Emergence of the E484K Mutation in SARS-CoV-2 Lineage B.1.1.345 in Upstate New York

Emil Lesho[1*#], Brendan Corey[2*], Francois Lebreton[2], Ana C. Ong[2], Brett E. Swierczewski[2], Jason W. Bennett[2], Edward E. Walsh[1,3], and Patrick Mc Gann[2#]

[1] Infectious Diseases Unit, Rochester General Hospital, Rochester, New York, USA

[2] Multidrug-Resistant Organism Repository and Surveillance Network (MRSN), Walter Reed Army Institute of Research, Silver Spring, Maryland, USA

[3] University of Rochester School of Medicine, Rochester, NY, USA

* Emil Lesho and Brendan Corey contributed equally to this article

#Corresponding author

Patrick Mc Gann, PhD

Walter Reed Army Institute of Research

503 Robert Grant Avenue, 2A36, Silver Spring, MD 20910, USA

Email: patrick.t.mcgann4.civ@mail.mil

Abstract

A SARS-CoV-2 B.1.1.345 variant carrying the E484K mutation was detected in four patients with no apparent epidemiological association from a hospital network in upstate New York. Subsequent analysis identified an additional eleven B.1.1.345 variants from this region between December 2020 and February 2021.

Keywords: SARS-CoV-2, E484K mutation, Lineage B.1.1.345

Text

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the etiological agent of coronavirus disease 2019 (COVID-19), has caused a global pandemic that is reshaping society. As part of an unprecedented global response from the scientific community the development and rapid deployment of multiple antibody-based countermeasures that target the viral spike protein has proceeded at an unprecedented pace.

Over the past six months, the emergence of three variants with mutations in the spike protein has raised serious concerns about the durability of the current suite of vaccines and immunotherapies. The variants, which are colloquially referred to as the United Kingdom (B.1.1.7 using the Pangolin nomenclature [1]), South African (B.1.351), and Brazilian (P.1) variants have been increasingly identified worldwide. The first variant (B.1.1.7) was identified in September 2020 and carries a N501Y mutation in the receptor-binding domain (RBD) of the spike protein [2]. Critically, recent studies have suggested this mutation potentially increases transmission and virulence [3, 4]. Shortly thereafter, this same mutation was identified in B.1.351 and P.1 [5, 6]. Both variants also carried an additional mutation (E484K) in the spike that can increase resistance to neutralization by many monoclonal antibodies (mAb), while most convalescent sera and mRNA vaccine-induced immune sera show reduced inhibitory activity [7, 8]. Notably, interim results from the Novavax adjuvented spike protein nanoparticle COVID-19 vaccine trial indicated an efficacy of 85.6% against B.1.1.7 infection (versus 95.6% for the original strain), but just 49.4% to 60% in a South Africa Phase 2b trial in which the B.1.351 strain was detected in the majority of COVID-19 events [9].

In the United States, there have been 8,337 cases with the B.1.1.7 variant identified in 51 jurisdictions, 266 cases with B.1.351 (29 jurisdictions), and 79 cases with P.1 (19 jurisdictions) as of this writing (https://www.cdc.gov/coronavirus/2019-

ncov/transmission/variant-cases.html; Last accessed March 26th 2021). In addition, several regional variants have been identified, such as the 20C-US lineage in the Midwest [10] and the B.1.2 lineage carrying the Q677P mutation in New Mexico and Louisiana [11]. However, these lack the E484K and N501Y mutations. In contrast, the recently described B.1.536 variant in New York City carries E484K [12]. Notably, this variant has been spreading at an alarming rate since November and by February 15th 2021 accounted for 12.3% of all cases in this catchment area. Similar to previous studies with variants carrying E484K, antibody neutralization by mAbs REGN10933, CB6, and LY-CoV555 are either impaired or abolished with this variant [12]. Furthermore, the neutralizing activities of convalescent plasma or post-vaccination sera were lower by 7.7-fold and 3.4-fold, respectively [12]. These data illustrate the alarming rise of this mutation, which has now been detected in 122 of 1,242 active lineages (as of this writing), with various degrees of frequency (https://outbreak.info/; last accessed March 26th, 2021).

In June 2020, a pilot study to investigate SARS-CoV-2 was initiated between the Multi-drug resistant organism Repository and Surveillance Network (MRSN) and the Rochester Regional Health System; a network consisting of 8 acute care hospitals, 9 urgent care centers, and 6 long-term care facilities in St Lawrence County and the 9-county Finger Lakes Region of NY. Every hospital-based laboratory in the network is capable of testing for COVID-19 infection and to date 438,107 samples have been tested with 25,468 distinct patients testing positive. For this study, samples were provided by five hospitals in the Finger Lakes region. RNA is extracted from SARS-CoV-2-positive nasopharyngeal swabs and transferred to the MRSN for sequencing (See Methods). To date, >150 samples spanning April 2020 to February 2021 have been analyzed, with 14 different Pangolin lineages identified (manuscript in preparation). Though notable mutations were found in the spike

proteins from these earlier samples, including D614G, S673T, Q677H, N679K, and P681H, the E484K and N501Y mutations were not identified.

The most recent shipment consisted of 20 samples from 19 patients collected between January 27th and February 7th, 2021. Upon analysis, the E484K mutation was detected in five samples collected from four patients, one male and three female, between January 27th and February 4th 2021 (Supplemental Table S1). The average age was 85 years (range 74-92) and, despite advanced age and multiple co-morbidities, all recovered. Three patients had mild courses with no specific anti-SARS-CoV-2 treatments. One patient received Remdesivir and Dexamethasone and did not require intensive care or endotracheal intubation. Notably, three patients lived in separate assisted living facilities with no known connections and the fourth resided at home. All patients reported no significant travel in the preceding 6 months and none had received a SARS-CoV-2 vaccine.

The five strains were initially assigned to Pangolin Lineage B.1.1.220 (NextStrain clade 20B; **Figure 1A**) along with an additional 124 SARS-CoV-2 genomes from New York State, including 12 that also carried the E484K mutation (Figure 1B, supplemental Table 1) . However, subsequent refinement by Pangolin reclassified all 129 genomes into 4 distinct lineages (B.1.1, B.1.1.116, B.1.1.345, and B.1.1.486; Figure 1B and Supplemental Table 1) and all samples from Upstate New York carrying the E484K mutation were assigned to B.1.1.345 (Figure 1B). In addition to the E484K mutation, B.1.1.345 carried 4 notable amino acid mutations when compared to the Wuhan-Hu-1 SARS-CoV-2 reference genome (NCBI GenBank Acc. MN908947), namely R203K, G204R, P314L and D614G. This latter mutation has become the most prevalent mutation world-wide and studies suggest it increases viral transmissibility and results in higher viral loads [13].

To better understand the relationship between the sixteen B.1.1.345 samples we performed a whole genome high-resolution single nucleotide polymorphism (SNP) based

comparison (Figure 1C;). Of the five samples identified in our network, the two samples from the same patient were genetically identical (0 SNPs) and the remaining three samples were separated by 1-4 SNPs (Figure 1C). The remaining eleven B.1.1.345 genomes were deposited by the Wadsworth Center (Albany, NY), a sentinel institute of the New York Department of Health, and were separated from our samples by 0-7 SNPs (Figure 1D). To date, B.1.1.345 was only been identified in the Finger Lakes region (Figure 1D), with the earliest sample collected on December 31st, 2020 (Supplemental Table 1; Figure 1D). In addition to B.1.1.345, a further 110 genomes previously identified as B.1.1.220 were reclassified as B.1.1.486 (Figure 1B, Supplemental Table 1). Ninety (82%) samples were collected from New York City and adjacent counties, with the remaining being collected across New York State. Interestingly, one sample (1098129) collected in New York city on February 18th, 2021 carried the E484K mutation, indicating that the E484K mutation has arisen independently in B.1.1.486 in the New York Region.

To the best of our knowledge, this is the first report of a Lineage B.1.1.345 bearing the important E484K mutation. It is notable that no B.1.1.345 variants were detected in our surveillance effort prior to December 2020 when it suddenly comprised 25 % of the late January and early February samples. The observed mutational convergence within the relatively small subset of Lineage B.1.1.345 genomes from New York State further indicates that codon S/484 is evolving under a strong degree of positive selection, consistent with studies showing the importance of this mutation in antibody evasion [4-8, 12]. Though the selective pressure on this position is likely multi-factorial, the expanded use of antibody therapy and increased vaccinations in the New York region may be playing a role in the emergence of these variants in the region. Notably, as of this writing (March 2021), only sixteen B.1.1.345 genomes (all from NY) have been uploaded to GISAID; additional B.1.1.345 sequences from this region are required to better understand the current distribution

and apparent emergence of this concerning variant. It will also be important to assess how the E484K mutation affects protection by vaccines against this specific variant that lacks the other mutations in the B1.351 and P1 strains. As the vaccine roll-out gathers pace across the United States, it is imperative that every means necessary are employed to rapidly identify the emergence of variants-of-concern and immediate action taken to limit their spread.

NOTES

Acknowledgements

We are deeply indebted to the clinical and research laboratories of the Rochester Regional Health System and ACM Global Laboratories, especially Maria Formica, Julie Freed, Chalece Holland, Jean Campbell, and Julia Nary.

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official, or reflecting the views of the Department of the Army or the Department of Defense. Some authors (Bennett, Mc Gann, Swierczewski) are employees of the U.S. Government and this work was prepared as part of their official duties.

Funding

This study was partially funded by the Defense Health Program (DHP) Operations & Maintenance (O&M).

Conflict of Interest Statement

EW reports grants for RSV epidemiology and grant for SARS CoV-2 epidemiology from Merck, grants for RSV vaccine trial in adults and grant for viral epidemiology in hospitalized adults from Janssen, grants for RSV and SARS CoV-2 vaccine trials in adults from Pfizer, and grant to evaluate gene expression in respiratory tract infections in hospitalized persons, including SARS CoV-2 from NIH-NIAID; payment/honoraria from Janssen for serving as a speaker for symposium on RSV in elderly; serves as DSMB for pandemic influenza vaccine trial for GSK.

All other authors declare no conflicts of interest.

References:

1. Rambaut A, Holmes EC, O'Toole A, *et al*. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol **2020**; 5(11): 1403-7.
2. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med **2020**; 26(4): 450-2.
3. Davies NG, Abbott S, Barnard RC, *et al*. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science **2021**.
4. Volz E, Mishra S, Chand M, *et al*. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature **2021**.
5. Faria NR, Mellan TA, Whittaker C, *et al*. Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. Science **2021**; Apr 14.
6. Tegally H, Wilkinson E, Giovanetti M, *et al*. Emergence of a SARS-CoV-2 variant of concern with mutations in spike glycoprotein. Nature **2021**.
7. Chen RE, Zhang X, Case JB, et al. Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. Nat Med **2021**.
8. Ho D, Wang P, Liu L, *et al*. Increased Resistance of SARS-CoV-2 Variants B.1.351 and B.1.1.7 to Antibody Neutralization. Res Sq **2021**.
9. Novavax. Novavax COVID-19 Vaccine Demonstrates 89.3% Efficacy in UK Phase 3 Trial. Press release: https://ir.novavax.com/news-releases/news-release-details/novavax-covid-19-vaccine-demonstrates-893-efficacy-uk-phase-3. **2021**.
10. Adrian A. Pater MSB, Christopher L. Barkau, Katy N. Ovington, Ramadevi Chilamkurthy, Mansi Parasrampuria, Seth B. Eddington, Abadat O. Yinusa, Adam A. White, Paige E. Metz, Rourke J. Sylvain, Madison M. Hebert, Scott W. Benzinger,

Koushik Sinha, Keith T. Gagnon. Emergence and Evolution of a Prevalent New SARS-CoV-2 Variant in the United States. bioRxiv **2021**. Available from: https://doi.org/10.1101/2021.01.11.426287

11. Hodcroft EB, Domman DB, Snyder DJ, et al. Emergence in late 2020 of multiple lineages of SARS-CoV-2 Spike protein variants affecting amino acid position 677. medRxiv **2021**. Available from: https://doi.org/10.1101/2021.02.12.21251658

12. Annavajhala MK, Mohri H, Zucker JE, et al. A Novel SARS-CoV-2 Variant of Concern, B.1.526, Identified in New York. medRxiv **2021**. Available from: https://doi.org/10.1101/2021.02.23.21252259

13. Korber B, Fischer WM, Gnanakaran S, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell **2020**; 182(4): 812-27 e19.

Figure Legend

Figure 1. Emergence of the E484K mutation in B.1.1.345 SARS-CoV-2 isolates from Upstate New York. (A) Time-scaled, global phylogeny of 3,960 SARS-CoV-2 genomes sampled between Dec 2019 and Mar 2021 (modified from Nextstrain (https://nextstrain.org) dashboard). A full table of acknowledgements for the data included in this phylogeny is available in the supplement (Table S3). Major clades (Nextstrain nomenclature), variants (literature) and lineages (Pangolin) are indicated. Strains are colored based on the presence of a glutamic acid (E, teal) or lysine (K, 243 orange) at position 484 of the spike protein. **(B)** SNP-based, unrooted phylogram of B.1.1.345, B.1.1.486, B.1.1 and B.1.1.6 variants (colored shading) collected from New York State between October 2020 and March 2021 (n = 129). Empty circles indicate strains with Glutamic acid (E) at residue 484 of the spike protein while filled circles indicate those with Lysine (K) at that position. Genomes obtained from GISAID (www.gisaid.org) (Table S1 and S2) are shown with a red circle. Genomes newly described in this report are shown in blue (note that only four out of five circles are visible as two samples from a single patient were identical). **(C)** SNP-based, mid-point rooted maximum-likelihood phylogeny of sixteen monophyletic B.1.1.345 strains carrying the E484K mutation. Both comparator (Table S2) and newly sequenced isolates (bold font) are labeled using their GISAID Accession ID. Where known, the county (colored label) and collection date (colored circle) are indicated. For both panels B and C, the number of SNPs separating each strain can be inferred from the respective scale. **(D)** Map of the Finger Lakes region indicating the origin of fourteen B.1.1.345 strains carrying the E484K mutation (county of origin was not available for 1 isolate and the other is a duplicate sample from one patient). Inset map of the entire New York State and landmark cities are provided for geographical context.

Figure 1