bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Myeloablation-associated deletion of ORF4 in a human coronavirus 229E infection

Alexander L. Greninger[1,2,#], Gregory Pepper[1], Ryan C. Shean[1], Anne Cent[1], Isabel Palileo[1], Jane M. Kuypers[1], Joshua T. Schiffer[2,3], Keith R. Jerome[1,2]

[1]Department of Laboratory Medicine, University of Washington, Seattle, WA

[2]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA

[3]Department of Medicine, University of Washington, Seattle, WA

[#]corresponding author, agrening@uw.edu

Abstract

We describe metagenomic next-generation sequencing (mNGS) of a human coronavirus 229E from a patient with AML and persistent upper respiratory symptoms, who underwent hematopoietic cell transplantation (HCT). mNGS revealed a 548-nucleotide deletion, which comprised the near entirety of the ORF4 gene, and no minor allele variants were detected to suggest a mixed infection. As part of her pre-HCT conditioning regimen, the patient received myeloablative treatment with cyclophosphamide and 12 Gy total body irradiation. Iterative sequencing and RT-PCR confirmation of 4 respiratory samples over the 4-week peritransplant period revealed that the pre-conditioning strain contained an intact ORF4 gene, while the deletion strain appeared just after conditioning and persisted over a 2.5-week period. This sequence represents one of the largest genomic deletions detected in a

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

human RNA virus and is the first description of large-scale viral mutation associated with myeloablation for HCT.

Importance

Unlike DNA viruses, most RNA viruses pack a considerable amount of function in a relatively small genome, often in the form of large polyproteins. RNA viruses can therefore rarely afford to lose genes. While retrospectively sequencing virus-positive respiratory specimens in our laboratory, we found a human coronavirus 229E strain with a large genomic deletion comprising nearly an entire gene. w The development of the deletion was temporally associated with the patient's myeloablation treatment. These data indicate that ORF4 is likely dispensable for in vivo human coronavirus 229E replication and a poor target for therapeutic intervention. They also show the potential for viral mutation associated with chemotherapy and irradiation treatment.

Introduction

Coronaviruses are positive-stranded RNA viruses with large genomes, ranging in size from 25-35kb. Although much attention has been paid to the epidemic potential and high mortality rates for MERS and SARS coronaviruses, the most common human coronaviruses (HCoV) include 229E, OC43, HKU1, and NL63 viruses (1). HCoV 229E and NL63 are alphacoronaviruses with genome lengths of 27-28kb, and HCoV HKU1 and OC43 are betacoronaviruses with genome lengths around 30-31kb (2). Few studies have been conducted on the genome evolution of the common HCoV species (3, 4).

HCoVs enact a major burden on hematopoietic cell transplant (HCT) patients with >10% cumulative incidence of HCoV infection among HCT by day 100 of transplant (5). HCoV in HCT patients is associated with prolonged shedding, with a median duration of 3

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

weeks (5). Conditioning regimens for HCT usually consist of chemotherapy and irradiation that may induce genomic alterations (6, 7). Treatment-related myeloid leukemias are well-known to be associated with radiation therapy or chemotherapeutic agents such as cyclophosphamide (8). Here we describe the generation of a large deletion in an RNA virus that was temporally associated with the myeloablative conditioning regimen for HCT.

Materials and Methods

Cell culture and qRT-PCR

The HCoV 229E type strain ATCC VR-740 was obtained and used as a positive control for cell culture. 200ul of each patient nasal swab in viral transport media (50ul 229E control/150ul 2% MEM) was inoculated onto MRC-5 cells and incubated at 37C in stationary rack for 30 minutes. After incubation, 1.2mL of 2% MEM was added to each tube and the tubes incubated on a rotating drum at 37C. Each day tubes were read for cytopathic effect and 50uL of supernatant was stored at -80C for qRT-PCR. 50uL culture supernatants were extracted on a Roche MagNA Pure LC 2.0 and RNA was eluted in 100uL water. 10uL of RNA was used for qRT-PCR with primers targeting the HCoV polymerase gene (9). The assay used detects the four common HCoV species with equal sensitivity and does not distinguish between species. qRT-PCR on patient samples was performed on 200uL of nasal swab in viral transport with the same extraction and PCR conditions.

Metagenomic next-generation sequencing and confirmatory RT-PCR

mNGS was performed as described previously (10). Briefly, 20uL of extracted RNA was treated with DNAse I and double stranded cDNA was created using random hexamers with SuperScript III and Sequenase v2.0 (Thermo Fisher). mNGS libraries were created using one-third volume NexteraXT reactions with 16 cycles of dual-indexed PCR

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

amplification. Libraries were sequenced on an Illumina MiSeq and sequencing data were analyzed to the alignment to the HCoV 229E reference genome (NC_002645) and visualized using Geneious v9.1 (11). Confirmatory gap junction RT-PCR was performed on 1uL of RNA extracted as above using 35 cycles following manufacturer's recommended conditions (Tm 55C) using quarter-volumes of the Qiagen One-Step RT-PCR kit and primers HCoV_24034F (5′ – TTGTTGTGAATCAACTAAACTTCC – 3′) and HCoV_24808R (5′ – ACACACCAGAGTAGTACATTAAC – 3′).

Results

The virus analyzed was from a female in her 40s with acute myelogeneous leukemia presenting for hematopoetic cell transplant. The previous year she had undergone 4 cycles of chemotherapy (G-CLAM, G-CLA, cytarabine, and decitabine-primed MEC). Her prior month's bone marrow evaluation showed 16% blasts by flow cytometry with normal cytogenetics. On day –12 prior to transplant, she was found to be infected by a coronavirus 229E with a cycle threshold (Ct) of 30 and persisted at the same Ct on day -9 (Figure 1A). Based on the patient's risk of relapsing leukemia and relative rarity of HCoV pneumonitis, the decision was made to continue with HCT despite ongoing viral shedding. On days -7 and -6 she received cyclophosphamide 60mg/kg, and on days -4, -3, and -2 she received a cumulative dose of twice-daily 2 Gy total body irradiation for a cumulative dose of 12 Gy (Figure 1A). A mismatched unrelated donor peripheral blood stem-cell transplant was performed on day 0. Her course was complicated by bacteremia on day 10 and she was discharged on day 20.

Four nasal swab samples over a four-week period were available from the patient. All qRT-PCR cycle thresholds (Ct) recovered over the four-week period were comparable, ranging between 31.3-33.5 (Table). mNGS of the Day 15 post-transplant specimen revealed

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

a large deletion that encompassed 548 of the 660 nucleotides of the ORF 4 gene (Figure 1B) with average coverage of 68X across the viral genome. Repeat library preparations and mNGS of the same specimen demonstrated the same deletion. No read was recovered within the area of the deletion and no minor alleles were present at >5% allele frequency, suggesting the presence of a clonal viral population with the genomic deletion. mNGS of the day -12 and -9 specimens from the same patient revealed a completely intact ORF4a/4b (Figure 1B). Other than the large deletion, no variants were observed between the HCoV 229E consensus genomes recovered from day -12, day -9, and day 15 specimens. No junctional reads across the 548-nucleotide deletion were found in either the day -12 or day -9 specimen. mNGS of the day 2 specimen also demonstrated the 548-nucleotide deletion with coverage across the genome consistent with RNA degradation (Figure S1). Confirmatory RT-PCR across the ORF 4a/4b gene demonstrated an intact gene in the pre-myeloablation specimens and a smaller PCR product consistent with the 548-nucleotide deletion in the day -2 and day 15 post-myeloablation specimens (Figure 1C).

Attempts to culture each of the patient's specimens on MRC-5 cells proved unsuccessful. No CPE was visualized for any of the patient's four clinical samples and HCoV 229E RNA levels in the culture supernatant were undetectable by qRT-PCR after inoculation (Figure S1X). CPE was evident by day 4 for the positive control HCoV 229E ATCC type strain and RNA levels increased by 5 Cts over 6 days (Figure S2).

Discussion

We describe metagenomic detection of a large deletion in a HCoV 229E strain temporally associated with a myeloablative regimen of cyclophosphamide and total body irradiation in an HCT patient. The ORF4-deleted virus maintained the same viral load over a 2.5 week period and was clonal by day 15 after transplant, suggesting that it was

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

replication competent in vivo. The ability of the virus to persist in vivo despite the near complete loss of ORF4 suggests that this accessory gene is completely dispensable for HCoV 229E replication and may outcompete the wild-type virus under certain conditions. Indeed, HCoV 229E type strain contains a 2-nucleotide deletion that splits ORF4 into ORF4a/4b, and the alpaca alphacoronavirus contains a 1-nucleotide insertion that also disrupts the reading frame (12, 13). Meanwhile, all other group1b coronaviruses including all known HCoV 229E clinical strains contain an intact ORF4 (12, 14). This is the first detection of a non-intact ORF4 in HCoV 229E associated with human infection.

The function of ORF4 in the alphacoronaviruses is unknown. The protein sequence shows no significant alignment to any protein in Protein Data Bank by HHPred (15). ORF4a of the HCoV 229E type strain has been suggested to be a viroporin that possessed ion channel activity (16).

Many potential mechanisms for development of the deletion exist. The chemotherapy and/or irradiation could have been directly toxic to the coronavirus genome. Cyclophosphamide is a well-known clastogen and has recently been associated with increased development of single nucleotide variations insertions and deletions (7, 17). Alternatively, the myeloablative regime could have removed immune pressure, allowing for the loss of ORF4. It seems that in certain circumstances, deletion of ORF4 could prove a selective advantage as it was the only allele present in the day 15 specimen.

This study has several limitations. We were unable to culture any of the patient's HCoV specimens to demonstrate ability to replicate in vitro. This is not necessarily surprising as HCoV 229E are not usually cultured in the clinical virology lab. We did not attempt reverse genetics of any alphacoronaviruses to definitively show ORF4 is dispensable for in vitro replication, although the ability of the type strain to replicate despite a 2-nucleotide deletion in ORF4 is consistent with this possibility.

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

In summary, we describe a large deletion in HCoV 229E encompassing almost an entire gene that was temporally associated with myeloblation. Little is known about the genome evolution of HCoV 229E as only 20 genomes are currently available in NCBI Genbank. Similarly sized (500-700 nucleotide) deletions have previously been found in hepatitis D virus in patients receiving antiviral therapy (18). As mNGS is increasingly performed in the clinical lab (19), much intriguing biology can be unlocked from clinical specimens, providing a continuing case for the sequencing clinical microbial genomes (20–22).

Accession numbers

These sequences are deposited in NCBI Genbank under the strain names SC677 day 15 (KY369909), SC379 day -12 (KY621348), and SC399 day -9 (KY674914).

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

1. Lepiller Q, Barth H, Lefebvre F, Herbrecht R, Lutz P, Kessler R, Fafi-Kremer S, Stoll-Keller F. 2013. High incidence but low burden of coronaviruses and preferential associations between respiratory viruses. J Clin Microbiol 51:3039–3046.

2. Forni D, Cagliani R, Clerici M, Sironi M. 2017. Molecular Evolution of Human Coronavirus Genomes. Trends Microbiol 25:35–48.

3. Dijkman R, Hoek L van der. 2009. Human Coronaviruses 229E and NL63: Close Yet Still So Far. J Formos Med Assoc 108:270–279.

4. Posthuma CC, Te Welthuis AJW, Snijder EJ. 2017. Nidovirus RNA polymerases: complex enzymes handling exceptional RNA genomes. Virus Res.

5. Milano F, Campbell AP, Guthrie KA, Kuypers J, Englund JA, Corey L, Boeckh M. 2010. Human rhinovirus and coronavirus detection among allogeneic hematopoietic stem cell transplantation recipients. Blood 115:2088–2094.

6. Gyurkocza B, Sandmaier BM. 2014. Conditioning regimens for hematopoietic cell transplantation: one size does not fit all. Blood 124:344–353.

7. Szikriszt B, Póti Á, Pipek O, Krzystanek M, Kanu N, Molnár J, Ribli D, Szeltner Z, Tusnády GE, Csabai I, Szallasi Z, Swanton C, Szüts D. 2016. A comprehensive survey of the mutagenic impact of common cancer cytotoxics. Genome Biol 17:99.

8. Godley LA, Larson RA. 2008. Therapy-related Myeloid Leukemia. Semin Oncol 35:418–429.

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

9. Kuypers J, Martin ET, Heugel J, Wright N, Morrow R, Englund JA. 2007. Clinical disease in children associated with newly described coronavirus subtypes. Pediatrics 119:e70-76.

10. Greninger AL, Zerr DM, Qin X, Adler AL, Sampoleo R, Kuypers JM, Englund JA, Jerome KR. 2017. Rapid Metagenomic Next-Generation Sequencing during an Investigation of Hospital-Acquired Human Parainfluenza Virus 3 Infections. J Clin Microbiol 55:177–182.

11. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinforma Oxf Engl 28:1647–1649.

12. Dijkman R, Jebbink MF, Wilbrink B, Pyrc K, Zaaijer HL, Minor PD, Franklin S, Berkhout B, Thiel V, van der Hoek L. 2006. Human coronavirus 229E encodes a single ORF4 protein between the spike and the envelope genes. Virol J 3:106.

13. Crossley BM, Mock RE, Callison SA, Hietala SK. 2012. Identification and Characterization of a Novel Alpaca Respiratory Coronavirus Most Closely Related to the Human Coronavirus 229E. Viruses 4:3689–3700.

14. Farsani SMJ, Dijkman R, Jebbink MF, Goossens H, Ieven M, Deijs M, Molenkamp R, van der Hoek L. 2012. The first complete genome sequences of clinical isolates of human coronavirus 229E. Virus Genes 45:433–439.

15. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:W244-248.

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

16. Zhang R, Wang K, Lv W, Yu W, Xie S, Xu K, Schwarz W, Xiong S, Sun B. 2014. The ORF4a protein of human coronavirus 229E functions as a viroporin that regulates viral production. Biochim Biophys Acta BBA - Biomembr 1838:1088–1095.

17. Povirk LF, Shuker DE. 1994. DNA damage and mutagenesis induced by nitrogen mustards. Mutat Res 318:205–226.

18. Hsu C-W, Chao M, Chen Y-C, Chang M-L, Huang S-F, Yeh C-T. 2015. Detection of hepatitis D virus RNA carrying large fragment deletions in patients with severe hepatitis B/D receiving oral antiviral therapy. J Med Virol 87:634–641.

19. Greninger AL, Chen EC, Sittler T, Scheinerman A, Roubinian N, Yu G, Kim E, Pillai DR, Guyard C, Mazzulli T, Isa P, Arias CF, Hackett J Jr, Schochetman G, Miller S, Tang P, Chiu CY. 2010. A Metagenomic Analysis of Pandemic Influenza A (2009 H1N1) Infection in Patients from North America. PLoS ONE 5:e13381.

20. Greninger AL, Bateman AC, Atienza EE, Wendt S, Makhsous N, Jerome KR, Cook L. 2016. Copy number heterogeneity of JC virus standards. J Clin Microbiol.

21. Greninger AL, Messacar K, Dunnebacke T, Naccache SN, Federman S, Bouquet J, Mirsky D, Nomura Y, Yagi S, Glaser C, Vollmer M, Press CA, Kleinschmidt-DeMasters BK, Klenschmidt-DeMasters BK, Dominguez SR, Chiu CY. 2015. Clinical metagenomic identification of Balamuthia mandrillaris encephalitis and assembly of the draft genome: the continuing case for reference genome sequencing. Genome Med 7:113.

22. Canuti M, van der Hoek L. 2014. Virus discovery: are we scientists or genome collectors? Trends Microbiol 22:229–231.

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure and Table Legends

Figure – Detection of a 548-nucleotide deletion in a HCoV 229E strain associated with myeloablation. A) Depiction of the case history surrounding myeloablation and coronavirus infection. A woman in her 40s with AML was found to be infected with a HCoV just before her HCT. Due to the patient's advanced disease, the decision was made to continue with the transplant. Four successive HCoV-positive nasal swab specimens were available from the patient; two from before treatment with myeloablative cyclophosphamide and total body irradiation and two from after. All days are reported relative to HCT. B) mNGS of three HCoV samples from the patient revealed the generation of a 548-nucleotide deletion in the ORF4a/4b gene. In addition to deleting 83% of the coding region from ORF4a/4b, the deletion resulted in a frame shift that resulted in a premature stop codon 10 amino acids before the reference genome stop codon. Primer binding sites are denoted in blue for confirmatory junction RT-PCR. C) Confirmatory junction RT-PCR of the four available HCoV-positive nasal swabs revealed the absence of the deletion before myeloablation and the presence of the deletion in specimens taken after myeloablation.

Supplemental Figure 1 – Deletion detected in Day -2 specimen. mNGS sequencing coverage is shown for Day -2 specimen. This specimen was collected on the last of three days of total body irradiation. The viral strain showed numerous areas of no coverage (highlighted in red) and PCR duplicates that could be consistent with low viral titer and PCR jackpotting or RNA degradation. 18 junctional reads that spanned the deletion in the Day 15 specimen were detected.

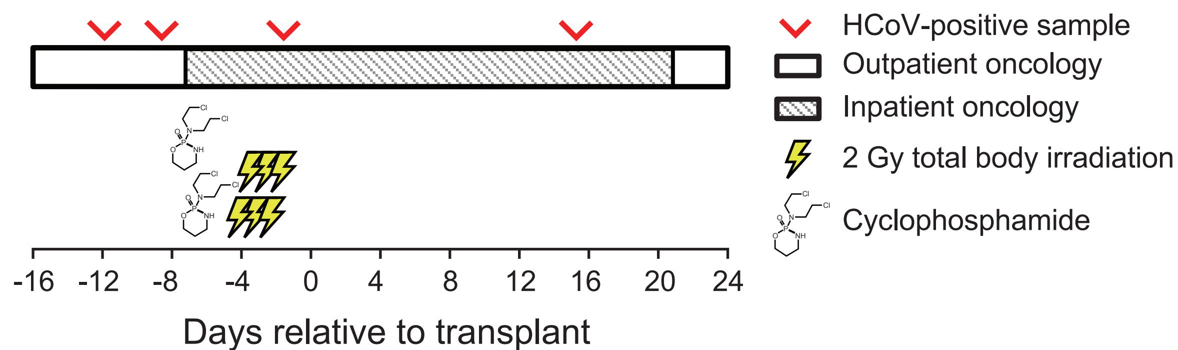
bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supplemental Figure 2 – The HCoV 229E strains sequenced in this study failed to grow in culture. None of the patient's strains of HCoV 229E grew in MRC5 cell culture as assessed by cytopathic effect or qRT-PCR. The patient's nasal swabs in viral transport media were inoculated onto MRC5 culture tubes for 30 minutes and supernatants were taken every day for 9 days. The qRT-PCR cycle threshold for each supernatant is depicted along with the original inoculum Ct. A HCoV 229E ATCC strain served as a positive control.

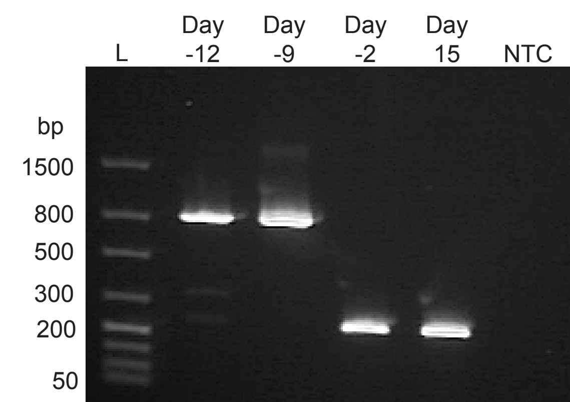
Table – Specimens sequenced in this study. Sequencing reads to HCoV 229E and total reads for each specimen are indicated, along with qRT-PCR cycle threshold and accession number of deposited genomes.

aCC-BY-NC 4.0 International license.

A



C



B

145X
27X average
Day -12
70X
16X average
Day -9
397X
68X average
Day 15
replicase polyprotein 1a
replicase polyprotein 1ab
replicase polyprotein 1ab
spike protein
N pro
124X
MALGLFTLQFVS--DVYTCFVNTQLLE*
nucleotide 23,910
24,127
24,676
24,865
spike protein
ORF 4a
ORF 4b
envelope protein

bioRxiv preprint doi: https://doi.org/10.1101/113423; this version posted March 3, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Day	HCoV Ct	HCoV reads	Total reads	Strain name	Accession
-12	31.5	7,257	4,395,609	SC379	KY621348
-9	33.5	3,733	30,166,065	SC399	KY674914
-2	32.7	5,064	11,185,156	SC475	N/A
15	31.3	11,990	1,594,894	SC677	KY369909

Table – Specimens sequenced in this study. Sequencing reads to HCoV 229E and total reads for each specimen are indicated, along with qRT-PCR cycle threshold and accession number of deposited genomes.