Supplementary Appendix

A Novel SARS-CoV-2 Variant of Concern, B.1.526, Identified in New York

Medini K. Annavajhala[1*], Hiroshi Mohri[2*], Jason E. Zucker[1], Zizhang Sheng[2], Pengfei Wang[2], Angela Gomez-Simmonds[1], David D. Ho[1,2,3#], Anne-Catrin Uhlemann[1#]

[1] Division of Infectious Diseases, Department of Internal Medicine, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA

[2] Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA.

[3] Department of Microbiology and Immunology, Columbia University Irving Medical Center, New York, NY, USA.

* Medini K. Annavajhala and Hiroshi Mohri contributed equally to this work.
#David D. Ho and Anne-Catrin Uhlemann contributed equally to this work.

Address correspondence to au2110@cumc.columbia.edu or dh2994@cumc.columbia.edu.

Materials and Methods.

Clinical cohort. This observational study took place at an academic quaternary care center in New York City. Nasopharyngeal swabs obtained as part of routine clinical care were tested by the Clinical Microbiology laboratory, and positive specimens were transferred to the Columbia University Biobank for inactivation and storage.

Electronic health records data extracted for this analysis included demographics, laboratory results, admission, discharge, and transfer dates, current and historical international classification of disease (ICD 9 and 10) codes extracted from the clinical data warehouse. This study was reviewed and approved by the Columbia University Institutional Review Board (protocol number AAAT0123).

PCR screening. To make the screening procedure faster we skipped the RNA purification step and instead prepared RNA using the heat inactivation method (*21*). 50 µl of nasal swab sample in VTM solution was transferred into 96-well PCR plates, covered with an adhesive aluminum foil (VWR 60941-076) and incubated at 95□C for 5 min using the PCR instrument. After the centrifugation of the plate at >2,100 x g for 5 min, 5 µl of the supernatant from each sample, which contains viral RNA, was used for the SNP assay.

The SNP assay consists of four steps as follows: reverse transcription (RT) of viral RNA, pre-read of the SNP assay, real-time PCR and post-read of the SNP assay. 5 µl of RNA from the supernatant was added to 15 µl of the single step RT-qPCR reaction mix, which consists of 5 µl of TaqPath 1-step RT-qPCR Master Mix, CG (4x) (ThermoFisher Scientific), 500 nM of forward and reverse primers, 120 nM of VIC-MGB probe, 50 nM of FAM-MGB probe, 1/2000 volume of ROX Reference Dye (Invitrogen) as the final concentration, and nuclease-free water to adjust the total reaction volume of 20 µl. Each reaction plate included 8 control wells, $5x10^6$ and $5x10^3$ copies of WA-1 (wild type), UK variant and South African variant, which are generated by PCR to match the variant sequences, and 2 wells with water as no template controls (NTC).

The primer pairs and probes used are as follows. For the SNP assay for position **501**, a primer pair of 501.F: 5'- GGT TTT AAT TGT TAC TTT CCT TTA CA-3' and 501.R: 5'-AGT TCA AAA GAA AGT ACT ACT ACT CTG TAT G-3' were used with two TaqMan probes (ThermoFisher Scientific), one for wild type, VIC.N501MGB: [VIC]-AA

CCC ACT AAT GGT-MGBNFQ and the other for variant type, FAM.Y501MGB: [FAM]-AAC CCA CTT ATG GT-MGBNFQ. For position **484**, a primer pair of 484.F: 5'-AGA GAG ATA TTT CAA CTG AAA TCT ATCAGG-3'and 484.R: 5'-GAA ACC ATA TGA TTG TAA AGG AAA GTA AC-3' were used with two probes, one for wild type, VIC.E484MGB: [VIC]-ATG GTG TTG AAG GT-MGBNFQ and the other for variant type, FAM.K484MGB: [FAM]-ATG GTG TTA AAG GT-MGBNFQ.

The reaction plate was subjected to 1) reverse-transcription reaction (RT) at the condition at 25□C for 2 min, at 50□C for 15 min and a hold at 4□C, 2) SNP assay (pre-read) at 60□C for 30 sec, 3) real-time PCR at 95□C for 20 sec followed by 50 cycles of two step PCR, at 95□C for 3 sec and at 60□C for 30 sec with the fast 7500 mode, followed by 4) SNP assay (post-read) at 60□C for 30 sec using ABI 7500 Fast Dx Real-Time PCR Instrument with SDS Software (ThermoFisher Scientific). The genotype at each key position for each sample was determined by reading the component signal of the amplification and the allelic discrimination analysis software in the program.

Whole genome sequencing. Isolates with cycle threshold (Ct) values below 35 were selected for sequencing using the ARTIC v3 low-cost protocol. Briefly, RNA was extracted using the Qiagen RNeasy Mini kit or Zymo DNA/RNA Mini kit. Reverse transcription was performed using LunaScript RT SuperMix (NEB). Tiling PCR was performed on the cDNA, and amplicons were barcoded using the Oxford Nanopore Native Barcoding Expansion 96 kit. Pooled barcoded libraries were then sequenced on an Oxford Nanopore MinION sequencer using R9.4.1 flow cells. Sequencing runs were monitored in real-time using RAMPART to ensure sufficient genomic coverage with minimal runtime. Consensus sequence generation was performed using the ARTIC bioinformatics pipeline. Genomes were manually curated by visually inspecting sequencing alignment files for verification of key residues.

Phylogenetic analysis. We downloaded 140 publicly available genome sequences from GISAID representing a range of SARS-CoV-2 lineages (Fig 1C). Public sequences and genomes from this study were aligned against the Wuhan-Hu-1 reference sequence using Muscle. The resultant alignment was masked to avoid erroneous inclusion of SNPs due

to sequencing errors as proposed by De Maio et al. using BedTools. IQtree was then used to generate a phylogenetic reconstruction with 1,000 bootstrap replicates. Interactive Tree of Life (iTOL) was used to visualize all phylogenetic tree figures.

Neutralization studies. We assayed the neutralizing activity of 12 RBD mAbs against E484K and WT (D614G) pseudoviruses in Vero E6 cells as previously described (*16*). We examined four mAbs with emergency use authorization (CB6, REGN10987, REGN10933 and LY-CoV555) as well as convalescent plasma (collected in Spring of 2020) and vaccinee (Moderna and Pfizer) sera as previously described (*10*). Briefly, Vero E6 cells (ATCC) were seeded in 96-well plates (2×10^4 cells per well). Pseudoviruses were incubated with serial dilutions of the test samples in duplicate or triplicate for 30 min at 37 °C. The mixture was added to cultured cells and incubated for an additional 24 h. Luminescence was measured using a Britelite plus Reporter Gene Assay System (PerkinElmer), and IC_{50} was defined as the dilution at which the relative light units were reduced by 50% compared with the virus control wells (virus + cells) after subtraction of the background in the control groups with cells only. The IC_{50} values were calculated using nonlinear regression in GraphPad Prism 8.0.

TABLE S1. PATIENT CHARACTERISTICS

PATIENT CHARACTERISTICS [N (%)]	Viral genotype		
	E484K (n=78)	Wildtype (n=771)	p-value
MALE SEX	38 (48.7)	361 (47.1)	0.9
AGE [MEAN (SD)]	58.1 (19.6)	52.5 (24.2)	0.049
RACE/ETHNICITY			0.8
BLACK	7 (9.1)	85 (11.1)	
HISPANIC/LATINO	34 (44.2)	356 (46.4)	
WHITE	13 (16.9)	102 (13.3)	
OTHER	23 (29.9)	224 (29.2)	
ANY HOSPITAL PRESENTATION	67 (85.9)	546 (70.8)	0.007
HIGHEST LEVEL OF CARE			0.01
OUTPATIENT	9 (11.8)	213 (28.1)	
ED	27 (35.5)	202 (26.6)	
INPATIENT, NOT ICU	36 (47.4)	285 (37.5)	
ICU	4 (5.3)	59 (7.8)	
DURATION OF ADMISSION [MEAN (SD)]	4.08 (5.7)	4.82 (9.2)	0.5

Supplementary Figures

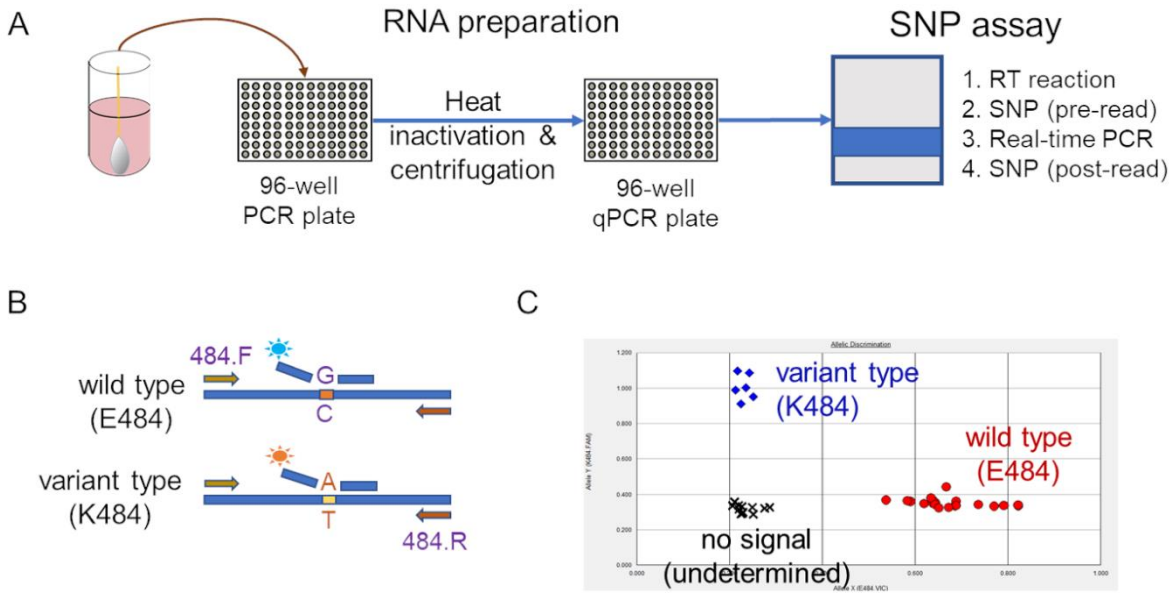

Figure S1. Rapid PCR-based screening assay protocol to identify samples harboring key substitutions. (A) Viral RNA is prepared by heat inactivation and centrifugation. The supernatant is then used for the SNP assay, which entails four steps: the reverse transcription (RT) reaction, SNP pre-read, real-time PCR, and SNP post-read. The whole protocol can be completed in two hours. (B) Genotype at targeted sites in COVID-19 viral RNA can be determined with two MGB probes, one for wild type (conjugated with VIC) and the other for variant type (conjugated with FAM). (C) Example signals for the variant type (K484; blue), the wild type (E484; red) and samples with no signal (black) are shown.

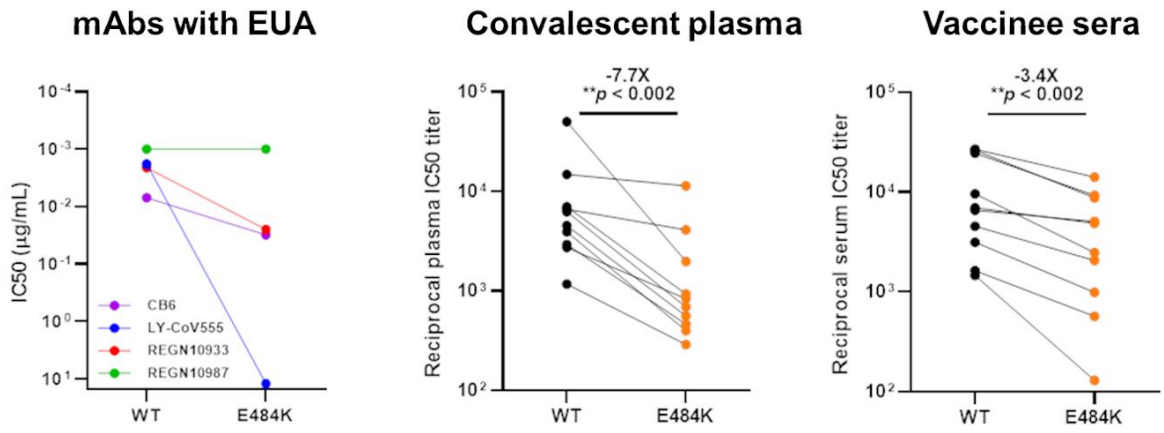

Figure S2. Neutralization studies of E484K pseudovirus compared to wildtype. We observed decreased neutralization of pseudovirus encoding S protein E484K with three monoclonal antibodies (mAbs) with emergency use authorization (EUA), REGN10933, CB6, LY-CoV555, but saw no effect on neutralizing activity of REGN10987. Neutralizing activities of convalescent plasma and vaccinee sera were lowered by 7.7-fold or 3.4-fold, respectively, against the E484K variant compared to wildtype pseudovirus.

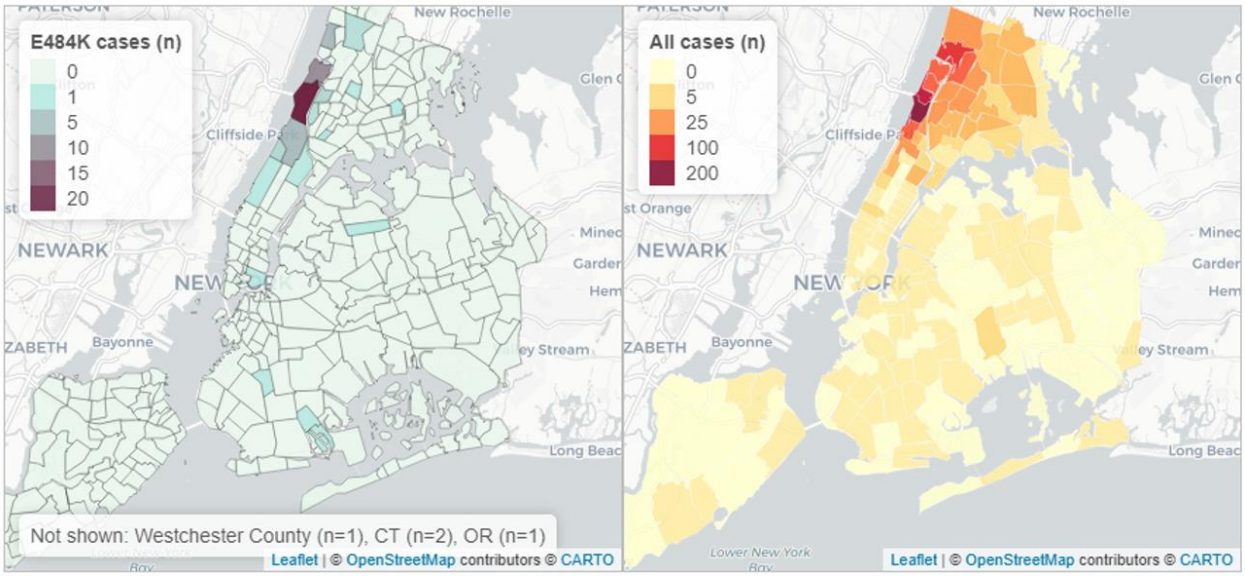

Figure S3. Geographic distribution of patients with E484K and wild-type SARS-CoV-2. (**A**) Patients with E484K isolates resided primarily in the neighborhoods surrounding our hospital center, although E484K samples were detected in patients from diverse neighborhoods across New York City. Four patients residing outside of the NYC-Yonkers region, in Westchester County (n=1), Connecticut (n=2), and Oregon (n=1), are not represented in this map. (**B**) Distribution of patients with both wild type and variant virus strains who tested positive for SARS-CoV-2 at our medical center between November 2020 and mid-February 2021.