Potential neuroinvasive and neurotrophic properties of SARS-CoV-2 in pediatric patients: Comparison of SARS-CoV-2 with non-segmented RNA viruses

Xiaodi Chen (xiaodi_chen@brown.edu)

Department of Pediatrics, Women & Infants Hospital of Rhode Island, The Alpert Medical School of Brown University https://orcid.org/0000-0002-1860-2814

Short Report

Keywords: severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), neuroinvasive and neurotrophic non-segmented RNA viruses, multiple sequence alignment, genome, phylogenetic tree, conserved domain, spike protein, pediatric patients

Posted Date: June 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-34610/v1

Version of Record: A version of this preprint was published at Journal of NeuroVirology on October 14th, 2020. See the published version at https://doi.org/10.1007/s13365-020-00913-5.

Abstract

The emerging severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is causing global health and economic crises. Infants and children can be infected, but are less likely to develop severe abnormalities including neurological disorders compared with adults. However, whether SARS-CoV-2 can directly cause neurological impairments in pediatric patients is not known. The possible evolutionary and molecular relationship between SARS-CoV-2 and non-segmented RNA viruses were examined with reference to neurological disorders in pediatric patients. SARS-CoV-2 appears to share similar functional domains and protein sequences with neuroinvasive and neurotropic RNA viruses and thus has the potential to cause neurological diseases in pediatric patients.

Introduction

Viral infections of the respiratory tract are one of the leading causes of morbidity and mortality in humans worldwide, especially in infants and children (Jartti *et al*, 2012). Respiratory viruses are able to enter brain (neuroinvasion) and infect nerve cells (neurotropism) in addition to causing respiratory diseases (Desforges *et al*, 2014). Coronaviruses (CoVs) are one of the most important respiratory viruses and contain a non-segmented RNA sequence. CoVs are phylogenetically classified into four genera: α-, β-, γ-, and δ-CoVs. Some members of CoVs such as severe acute respiratory syndrome coronavirus (SARS-CoV) and human CoV strains 229E (HCoV-229E), OC43 (HCoV-OC43), and HKU1 (HCoV-HKU1) can predispose to neurological injury (Bergmann *et al*, 2006). HCoV-OC43 and HCoV-HKU1 infections have especially been shown result in nervous system injury in pediatric patients (Principi *et al*, 2010).

Much attention has been given recently to the novel SARS-CoV-2 and its related coronavirus disease 2019 (COVID-19). Neurological syndromes including abnormalities in smell and taste, stroke, and acute necrotizing hemorrhagic encephalopathy have been observed in SARS-CoV-2 infected patients (Beyrouti *et al*, 2020; Poyiadji *et al*, 2020; Xydakis *et al*, 2020). Although these clinical findings suggest that SARS-CoV-2 could have neuroinvasive and neurotropic potential, whether SARS-CoV-2 plays a direct causative role remains to be determined. Moreover, SARS-CoV-2-related neurological abnormalities mostly occur in severe cases, in which virus-induced immune system hyperactivity, the "cytokine storm," contributes to disease severity (Wu *et al*, 2020). SARS-CoV-2 infection is less likely to be symptomatic or result in severe disease in infants and children compared with adult patients (Zimmermann and Curtis, 2020). Furthermore, whether SARS-CoV-2 can be detected in nervous system and directly predispose to neurological abnormalities in infants and children have not been reported up to April 2020. Furthermore, very limited information is available to describe the possible evolutionary and molecular relationships of SARS-CoV-2 with other neuroinvasive and neurotropic RNA viruses that have the potential to result in infection in pediatric patients.

The purpose of the present study was to compare phylogenetically the whole-genome sequences of SARS-CoV-2 and non-segmented RNA viruses including CoVs and other viruses that have the potential to infect the nervous system of infants and children with use of bioinformatics methodology. The conserved

domains (CDs) of SARS-CoV-2 and multiple sequence alignment (msa) methods were used to compare selected CDs that exist in both CoVs and members of other RNA viral families. Finally, the surface spike (S) glycoprotein and its protease cleavage sites were aligned among the CoVs to investigate their potential contribution to neurovirulence. The S protein of SARS-CoV-2 consists of two functional domains: the S1 protein that binds to host cell receptor angiotensin-converting enzyme 2 (ACE2), and the S2 protein that mediates viral and membrane fusion (Lan *et al*, 2020). The virus requires S protein priming by a cellular protease, transmembrane Serine protease 2 (TMPRSS2) for entry and membrane fusion after binding to the ACE2. TMPRSS2 cleaves the S protein at the S1/S2 boundary or within S2 subunit (Hoffmann *et al*, 2020; Lan *et al*, 2020).

Methods

The whole-genome sequences of 35 non-segmented RNA viruses including 13 CoVs (Table) were retrieved from National Center for Biotechnology Information (NCBI, Bethesda, MD, USA) for the purpose of phylogenetic analysis, which was conducted with MEGAX (Penn State University, PA, USA). All genomic sequences were aligned with the ClustalW algorithm and phylogenetic prediction inferred by the maximum likelihood method and Tamura-Nei model (Kumar *et al*, 2018; Tamura and Nei, 1993). The trees were visualized by Dendroscope-3 (University of Tübingen, Baden-Württemberg, Germany) (Huson and Scornavacca, 2012). The searches of each genomic open reading frame (ORF) and the CDs encoded by ORFs were performed through the ORF Finder and Conserved Domain Database webservers (NCBI), respectively (Marchler-Bauer *et al*, 2015). RStudio (RStudio, Inc., Boston, MA, USA) with msa package was used (Bodenhofer *et al*, 2015) for multiple protein sequence alignment. The results were visualized by RStudio and LaTeX with TEXshade package (Beitz, 2000).

Results

SARS-CoV-2 is genetically distant from other non-segmented RNA viruses.

Phylogenetic analyses were performed to understand the evolutionary relationship between SARS-CoV-2 and several neuroinvasive and neurotropic CoVs, which include CoVs known to infect humans such as SARS-CoV, MERS-CoV, HCoV-HKU1, HCoV-OC43, and HCoV-229E. CoVs that do not infect humans such as murine hepatitis virus strain JHM (MHV-JHM), porcine hemagglutinating encephalomyelitis virus (PHEV) and feline coronavirus (FCoV) (Bergmann *et al*, 2006) were also analyzed for the purpose of comparison. In addition, CoVs known to infect human hosts that do not exhibit evidence of detection in the nervous system such as HCoV-NL63 (Desforges *et al*, 2014) and CoVs that do not infect humans or the nervous system such as transmissible gastroenteritis virus (TGEV), avian infectious bronchitis virus, and sparrow delta-coronavirus were also analyzed to serve as reference viruses. The Figure 1A contains a rectangular phylogram indicating four of the CoV genera. The genus β-CoVs contain HCoV-HUK1, MHV-JHM, HCoV-OC43, PHEV, SARS-CoV, and MERS-CoV, which affect the nervous system. The SARS-CoV-2 is the closest relative to SARS-CoV because the branch lengths of the phylogram are proportional to the amount of inferred evolutionary change. Both of these viruses represent parental sequences to the HCoV-

HUK1, MHV-JHM, HCoV-OC43, and PHEV viruses, and descendant sequences of the MERS-CoV virus. Figure 1B represents a circular phylogram demonstrating the evolutionary relationship between SARS-CoV-2 and 23 non-segmented neuroinvasive and neurotropic RNA viruses known to infect humans (Messacar *et al*, 2018). Nineteen of these RNA viruses (red-labelled) are also known to infect infants and children (Messacar *et al*, 2018). This phylogram demonstrates that CoVs alone are one of the three major clades. CoVs are evolutionarily distant from, and do not evolve from a common ancestor of other members of viral families.

SARS-CoV-2 conserved domains are found in other neuroinvasive and neurotropic RNA viruses.

Figure 2A is a schematic diagram showing some important SARS-CoV-2s ORFs and their encoded CDs. Similar to other CoVs, SARS-CoV-2 contains at least six ORFs in its genome with their encoded nonstructural (nsp), structural and accessory proteins. Four main structural proteins, spike, membrane (or matrix), envelope, and nucleocapsid proteins, are encoded by ORFs 26, 5, 75, and 31, respectively, near the 3'-terminus. SARS-CoV-2 has two functional CDs, microdomain and viroporin, which can be found in the other neuroinvasive and neurotropic RNA viruses, and also contains the common RNA viral domains such as RNA-directed RNA polymerase (RdRp). The microdomain is encoded by ORF9 and can be found in SARS-CoV, HCoV-HKU1, rubella, and eastern equine encephalitis virus. The viroporin is encoded by ORF4 and can be found in SARS-CoV, coxsackievirus, echovirus, and poliovirus. Figure 2B and 2C show the msa results of macrodomain and viroporin, respectively. In macrodomain, SARS-CoV2 share 83.1% identity and 90.0% similarity with SARS-CoV, 35.6% identity and 48.5% similarity with HCoV-HKU1, 26.7% identity and 34.6% similarity with rubella, and 29.3% identity and 41.3% similarity with eastern equine encephalitis virus. In viroporin, SARS-CoV2 share 82.2% identity and 90.3% similarity with SARS-CoV, 19.0% identity and 32.0% similarity with coxsackievirus, 20.0% identity and 31.0% similarity with echovirus, and 19.3% identity and 33.6% similarity with poliovirus.

Diversity of spike protein may determine the severity of SARS-CoV-2 infection.

As the S protein is a major neurovirulent factor of several CoVs, there was a key focus on aligning the S protein among different CoVs (Miura *et al*, 2008; Phillips *et al*, 2002). Fingerprint analysis was implemented for the receptor binding domain (RBD) of S1 (Figure 3A) and S2 proteins (Figure 3B) in order to gain an overview of the sequence identity and similarity among the different CoVs. The complete sequence is depicted in one single line, and the amino acid residues are presented as colored vertical lines. Higher similarities correspond to darker vertical lines. As shown in Figures 3A and 3B, the S1 RBD is less conserved than that of S2. In the S1 RBD, SARS-CoV-2 shares 73.4% identity and 82.8% similarity with SARS-CoV, 21.2% identity and 33.4% similarity with HCoV-OC43, 19.8% identity and 33.3% similarity with MHV-JHM, and 19.4% identity and 28.9% similarity with HCoV-HKU1. All of these viruses possess neuroinvasive and neurotropic properties. However, the S1 RBD has only 10.6% identity and 22.6% similarity to that of HCoV-NL63, and 10.9% identity and 22.3% similarity to that of TGEV. Both of these viruses have not been reported to infect the human nervous system thus far. In the S2 protein, SARS-CoV-2 shares 89.4% identity and 95.5% similarity with SARS-CoV, 42.3% identity and 59.7% similarity with

HCoV-OC43, 41.3% identity and 59.7% similarity with MHV-JHM, and 39.1% identity and 57.2% similarity with HCoV-HKU1, but here SARS-CoV-2 shares relatively less identity and similarity with HCoV-NL63 (32.5% identity and 48.8% similarity) and TGEV (32.2% identity and 47.6% similarity). Figure 3C shows the protease cleavage sites at the S1/S2 boundary and within the S2 protein. The consensus motif of the cleavage sites at the S1/S2 boundary (sites 1 and 2) are much less conserved than the motifs within the S2 protein. The cleavage site 1 of S1/S2 boundary of SARS-CoV-2, SARS-CoV, MHV-JHM, HCoV-HKU1, HCoV-OC43, PHEV, and MERS-CoV is prone to mutation, and the site 1 in HCoV-299E, HCoV-NL63, TGEV, and FCoV does not contain a consensus motif. Furthermore, the cleavage site 2 at the S1/S2 boundary of SARS-CoV-2 shows a similar motif to SARS-CoV, MHV-JHM, HCoV-HKU1, HCoV-OC43, and PHEV, but not to HCoV-299E, HCoV-NL63, TGEV, FCoV, and MERS-CoV. These results suggest that SARS-CoV2 binding and priming host receptor probably share more similar mechanisms with those from neuroinvasive and neurotropic CoVs rather than the other CoVs.

Discussion

The primary objectives of the current study were to determine the possible evolutionary and molecular relationships between SARS-CoV-2 and non-segmented RNA viruses, especially the viruses that can infect the nervous system in infants and children. Furthermore, the consensus sequence motifs of S protein and its protease cleavage sites were focused on to discover their potential roles in neurovirulence.

As shown in Figure 1A, the SARS-CoV-2 is a member of genus β-CoVs. Two of the members (HCoV-HKU1 and HCoV-OC43) have been reported result in infections of the nervous system in pediatric patients. Therefore, it remains possible that SARS-CoV-2 is also neuroinvasive, neurotropic, and even neurovirulent in infants and children because neurological impairment has been reported in SARS-CoV-2 infected patients (Mao *et al*, 2020; Xydakis *et al*, 2020). Although SARS-CoV-2 is genetically distant from the other members of the RNA viral families, its macrodomain and viroporin can also be found in the RNA viruses, which infect the nervous system of infants and children. In the present investigation, the macrodomain was found in CoVs including SARS-CoV, HCoV-HKU1 and MHV, and togaviruses such as rubella and east equine encephalitis viruses. The macrodomain plays an important role in viral replication and pathogenesis, and its activity in togaviruses has been shown to affect neurovirulence in mice (Abraham *et al*, 2020). Viroporin was detected in SARS-CoV-2, SARS-CoV, and picornaviruses such as coxsakie-, echo-, and polio-viruses, which is consistent with previous reports (Nieva *et al*, 2012). During SARS-CoV infections, viroporin can activate caspase-1 by activating the NLRP3 inflammasome (Chen *et al*, 2019). Caspase-1 cleaves pro-interleukin (IL)-1β to mature IL-1β. As a major proinflammatory cytokine, IL-1β facilitates neuroinvasion by disrupting blood-brain barrier (BBB) integrity (Miner and Diamond, 2016). Although the exact functions of macrodomain and viroporin in SARS-CoV-2 have not been elucidated, the sequence identity and similarity suggest that SARS-CoV-2 may share similar mechanisms with neuroinvasive and neurotropic viruses to infect the nervous system in pediatric patients.

The present analysis also showed that the S1 RBD and the protease cleavage sites at the S1/S2 boundary are much less well conserved compared with the S2 protein and the protease cleavage site at

S2 among CoVs, respectively. These findings are consistent with previous reports (Perlman and Wheeler, 2016) and suggest that each CoV may need a specific binding receptor and protease to bind to its target cells. The high diversity of S1 RBD and cleavage sites at S1/S2 boundary may also determine the severity of the injury to host resulting from the viral infection. However, both SARS-CoV-2 and SARS-CoV cell entries require ACE2 binding and TMPRSS2 priming (Hoffmann *et al*, 2020; Li *et al*, 2003). This is probably because they share high sequence identity and similarity of S1 RBD and protease cleavage sites. ACE2 receptor expression can be found in human brain and brain-derived microvascular endothelial cells (Li *et al*, 2007), whereas the TMPRSS2 gene appears to be low or absent in the human brain (Glowacka *et al*, 2011; Vaarala *et al*, 2001). Furthermore, TMPRSS2 expression in human immature BBB endothelial cells has not been reported. This could explain the reason that SARS-CoV-2 infection is usually mild in the nervous system of n adults and children.

Conclusions

In summary, although there is no evidence of SARS-CoV-2 directly causing any known human neuropathology, SARS-CoV-2 shares some close molecular and structural similarity to neuroinvasive and neurotropic non-segmented RNA viruses. This leads to speculation about possible involvement of SARS-CoV-2 in causing neurological abnormalities in pediatric patients.

Declarations

The authors declare that there is no conflict of interest.

Acknowledgement

I gratefully acknowledge the helpful suggestions and editorial assistance of Barbara S. Stonestreet, M.D. (Professor of Pediatrics, Department of Pediatrics, Women & Infants Hospital of Rhode Island, USA), and Yuqi Wu (Department of Pediatrics, Women & Infants Hospital of Rhode Island, USA).

References

Abraham R, McPherson RL, Dasovich M, Badiee M, Leung AKL, Griffin DE (2020). Both ADP-Ribosyl-Binding and Hydrolase Activities of the Alphavirus nsP3 Macrodomain Affect Neurovirulence in Mice. mBio 11.

Beitz E (2000). TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. Bioinformatics 16: 135-9.

Bergmann CC, Lane TE, Stohlman SA (2006). Coronavirus infection of the central nervous system: host-virus stand-off. Nat Rev Microbiol 4: 121-32.

Beyrouti R, Adams ME, Benjamin L, Cohen H, Farmer SF, Goh YY, Humphries F, Jager HR, Losseff NA, Perry RJ, Shah S, Simister RJ, Turner D, Chandratheva A, Werring DJ (2020). Characteristics of ischaemic stroke associated with COVID-19. J Neurol Neurosurg Psychiatry.

Bodenhofer U, Bonatesta E, Horejs-Kainrath C, Hochreiter S (2015). msa: an R package for multiple sequence alignment. Bioinformatics 31: 3997-9.

Chen IY, Moriyama M, Chang MF, Ichinohe T (2019). Severe Acute Respiratory Syndrome Coronavirus Viroporin 3a Activates the NLRP3 Inflammasome. Front Microbiol 10: 50.

Desforges M, Le Coupanec A, Brison E, Meessen-Pinard M, Talbot PJ (2014). Neuroinvasive and neurotropic human respiratory coronaviruses: potential neurovirulent agents in humans. Adv Exp Med Biol 807: 75-96.

Glowacka I, Bertram S, Muller MA, Allen P, Soilleux E, Pfefferle S, Steffen I, Tsegaye TS, He Y, Gnirss K, Niemeyer D, Schneider H, Drosten C, Pohlmann S (2011). Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. J Virol 85: 4122-34.

Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Muller MA, Drosten C, Pohlmann S (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 181: 271-280 e8.

Huson DH, Scornavacca C (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst Biol 61: 1061-7.

Jartti T, Jartti L, Ruuskanen O, Soderlund-Venermo M (2012). New respiratory viral infections. Curr Opin Pulm Med 18: 271-8.

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

Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, Wang X (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 581: 215-220.

Li J, Gao J, Xu YP, Zhou TL, Jin YY, Lou JN (2007). [Expression of severe acute respiratory syndrome coronavirus receptors, ACE2 and CD209L in different organ derived microvascular endothelial cells]. Zhonghua Yi Xue Za Zhi 87: 833-7.

Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426: 450-4.

Mao L, Jin H, Wang M, Hu Y, Chen S, He Q, Chang J, Hong C, Zhou Y, Wang D, Miao X, Li Y, Hu B (2020). Neurologic Manifestations of Hospitalized Patients With Coronavirus Disease 2019 in Wuhan, China. JAMA Neurol.

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015). CDD: NCBI's conserved domain database. Nucleic Acids Res 43: D222-6.

Messacar K, Fischer M, Dominguez SR, Tyler KL, Abzug MJ (2018). Encephalitis in US Children. Infect Dis Clin North Am 32: 145-162.

Miner JJ, Diamond MS (2016). Mechanisms of restriction of viral neuroinvasion at the blood-brain barrier. Curr Opin Immunol 38: 18-23.

Miura TA, Travanty EA, Oko L, Bielefeldt-Ohmann H, Weiss SR, Beauchemin N, Holmes KV (2008). The spike glycoprotein of murine coronavirus MHV-JHM mediates receptor-independent infection and spread in the central nervous systems of Ceacam1a-/- Mice. J Virol 82: 755-63.

Nieva JL, Madan V, Carrasco L (2012). Viroporins: structure and biological functions. Nat Rev Microbiol 10: 563-74.

Perlman S, Wheeler DL (2016). Neurotropic Coronavirus Infections. Neurotropic Viral Infections: Volume 1: Neurotropic RNA Viruses: 115-148.

Phillips JJ, Chua MM, Rall GF, Weiss SR (2002). Murine coronavirus spike glycoprotein mediates degree of viral spread, inflammation, and virus-induced immunopathology in the central nervous system. Virology 301: 109-20.

Poyiadji N, Shahin G, Noujaim D, Stone M, Patel S, Griffith B (2020). COVID-19-associated Acute Hemorrhagic Necrotizing Encephalopathy: CT and MRI Features. Radiology: 201187.

Principi N, Bosis S, Esposito S (2010). Effects of coronavirus infections in children. Emerg Infect Dis 16: 183-8.

Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10: 512-26.

Vaarala MH, Porvari KS, Kellokumpu S, Kyllonen AP, Vihko PT (2001). Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. J Pathol 193: 134-40.

Wu Y, Xu X, Chen Z, Duan J, Hashimoto K, Yang L, Liu C, Yang C (2020). Nervous system involvement after infection with COVID-19 and other coronaviruses. Brain Behav Immun.

Xydakis MS, Dehgani-Mobaraki P, Holbrook EH, Geisthoff UW, Bauer C, Hautefort C, Herman P, Manley GT, Lyon DM, Hopkins C (2020). Smell and taste dysfunction in patients with COVID-19. Lancet Infect Dis.

Zimmermann P, Curtis N (2020). Coronavirus Infections in Children Including COVID-19: An Overview of the Epidemiology, Clinical Features, Diagnosis, Treatment and Prevention Options in Children. The Pediatric Infectious Disease Journal 39: 355-368.

Table

Table: Neuroinvasive and neurotrophic non-segmented RNA viruses analyzed in this study

Viral family	RNA viruses	RNA structure	Accession Nr#
Picornaviruses	Poliovirus 1	SS ⊕ linear, Icosahedral capsid, no envelope	FJ769385
	Coxsackievirus B1		JX976769
	Echovirus 2		AF465518
	Human parechovirus 3 (HPeV-3)		AJ889918
Flaviviruses	Japanese encephalitis virus	SS ⊕ linear, Icosahedral capsid, no envelope	KY927818
	West Nile virus		AF404756
	St. Louis encephalitis virus		KM267635
	Tick-borne encephalitis virus		FJ572210
	Zika virus		KX893855
Retroviruses	Human immunodeficiency virus (HIV)-1	SS ⊕ linear, conical capsid, envelope	AY352275
	Human T-cell leukemia virus type 1	SS ⊕ linear, Icosahedral capsid, envelope	MH399769
Togaviruses	Eastern equine encephalitis virus	SS ⊕ linear, Icosahedral capsid, envelope	MK028842
	Western equine encephalitis virus		MN477208
	Venezuelan equine encephalitis virus		L01442
	Chikungunya virus		MH229986
	Rubella		NC_001545
Coronavirus	Severe acute respiratory syndrome (SARS)-CoV-2	SS ⊕ linear, helical capsid, envelope	NC_045512
	SARS-CoV		NC_004718
	Middle East Respiratory Syndrome (MERS)-CoV		NC_019843
	HCoV-229E		KU291448
	HCoV-OC43		KU131570
	HCoV-HKU1		NC_006577
	Murine hepatitis virus (MHV)-JHM‡		AC_000192
	Feline coronavirus (FCoV) ‡		DQ010921

	Porcine hemagglutinating encephalomyelitis (PHEV) ‡		KY419112
	HCoV-NL63*		NC_005831
	Transmissible gastroenteritis virus (TGEV)* ‡		KX900411
	Avian infectious bronchitis virus* ‡		NC_001451
	Sparrow delta-coronavirus* ‡		MG812378
Paramyxoviruses	Respiratory syncytial virus	SS ⊖ linear, helical capsid, envelope	NC_001803
	Measles		K01711
	Mumps		AB470486
	Hendra virus		JN255805
	Nipah virus		NC_002728
Rhabdoviruses	Rabies	SS ⊖ linear, helical capsid, envelope	NC_001542

‡ The detection of virus has not been reported in human. * The detection of virus has not been reported in CNS.

Figures

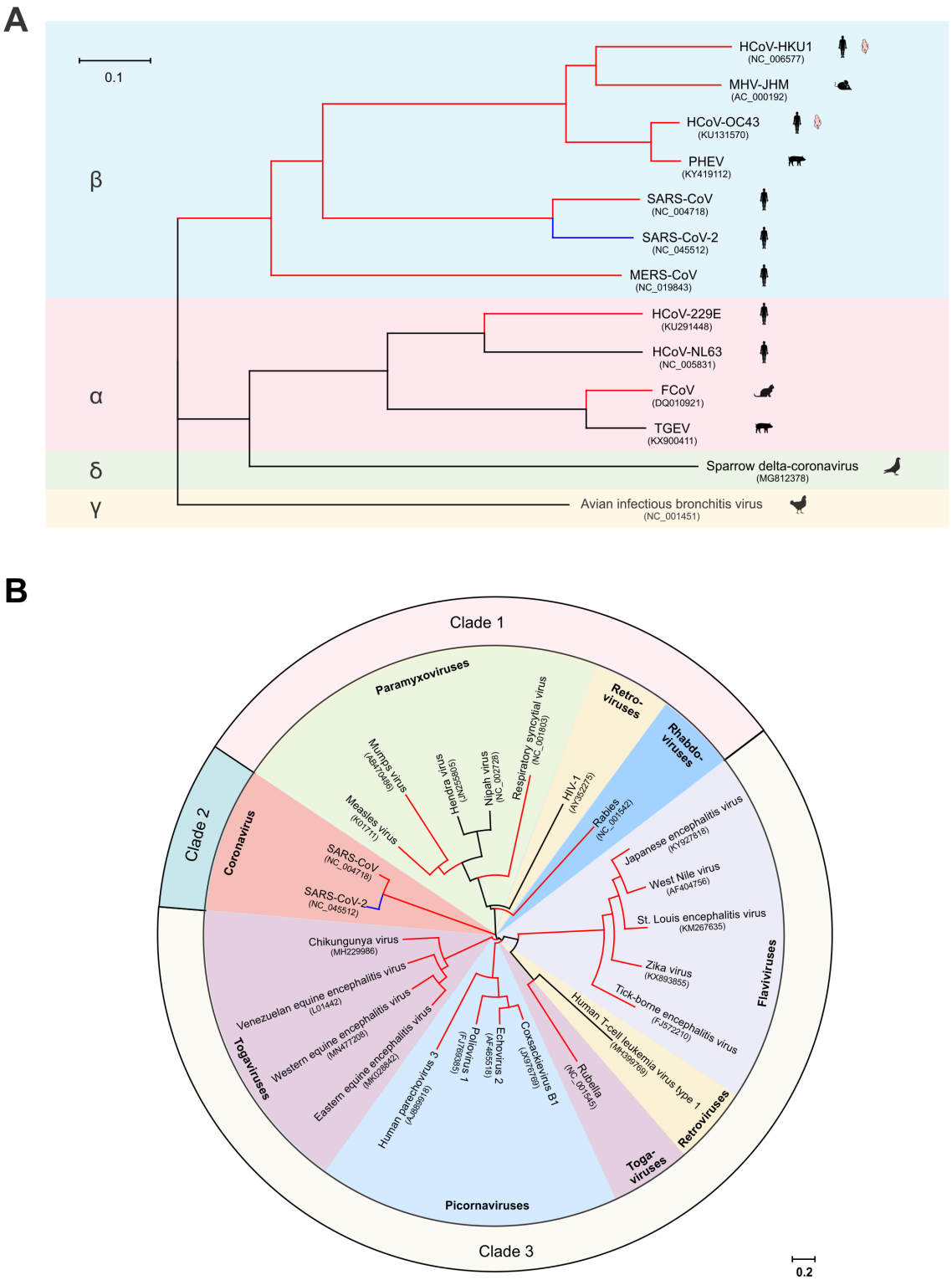

Figure 1

Evolutionary history of SARS-CoV-2 compared to selected CoVs (A) and neuroinvasive and neurotropic non-segmented RNA viruses (B). Evolutionary analyses were conducted in MEGAX. The nucleotide sequences were aligned with ClustalW and the results were visualized by Dendroscope-3. The accession numbers for the viruses studied in this present study are shown. A) The rectangular phylogram shows that SARS-CoV-2 belongs to genus β-CoVs and is genetically most close to SARS-CoV. This analysis

involved 13 nucleotide sequences and included a total of 32927 positions in the final dataset. The highest log likelihood of the tree is -363492.55. The branch of SARS-CoV-2 was labelled with blue color, whereas the branches of other CoVs infecting nervous system were labelled with red color. The CoVs (HCoV-HKU1 and HCoV-OC43) that associate with pediatric neurological disorders are indicated. The distance scale bar suggests a 0.1 (10%) genetic variation for the length of the scale between sequences. B) The circular phylogram shows evolutionary relationship of SARS-CoV-2 with human-infecting neuroinvasive and neurotropic non-segmented RNA viruses. Three major clades are shown. The CoVs (clade 2) do not have same ancestor with other RNA viral family. The highest log likelihood of the tree is -362743.38. The branch of SARS-CoV-2 was labelled with blue color, whereas the branches of other viruses infecting infants and children were labelled with red color. This analysis involved 24 nucleotide sequences and contained a total of 30543 positions in the final dataset. The distance scale bar suggests a 0.2 (20%) genetic variation for the length of the scale between sequences.

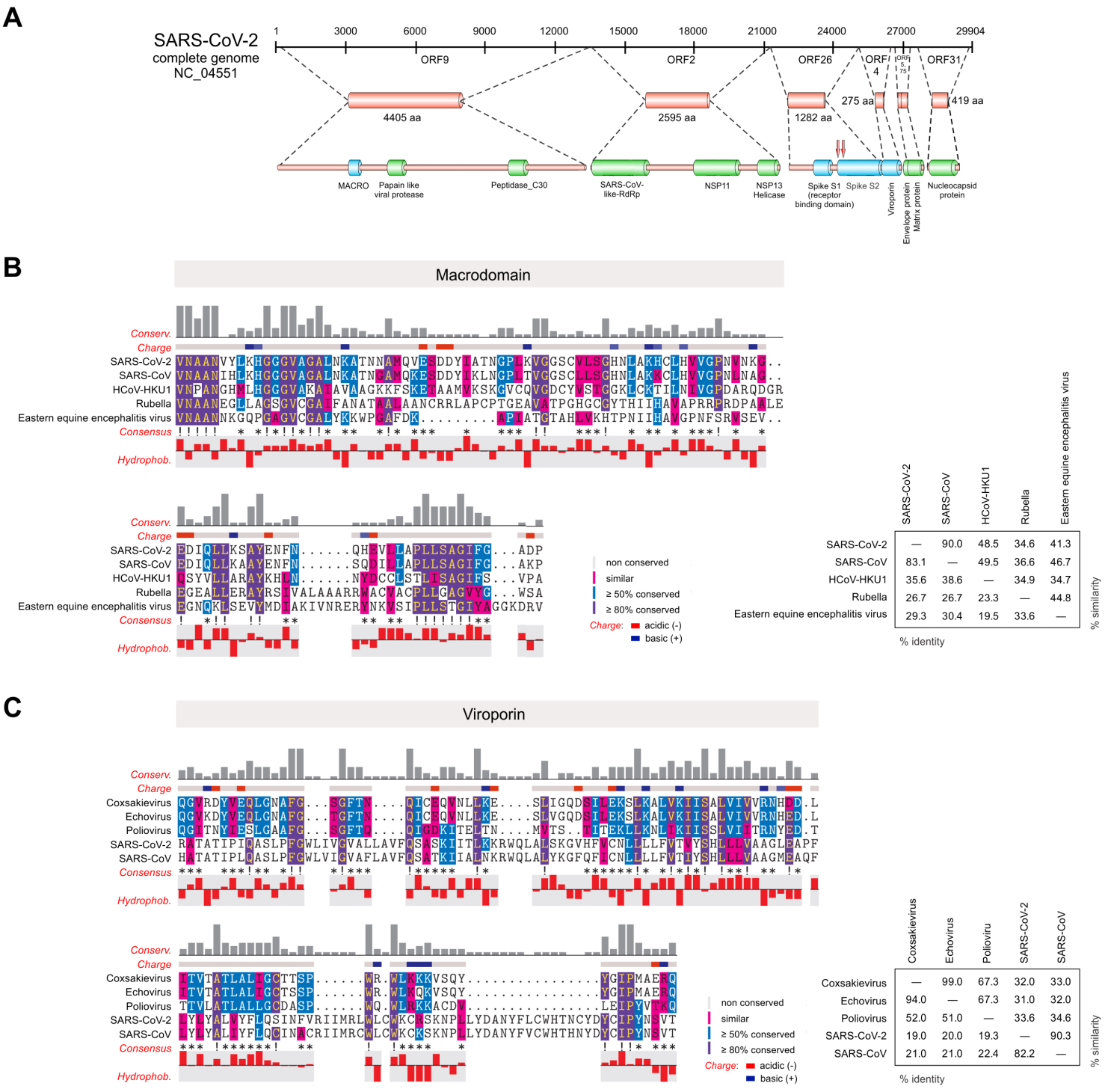

Figure 2

A) Schematic representation of SARS-CoV-2 complete genome (accession number: NC_04551) and selected ORF and encoded conserved domains. Macrodomain, S protein, and viroporin, investigated in the present study, are highlighted in blue color. The arrows indicate the protease priming sites on S protein. B) The macrodomain sequence of SARS-CoV-2 was compared to SARS-CoV, HCoV-HKU1, rubella, and eastern equine encephalitis virus. C) The viroporin sequence of SARS-CoV-2 was compared to SARS-CoV, coxsakievirus, echovirus, and poliovirus. RStudio with msa package including ClustalW command was

used for multiple sequence alignment. The results were visualized by RStudio and LaTeX with TEXshade package. All identical residues at a position were shaded in blue or purple if the number of matching residues is higher than 50% or 80%, respectively. The residues that are not identical but similar to the consensus sequence were shaded in red. Further, the degree of protein sequence conservation and amino acid properties such as charge and hydrophobicity were shown as color scales and bar graph along the alignment. On the top of the plot, residue conservation was shown as bars and the charge of amino acid side chain was shown as color scales (red: acidic; blue: basic). Hydrophobicity was shown at the bottom of the plot (upper red box: hydrophobic; underside box: hydrophilic). The degree of similarity and identity between all sequences in the alignment were shown in tables.

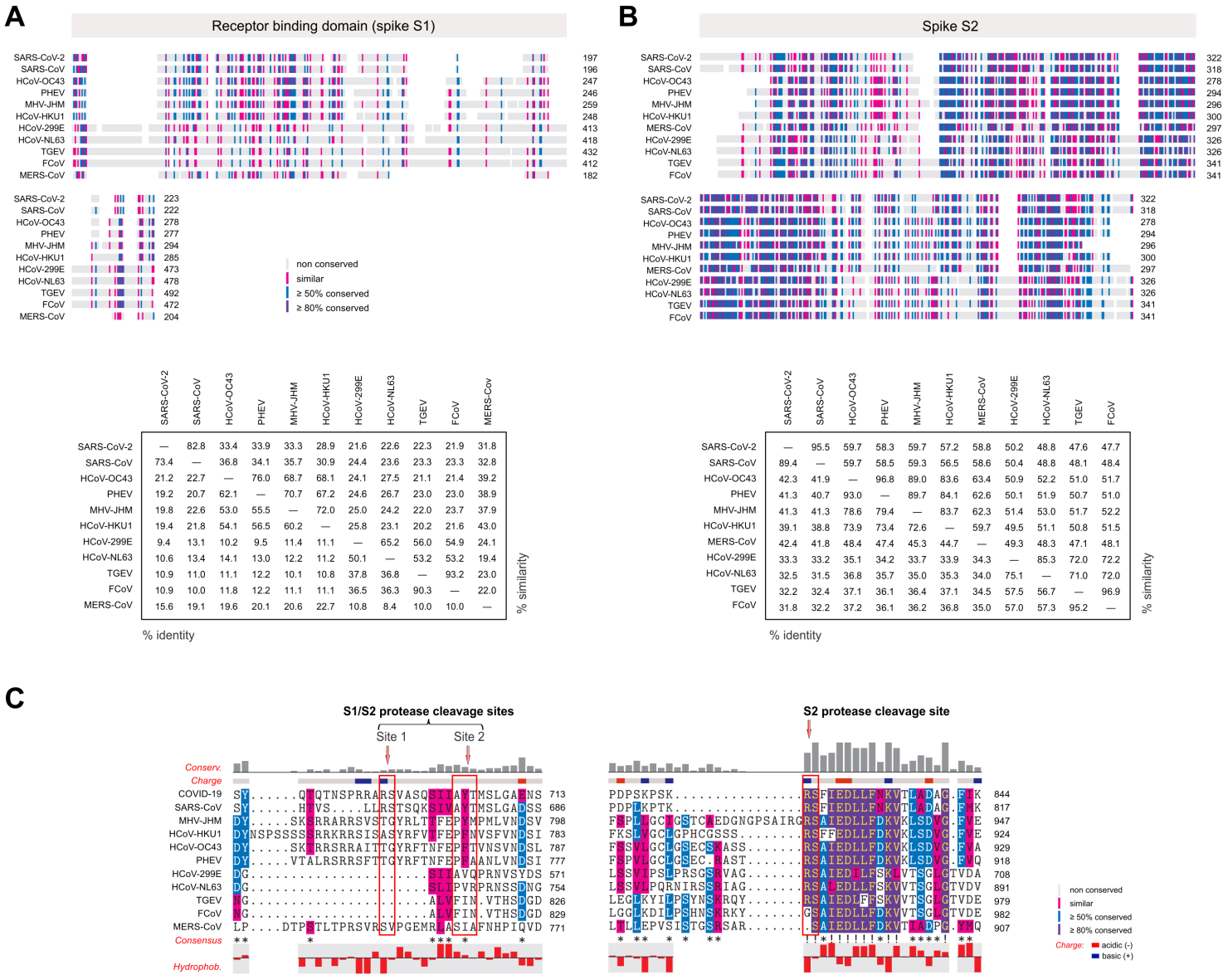

Figure 3

Multiple sequence alignments of S1 RBD, S2 protein, and S protein protease cleavage sites among CoVs. An overview of sequence similarities of S1 RBD (A) and S2 protein (B) was implemented by fingerprint plots, which depict the complete sequence in one single line. The residues were presented as colored

vertical lines (red: similar; blue: ≥ 50% conserved; purple: ≥ 80% conserved). S1 RBD showed much less conserved than S2 protein. The degree of similarity and identity between all sequences in the alignment were shown in tables. C) The protease cleavage sites of S protein were compared among CoVs with same msa methodology as described in Figure 2B and 2C. There are two cleavage sites at S1/S2 boundary. However, both are less conserved than the cleavage site at S2 protein.