acid signature pattern analysis: residues are numbered based on the ORF1a/b absolute amino acid position (starting from number 1) in the reference sequence JQ065043 (Pig/China). The only residues with 100% frequency in Haitian strains mapped in the ORF1a/b and Spike regions. **b)** Spike glycoprotein trimer structure.

Different colors are assigned to each monomer. Amino acid residues were numbered relative to their position in the Spike glycoprotein of the reference sequence. Red indicates residue 550, where Haitian sequences have a Valine to Alanine mutation. The other mutation (Proline to Leucine at position 38) was not resolved in the structure.

with uridylylate-specific activity³⁴. While the protein is not necessary for RNA synthesis, it is necessary in coronaviruses to escape recognition of dsRNA intermediates by the host³⁵. PDCoV Nsp15 inhibits the induction of IFN- β , the main intestinal antiviral cytokine, by preventing

interferon regulatory factors IRF1 nuclear translocation³⁶. The last two mutations in the Hu-PDCoV-specific signature map in the Spike glycoprotein. The first one, P8A, in the N terminal domain of the glycoprotein, is not resolved in the known crystallographic structure, possibly because the segment is too flexible to be seen by cryoEM. The second one, V550A, is located in the S1 subunit (between the RBD and the cleavage site between S1 and S2) on a short beta sheet forming intramolecular contact with a neighboring loop (Fig. 4b). The V550A change observed in Hu-PDCoV (removal of two methyl groups) is present at relatively low frequency in other Asian PDCoVs, neither of which displays the additional amino acid changes observed in the Haitian strains. This change, albeit minor, eliminates specific Van der Waals contact (with Proline at position 535 and the backbone carbonyl at position 532), potentially enhancing protein flexibility and dynamic movement of S1. Since mutations that prevent intermolecular spike protein interactions between S1 and S2 of SARS-CoV-2 variant B.1.1.7 have been observed (A570D, D614G and S982A)^{37,38}, V550A may represent a common mechanism that enhances dynamic movements, accelerating virus membrane fusion events and transmission.

Conclusions

To our knowledge, this is the first report of PDCoV infection in humans. The virus was identified in blood plasma samples, consistent with systemic dissemination. While the children in this study did not report diarrhea, two noted abdominal pain, which would be consistent with intestinal involvement; both also reported cough, which may reflect upper respiratory involvement. Since we could not sample from animals in the two regions where the virus was detected, we cannot definitively identify the zoonotic route of the human infections. However, we would note that pigs (often free-roaming) are not uncommon in rural and urban areas of Haiti, and our findings are consistent with a virus maintained in the swine population and capable of successful spillover in humans thanks to the independent emergence of specific amino acid changes through convergent evolution. Further *in vitro* experiments will be necessary for a functional characterization of the Haitian mutants. Yet, the presence of a unique amino acid signature involving both Nsp15 protein and the S glycoprotein, which evolved in parallel along distinct phylogenetic lineages, does suggest a role of these amino acid changes in selection driven adaptation to infection of human hosts.

In pigs, PDCoV is a recently recognized porcine coronavirus which has demonstrated an ability to spread rapidly in animal populations at a global level. Our data highlight the potential for movement of this and other coronavirus strains into human populations, potentially in rural or less-developed regions where contact with domestic animals is common. The children infected with this virus had only mild illness; at this point in time, we do not find evidence that Hu-PDCoV represents a major human health threat. It is of interest, however, that Haiti, despite high levels of infection with SARS-CoV-2³⁹, appears to have had less severe COVID-19 diseases than initially expected⁴⁰. Indeed, one hypothesized reason for a decrease in severe SARS-CoV-2 cases in the general population would be the presence of pre-existing immunity to coronavirus strains other than SARS-CoV-2, which may have entered human populations.

Supplemental Figures and Tables

Supplemental Table S1 | Accession numbers of deltacoronaviruses accessed from NCBI.

Supplemental Figure S1 | Neighbor net plot of PDCoV fragments and assessment of phylogenetic signal. Neighbor nets were inferred from pair-wise p-distances of 47 genomes for the **a)** major and **b)** minor genome fragment identified by RDP4 (see Methods). Recombination signal was assessed within each fragment by the PHI test. No significant evidence of recombination was found in either fragment (PHI test $p > 0.05$). The only change in topology, in the split decomposition networks shown in panel a and b, concerned two Chinese sequences (highlighted by the red circle) displaying mixed ancestry possibly due to homoplasmy. These sequences were also removed before performing any further phylogenetic analyses. **c)** Likelihood mapping of the final 109 sequences alignment (see Methods) showing extremely low phylogenetic noise (4.6% in the center of the triangle), as required for reliable phylogeny inference.

Supplemental Figure S2 | Analysis of the temporal signal with TempEst. Root-to-tip distance (y-axis) vs. sampling time linear regression in the ML likelihood phylogeny inferred from: **a)** All PDCoV sequences in the final data set ($n = 109$) including sparrow outgroup sequences **b)** PDCoV sequences without outgroups ($n = 105$), with sequences in the red circle belonging to the South East Asia clade (see figure 1) showing a clear departure from the strict molecular clock model, and **c)** PDCoV sequences after removal of the South East Asia clade ($n = 94$), showing greatly improved clock signal ($R^2 = 0.68$).

References

- 1 Su, S. *et al.* Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol* **24**, 490-502, doi:10.1016/j.tim.2016.03.003 (2016).
- 2 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269, doi:10.1038/s41586-020-2008-3 (2020).
- 3 Ye, Z. W. *et al.* Zoonotic origins of human coronaviruses. *Int J Biol Sci* **16**, 1686-1697, doi:10.7150/ijbs.45472 (2020).
- 4 Fehr, A. R. & Perlman, S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* **1282**, 1-23, doi:10.1007/978-1-4939-2438-7_1 (2015).
- 5 Graham, R. L. & Baric, R. S. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J Virol* **84**, 3134-3146, doi:10.1128/JVI.01394-09 (2010).

- 6 Li, W. *et al.* Broad receptor engagement of an emerging global coronavirus may potentiate its diverse cross-species transmissibility. *Proc Natl Acad Sci U S A* **115**, E5135-E5143, doi:10.1073/pnas.1802879115 (2018).
- 7 He, W. T. *et al.* Genomic Epidemiology, Evolution, and Transmission Dynamics of Porcine Deltacoronavirus. *Mol Biol Evol* **37**, 2641-2654, doi:10.1093/molbev/msaa117 (2020).
- 8 Tagliamonte, M. S. *et al.* Multiple Recombination Events and Strong Purifying Selection at the Origin of SARS-CoV-2 Spike Glycoprotein Increased Correlated Dynamic Movements. *Int J Mol Sci* **22**, doi:10.3390/ijms22010080 (2020).
- 9 Korotkova, E. *et al.* Pressure for Pattern-Specific Intertypic Recombination between Sabin Polioviruses: Evolutionary Implications. *Viruses* **9**, doi:10.3390/v9110353 (2017).
- 10 Xiao, Y. *et al.* RNA Recombination Enhances Adaptability and Is Required for Virus Spread and Virulence. *Cell Host Microbe* **19**, 493-503, doi:10.1016/j.chom.2016.03.009 (2016).
- 11 Woo, P. C. *et al.* 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**, 3995-4008, doi:10.1128/JVI.06540-11 (2012).
- 12 Wang, L., Byrum, B. & Zhang, Y. Detection and genetic characterization of deltacoronavirus in pigs, Ohio, USA, 2014. *Emerg Infect Dis* **20**, 1227-1230, doi:10.3201/eid2007.140296 (2014).
- 13 Jung, K., Hu, H. & Saif, L. J. Porcine deltacoronavirus infection: Etiology, cell culture for virus isolation and propagation, molecular epidemiology and pathogenesis. *Virus Res* **226**, 50-59, doi:10.1016/j.virusres.2016.04.009 (2016).
- 14 Ma, Y. *et al.* Origin, evolution, and virulence of porcine deltacoronaviruses in the United States. *mBio* **6**, e00064, doi:10.1128/mBio.00064-15 (2015).
- 15 Wang, Q., Vlasova, A. N., Kenney, S. P. & Saif, L. J. Emerging and re-emerging coronaviruses in pigs. *Curr Opin Virol* **34**, 39-49, doi:10.1016/j.coviro.2018.12.001 (2019).
- 16 Wang, B. *et al.* Porcine Deltacoronavirus Engages the Transmissible Gastroenteritis Virus Functional Receptor Porcine Aminopeptidase N for Infectious Cellular Entry. *J Virol* **92**, doi:10.1128/JVI.00318-18 (2018).
- 17 Boley, P. A. *et al.* Porcine Deltacoronavirus Infection and Transmission in Poultry, United States(1). *Emerg Infect Dis* **26**, 255-265, doi:10.3201/eid2602.190346 (2020).
- 18 CIA. *The world factbook - Haiti*, <<https://www.cia.gov/the-world-factbook/countries/haiti/>> (
- 19 Alexander, F. C. Experiences with African swine fever in Haiti. *Ann N Y Acad Sci* **653**, 251-256, doi:10.1111/j.1749-6632.1992.tb19654.x (1992).
- 20 Gaertner, P. *WHETHER PIGS HAVE WINGS: African Swine Fever Eradication and Pig Repopulation in Haiti.*, <<http://faculty.webster.edu/corbetre/haiti/misctopic/pigs/gaertner.htm>> (
- 21 Beau De Rochars, V. E. M. *et al.* Spectrum of outpatient illness in a school-based cohort in Haiti, with a focus on diarrheal pathogens. *Am J Trop Med Hyg* **92**, 752-757, doi:10.4269/ajtmh.14-0059 (2015).

- 22 Beau De Rochars, V. M. *et al.* Isolation of Coronavirus NL63 from Blood from Children in Rural Haiti: Phylogenetic Similarities with Recent Isolates from Malaysia. *Am J Trop Med Hyg* **96**, 144-147, doi:10.4269/ajtmh.16-0585 (2017).
- 23 Bonny, T. S. *et al.* Complete Genome Sequence of Human Coronavirus Strain 229E Isolated from Plasma Collected from a Haitian Child in 2016. *Genome Announc* **5**, doi:10.1128/genomeA.01313-17 (2017).
- 24 Kim, Y. Y. *et al.* Acute Respiratory Illness in Rural Haiti. *Int J Infect Dis* **81**, 176-183, doi:10.1016/j.ijid.2019.02.003 (2019).
- 25 Lednicky, J. *et al.* Zika Virus Outbreak in Haiti in 2014: Molecular and Clinical Data. *PLoS Negl Trop Dis* **10**, e0004687, doi:10.1371/journal.pntd.0004687 (2016).
- 26 Lednicky, J. *et al.* Mayaro Virus in Child with Acute Febrile Illness, Haiti, 2015. *Emerg Infect Dis* **22**, 2000-2002, doi:10.3201/eid2211.161015 (2016).
- 27 Pan, M. *et al.* Collection of Viable Aerosolized Influenza Virus and Other Respiratory Viruses in a Student Health Care Center through Water-Based Condensation Growth. *mSphere* **2**, doi:10.1128/mSphere.00251-17 (2017).
- 28 Lednicky, J. A. *et al.* Keystone Virus Isolated From a Florida Teenager With Rash and Subjective Fever: Another Endemic Arbovirus in the Southeastern United States? *Clin Infect Dis* **68**, 143-145, doi:10.1093/cid/ciy485 (2019).
- 29 Hu, H., Jung, K., Wang, Q., Saif, L. J. & Vlasova, A. N. Development of a one-step RT-PCR assay for detection of pancoronaviruses (alpha-, beta-, gamma-, and delta-coronaviruses) using newly designed degenerate primers for porcine and avian fecal samples. *J Virol Methods* **256**, 116-122, doi:10.1016/j.jviromet.2018.02.021 (2018).
- 30 Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**, 254-267, doi:10.1093/molbev/msj030 (2006).
- 31 Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* **1**, vev003, doi:10.1093/ve/vev003 (2015).
- 32 Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* **2**, vew007, doi:10.1093/ve/vew007 (2016).
- 33 Korber, B. & Myers, G. Signature pattern analysis: a method for assessing viral sequence relatedness. *AIDS research and human retroviruses* **8**, doi:10.1089/aid.1992.8.1549 (1992).
- 34 Ivanov, K. A. *et al.* Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc Natl Acad Sci U S A* **101**, 12694-12699, doi:10.1073/pnas.0403127101 (2004).
- 35 Deng, X. *et al.* Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc Natl Acad Sci U S A* **114**, E4251-E4260, doi:10.1073/pnas.1618310114 (2017).
- 36 Liu, S. *et al.* Porcine deltacoronavirus (PDCoV) infection antagonizes interferon-lambda1 production. *Vet Microbiol* **247**, 108785, doi:10.1016/j.vetmic.2020.108785 (2020).
- 37 Public Health England. Investigation of novel SARS-CoV-2 variant: variant of concern 202012/01, technical briefing 3. London, United Kingdom: Public Health England (2020).
- 38 Rambaut, A., Loman, N., Pybus, O., Barclay, W., Barrett, J., Carabelli, A., Connor, T., Peacock, T., Robertson, D.L., Volz, E. . Preliminary genomic characterisation of an

- emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *COVID-19 Genomics Consortium UK (CoG-UK)* (2020).
- 39 Rouzier, V., Liautaud, B. & Deschamps, M. M. Facing the Monster in Haiti. *N Engl J Med* **383**, e4, doi:10.1056/NEJMc2021362 (2020).
- 40 Bazell, R. Looking for the light in Haiti. *Science* **369**, 755-756, doi:10.1126/science.369.6505.755 (2020).

Acknowledgments: Work was supported in part by a grant (R01-AI123657S1) from NIAID to JGM.

Author contributions: JAL supervised virology work, identified the viruses subsequent to performing Sanger sequencing and assembly of the sequences, and contributed to manuscript writing; MST performed genomic and evolutionary analyses and wrote the manuscript; SKW performed virology and molecular work, including GenMark analyses; MAE performed virology and molecular work and coordinated sample shipments and communications between stakeholders in Haiti and the University of Florida; MMA, CJS, TSB, and JCL performed virology and molecular work; TT and SC collected plasma samples, performed medical exams, and documented the findings thereof; CM contributed to bioinformatic analyses and edited the manuscript; VMBDR provided medical oversight over the project and established the internal and external processes including obtaining the necessary permissions for this project; DAO carried out the structural analysis; JAL, MS and GM designed the study and contributed to manuscript writing.

Data availability statement: GenBank accession numbers for sequence data are included in Table 1. Supplementary Table S1 includes a listing of accession numbers of deltacoronaviruses accessed from NCBI for the phylogenetic studies.

Methods

Clinical Sample Collection

From 2012 to 2020, our research group monitored a cohort of approximately 1250 school children in the Gressier region of Haiti²¹. Children attended one of four schools operated by the Christianville Foundation and had free access to medical care through a school-based clinic. Children presenting to the clinic with an acute undifferentiated febrile illness, defined as a history of fever and/or a measured temperature > over 37.5°C in the clinic with no localizing symptoms or signs (i.e., no respiratory, skin, or urinary symptoms or signs) were invited to enroll in an ongoing study of acute febrile illness; written informed consent for study participation was obtained from parents, with assent from children. After enrollment, clinic health care providers recorded clinical data in a study questionnaire and a sample of venous blood (1-3 ml) was collected in an acid-citrate-dextrose blood collection tube. The blood samples were subsequently centrifuged to pellet the platelets, red blood cells, and white blood cells, and the resulting plasma was aseptically transferred to cryovials and stored at -80°C for subsequent analysis. Appropriate medical care based on clinical presentation and laboratory studies was provided to study participants by clinic health care providers. Data on identification of arboviruses and other virus species among children participating in the study have been previously reported^{22,23,25,26,41-44}. Studies were approved by the Institutional Review Board (IRB) at the University of Florida and the Haitian National IRB. As the study was done in young children the amount of plasma collected was limited, and samples have, in most instances, been exhausted, due to the range of studies initially conducted on the samples while screening for other pathogens. IRB restrictions limit our ability to share samples outside of our institution.

Virus Identification and Sequencing

Attempts at next-generation sequencing using an Illumina MiSeq platform generated minimal coverage, so we sorted to Sanger Sequencing using the primer system outlined by Liang et al⁴⁵, with one addition: to obtain the 5' ends of the viral genomes, a Rapid Amplification of cDNA Ends (RACE) kit was used per the manufacturer's protocols (Life Technologies, Carlsbad, CA, USA), and the resulting amplicons TA-cloned into plasmids and sequenced. PCR amplicons for Sanger Sequencing were amplified using AccuScript High-Fidelity reverse transcriptase (Agilent Technologies, Inc., Santa Clara, CA) in the presence of SUPERase-In RNase inhibitor (Ambion,

Austin, TX), followed by PCR with Q5 DNA polymerase (New England Biolabs. They were next purified using a QIAquick PCR purification kit (Qiagen Inc., Germantown, MD) before TA-cloning. The inserts in the plasmids were subsequently sequenced bidirectionally using a gene-walking approach, based on obtaining at least 800 bp or non-ambiguous sequence. Briefly, pairs of non-overlapping primers and Q5 polymerase were used to produce 42 separate amplicons corresponding to the PDCoV genome, and each amplicon was Sanger sequenced bidirectionally.

Sequence data assembly

The identity of the whole genome sequences was confirmed *via* online BLAST⁴⁶ of the nr/nt NCBI database. Following positive identification, available PDCoV sequences from pigs were downloaded from NCBI (www.ncbi.nlm.nih.gov), together with closely related sparrow deltacoronavirus sequences⁴⁷ to be used as outgroups in the phylogenetic analysis (see below). The final full genomes data set assembled included 104 PDCoV genomes from pigs, 4 from sparrows (Supplemental Table S1), and 3 newly sequenced Hu-PDCoV strains, which were aligned with MAFFT⁴⁸ v.7.407. Potential recombinant origin of Haitian sequences was assessed using an algorithm based on the PHI test^{49,50} implemented SplitsTree4³⁰ v.4.14.8 and with the RDP4³¹ software package v.4.97. RDP4 detected several possible recombination events in the full data set (although homoplasmy could not be ruled out), resulting in at least two major breakpoints at nucleotide position 14,874 – 17,234 (numbered according to reference sequence JQ065043) that involved 60 strains from Chinese pigs belonging to sequence clusters unrelated to the new human isolates. After such strains were removed from the alignment, together with the outgroup sequences, the PHI test on the remaining subset of 47 sequences (44 from pigs + 3 new human strains) still detected evidence of recombination ($p < 0.001$). Split decomposition analysis using SplitsTree5 of the major (1 – 14873/17235 – 25,413) and minor recombinant fragment (14,874 – 17,234) identified two additional Chinese sequences (KY513724 and MN942260) with mixed ancestry, possibly due to homoplasmy, that were removed from the full data set. However, since the other Chinese pig strains of possible recombinant origin clustered in clades clearly distinct from the ones including the human isolates, and our goal was to assess the most recent common ancestor between human and most closely related pig strains, a final alignment (available from the authors upon request) including 109 sequences (4 outgroup sequences from sparrow, 102 from pigs and 3 from human patients) was chosen to carry out subsequent phylogenetic analyses.

Phylogenetic and amino acid signature analysis

Phylogenetic signal was verified using likelihood mapping⁵¹, as implemented in IQTREE v.2.0.6⁵². Maximum likelihood tree was calculated using the same version of IQTREE, with the best fitting nucleotide substitution model according to the Bayesian Information Criterion (BIC) and 1,000 bootstrap replicates. The correlation between root-to-tip genetic divergence and sampling dates to assess clock signal of the alignment were performed with TempEst³² prior to Bayesian phylodynamic analyses. The time-scaled tree was calculated using the Bayesian phylodynamic inference framework in BEAST v. 1.10.4⁵³;

Markov chain Monte Carlo (MCMC) samplers were run for 200 million generations, with sampling every 20,000 generations, to ensure proper mixing, which was assessed by calculating the effective sampling size (ESS) of each parameter estimate. The HKY nucleotide substitution model was used with empirical base frequencies and gamma distributions of site-specific rate heterogeneity⁵⁴. The molecular clock was calibrated by enforcing a strict clock and choosing either a constant size or a Bayesian Skyline Plot demographic prior⁵⁵. A maximum clade credibility tree was inferred from the posterior distribution of trees using TreeAnnotator, specifying a burn-in of 104 million and median node heights. The MCC was edited graphically using *ggtree*⁵⁶⁻⁵⁸. MCMC runs with different demographic priors gave the same result.

Signature pattern analysis of the Haitian strains as compared to the rest of the downloaded sequences was performed with online version of VESPA³³; the PDCoV reference sequence JQ065043.2 was used as a guide for the codon coordinates. The spike glycoprotein three-dimensional structure PDB 6b7n⁵⁹ was used as a base, and figures were generated using PyMol⁶⁰.

Methods References

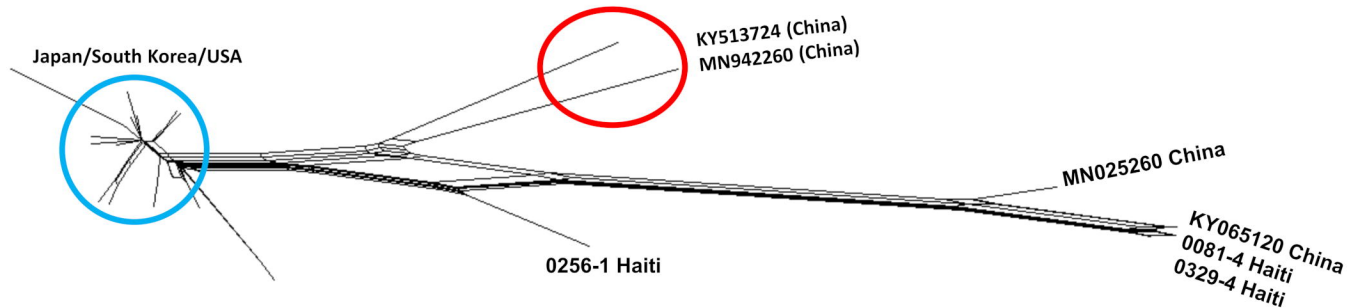
- 21 Beau De Rochars, V. E. M. *et al.* Spectrum of outpatient illness in a school-based cohort in Haiti, with a focus on diarrheal pathogens. *Am J Trop Med Hyg* **92**, 752-757, doi:10.4269/ajtmh.14-0059 (2015).
- 22 Beau De Rochars, V. M. *et al.* Isolation of Coronavirus NL63 from Blood from Children in Rural Haiti: Phylogenetic Similarities with Recent Isolates from Malaysia. *Am J Trop Med Hyg* **96**, 144-147, doi:10.4269/ajtmh.16-0585 (2017).

- 23 Bonny, T. S. *et al.* Complete Genome Sequence of Human Coronavirus Strain 229E Isolated from Plasma Collected from a Haitian Child in 2016. *Genome Announc* **5**, doi:10.1128/genomeA.01313-17 (2017).
- 25 Lednicky, J. *et al.* Zika Virus Outbreak in Haiti in 2014: Molecular and Clinical Data. *PLoS Negl Trop Dis* **10**, e0004687, doi:10.1371/journal.pntd.0004687 (2016).
- 26 Lednicky, J. *et al.* Mayaro Virus in Child with Acute Febrile Illness, Haiti, 2015. *Emerg Infect Dis* **22**, 2000-2002, doi:10.3201/eid2211.161015 (2016).
- 30 Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**, 254-267, doi:10.1093/molbev/msj030 (2006).
- 31 Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* **1**, vev003, doi:10.1093/ve/vev003 (2015).
- 32 Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* **2**, vew007, doi:10.1093/ve/vew007 (2016).
- 33 Korber, B. & Myers, G. Signature pattern analysis: a method for assessing viral sequence relatedness. *AIDS research and human retroviruses* **8**, doi:10.1089/aid.1992.8.1549 (1992).
- 41 Ball, J. D. *et al.* Clinical and Epidemiologic Patterns of Chikungunya Virus Infection and Coincident Arboviral Disease in a School Cohort in Haiti, 2014-2015. *Clin Infect Dis* **68**, 919-926, doi:10.1093/cid/ciy582 (2019).
- 42 Lednicky, J. A. *et al.* Emergence of Madariaga virus as a cause of acute febrile illness in children, Haiti, 2015-2016. *PLoS Negl Trop Dis* **13**, e0006972, doi:10.1371/journal.pntd.0006972 (2019).
- 43 Blohm, G. *et al.* Mayaro as a Caribbean traveler: Evidence for multiple introductions and transmission of the virus into Haiti. *Int J Infect Dis* **87**, 151-153, doi:10.1016/j.ijid.2019.07.031 (2019).
- 44 ElBadry, M. *et al.* Isolation of an Enterovirus D68 from Blood from a Child with Pneumonia in Rural Haiti: Close Phylogenetic Linkage with New York Strain. *Pediatr Infect Dis J* **35**, 1048-1050, doi:10.1097/INF.0000000000001283 (2016).
- 45 Liang, Q., Li, B., Zhang, H. & Hu, H. Complete Genome Sequences of Two Porcine Deltacoronavirus Strains from Henan Province, China. *Microbiol Resour Announc* **8**, doi:10.1128/MRA.01517-18 (2019).
- 46 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410, doi:10.1016/s0022-2836(05)80360-2 (1990).
- 47 Chen, Q. *et al.* The emergence of novel sparrow deltacoronaviruses in the United States more closely related to porcine deltacoronaviruses than sparrow deltacoronavirus HKU17. *Emerging microbes & infections* **7**, doi:10.1038/s41426-018-0108-z (2018).
- 48 Nakamura, T., Yamada, K., Tomii, K. & Katoh, K. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics (Oxford, England)* **34**, doi:10.1093/bioinformatics/bty121 (2018).
- 49 Bruen, T. C., Philippe, H. & Bryant, D. A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**, 2665-2681, doi:10.1534/genetics.105.048975 (2006).

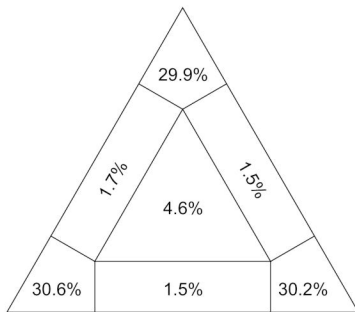
- 50 Salemi, M., Gray, R. R. & Goodenow, M. M. An exploratory algorithm to identify intra-host recombinant viral sequences. *Mol Phylogenet Evol* **49**, 618-628, doi:10.1016/j.ympev.2008.08.017 (2008).
- 51 Strimmer, K. & von Haeseler, A. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc Natl Acad Sci U S A* **94**, 6815-6819 (1997).
- 52 Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).
- 53 Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology* **7**, 214, doi:10.1186/1471-2148-7-214 (2007).
- 54 Hasegawa, M., Kishino, H. & Yano, T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* **22**, 160-174, doi:10.1007/BF02101694 (1985).
- 55 Hall, M. D., Woolhouse, M. E. & Rambaut, A. The effects of sampling strategy on the quality of reconstruction of viral population dynamics using Bayesian skyline family coalescent methods: A simulation study. *Virus Evol* **2**, vew003, doi:10.1093/ve/vew003 (2016).
- 56 Yu, G. Using ggtree to Visualize Data on Tree-Like Structures. *Curr Protoc Bioinformatics* **69**, e96, doi:10.1002/cpbi.96 (2020).
- 57 Yu, G., Lam, T. T., Zhu, H. & Guan, Y. Two Methods for Mapping and Visualizing Associated Data on Phylogeny Using Ggtree. *Mol Biol Evol* **35**, 3041-3043, doi:10.1093/molbev/msy194 (2018).
- 58 Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**, 28-36, doi:https://doi.org/10.1111/2041-210X.12628 (2017).
- 59 Shang, J. *et al.* Cryo-Electron Microscopy Structure of Porcine Deltacoronavirus Spike Protein in the Prefusion State. *Journal of virology* **92**, doi:10.1128/JVI.01556-17 (2018).
- 60 Schrodinger, L. *The PyMOL Molecular Graphics System, Version 2.0* (2015).

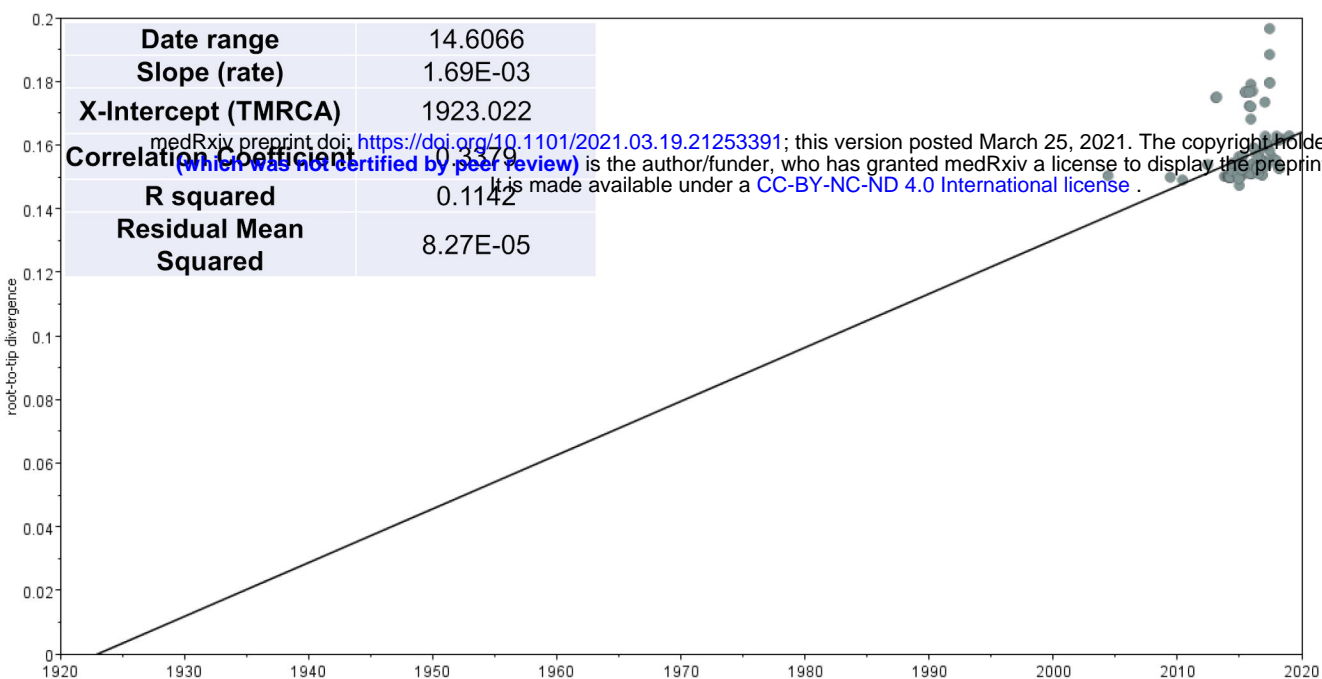
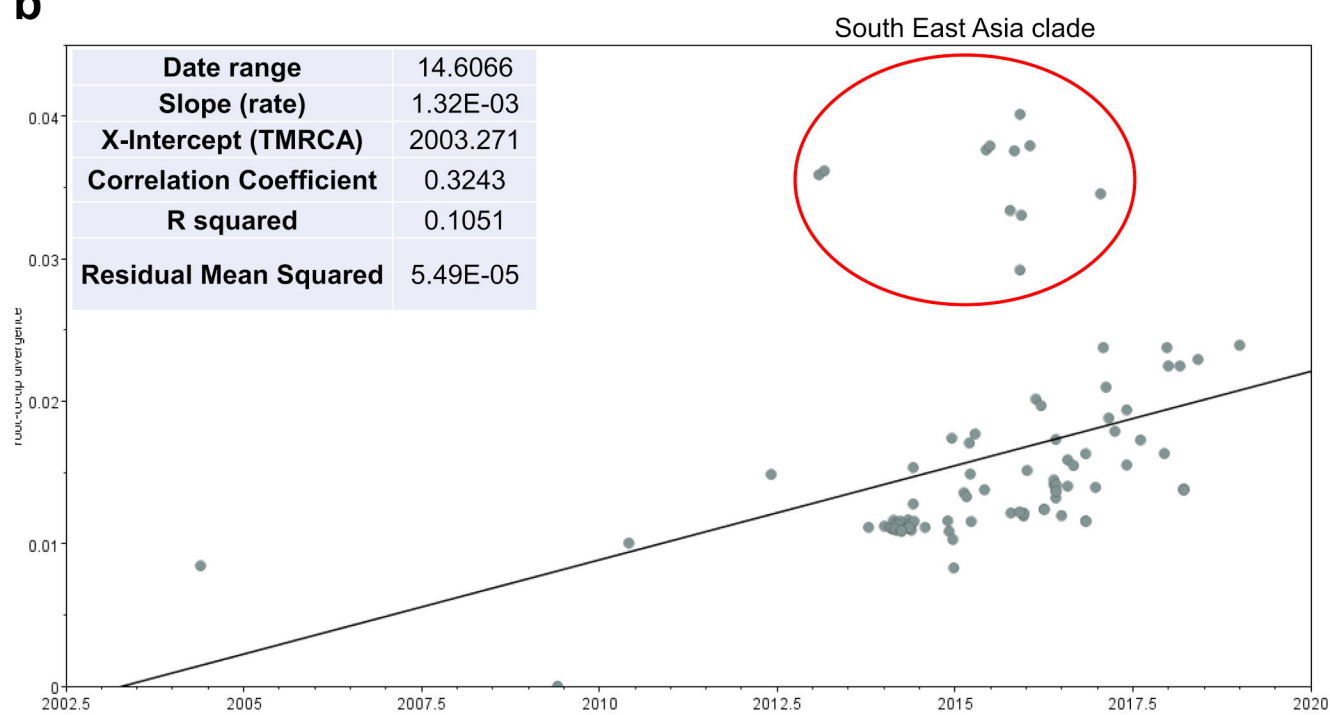
a

Phi test for recombination: $p = 0.13$

b

Phi test for recombination: $p = 0.27$

c

a**b****c**