Investigation of Virus Evolution in SARS-CoV-2 Clinical Samples From Experimentally Infected Cats

Maddie Butterfield[ab], David Meekins[ab], Chester McDowell[ab], Juergen A. Richt[ab]

Department of Medicine/Pathobiology[a], Center of Excellence for Emerging and Zoonotic Animal Diseases[b],
College of Veterinary Medicine, Kansas State University, Manhattan, Kansas, USA

C · E · E · Z · A · D
Center of Excellence
for Emerging and Zoonotic Animal Diseases
A Department of Homeland Security
Center of Excellence

Abstract

Understanding the mutation frequency of the genome of RNA viruses such as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is of particular importance for disease control. High mutation frequency usually found in RNA viruses results in accumulation of genetic mutations, potentially altering the virulence and tropism of the virus. In this study, we evaluated the replication fidelity of the SARS-CoV-2 genome by determining: (a) the stability of the SARS-CoV-2 isolate USA-WA1/2020 genome after replication in experimentally infected cats and (b) the uniqueness and quantity of SARS-CoV-2 viral quasi-species in feline samples compared to the original USA-WA1/2020 inoculum. Samples (nasal, rectal, and oropharyngeal swabs, bronchoalveolar lavage fluid and mucus) were obtained from cats experimentally infected via the intranasal/oral route using the human SARS-CoV-2 isolate USA-WA1/2020. SARS-CoV-2-RNA was isolated and sequenced via next-generation sequencing (NGS) from 28 feline samples and the VeroE6 cell passaged inoculum. Consensus sequences were extracted from feline samples and compared to the reference USA-WA1/2020 inoculum sequence to determine changes in nucleotide and amino acid sequences. The SARS-CoV-2 N-, E- and Orf1ab-gene regions were analyzed (1,254 bps, 222 bps, and 21,289 bps, respectively) in detail, which represent approx. 80% of the SARS-CoV-2 genome. Feline consensus sequences deviated from the reference sequence only in the Orf1ab gene region with an average of 3.21 ±1.10 nucleotide changes (±1SD) per consensus sequence compared to the reference sequence. The mutations found after cross-species transmission to cats will be further analyzed for their impact on diagnostic tools and vaccines for all animal species. A detailed understanding of mutational "hotspots" in SARS-CoV-2, particularly after cross-species adaptation, will aid in the development of effective therapeutics and vaccinations and improved veterinary diagnostics and care.

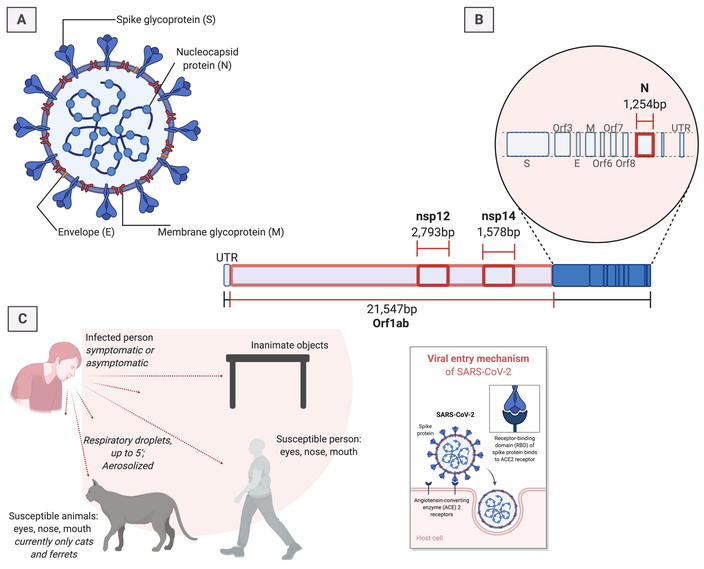

Figure 1: SARS-CoV-2 virion, genome organization, and transmission – A. SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus of the genus *Betacoronavirus* within family *Coronoviridae.* Human pathogens MERS-CoV and SARS-CoV also belong to the genus *Betacoronavirus.* **B.** The SARS-CoV-2 genome is ~30kb and encodes 27 proteins[1], including four structural proteins: the spike glycoprotein (S) mediates receptor binding and viral entry[2], the envelope (E) protein plays a role in viral assembly and budding and envelope development[3]. The membrane (M) glycoprotein defines the shape of the virion and mediates production of new viral particles[4], and the nucleocapsid (N) protein forms organizational complexes with the RNA genome as well as is critical to genomic replication[5]. Global SARS-CoV-2 diagnostic RT-PCR assays include screens for RNA from the N-gene region, the nsp-12 (RNA-dependent RNA polymerase) and nsp-14 regions within the orf1ab region, and the Orf1ab region itself.[6] These regions (N gene, nsp-12, nsp-14, Orf1ab) are therefore included in the areas of interest within this study. **C.** SARS-CoV-2, the pathogen responsible for COVID-19, is transmitted through respiratory droplets and aerosol. Susceptible individuals become infected through inhalation of virus or through the eyes, or exposure to infected fomites. The virus enters host cells when the receptor-binding domain of the spike (S) protein binds to angiotensin-converting enzyme (ACE2) receptors on host tissues. Animals have shown to be susceptible to SARS-CoV-2 from human infection, though only domestic and "big" cats, dogs, and mink have shown natural susceptibility at this time[7].

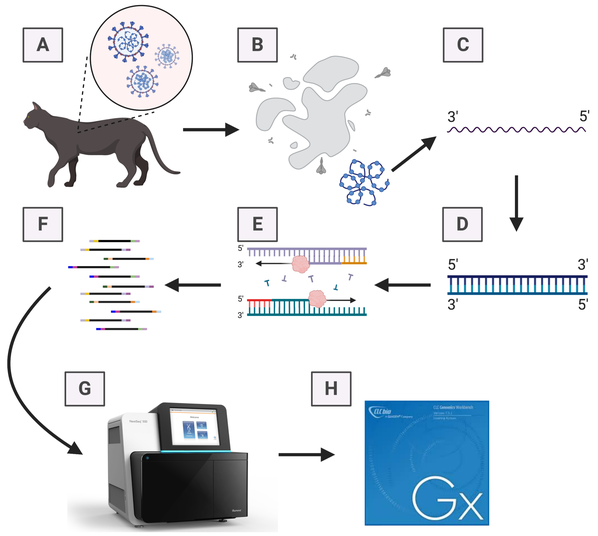

Figure 2: Preparation and sequencing of SARS-CoV-2 experimentally-infected cat samples – A. Samples were collected from felines after intranasal/oral challenge with the human SARS-CoV-2 isolate USA-WA1/2020. Samples included rectal, oropharyngeal, and nasal swabs, mucus samples, and bronchoalveolar lavage fluid (BALF). **B.** Sample material was lysed and **C.** viral RNA was extracted using GeneReach Taco RNA Extraction Kit (GeneReach Biotechnology Corp., Taichung City, Taiwan), reverse-transcribed **(D)** and amplified **(E)** via PCR using a tiled-primer system to amplify the entire SARS-CoV-2 genome. A sequencing library was prepared **(F)** using Nextera XT library preparation kit (Illumina, Inc., San Diego, CA, USA). Next-generation sequencing was completed **(G)** using a standard protocol on the Illumina NextSeq with 150bp paired-end reads. Sequences were quality-controlled using a standard NextSeq protocol and resulting data were imported into CLC Genomics Workbench 7.5 **(H).** Reads were mapped to the USA-WA1/2020 reference sequence and a consensus sequence was extracted from each sample. Consensus sequences were then analyzed in comparison to the reference USA-WA1/2020 sequence.

Conclusions:

- SARS-CoV-2 is an emerging zoonotic disease with global distribution and has caused over 14 million infections and over 600 thousand deaths worldwide to date.
- The genome of SARS-CoV-2 encodes 27 proteins and 4 structural proteins; 4 viral genes are used globally as targets for diagnostic screening: the nucleoprotein, the RNA-dependent RNA polymerase, nsp14, and Orf1ab.
- After cross-species transmission of a human SARS-CoV-2 isolate to cats, the regions encoding the nucleocapsid region and RNA-dependent RNA polymerase (N region and nsp12, respectively) remained 100% identical to the corresponding USA-WA1/2020 reference sequence regions.
- All (100%) feline sample consensus sequences had an identical single-nucleotide synonymous mutation C18060T in nsp14.
- Orf1ab had an average of 3.21 ±1.10 nucleotide changes (±1SD) per consensus sequence among 28 feline samples.
- 93.75% (15/16) of rectal swabs had an identical nonsynonymous mutation at nucleotide location 21304, which results in an AA substitution from R (arg) to S (ser).
- The identified mutations, particularly at nucleotide location 18060 and 21304, will be further analyzed for their impact on diagnostic tools, therapeutics and vaccines.

Acknowledgements & Support

This work was supported by the Elanco Veterinary Scholars Program, NBAF Transition Funds from the State of Kansas, KSU Internal Funds, the NIAID Centers of Excellence for Influenza Research and Surveillance under contract nos. HHSN 272201400006C., the DHS Center of Excellence for Emerging and Zoonotic Animal Diseases under grant no. 2010-ST061-AG0001, and the Kansas State Veterinary Research Scholars Program. We would like to thank Dr. Jessie Trujillo, Emily Gilbert-Esparza, and the entire Richt lab for their tireless support of this project.

References:

1: Wu A, Peng Y, Huang B, et al. 2020. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host Microbe.* 27(3):325-328. doi:10.1016/j.chom.2020.02.001
2: Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet.* 2020;395(10224):565-574. doi:10.1016/S0140-6736(20)30251-8
3: Nieto-Torres JL, Dediego ML, Alvarez E, et al. 2011. Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. *Virology.* 415(2):69-82. doi:10.1016/j.virol.2011.03.029
4: Neuman BW, Kiss G, Kunding AH, et al. 2011. A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol.* 174(1):11-22. doi:10.1016/j.jsb.2010.11.021
5: McBride R, van Zyl M, Fielding BC. 2014. The coronavirus nucleocapsid is a multifunctional protein. *Viruses.* 6(8):2991-3018. doi:10.3390/v6082991
6: Udugama B, Kadhiresan P, Kozlowski HN, et al. 2020. Diagnosing COVID-19: The Disease and Tools for Detection. *ACS Nano.* 14(4):3822-3835. doi:10.1021/acsnano.0c02624
7: Shi J, Wen Z, Zhong G, et al. 2020. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science.* 368(6494):1016-1020. doi:10.1126/science.abb7015
8: Smith EC. 2017. The not-so-infinite malleability of RNA viruses: Viral and cellular determinants of RNA virus mutation rates. PLoS Pathog 13(4): e1006254

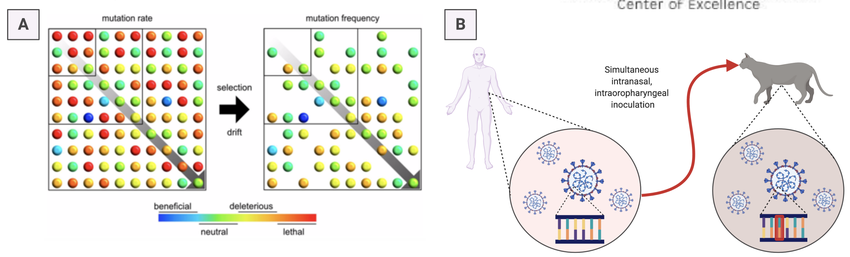

Figure 3: Viral Quasi-Species Hypothesis: Cross-species transmission of viruses may result in genetic adaptation due to novel selection pressures – A. Infinite replication of RNA viruses greatly increases the probability of nucleotide mutations occurring along the RNA genome[8]. These mutations can range from beneficial alterations to lethal consequences for the virus. Mutations are subjected to selective pressures over the course of many replication cycles due to host conditions, therefore an accumulation of mutations may result in a change in the virulence and tropism of a virus. **B.** When a defined, NGS-sequenced human SARS-CoV-2 strain (i.e. USA-WA1/2020) is introduced to a novel feline host, we hypothesize that the cross-species transmission will result in mutations in the viral genome arising from viral adaptation to the new feline host (A: Figure 1 from Smith et al. 2017 [8].)

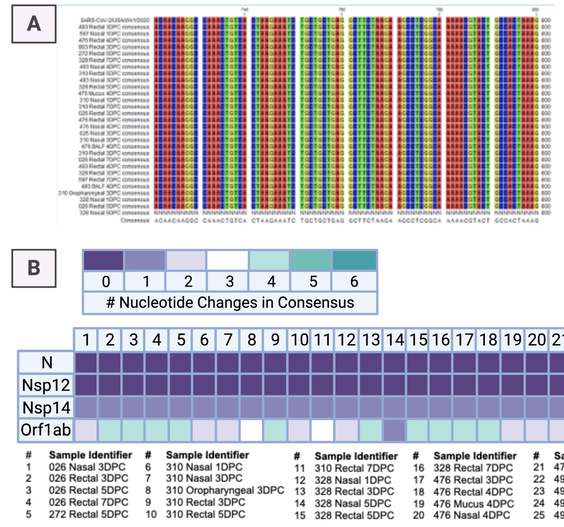

Figure 4: Nucleotide consensus alignment to reference sequence and pairwise analysis of regions N, Nsp12, Nsp14, and Orf1ab - A. Using CLC Genomics Workbench 7.5, consensus strands were extracted from each sample. N regions (1,254bp), nsp12 regions (2,793bp), nsp14 regions (1,578bp), and orf1ab (21,289bp) regions from all consensus sequences were aligned to their respective regions in the reference SARS-CoV-2/USA-WA1/2020 isolate. **B.** A pairwise alignment analysis was run on these nucleotide alignments. The N and nsp12 regions had 100% homology to the respective consensus sequences of USA-WA1/2020. The nsp14 region had 1 identical mutation in every consensus sequence compared to the reference sequence. The orf1ab region varied in mutation numbers, ranging from 1 to 6 with an average of 3.21 ±1.10 nucleotide changes (±1SD) per consensus sequence.

Location Along Genome (NT)	Region in Genome	Mutation	Percent (Number) Consensus Sequences Present	AA Change
686	Orf1ab	A→G	6.89% (2)	V→V
3993	Orf1ab	TG→CA	3.45% (1)	CG→CA
6869	Orf1ab	A→T	3.45% (1)	S→S
8782	Orf1ab	C→T	96.55% (28)	A→V
18060	Orf1ab, nsp14	C→T	100% (29)	L→L
21304	Orf1ab, nsp13	C→A	51.72% (15)*	R→S
			*93.75% (15/16) rectal swabs	

Table 1: Description of nucleotide mutations along 29 consensus sequences from cat samples – Mutations were only present in the consensus sequence if present in over 50% of reads from the sample. The C→T mutation at 18060 in the nsp14 region was present in 100% of sample consensus sequences. This mutation is a synonymous mutation, maintaining a leucine amino acid (L→L). The C→A mutation at 21304 was present in only 51.72% of sample consensus sequences but was present in 93.75% of rectal swab sample consensus sequences. This mutation is a nonsynonymous mutation, changing a positively-charged arginine amino acid to an electrically-uncharged serine (R→S).