medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Dominance of Alpha and Iota variants in SARS-CoV-2 vaccine breakthrough infections in New York City

Author Affiliations:

Ralf Duerr, MD, PhD[1*], Dacia Dimartino[2], PhD, Christian Marier[2], Paul Zappile[2], Guiqing Wang, MD, PhD[3], Jennifer Lighter, MD[4], Brian Elbel, PhD[5], Andrea Troxel, PhD[5] and Adriana Heguy, PhD [2,3*]

1. Department of Microbiology, NYU Grossman School of Medicine
2. Genome Technology Center, Office of Science and Research, NYU Langone Health
3. Department of Pathology, NYU Grossman School of Medicine
4. Department of Pediatric Infectious Diseases, NYU Grossman School of Medicine
5. Department of Population Health, NYU Grossman School of Medicine

* Co-corresponding authors

Conflict of Interest Statement: The authors declare that no conflict of interest exists

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

The efficacy of COVID-19 mRNA vaccines is high, but breakthrough infections still occur. We compared the SARS-CoV-2 genomes of 67 breakthrough cases after full vaccination with BNT162b2 (Pfizer/BioNTech), mRNA-1273 (Moderna), or JNJ-78436735 (Janssen) to unvaccinated controls (February-April 2021) in metropolitan New York, including their phylogenetic relationship, distribution of variants, and full spike mutation profiles. Their median age was 45 years; seven required hospitalization and one died. Most breakthrough infections (54/67) occurred with B.1.1.7 (Alpha) or B.1.526 (Iota). Among the 7 hospitalized cases, 5 were infected with B.1.1.7, including 1 death. Both unmatched and matched statistical analyses considering age, sex, vaccine type, and study month as covariates supported the null hypothesis of equal variant distributions between vaccinated and unvaccinated in chi-squared and McNemar tests ($p>0.1$) highlighting a high vaccine efficacy against B.1.1.7 and B.1.526. There was no clear association among breakthroughs between type of vaccine received and variant. In the vaccinated group, spike mutations in the N-terminal domain and receptor-binding domain that have been associated with immune evasion were overrepresented. The evolving dynamic of SARS-Co-V2 variants requires broad genomic analyses of breakthrough infections to provide real-life information on immune escape mediated by circulating variants and their spike mutations.

INTRODUCTION

The novel betacoronavirus SARS-CoV-2 arose as a new human pathogen at the end of 2019, and rapidly spread to every corner of the globe, causing a pandemic of enormous proportions, with 179 million cumulative infections and almost 4 million deaths counted worldwide as of mid-June 2021 [1]. As soon as the causative agent was identified and sequenced [2], massive efforts to develop vaccines were initiated and tested simultaneously in multiple clinical trials. These efforts led to the rapid deployment of several of these vaccines by December 2020, less than a year after the first SARS-CoV-2 viral sequence had been determined. In the United States, vaccination started in December 2020, using two novel mRNA-based vaccines, BNT162b2 (Pfizer/BioNTech), and mRNA-1273 (Moderna), both using the SARS-CoV-2 spike mRNA sequence of the original Wuhan variant. The Janssen COVID-19 vaccine, JNJ-78436735, was also deployed shortly after the mRNA vaccines, and these vaccines were shown to have high efficacy in clinical trials [3-5]. With millions of people vaccinated in multiple countries, they have proven to be highly effective in the real world [6-12], although a very small percentage of breakthrough infections have occurred, as expected [13-21]. Contemporaneously with the vaccination efforts in several countries, new SARS-CoV-2 variants emerged, four of which were designated by the WHO as variants of concern (VOC), B.1.1.7 or Alpha, B.1.351 or Beta, P.1 or Gamma, and B.1.617.2 or Delta, which arose originally in the UK, South Africa, Brazil, and India, respectively [22, 23]. VOCs are classified as such if they are more transmissible, cause more severe disease or a significant reduction in neutralization by antibodies generated via previous infection or after vaccination. In addition, there are variants of interest (VOI), which carry mutations that have been associated with changes to receptor binding, reduced neutralization by anti-SARS-CoV-2 antibodies, or may increase transmissibility or severity. One of these variants is B.1.526 (Iota),

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

which arose in New York City in late December 2020 [24, 25]. Although it is not yet clear that any particular VOC and VOI is associated with vaccine breakthrough, data from Israel suggest that there might be higher vaccine breakthrough rates with the B.1.351 variant [17]. In addition, it is possible that certain amino acid mutations in spike, irrespective of affiliation to a specific VOC or VOI, are critical for vaccine breakthrough, but this has only been rudimentarily studied [18]. There is scarcity of data about determinants of vaccine breakthrough. The few reported studies have included breakthrough cases after first or second immunization, however, breakthrough cases after full vaccination remained moderate or low [13-21]. Thus, comprehensive studies with site-specific mutation analyses are needed on a larger number of fully vaccinated individuals from different geographic regions. Here, we added such data from the NY metropolitan area. We carried out full SARS-CoV-2 genome sequencing of SARS-CoV-2-positive individuals, 14 days after their completed vaccination series with mostly Pfizer and Moderna vaccines in the multi-center NYU Langone Health System. Analyses included statistical comparison of variant distribution as well as mutation rates at every residue in spike between vaccinated and unvaccinated individuals.

METHODS

Study design and sample collection. The design is an observational case-control study of SARS-CoV-2 vaccine breakthrough infections in the NYU Langone Health system, a large healthcare system in the New York City metro area, with primary care hospitals located in Manhattan (New York County), Brooklyn (Kings County) and Nassau County (Long Island). The case group consisted of individuals who tested positive by real-time RT-PCR for SARS-CoV-2 RNA regardless of cycle threshold (Ct), any time after 14 days of inoculation with the second dose of BNT162b2 (Pfizer/BioNTech) or mRNA-1273 (Moderna) vaccines, or with the single dose COVID Janssen vaccine, according to our electronic health records (EHR). The control group consisted of full-genome sequenced SARS-CoV-2 positive cases in our health system who were randomly selected for SARS-CoV-2 genomic surveillance, had Ct ≤30, and were collected in the same time period as the breakthrough infections. Nasopharyngeal swabs were sampled from individuals suspected to have an infection with SARS-CoV-2 as part of clinical diagnostics or hospital admittance. Samples were collected in 3 mL viral transport medium (VTM; Copan universal transport medium or equivalent). Clinical testing was performed using various FDA emergency use authorization (EUA)–approved platforms for detection of SARS-CoV-2, i.e., the Roche Cobas 6800 SARS-CoV-2 (90% of the samples in this study) and Cepheid Xpert SARS-CoV-2 or SARS-CoV-2/Flu/RSV assays.

RNA extraction, cDNA synthesis, library preparation and sequencing

RNA was extracted from 400 μl of each nasopharyngeal swab specimens using the MagMAX™ Viral/Pathogen Nucleic Acid Isolation Kit on the KingFisher flex system (Thermo Fisher Scientific) following the manufacturer's instructions. Total RNA (11 μl) was converted to first strand cDNA by random priming using the Superscript IV first-strand synthesis system (Invitrogen, ref#

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

180901050). Libraries were prepared using Swift Normalase Amplicon SARS-CoV-2 Panel (SNAP) and SARS-CoV-2 additional Genome Coverage Panel (Cat# SN-5X296 core kit, 96rxn), using 10 µl of first strand cDNA, following the manufacturer's instructions [26]. Final libraries were run on Agilent Tapestation 2200 with high sensitivity DNA Screentape to verify the amplicon size of about 450 bp. Normalized pools were run on the Illumina NovaSeq 6000 system with the SP 300 cycle flow cell. Run metrics were paired-end 150 cycles with dual indexing reads. Typically, two pools representing two full 96 well plates (192 samples) were sequenced on each SP300 NovaSeq flow cell.

Sequenced read processing

Sequencing reads were demultiplexed using the Illumina bcl2fastq2 Conversion Software v2.20 and adapters and low-quality bases were trimmed with Trimmomatic v0.36 [27]. BWA v0.7.17 [28] was utilized for mapping reads to the SARS-CoV-2 reference genome (NC_045512.2, wuhCor1) and mapped reads were soft-clipped to remove SNAP tiled primer sequences using Primerclip v0.3.8 [29]. BCFtools v1.9 [30] was used to call mutations and assemble consensus sequences, which were then assigned phylogenetic lineage designations according to PANGO nomenclature [31]. Sequences that did not yield a near-complete viral genome (<23,000bp, <4000x coverage) were discarded from further analysis.

Phylogenetic analyses

SARS-CoV-2 full genome sequences were aligned using Mafft v.7 [32]. The alignment was cropped to base pairs 202-29657 according to the Wuhan-Hu-1 reference to remove N- and C-terminal regions with unassigned base pairs. Maximum likelihood IQ trees were performed using the IQ-TREE XSEDE tool, multicore version 2.1.2, on the Cipres Science gateway v.3.3 [33]. GTR+I+G was chosen as the best-fit substitution model. Support values were generated with 1,000 bootstrap replicates and the ultrafast bootstrapping method. Phylogenetic trees were visualized in Interactive Tree Of Life (iTOL) v.6 [34]. The tree was constructed using 67 vaccine breakthrough and 1,187 unvaccinated control SARS-CoV-2 sequences from our NYU cohort (greater NYC area) together with 3,802 global reference sequences for a total of 5,056 SARS-CoV-2 genomic sequences. The reference sequences were retrieved from a Nextstrain build with North America-focused global subsampling [4] and included 1,361 other US sequences.

Mutation analysis

Spike amino acid counts and calculations of site-specific mutation frequencies compared to Wuhan-Hu-1 as reference were done on MAFFT-aligned SARS-CoV-2 sequences using R[35] and R Studio[36] using scripts based on the seqinr and tidyverse (dplyr, lubridate, magrittr) packages. The calculations exclusively considered residues that were accurately covered by sequencing, i.e., non-ambiguous characters and gaps. Fisher exact tests and multiplicity corrections (Benjamini-Hochberg) were done in Program R/RStudio. Multiplicity corrected P values (q) <0.05 were considered significant. Highlighter analyses were performed on MAFFT-aligned SARS-CoV-2 amino acid sequences of the spike genomic region. SARS-CoV-2 vaccine breakthrough sequences were compared to Wuhan-Hu-1 as master using the Highlighter tool provided by the Los Alamos HIV sequence database [37].

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Statistical analysis

To address confounding and other sources of bias arising from the use of observational data, we estimated a propensity score for the likelihood of full vaccination, and matched vaccinated to unvaccinated patients, including age, sex, county of residence, and study month (February, March, April 2021) as covariates [38]. Propensity-score matching was implemented using the nearest neighbor strategy using a 1:1 ratio without replacement, with the MatchIt algorithm in RStudio version 1.4.1106 [39]. Before and after matching, we evaluated the presence of three Pango lineages, B.1.1.7, B.1.526 and P.1 compared to all others lineages combined. On unmatched data, we calculated a Pearson chi-squared test statistic; on matched data we calculated McNemar's test. The tables used for these calculations are in Supplementary Material (**Supplementary Tables 1 and 2**). Comparisons of Ct values and mutation counts between groups were made using non-parametric Mann-Whitney tests or Kruskal-Wallis tests with Dunn's multiplicity correction.

Study Approval. This study was approved by the NYU Langone Health Institutional Review Board, protocol number i21-00493.

RESULTS

Alpha and Iota variants dominate vaccine breakthrough infections in metropolitan New York

A total of 126,367 fully vaccinated individuals were recorded in our electronic health records by April 30th, 2021, of whom the majority (123,511, 98%) were vaccinated with mRNA vaccines (Pfizer/BioNTech = 103,166 and Moderna = 20,345), and the rest (2,856) with the adenovirus-based Janssen COVID-19 vaccine, administered as single dose. We recorded 86 cases of vaccine breakthrough infection (69x Pfizer/BioNTech, 16x Moderna, and 1x Janssen) between February 1st and April 30th of 2021, representing 1.2% of the total SARS-CoV-2 positive cases in our healthcare system and 0.07% of the fully vaccinated population in our medical records. Out of the 86 cases, 67 cases (80%) yielded full SARS-CoV2 genomes (54x Pfizer/BioNTech, 12x Moderna, and 1x Janssen) that passed quality control (QC) and allowed us to determine the Pango lineage and mutations across the viral genome, including the spike gene. As expected, the 19 excluded breakthrough cases with low genome coverage and failed QC had significantly higher Ct values (median: 35, range: 31-42) than the 67 samples with full viral genome coverage ($P<0.0001$, Mann Whitney test), which had Ct values below 34 (median: 24, range: 13-34) (**Figure 1A**).

Although B.1.1.7 (Alpha) infection has been associated with overall higher viral load/lower Ct values [40, 41], the Ct values in our recorded breakthrough infections were not significantly different for B.1.1.7 compared to B.1.526 (Iota) or all other variants (**Figure 1B**).

Among the 67 COVID-19 cases post-vaccination with adequate SARS-CoV-2 genome coverage, the median age was 45 years, 33 were male and 24 were female. Seven required hospitalization for COVID-19, among whom there was one death in an elderly patient with multiple comorbidities who already was on in-home oxygen previous to post-vaccination COVID-19 infection and had a lengthy stay at the ICU (**Table 1, Supplementary Table 3**).

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The distribution of PANGO lineages in the 67 vaccine breakthroughs was as follows: 26 (38.8%) had B.1.1.7, 28 (41.8%) had B.1.526 (including sublineages B.1.526.1 and B.1.526.2), 1 (0.2%) had P.1, and 12 (17.9%) had other variants (**Supplementary Figure 1**). Among the 1,187 sequences from the group of unvaccinated patients, 345 (29.1%) had B.1.1.7, 487 (41.0%) had B.1.526 (including sublineages B.1.526.1, B.1.526.2, and B.1.526.3), 12 (1.0%) had P.1, and 343 (28.9%) had other variants. Of the seven hospitalizations, five were infected with B.1.1.7, including the fatal case, and two with B.1.526.

VOCs and VOIs are equally distributed in vaccinated and unvaccinated infections.
To compare vaccinated breakthrough infections with unvaccinated controls statistically, we included 1,187 unvaccinated individuals from the same study cohort who became SARS-CoV-2-infected in the same study months (February-April 2021) as the vaccine breakthrough cases. A chi-squared test for rejecting the null hypothesis of equal Pango lineage distributions (B.1.1.7, B.1.526, and other variants combined) between vaccinated and unvaccinated patients resulted in a p-value of 0.18.

To address confounding and other sources of bias arising from the use of observational data, we estimated a propensity score for the likelihood of full vaccination [38], and successfully matched all 67 vaccinated to unvaccinated patients, including age, sex, county of residence, and study month (February, March, and April 2021) as covariates. The standardized mean difference between the matched pairs was 0.0125, reduced by 99.7 percent from 0.590 prior to matching.

Supplementary Table 1 shows the distribution of variants in the matched pairs. McNemar's test of the null hypothesis of equal distributions between vaccinated and unvaccinated patients, assessing the three VOC/VOI separately, could not be calculated due to sparse data. When we collapsed the table to reflect all VOCs/VOIs compared to other variants, McNemar's test resulted in a p-value of 0.82 (**Supplementary Table 2**). Thus, vaccinated and unvaccinated individuals in the metropolitan New York area are similarly affected by the regionally circulating VOCs and VOIs. In addition, there is no clear association among vaccinated patients between type of vaccine received and Pango lineage (chi-squared test, p=0.36).

Widespread phylogenetic dispersal of vaccine breakthrough sequences among unvaccinated controls
As a way to ascertain potential bias in our sampling, we carried out a phylogenetic analysis of our 67 breakthrough sequences in the context of 1,187 unvaccinated SARS-CoV-2-positive controls (selected randomly as part of our greater New York catchment area genomic surveillance) together with sub-sampled sequences from the United States as well as globally, the latter two groups based on Nextstrain builds of GISAID sequences (**Figure 2**). The main variants circulating in the New York area between the months of February and April 2021 were B.1.1.7 and B.1.526 (purple). Accordingly, our vaccine breakthrough samples (orange branch symbols and gray rays) mostly engaged B.1.1.7 and B.1.526 clades and were interspersed among the unvaccinated controls as well as other United States sequences. There was no evidence of extensive clustering that might indicate onward transmissions or transmission

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

chains of vaccine breakthrough infections. Instead, they were widely distributed and appeared to mostly involve independent clusters of infections.

Enrichment of NTD deletions and RBD escape mutations in vaccine breakthrough compared to unvaccinated control sequences.

To screen whether vaccine breakthrough preferentially occurred with distinct vaccine escape mutations, we performed a comparative analysis of spike mutations between case and control groups. In our aligned data set of 67 vaccine breakthrough and 1,187 unvaccinated control sequences, spike mutations (compared to Wuhan-Hu-1) occurred at 233 amino acid residues. While most of these sites were more frequently mutated in control cases (191 sites), and one site (D614) being equally mutated in both groups (100% D614G), 41 sites exhibited increased mutation rates in the vaccine breakthrough group (**Figure 3A**). Interestingly, the degree of enrichment (Δ mut) was higher at the 41 breakthrough-enriched sites compared to the 191 sites enriched in controls. Contributing factors presumably included random mutations in the absence of immune pressure in controls, adaptive selection of immune escape mutations in vaccine breakthroughs, but also uneven case numbers in both groups. When we disregarded unique mutations per data set in our calculations, the mutation analysis yielded 21 distinct spike sites with enriched mutations in breakthrough infections (**Figure 3A, Supplementary Table 4**). Although individual sites did not achieve significance in Fisher's exact tests, the array of sieved mutation sites drew a striking pattern of N-terminal domain (NTD) deletions (ΔV69-H70 and ΔY144), receptor-binding domain (RBD) mutations (K417, S477, E484, and N501), a mutation near the furin binding site (P681X), and also C-terminal mutations (T1027X and D1118X), all of which have been associated with enhanced immune evasion or ACE2 receptor binding, and/or recurring mutations in emerging VOCs/VOIs [22, 42].

DISCUSSION

Our data from a large metropolitan healthcare system in the greater New York City area underline a high vaccine efficacy in fully vaccinated individuals, 14 days or longer post-last dose of vaccination. Efficacy is maintained against several circulating variants including VOCs and VOIs. Compared to the large number of SARS-CoV-2 infections among unvaccinated individuals, the recorded breakthrough cases between February and April 2021 (n=67) remained at~1% of total infections, with the caveat that both breakthrough cases as well as unvaccinated controls were not exhaustively screened and covered.

Despite the overall effectiveness of vaccination, our full spike mutation analysis revealed a broad set of spike mutations (n=21) to be elevated in the vaccine breakthrough group. It indicates that adaptive selection is in progress that may subsequently come into full effect. At this point, the breakthrough cases and differences in mutation rates between case and control groups are still too low to draw meaningful conclusions. However, the overrepresentation of functionally important spike mutations including NTD deletions ΔH69-V70 and ΔY144 together with RBD mutations K417X, S477X, E484X, and N501X as well as P681X (near the S1/S2 interface) in the absence of a clear elevation of VOCs or VOIs may indicate a starting sieve effect at individual or combinations of functional mutations. Spike mutations and deletions

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

reported to confer neutralization escape *in vitro*[23, 43, 44] might thus be responsible for a sieve effect in a real-life situation, i.e., among vaccinated individuals.

During the time of our sample and data collection, there were two major variants arising in the New York City metro area constituting about two thirds of all cases, B.1.1.7, which was first reported in the UK [45], and B.1.526, which arose in New York City around December of 2020 [24, 25]. B.1.1.7 was deemed a variant of concern by the WHO and CDC and was associated with higher viral loads in infected individuals [40, 41], enhanced epidemiological spread [45-47], an extended window of acute infection [48], less effective innate and adaptive immune clearance [49], and increased death rates in elderly patients and/or individuals with comorbidities [50-53].
The B.1.1.7 variant, e.g., through the RBD N501Y mutation but also through NTD deletions, acquired an enhanced affinity to ACE2, higher infectivity and virulence [54-57], while maintaining sensitivity to neutralization though with slightly impaired nAb titers [54, 58-60]. The B.1.526 variant and its derivatives were designated VOI because of the presence of RBD mutations such as E484K or S477N that favor immune evasion [23, 25].

To study whether the more transmissible B.1.1.7 or B.1.526 because of its critical RBD mutations and prevalence/place of origin in our city were overrepresented in the breakthrough cases, we performed a comparative statistical analysis. Extending a few other studies reported so far [13-21], we focused on a strict threshold of 14 days post last dose of vaccination, and we performed both unmatched and matched analyses side-by-side. We adjusted the confounding factors sex, age, study month, and residence area, though we are aware that other confounding factors such as biased sampling or behavioral differences between groups might also play a role.
Our finding that B.1.1.7 and B.1.526 are not preferentially represented in the vaccine breakthrough group but distributed at similar proportions as other variants in case and control groups suggests that the studied mRNA vaccines are comparably effective against these predominant variants in the NYC area. This is in agreement with recent data from Israel evaluating B.1.1.7 in this context [17]. A sieve effect of the B.1.526 variant has not been studied in this detail, except for two studies with small sample sizes of n=2 and n=11 that did not allow strong conclusions [14, 19]. A recent CDC study reported a total number of 10,262 SARS-CoV-2 vaccine breakthrough infections across the USA as of end-April 2021 causing hospitalization in 10% of cases (n=995) [20], a rate identical to our study.

Interestingly, the majority of our breakthrough infections had low Ct values (Ct <35), implying a moderate to high viral load in these cases, at least in the nasopharynx from where our samples were collected. These high viral loads suggest the possibility of transmission to others [61]. Although our phylogenetic analysis does not provide evidence of transmission to others from our breakthrough cases at this time, this should be expected with the growing number of breakthrough cases in the next months. This may have significant epidemiologic and clinical significance if transmissible breakthrough infections carry specific spike mutations associated with immune evasion.

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In conclusion, our data indicate that vaccine breakthrough in fully mRNA-vaccinated individuals is not more frequent with the VOC Alpha or the NY local VOI Iota, which underlines the high efficacy of mRNA COVID-19 vaccines against different variants. The Alpha variant appears overrepresented in the hospitalized group, though the low case numbers of hospitalizations do not allow firm conclusions. The increased presence of mutations in key regions of the spike protein in the vaccine breakthrough group is of potential concern and requires continued monitoring. Genomic surveillance of vaccine breakthrough cases should be carried out on a broader scale throughout the United States and the entire word to achieve higher case numbers and the inclusion of different VOCs and VOIs.

Author contributions:

Ralf Duerr, MD, PhD designed the study, analyzed data and wrote the manuscript; Dacia Dimartino, PhD generated sequencing data; Christian Marier analyzed genomic data; Paul Zappile generated sequencing data; Guiqing Wang, MD, PhD collected samples and performed clinical SARS-CoV-2 detection; Jennifer Lighter, MD reviewed clinical information; Brian Elbel, PhD monitored epidemiological data and generated lists of breakthrough infections; Andrea Troxel, PhD performed statistical analyses; Adriana Heguy, PhD designed the study, generated genomic data and wrote the manuscript. All authors reviewed and edited the manuscript.

Acknowledgements. The authors wish to thank Vincent Setang and NYU Langone Health DataCore for support extracting the data for this study from clinical databases, Dr. Maria Agüero-Rosenfeld and the clinical laboratory technicians for assistance in testing, saving and retrieving specimens, especially Joanna Fung for her assistance with RNA extraction project. We also thank Dr. Joan Cangiarella for her continuous support of genomic surveillance for SARS-CoV-2 at NYULH, including providing institutional funding for this study. We are grateful to the authors and submitting laboratories who deposited data in GISAID, in particular to those whose sequences we used to create the phylogenetic tree. These are all named in Supplemental Table 5. The NYULH Genome Technology Center is partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center.

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Tables.

Table 1. Vaccine breakthrough study population

Sample ID (arbitrary #)	Days after completion of vaccination series	Vaccine	Hospitalization	Sex	Age range	Ct	Pango Lineage
P21-2388	31	Pfizer	No	M	31-40	18	A.2.5.2
P21-0524	28	Pfizer	No	M	21-30	33	B.1.526.2
P21-1296	40	Pfizer	No	F	41-50	31	B.1.526
P21-2376	27	Pfizer	No	F	21-30	19	B.1.526.1
P21-1416	45	Pfizer	No	F	41-50	27	B.1.2
P21-0852	54	Pfizer	No	F	21-30	22	B.1.1.7
P21-1038	22	Moderna	No	F	71-80	27	B.1.2
P21-2765	28	Moderna	Yes (COVID)	M	81-90	17	B.1.526
P21-1413	67	Pfizer	No	M	31-40	23	B.1.526
P22-1140	23	Pfizer	Yes (COVID)	M	61-70	15	B.1.526.2
P21-1407	49	Pfizer	No	M	31-40	28	B.1.1.519
P21-2366	57	Pfizer	Yes (not COVID)	M	71-80	24	B.1.526.2
P21-2383	17	Pfizer	No	M	41-50	33	B.1.526
P21-1292	33	Pfizer	Yes (COVID)	F	71-80	20	B.1.1.7
P21-1293	58	Pfizer	No	F	31-40	27	B.1.526.2
P21-1234	21	Pfizer	No	M	71-80	25	B.1.526
P21-1406	65	Pfizer	No	F	31-40	13	B.1.526.1
P21-2368	52	Pfizer	No	F	21-30	29	B.1.1.7
P21-1412	62	Pfizer	No	F	51-60	22	B.1.526.1
P21-1411	54	Pfizer	No	F	51-60	30	B.1.111
P21-1405	58	Pfizer	No	F	31-40	25	B.1.526
P21-2373	78	Pfizer	No	M	31-40	33	B.1.1.7
P21-1394	70	Pfizer	No	F	21-30	32	B.1.1.7
P21-1396	70	Pfizer	No	F	31-40	27	B.1.1.7
P21-2365	57	Pfizer	No	F	21-30	27	B.1.526.1
P21-2379	65	Pfizer	No	F	41-50	22	B.1.1.7
P21-1403	72	Pfizer	No	F	21-30	27	B.1.526
P21-1404	73	Pfizer	No	F	41-50	17	B.1.1.7
P21-2380	70	Pfizer	No	M	31-40	33	B.1.526.1
P21-1362	61	Pfizer	No	F	41-50	20	B.1.526.1
P21-1595	53	Pfizer	No	F	71-80	18	B.1.1.7

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

P21-1415	42	Moderna	No	M	61-70	24	B.1.1.7
P21-1556	85	Pfizer	No	M	41-50	19	B.1.526.2
P21-1993	50	Moderna	Yes (COVID)	M	81-90	23	B.1.1.7
P21-2384	50	Moderna	No	F	71-80	24	B.1.1.7
P21-1596	57	Moderna	No	M	71-80	34	B.1.1
P21-2191	68	Pfizer	No	M	31-40	18	B.1.526
P21-2192	23	Janssen	Yes (COVID)	M	41-50	25	B.1.1.7
P21-2738	49	Moderna	Yes (not COVID)	M	71-80	28	B.1.1.7
P21-1574	22	Pfizer	No	F	21-30	23	B.1
P21-1985	76	Pfizer	No	M	21-30	23	B.1.1.7
P21-2370	48	Pfizer	No	M	71-80	31	B.1.526
P21-1984	75	Pfizer	No	M	51-60	27	B.1.526.2
P21-1970	71	Moderna	No	F	31-40	22	B.1.1.7
P21-1976	89	Pfizer	No	M	31-40	19	B.1.526.1
P21-1992	76	Pfizer	No	M	41-50	21	B.1.1.7
P21-1964	78	Pfizer	No	M	51-60	33	B.1.1.7
P21-2385	72	Pfizer	No	M	51-60	29	B.1.1.7
P21-2096	88	Pfizer	No	F	31-40	33	B.1.575
P21-2080	60	Moderna	Yes (not COVID)	M	71-80	24	B.1.1.7
P21-2193	94	Pfizer	No	F	41-50	25	P.1
P21-2364	59	Pfizer	No	M	51-60	31	B.1.1.7
P21-2378	20	Pfizer	No	M	61-70	25	B.1.1
P21-2382	26	Pfizer	No	M	61-70	18	B.1.1.7
P21-2105	97	Pfizer	No	M	21-30	20	B.1.1.7
P21-2372	66	Pfizer	No	M	21-30	21	B.1.526
P21-2137	67	Moderna	No	F	31-40	18	B.1.526
P21-2143	55	Pfizer	Yes (not COVID)	F	81-90	20	B.1.526
P21-2371	54	Moderna	No	F	61-70	20	B.1.429
P21-2764	27	Pfizer	Yes (COVID)	F	51-60	31	B.1.1.7
P21-2124	61	Pfizer	No	F	21-30	24	B.1.526
P21-2375	105	Pfizer	No	M	51-60	30	B.1.526
P21-2381	64	Moderna	Yes (not COVID)	M	71-80	22	B.1.575
P21-2386	31	Pfizer	Yes (COVID)	F	51-60	22	B.1.1.7

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

P21-2594	102	Pfizer	No	F	31-40	18	B.1
P21-2573	100	Pfizer	No	F	31-40	19	B.1.1.7
P21-2592	99	Pfizer	No	F	21-30	25	B.1.526

Ct=cycle threshold from real-time RT-PCR SARS-CoV-2 test

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figures.

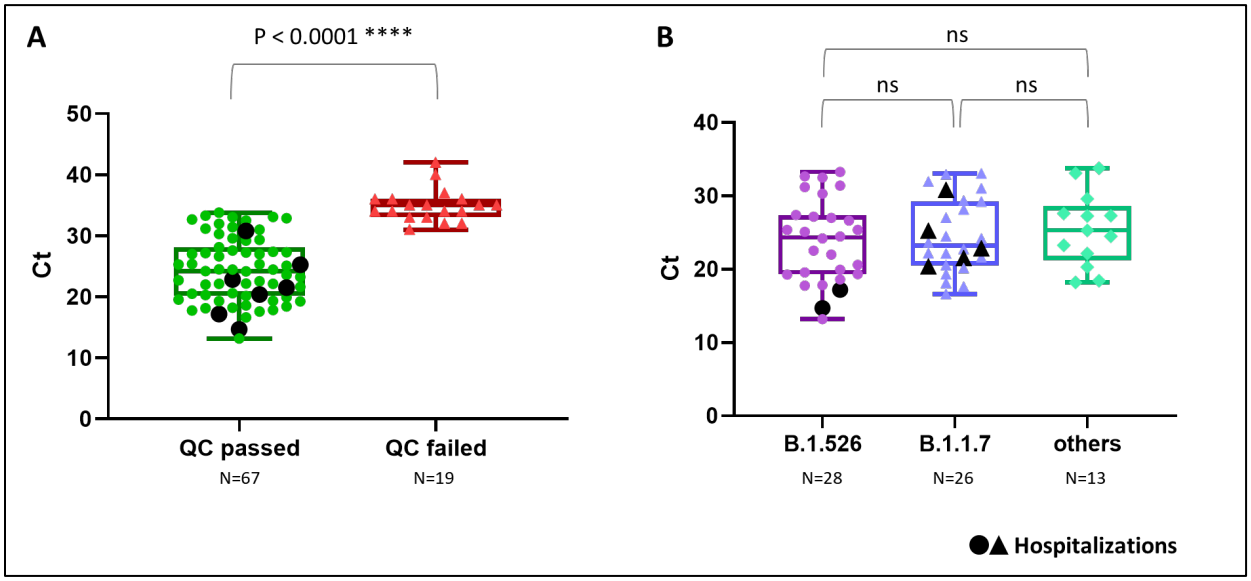


Figure 1. RT-qPCR Ct values in post-vaccine breakthrough infections. **A)** Ct plots of samples that yielded a full genome with sufficient coverage to determine lineage and mutations (QC passed = >23,000 bp and >4000X coverage; QC failed = <23,000 bp and <4000X coverage). **B)** Ct plots of all samples that passed QC, by lineages, B.1.1.7 (Alpha) and B.1.526 (Iota), compared to all others. Hospitalized cases are shown in larger sized black symbols. QC: quality control.

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

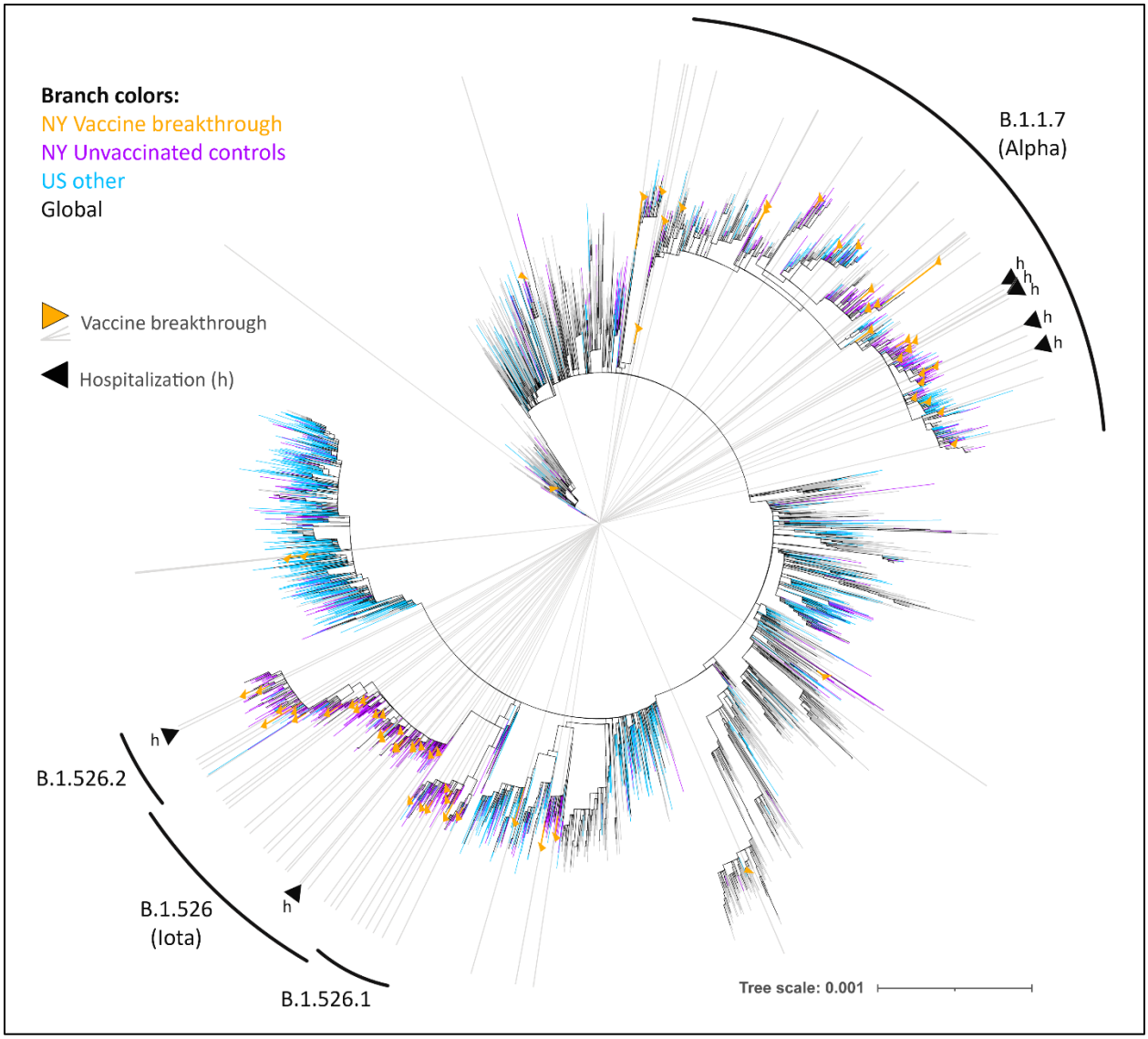


Figure 2. Maximum likelihood tree of SARS-CoV-2 vaccine breakthrough, unvaccinated matched control, and global reference sequences.

IQ tree of 5,056 SARS-CoV-2 full genome sequences (base pairs 202-29,657 according to Wuhan-Hu-1 as reference), including 67 vaccine breakthrough (orange) and 1,187 unvaccinated control SARS-CoV-2 sequences from our NYU cohort (greater NYC area) (purple) together with 1,361 other US (cyan) and 2,441 non-US global reference sequences (black). The tree was generated with a GTR+I+G substitution model and 1,000 bootstrap replicates, and the substitution scale of the tree is indicated at the bottom right. The branches of the tree are colored as indicated. Vaccine breakthrough sequences are additionally highlighted by orange triangles as branch symbols and gray rays radiating from the root to the outer rim of the tree. Hospitalizations due to COVID-19 among the vaccine breakthrough infections are indicated by black triangles (h). The variants responsible for most vaccine breakthrough infections in our study cohort are labeled with respective Pango lineages (WHO classification in parenthesis).

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

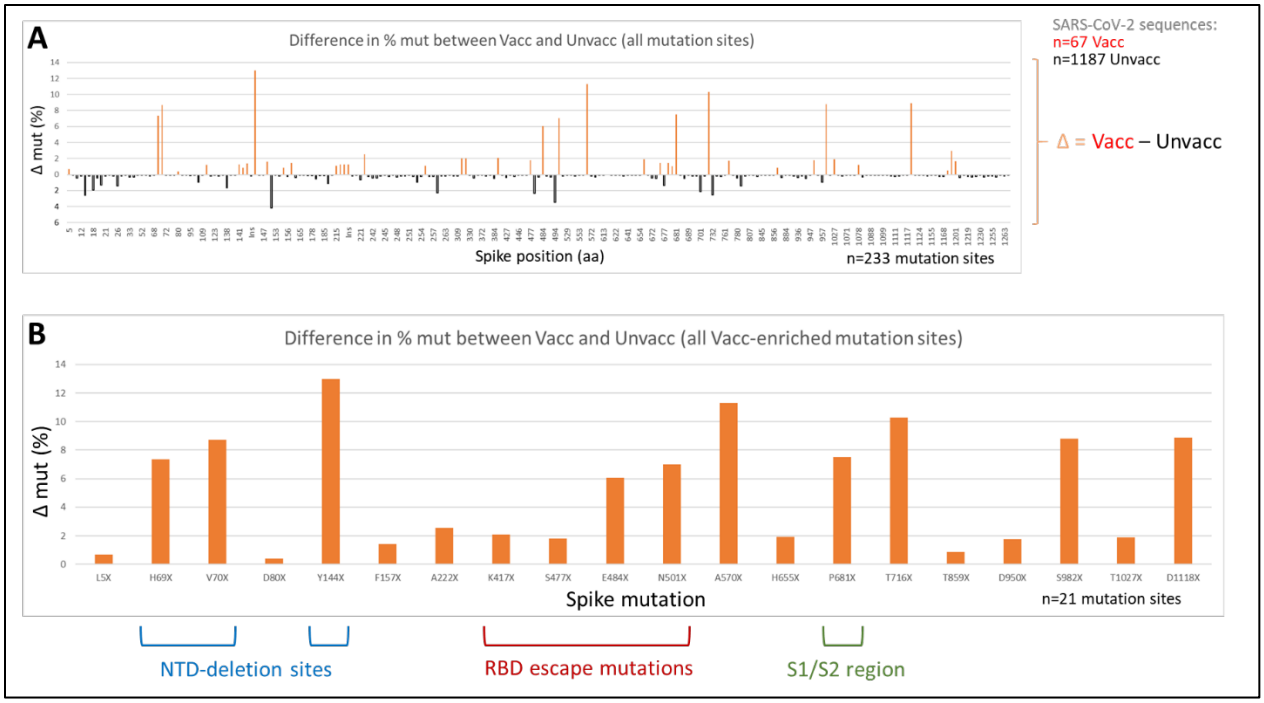


Figure 3. Site-specific spike mutation analysis in SARS-CoV-2 vaccine breakthrough sequences compared to unvaccinated matched controls.

A) Comparison of site-specific amino acid mutation (mut) frequencies in spike in 67 vaccine breakthrough sequences (Vacc) compared to 1,187 unvaccinated matched controls (Unvacc) from the same cohort. The Wuhan-Hu-1 sequence served as reference to call mutations per site, and all spike mutation sites of the study sequences are shown along the x-axis according to their spike position (n=233). The mirror plot displays differences of mutation frequencies between Vacc and Unvacc groups; orange bars (top) refer to higher mutation rates in Vacc sequences, whereas black bars (bottom) refer to higher mutation rates in Unvacc matched controls. **B**) Enrichment of spike mutations in SARS-CoV-2 vaccine breakthrough sequences. All sites with greater spike mutation rates in Vacc compared to Unvacc controls are shown; unique occurrences of mutations were disregarded. Mutation sites in the spike N-terminal domain (NTD), receptor binding domain (RBD), and S1/S2 interface region that have been associated with variants of concern and/or neutralization immune escape are highlighted.

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Supplementary Tables.

Supplementary Table 1: SARS-CoV-2 breakthrough and unvaccinated control infections in matched pairs in New York.

		Vaccinated				
		B.1.1.7	B.1.526	P.1	other	Total
Unvaccinated	B.1.1.7	11	7	0	7	25
	B.1.526	12	14	1	4	31
	P.1	0	1	0	0	1
	other	3	6	0	1	10
	Total	26	28	1	12	67

Supplementary Table 2: SARS-CoV-2 breakthrough and unvaccinated control infections in matched pairs in New York

		Vaccinated		
		VOC/VOI	other	Total
Unvaccinated	VOC/VOI	46	11	57
	other	9	1	10
	Total	55	12	67

Supplementary Table 3: Full specifications of breakthrough infections

Sample ID (arbitrary #)	Dates after completion of vaccination series	Vaccine	Hospitalization	Sex	Age range	Ct for TARGET 1 (orf1)	GISAID Accession	GISAID Virus Name	GISAID Clade	Pango Lineage (Updated 06/11)	AA Substitutions
P21-2388	31	Pfizer	not hospitalized	M	31-40	18.45	EPI_ISL_1922872	NYULH1636	S	A.2.5.2	Spike D215A, Spike D614G, Spike G142del, Spike ins215AGY, Spike L141del, Spike L452R, Spike V143del, N M234I, N P365S, N P383L, N S197L, NS3 S74F, NS7a A105S, NS8 L84S, NSP1 L4F, NSP3 K839E, NSP4 F308Y, NSP4 T492I, NSP6 H11Q, NSP12 M629T
P21-0524	28	Pfizer	not hospitalized	M	21-30	33.3	EPI_ISL_1423379	NYULH1002	GH	B.1.526.2	Spike A701V, Spike D253G, Spike E484K, Spike L5F, Spike T95I, N M234I, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1296	40	Pfizer	not hospitalized	F	41-50	31.42	EPI_ISL_1423503	NYULH997	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike L5F, Spike T95I, N H300Y, N I94V, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-2376	27	Pfizer	not hospitalized	F	21-30	19.28	EPI_ISL_1922862	NYULH1625	GH	B.1.526.1	Spike D80G, Spike D614G, Spike F157S, Spike L452R, Spike T859N, Spike Y144del, N M234I, N T205I, NS3 P42L, NS3 P104L, NS3 Q57H, NS7a L116F, NS8 A51S, NS8 T11I, NSP2 T85I, NSP3 A861S, NSP4 A446V, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP6 V278I, NSP12 P323L
P21-1416	45	Pfizer	not hospitalized	F	41-50	27.32	EPI_ISL_1532785	NYULH1089	GH	B.1.2	Spike D614G, N P13L, N P67S, N P199L, NS3 G172V, NS3 Q57H, NS8 S24L, NSP2 T85I, NSP2 T547I, NSP3 P968L, NSP5 L89F, NSP6 T181I, NSP12 P323L, NSP12 V354L, NSP13 Q275L, NSP14 N129D, NSP16 R216C
P21-0852	54	Pfizer	not hospitalized	F	21-30	22.18	EPI_ISL_1262524	NYULH656	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, M G89S, N A398V, N D3L, N G204R, N R203K, N S235F, NS3 L15F, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP4 F17L, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 K460R, NSP14 S171G
P21-1038	22	Moderna	not hospitalized	F	71-80	27.27	EPI_ISL_1300181	NYULH748	GH	B.1.2	Spike D614G, N D377Y, N P67S, N P199L, NS3 G172V, NS3 Q57H, NS8 S24L, NSP2 T85I, NSP3 D59G, NSP3 M829I, NSP4 M458I, NSP5 L89F, NSP8 A14V, NSP12 K59R, NSP12 P323L, NSP14 N129D, NSP16 R216C
P21-2765	28	Moderna	Hospitalized for COVID	M	81-90	17.2	EPI_ISL_2400382	NYULH1816	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike L5F, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP5 T21I, NSP6 F108del, NSP6 G107del, NSP6 L37F, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1413	67	Pfizer	not hospitalized	M	31-40	22.56	EPI_ISL_1532783	NYULH1087	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike L5F, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 P104L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H, NSP14 A1S, NSP16 P236L
P22-1140	23	Pfizer	Hospitalized for COVID	M	61-70	14.70	EPI_ISL_1322825	NYULH877	GH	B.1.526.2	Spike D253G, Spike D614G, Spike Q957R, Spike S477N, Spike T95I, N A414V, N P13L, N S202R, NS3 P42L, NS3 Q57H, NS7a L116F, NS8 T11I, NSP2 T85I, NSP3 G1128S, NSP3 T1022I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H, NSP13 V521I
P21-1407	49	Pfizer	not hospitalized	M	31-40	27.6	EPI_ISL_1532784	NYULH1084	GR	B.1.1.519	Spike D614G, Spike P681H, Spike T478K, Spike T732A, N G204R, N R203K, NSP1 G112C, NSP3 P141S, NSP3 V1704F, NSP4 T492I, NSP6 I49V, NSP6 L37F, NSP9 T35I, NSP12 P323L, NSP15 L162F
P21-2366	57	Pfizer	hospitalized but not for covid	M	71-80	24.23	EPI_ISL_1922855	NYULH1618	GH	B.1.526.2	Spike D253G, Spike D614G, Spike Q957R, Spike S477N, Spike T95I, N P13L, N S202R, NS3 P42L, NS3 Q57H, NS7a L116F, NS8 T11I, NSP2 T85I, NSP3 A1324T, NSP3 G1128S, NSP3 T169I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP9 D78N, NSP12 P227S, NSP12 P323L, NSP13 Q88H, NSP15 H337Y
P21-2383	17	Pfizer	not hospitalized	M	41-50	32.69	EPI_ISL_1922868	NYULH1632	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 M601I, NSP12 P323L, NSP13 Q88H
P21-1292	33	Pfizer	Hospitalized for COVID	F	71-80	20.38	EPI_ISL_1423499	NYULH993	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 M1441I, NSP3 T183I, NSP5 D34G, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 M519L, NSP12 P227L, NSP12 P323L, NSP13 A509V
P21-1293	58	Pfizer	not hospitalized	F	31-40	27.40	EPI_ISL_1423500	NYULH994	GH	B.1.526.2	Spike D253G, Spike D614G, Spike L5F, Spike Q957R, Spike S477N, Spike T95I, N P13L, N S202R, NS3 P42L, NS3 Q57H, NS7a L116F, NS7a R78C, NS8 T11I, NSP2 T85I, NSP3 G1128S, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1234	21	Pfizer	not hospitalized	M	71-80	25.39	EPI_ISL_1423443	NYULH937	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike L5F, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP1 H110Y, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1406	65	Pfizer	not hospitalized	F	31-40	13.2	EPI_ISL_1532774	NYULH1083	GH	B.1.526.1	Spike D80G, Spike D614G, Spike D950H, Spike F157S, Spike L452R, Spike T859N, Spike Y144del, N M234I, N T205I, NS3 P42L, NS3 P104L, NS3 Q57H, NS8 A51S, NS8 T11I, NSP2 T85I, NSP4 A446V, NSP4 K399E, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP6 V278I, NSP12 P323L
P21-2368	52	Pfizer	not hospitalized	F	21-30	29.21	EPI_ISL_1922856	NYULH1619	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike K1191N, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS3 P240S, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP8 Q24R, NSP12 P323L, NSP13 K460R, NSP16 K160R, NSP16 T140I
P21-1412	62	Pfizer	not hospitalized	F	51-60	22.01	EPI_ISL_2202740	NYULH1086	GH	B.1.526.1	Spike D80G, Spike D614G, Spike D950H, Spike F157S, Spike L452R, Spike T859N, Spike Y144del, N E323K, N M234I, N T205I, NS3 P42L, NS3 P104L, NS3 Q57H, NS8 A51S, NS8 T11I, NSP2 T85I, NSP3 A861S, NSP4 A446V, NSP4 K399E, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP6 V278I, NSP12 P323L, NSP13 A469T

P21-1411	54	Pfizer	not hospitalized	F	51-60	29.62	EPI_ISL_1532768	NYULH1085	GH	B.1.111	Spike D614G, Spike N679K, N K387R, N T205I, NS3 G44E, NS3 Q57H, NS3 T223I, NS8 L57S, NSP1 I114T, NSP12 P323L
P21-1405	58	Pfizer	not hospitalized	F	31-40	25.11	EPI_ISL_1532782	NYULH1082	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike L5F, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H, NSP14 T428A
P21-2373	78	Pfizer	not hospitalized	M	31-40	33.06	EPI_ISL_1922860	NYULH1623	G	B.1.1.7	Spike A222S, Spike H69del, Spike V70del, Spike Y144del, M G89S, N D3L, N D401Y, NS3 L15F, NS8 S24L, NSP3 A890D, NSP4 F17L, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP13 K460R, NSP14 S171G, NSP16 A83T
P21-1394	70	Pfizer	not hospitalized	F	21-30	31.96	EPI_ISL_1654609	NYULH1256	GR	B.1.1.7	Spike A570D, Spike D614G, Spike H69del, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP6 R138del, NSP6 V139del, NSP6 W140del, NSP13 A509V, NSP13 P234L
P21-1396	70	Pfizer	not hospitalized	F	31-40	27.01	EPI_ISL_1654610	NYULH1258	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, E L21F, N D3L, N G204R, N R203K, N S235F, NS7a I103del, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP2 K489E, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 A399T, NSP12 P323L, NSP16 S241P
P21-2365	57	Pfizer	not hospitalized	F	21-30	26.66	EPI_ISL_1922854	NYULH1617	GH	B.1.526.1	Spike D80G, Spike D614G, Spike D950H, Spike F157S, Spike L452R, Spike T859N, Spike Y144del, N M234I, N T205I, NS3 P42L, NS3 P104L, NS3 Q57H, NS8 A51S, NS8 T11I, NSP2 T85I, NSP4 A446V, NSP4 K399E, NSP4 L438P, NSP5 F223I, NSP5 R222Q, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP6 V278I, NSP7 Q63R, NSP12 P323L
P21-2379	65	Pfizer	not hospitalized	F	41-50	22.16	EPI_ISL_1922864	NYULH1628	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N148T, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS3 L15F, NS8 G8R, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 S126L, NSP3 T183I, NSP4 F17L, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 K460R
P21-1403	72	Pfizer	not hospitalized	F	21-30	27.17	EPI_ISL_1532731	NYULH1080	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike Q1201L, Spike T95I, N M234I, N P199L, NS3 A33S, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP3 T1189I, NSP4 L438P, NSP4 S218F, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H, NSP16 L163F
P21-1404	73	Pfizer	not hospitalized	F	41-50	16.63	EPI_ISL_1532746	NYULH1081	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike L5F, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS3 T151I, NS7a Q62stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP2 L137F, NSP3 A890D, NSP3 I1412T, NSP3 P340L, NSP3 T183I, NSP3 T1303I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP12 T739I
P21-2380	70	Pfizer	not hospitalized	M	31-40	32.53	EPI_ISL_1922865	NYULH1629	GH	B.1.526.1	Spike D80G, Spike D614G, Spike D950H, Spike F157S, Spike L452R, Spike T859N, Spike Y144del, N M234I, N T205I, NS3 P42L, NS3 Q57H, NS8 A51S, NS8 F3L, NS8 T11I, NSP2 T85I, NSP4 A446V, NSP4 K399E, NSP4 L438P, NSP5 P96S, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP6 V278I, NSP12 P323L, NSP14 V398L
P21-1362	61	Pfizer	not hospitalized	F	41-50	19.94	EPI_ISL_1532753	NYULH1061	GH	B.1.526.1	Spike D80G, Spike D614G, Spike D950H, Spike F157S, Spike L452R, Spike T323I, Spike T676I, Spike T859N, Spike Y144del, N M234I, N T205I, NS3 P42L, NS3 P104L, NS3 Q57H, NS8 A51S, NS8 T11I, NSP2 T85I, NSP3 A861S, NSP4 A446V, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP6 V278I, NSP12 P323L, NSP16 K160R
P21-1595	53	Pfizer	not hospitalized	F	71-80	17.6	EPI_ISL_1576946	NYULH1178	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike P681H, Spike S982A, Spike T716I, N D3L, N G204R, N R203K, N S235F, NS8 T11I, NSP2 L550F, NSP3 A890D, NSP3 I1412T, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L
P21-1415	42	Moderna	not hospitalized	M	61-70	23.61	EPI_ISL_1532745	NYULH1088	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS3 P240S, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 E95D, NSP3 I1412T, NSP3 M1441I, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L, NSP13 T588I
P21-1556	85	Pfizer	not hospitalized	M	41-50	18.62	EPI_ISL_1576907	NYULH1139	GH	B.1.526.2	Spike D253G, Spike D614G, Spike L5F, Spike Q957R, Spike S477N, Spike T95I, N A414V, N P13L, N S202R, NS3 P42L, NS3 Q57H, NS7a L116F, NS8 T11I, NSP2 T85I, NSP3 A1179T, NSP3 G1128S, NSP3 T1022I, NSP4 L438P, NSP4 T319I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1993	50	Moderna	Hospitalized for COVID	M	81-90	22.85	EPI_ISL_1669821	NYULH1262	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS3 P240S, NS3 T9I, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 E95D, NSP3 K963R, NSP3 M1441I, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L, NSP16 A34V
P21-2384	50	Moderna	not hospitalized	F	71-80	24.49	EPI_ISL_1922869	NYULH1633	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS3 P240S, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 E95D, NSP3 I1412T, NSP3 K963R, NSP3 M1441I, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L
P21-1596	57	Moderna	not hospitalized	M	71-80	33.8	EPI_ISL_1576948	NYULH1179	GR	B.1.1	Spike A570D, Spike D614G, Spike P681H, N D3L, N G204R, N R203K, NS8 A51S, NS8 T11I, NSP6 F108del, NSP6 G107del, NSP6 S106del
P21-2191	68	Pfizer	not hospitalized	M	31-40	17.86	EPI_ISL_1790045	NYULH1414	GH	B.1.526	Spike D614G, N M234I, N P168S, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP5 P241S, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 G54C
P21-2192	23	Janssen (J&J)	Hospitalized for COVID	M	41-50	25.25	EPI_ISL_1790046	NYULH1415	GR	B.1.1.7	Spike A570D, Spike D614G, Spike P681H, Spike T716I, N D3L, N G204R, N R203K, N S235F, NS8 K68stop, NS8 Q27stop, NS8 Y73C, NSP6 R138del, NSP6 V139del, NSP6 W140del
P21-2738	49	Moderna	hospitalized but not for covid	M	71-80	28.14	EPI_ISL_2400287	NYULH1814	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike K1191N, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A225V, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP8 Q24R, NSP12 P323L, NSP13 K460R, NSP14 D324E, NSP16 K160R
P21-1574	22	Pfizer	not hospitalized	F	21-30	23.28	EPI_ISL_1576925	NYULH1157	G	B.1	Spike D253G, Spike D614G, Spike S477N, Spike T95I, N G204R, N R203K, N S235F, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP3 G1128S, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H

P21-1985	76	Pfizer	not hospitalized	M	21-30	22.74	EPI_ISL_1654640	NYULH1248	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, M G89S, N A398V, N D3L, N G204R, N R203K, N S235F, NS3 L15F, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP4 F17L, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP8 Q73K, NSP12 P323L, NSP13 K460R
P21-2370	48	Pfizer	not hospitalized	M	71-80	31.20	EPI_ISL_1922857	NYULH1620	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1984	75	Pfizer	not hospitalized	M	51-60	27.03	EPI_ISL_1654639	NYULH1247	GH	B.1.526.2	Spike D253G, Spike D614G, Spike L5F, Spike Q957R, Spike S477N, Spike T95I, N P13L, N S202R, NS3 P42L, NS3 Q57H, NS7a L116F, NS7b H37L, NS8 T11I, NSP2 T85I, NSP2 T434I, NSP3 G1128S, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1970	71	Moderna	not hospitalized	F	31-40	21.63	EPI_ISL_1669802	NYULH1260	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS3 A99T, NS3 V90F, NS7b T40I, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 K460R, NSP14 E347G
P21-1976	89	Pfizer	not hospitalized	M	31-40	19.36	EPI_ISL_1669803	NYULH1261	GH	B.1.526.1	Spike D80G, Spike D614G, Spike D950H, Spike F157S, Spike L452R, Spike S155R, Spike T859N, Spike Y144del, N M234I, N T205I, NS3 P42L, NS3 P104L, NS3 Q57H, NS8 A51S, NS8 T11I, NSP2 T85I, NSP3 A861S, NSP4 A446V, NSP4 K399E, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP6 V278I, NSP12 P323L, NSP12 T806I, NSP13 T239I
P21-1992	76	Pfizer	not hospitalized	M	41-50	20.54	EPI_ISL_1654647	NYULH1255	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N G204R, N R203K, N S235F, NS8 S67F, NS8 T11I, NSP1 N178S, NSP2 L550F, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-1964	78	Pfizer	not hospitalized	M	51-60	32.91	EPI_ISL_1669801	NYULH1259	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, M G89S, N A398V, N D3L, N G204R, N R203K, N S235F, NS3 L15F, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP4 F17L, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP8 Q73K, NSP12 P323L, NSP13 K460R
P21-2385	72	Pfizer	not hospitalized	M	51-60	29.35	EPI_ISL_1922870	NYULH1634	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I310T, NSP3 I1412T, NSP3 M1441I, NSP3 P108L, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L
P21-2096	88	Pfizer	not hospitalized	F	31-40	33.12	EPI_ISL_1789975	NYULH1340	GH	B.1.575	Spike D614G, Spike P681H, Spike T678I, Spike T716I, M I82T, N T205I, NS3 Q57H, NS7a L96F, NSP4 T492I, NSP6 T181I, NSP12 P323L, NSP13 T127I, NSP15 L162V
P21-2080	60	Moderna	hospitalized but not for covid	M	71-80	24.18	EPI_ISL_1789962	NYULH1325	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike K1191N, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A225V, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP8 Q24R, NSP12 P323L, NSP13 K460R, NSP14 D324E, NSP16 K160R
P21-2193	94	Pfizer	not hospitalized	F	41-50	24.51	EPI_ISL_1790047	NYULH1416	GR	P.1	Spike D614G, Spike H655Y, Spike K417T, Spike T1027I, Spike V1176F, N G204R, N P80R, N R203K, NS3 S253P, NS8 E92K, NSP3 K977Q, NSP3 P822S, NSP6 F108del, NSP6 G107del, NSP6 K4R, NSP6 S106del, NSP12 P94L, NSP12 P323L, NSP13 E341D, NSP14 A119V
P21-2364	59	Pfizer	not hospitalized	M	51-60	31.03	EPI_ISL_1922853	NYULH1616	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP3 T1127I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L, NSP14 P451S
P21-2378	20	Pfizer	not hospitalized	M	61-70	25.32	EPI_ISL_1922863	NYULH1627	GR	B.1.1	Spike A222V, Spike D614G, Spike G769V, Spike S255F, N G204R, N Q418H, N R203K, N S187L, NSP2 A306V, NSP3 S1656A, NSP12 P323L, NSP13 G439R, NSP14 P412H
P21-2382	26	Pfizer	not hospitalized	M	61-70	18.13	EPI_ISL_1922867	NYULH1631	GR	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 C61F, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L
P21-2105	97	Pfizer	not hospitalized	M	21-30	20.16	EPI_ISL_2202770	NYULH1344	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, N D3L, N G204R, N R14C, N R203K, N S235F, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 P198L, NSP6 S106del, NSP12 P227L, NSP12 P323L
P21-2372	66	Pfizer	not hospitalized	M	21-30	20.61	EPI_ISL_1922859	NYULH1622	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 P104L, NS3 Q57H, NS8 T11I, NSP2 A386S, NSP2 T85I, NSP3 S1588L, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-2137	67	Moderna	not hospitalized	F	31-40	17.78	EPI_ISL_1790005	NYULH1374	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike L5F, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 A65V, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-2143	55	Pfizer	hospitalized but not for covid	F	81-90	19.58	EPI_ISL_1790010	NYULH1379	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike L5F, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS6 D30del, NS6 F22del, NS6 I26del, NS6 K23del, NS6 L29del, NS6 N28del, NS6 S25del, NS6 V24del, NS6 W27del, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 I210T, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-2371	54	Moderna	not hospitalized	F	61-70	20.33	EPI_ISL_1922858	NYULH1621	GH	B.1.429	Spike D614G, Spike L452R, N T205I, NS3 Q57H, NSP2 T85I, NSP3 V1331F, NSP5 K90R, NSP9 I65V, NSP12 P323L, NSP13 D260Y, NSP14 S418C, NSP16 S33N
P21-2764	27	Pfizer	Hospitalized for COVID	F	51-60	30.82	EPI_ISL_2400336	NYULH1815	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, E S68F, N D3L, N G204R, N R203K, N S235F, NS3 D210Y, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 M1441I, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L, NSP13 A509V
P21-2124	61	Pfizer	not hospitalized	F	21-30	24.47	EPI_ISL_1789994	NYULH1362	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike K417N, Spike L5F, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H

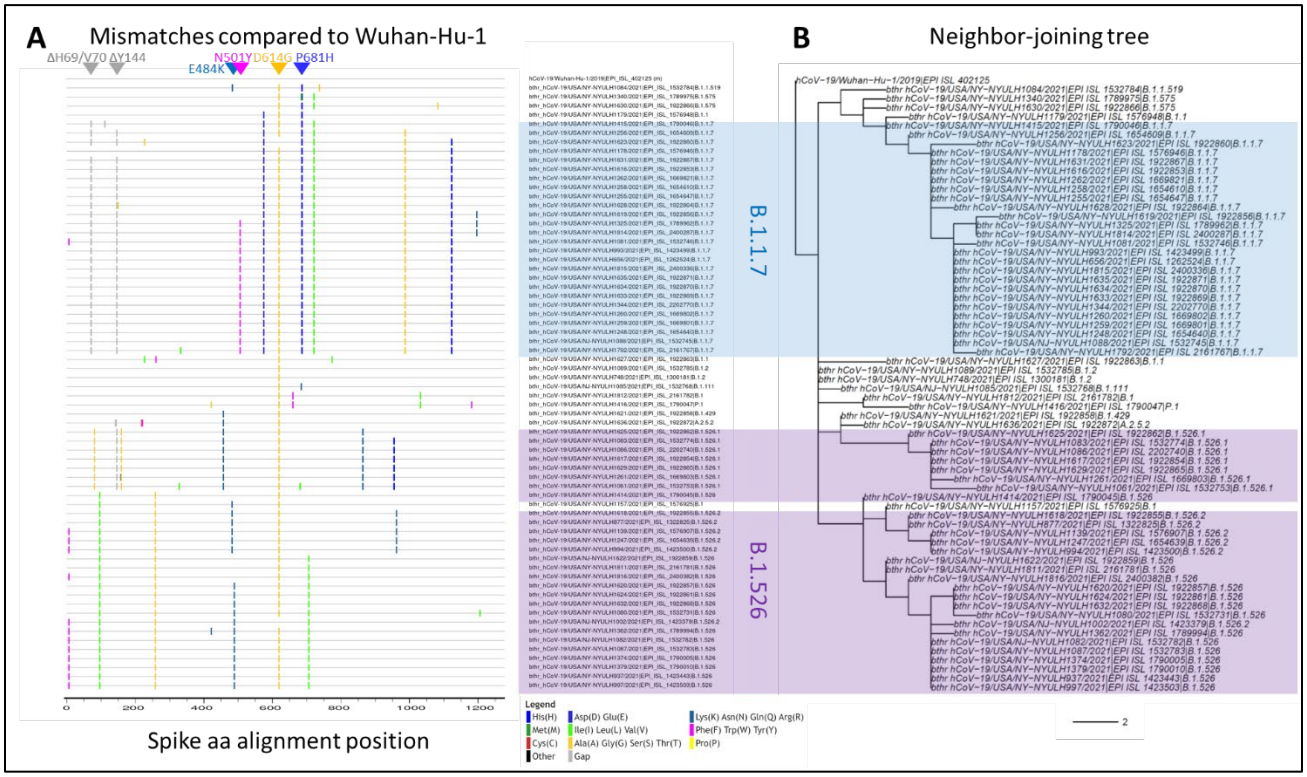
P21-2375	105	Pfizer	not hospitalized	M	51-60	30.32	EPI_ISL_1922861	NYULH1624	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike E484K, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H
P21-2381	64	Moderna	hospitalized but not for covid	M	71-80	22.14	EPI_ISL_1922866	NYULH1630	GH	B.1.575	Spike A1078S, Spike D614G, Spike P681H, Spike T716I, M I82T, N G96C, N K127R, N L224F, N T205I, NS3 Q57H, NSP2 S122F, NSP2 T85I, NSP3 A333S, NSP4 T492I, NSP6 T181I, NSP12 A97V, NSP12 P323L, NSP13 P47S
P21-2386	31	Pfizer	Hospitalized for COVID	F	51-60	21.52	EPI_ISL_1922871	NYULH1635	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike Y144del, E S68F, N D3L, N G204R, N R203K, N S235F, NS3 D210Y, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 M1441I, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L, NSP13 A509V
P21-2594	102	Pfizer	not hospitalized	F	31-40	18.22	EPI_ISL_2161782	NYULH1812	G	B.1	Spike D614G, Spike H655Y, N T205I, NS3 S253P, NS3 V256F, NS8 E92K, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP13 E341D
P21-2573	100	Pfizer	not hospitalized	F	31-40	19.32	EPI_ISL_2161767	NYULH1792	GRY	B.1.1.7	Spike A570D, Spike D614G, Spike D1118H, Spike H69del, Spike N501Y, Spike P681H, Spike S982A, Spike T716I, Spike V70del, Spike V327I, Spike Y144del, N D3L, N G204R, N R203K, N S235F, NS8 K68stop, NS8 Q27stop, NS8 R52I, NS8 Y73C, NSP3 A890D, NSP3 I1412T, NSP3 T183I, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P227L, NSP12 P323L, NSP14 P451S
P21-2592	99	Pfizer	not hospitalized	F	21-30	25.37	EPI_ISL_2161781	NYULH1811	GH	B.1.526	Spike A701V, Spike D253G, Spike D614G, Spike T95I, N M234I, N P199L, NS3 P42L, NS3 P262S, NS3 Q57H, NS8 T11I, NSP2 T85I, NSP3 P1469S, NSP3 S1396T, NSP4 L438P, NSP6 F108del, NSP6 G107del, NSP6 S106del, NSP12 P323L, NSP13 Q88H

Supplementary Table 4: Spike mutation statistics at 21 sites with enriched mutation rates in vaccine breakthrough infections in New York.

Spike pos	5	69	70	80	144	157	222	417	477	484	501	570	655	681	716	859	950	982	1027	1118	1191
SARS-CoV-2/Wuhan-Hu-1	L	H	V	D	Y	F	A	K	S	E	N	A	H	P	T	T	D	S	T	D	K
spike mutation	L5X	**H69X**	**V70X**	D80X	**Y144X**	F157X	A222X	**K417X**	**S477X**	**E484X**	**N501X**	A570X	H655X	**P681X**	T716X	T859X	D950X	S982X	T1027X	D1118X	K1191X
important sites		**del**			**del**			**RBD (aa 331 - 524)**						**VOCs (UK & India)**							
Unvacc-ctr WT (%)	62.42	69.47	70.82	89.63	59.52	89.36	99.32	98.5	89.31	76.06	69.51	70.99	98.9	62.06	68.49	90.1	90	70.91	98.9	70.78	98.11
Unvacc-ctr mut (%)	37.58	30.53	29.18	10.37	40.48	10.64	0.68	1.5	10.69	23.94	30.49	29.01	1.1	37.94	31.51	9.9	10	29.09	1.1	29.22	1.89
Vacc WT (%)	61.76	62.12	62.12	89.23	46.55	87.93	96.77	96.43	87.5	70	62.5	59.7	96.97	54.55	58.21	89.23	88.24	62.12	97.01	61.9	95.16
Vacc mut (%)	38.24	37.88	37.88	10.77	53.45	12.07	3.23	3.57	12.5	30	37.5	40.3	3.03	45.45	41.79	10.77	11.76	37.88	2.99	38.1	4.84
Difference: Vacc (%) minus Unvacc-ctrl (%) (enrichment)	0.66	7.35	8.7	0.4	12.97	1.43	2.55	2.07	1.81	6.06	7.01	11.29	1.93	7.51	10.28	0.87	1.76	8.79	1.89	8.88	2.95
p (Fisher)	1	0.22	0.17	0.84	0.06	0.67	0.09	0.23	0.64	0.35	0.38	0.05	0.19	0.24	0.08	0.83	0.63	0.13	0.19	0.16	0.13
q (multiplicity correction, BH)	1	0.39	0.39	0.88	0.39	0.78	0.39	0.39	0.78	0.53	0.54	0.39	0.39	0.39	0.39	0.88	0.78	0.39	0.39	0.39	0.39

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Supplementary Figures.



Supplementary Figure 1. Spike mutation patterns in vaccine breakthrough sequences.
A) Highlighter plot showing spike amino acid mutations of 67 vaccine breakthrough SARS-CoV-2 sequences compared to the Wuhan-Hu-1 reference sequence as master (top line). Mutations are shown as ticks, color-coded according to the legend at the bottom. Key mutations are indicated by triangles and labeled. Study sequences are sorted according to the Neighbor-joining tree to the right (**B**). The scale bar at the bottom indicates the branch length of two substitutions. B.1.1.7 and B.1.526 sequences are highlighted in blue and purple, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Supplementary Methods

cDNA synthesis, library preparation and sequencing

Total RNA (11 μl) was converted to first strand cDNA by random priming using the Superscript IV first-strand synthesis system (Invitrogen, ref# 180901050). Random priming was performed as follows: 65˚C for 5 mins, cooled on ice followed by addition of mix to run cDNA first strand synthesis 23˚C 10mins, 50˚C for 1 hour, 80˚C for 10 mins, and 4˚C hold. Libraries were prepared using Swift Normalase Amplicon Panel (SNAP) SARS-CoV-2 and Sar-Cov-2 additional Genome Coverage (Cat# SN-5X296 core kit, 96rxn), using 10 μl of first strand cDNA, following the manufacturer's instruction: https://swiftbiosci.com/wp-content/uploads/2021/06/PRT-028-Swift-Normalase-Amplicon-Panels-SNAP-SARS-CoV-2-Panels-Rev-9.pdf). First step incorporated tiled primer pairs which target and enrich for the entire 29.9kb covid-19 viral genome (NCBI Reference Sequence NC_045512.2) during multiplex PCR. Multiplex PCR was run as follows: 98˚C 30 secs, 4 repeating cycles of 98˚C/ 10secs, 61˚C/5mins, 65˚C/1min. 24 repeating cycles of 98˚C/10secs, 64˚C/1min followed by 1 cycle of 65˚C for 1 min and held at 4 ˚C. The multiplexed samples were cleaned with 1x volume room temperature Ampure XP beads (Beckman Coulter, #A63882) for 5 mins, washed with 80% ethanol twice, and beads were resuspended in 17.4ul of TE buffer (part of swift kit). Indexing reagent mix was added to the resuspended beads. One unique SNAP UD dual indexing primer pair was added to each sample. Indexing PCR was run at 37˚C/20mins, 98˚C/30secs, 9 repeating cycles of 98˚C/10secs, 60˚C/30secs, 66˚C/1min. After indexing PCR was completed, samples were cleaned with 32.5ul of PEG NaCl solution (provided in swift kit, 5min bind time, 80% ethanol wash x2) and eluted in 20ul of post PCR TE buffer. Final libraries were run on Agilent Tapestation 2200 with high sensitivity DNA Screentape to verify amplicon size of about 450 bp. Samples were pooled according to the enzymatic normalase step. The loading concentration for the normalized pool was verified on QPCRusing the Kapa library quantification complete kit (Kapa-Roche, #KK4824) run on the Bio-Rad CFX384 Real-time system. Normalized pools were run on the Illumina NovaSeq 6000 system with the SP 300 cycle flow cell. Run metrics were paired end 150 cycles with dual indexing reads. Typically, 2 pools representing 2 full 96 well plates (192 samples) were sequenced on each SP300 NovaSeq flow cell.

Sequenced read processing

Binary base call files outputted by the Illumina NovaSeq 6000 sequencing instrument (RTA Version 3.4.4) were converted to FASTQ format and demultiplexed according to the unique dual-index adapter sequences of each sample using the Illumina bcl2fastq2 Conversion Software v2.20. Adapters and low quality bases were trimmed using Trimmomatic v0.36 (Bolger et al. 2014) with the settings 'ILLUMINACLIP:trimmomatic.fa:2:30:10:1:true TRAILING:5 SLIDINGWINDOW:4:15 MINLEN:35'. The BWA-MEM algorithm of the alignment program, BWA v0.7.17 (Li and Durbin 2009), was utilized for mapping reads of each sample to the SARS-CoV-2 reference genome (NC_045512.2, wuhCor1) with the '-M' flag for marking shorter split hits as secondary and the parameter '-U' set to a 17 alignment score penalty for an unpaired read pair. Name-sorted mapped reads in SAM alignment format were soft-clipped to remove primer sequences specific to the short overlapping (tiled) amplicon design of the Swift Normalase

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Amplicon SARS-CoV-2 Panel using Primerclip v0.3.8 (Swift Biosciences: https://github.com/swiftbiosciences/primerclip), and the Picard v2.18.20 "AddOrReplaceReadGroups" tool (Broad Institute: http://broadinstitute.github.io/picard/) was used in converting the primer-trimmed SAM files to coordinate-sorted and indexed BAM format with read-groups added for downstream analysis. The variant caller, BCFtools v1.9 (Li et al. 2009), was utilized to detect genetic mutations of the collected viral samples from BAM input:

```
bcftools mpileup -r NC_045512.2 --count-orphans --no-BAQ --max-depth 50000 --max-idepth 500000 --annotate FORMAT/AD,FORMAT/ADF,FORMAT/ADR,FORMAT/DP,FORMAT/SP,INFO/AD,INFO/ADF,INFO/ADR --output-type v |
bcftools call --ploidy 1 --keep-alts --multiallelic-caller --output-type v |
bcftools norm --multiallelics -any --output-type v |
bcftools filter -i "(DP4[0]+DP4[1])<(DP4[2]+DP4[3]) && ((DP4[2]+DP4[3])>0) & FORMAT/AD[0:1]>20" --output-type v |
bcftools filter -e "IMF<0.5" --output-type v
```

To generate viral sequences, the output in VCF format was applied to the reference sequence using `bcftools consensus` while the BEDTools v2.30.0 utilities (Quinlan and Hall 2010), "genomecov" (with the '-bga' option to report depth at each base including zero coverage sites), "merge" (for creating a BED file of low depth sites with <1000x coverage to be masked), and "maskfasta" in conjunction with the multiple sequence alignment program, MAFFT v4.471 (Katoh et al. 2002) in '--auto' mode, were used in the production of final consensus sequences with all bases below 1000x coverage masked and both non-targeted ends of the sequences trimmed. Sequences with fewer than 23,000 bp with at least 4000x coverage were discarded from further analysis for having not yielded a near-complete SARS-CoV-2 genome, resulting in 1,254 final sequences suitable for downstream analysis (Supplemental Table). Each assembled SARS-CoV-2 genome sequence was assigned the most likely phylogenetic lineage according to Phylogenetic Assignment of Named Global Outbreak Lineages (PANGO) nomenclature (Rambaut et al. 2020) using the PANGO lineage tool (version: 2021-06-05).

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. Worldometer. COVID-19 Coronavirus Pandemic. Accessed June 15th, 2021, https://www.worldometers.info/coronavirus/
2. Wu F, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. Mar 2020;579(7798):265-269. doi:10.1038/s41586-020-2008-3
3. Baden LR, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N Engl J Med*. Feb 4 2021;384(5):403-416. doi:10.1056/NEJMoa2035389
4. Polack FP, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med*. Dec 31 2020;383(27):2603-2615. doi:10.1056/NEJMoa2034577
5. Sadoff J, et al. Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. *N Engl J Med*. Apr 21 2021;doi:10.1056/NEJMoa2101544
6. Dagan N, et al. BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting. *N Engl J Med*. Feb 24 2021;doi:10.1056/NEJMoa2101765
7. Daniel W, et al. Early Evidence of the Effect of SARS-CoV-2 Vaccine at One Medical Center. *New England Journal of Medicine*. 2021;doi:10.1056/NEJMc2102153
8. Thompson MG, et al. Interim Estimates of Vaccine Effectiveness of BNT162b2 and mRNA-1273 COVID-19 Vaccines in Preventing SARS-CoV-2 Infection Among Health Care Personnel, First Responders, and Other Essential and Frontline Workers - Eight U.S. Locations, December 2020-March 2021. *MMWR Morb Mortal Wkly Rep*. Apr 2 2021;70(13):495-500. doi:10.15585/mmwr.mm7013e3
9. Rossman H, et al. COVID-19 dynamics after a national immunization program in Israel. *Nat Med*. Apr 19 2021;doi:10.1038/s41591-021-01337-2
10. Haas EJ, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. *Lancet*. May 15 2021;397(10287):1819-1829. doi:10.1016/S0140-6736(21)00947-8
11. Chodick G, et al. The effectiveness of the TWO-DOSE BNT162b2 vaccine: analysis of real-world data. *Clin Infect Dis*. May 17 2021;doi:10.1093/cid/ciab438
12. Pilishvili T, et al. Interim Estimates of Vaccine Effectiveness of Pfizer-BioNTech and Moderna COVID-19 Vaccines Among Health Care Personnel - 33 U.S. Sites, January-March 2021. *MMWR Morb Mortal Wkly Rep*. May 21 2021;70(20):753-758. doi:10.15585/mmwr.mm7020e2
13. Vanker A, et al. Adverse Outcomes Associated with SARS-CoV-2 variant B.1.351 Infection in Vaccinated Residents of a Long Term Care Home, Ontario, Canada. *Clin Infect Dis*. Jun 6 2021;doi:10.1093/cid/ciab523
14. Thompson CN, et al. Rapid Emergence and Epidemiologic Characteristics of the SARS-CoV-2 B.1.526 Variant - New York City, New York, January 1-April 5, 2021. *MMWR Morb Mortal Wkly Rep*. May 14 2021;70(19):712-716. doi:10.15585/mmwr.mm7019e1
15. Teran RA, et al. Postvaccination SARS-CoV-2 Infections Among Skilled Nursing Facility Residents and Staff Members - Chicago, Illinois, December 2020-March 2021. *MMWR Morb Mortal Wkly Rep*. Apr 30 2021;70(17):632-638. doi:10.15585/mmwr.mm7017e1

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

16. Pollett SD, et al. The SARS-CoV-2 mRNA vaccine breakthrough infection phenotype includes significant symptoms, live virus shedding, and viral genetic diversity. *Clin Infect Dis*. Jun 12 2021;doi:10.1093/cid/ciab543
17. Kustin T, et al. Evidence for increased breakthrough rates of SARS-CoV-2 variants of concern in BNT162b2-mRNA-vaccinated individuals. *Nat Med*. Jun 14 2021;doi:10.1038/s41591-021-01413-7
18. Jacobson KB, et al. Post-vaccination SARS-CoV-2 infections and incidence of the B.1.427/B.1.429 variant among healthcare personnel at a northern California academic medical center. *medRxiv*. Apr 24 2021;doi:10.1101/2021.04.14.21255431
19. Hacisuleyman E, et al. Vaccine Breakthrough Infections with SARS-CoV-2 Variants. *N Engl J Med*. Jun 10 2021;384(23):2212-2218. doi:10.1056/NEJMoa2105000
20. Cdc Covid- Vaccine Breakthrough Case Investigations Team. COVID-19 Vaccine Breakthrough Infections Reported to CDC - United States, January 1-April 30, 2021. *MMWR Morb Mortal Wkly Rep*. May 28 2021;70(21):792-793. doi:10.15585/mmwr.mm7021e3
21. Cavanaugh AM, et al. COVID-19 Outbreak Associated with a SARS-CoV-2 R.1 Lineage Variant in a Skilled Nursing Facility After Vaccination Program - Kentucky, March 2021. *MMWR Morb Mortal Wkly Rep*. Apr 30 2021;70(17):639-643. doi:10.15585/mmwr.mm7017e2
22. Duerr R, et al. SARS-CoV-2 Portrayed Against HIV: Contrary Viral Strategies in Similar Disguise. *Preprints*. 2021;2021040193. doi:10.20944/preprints202104.0193.v1
23. Wang R, et al. Analysis of SARS-CoV-2 variant mutations reveals neutralization escape mechanisms and the ability to use ACE2 receptors from additional species. *Immunity*. 2021/06/08/ 2021;doi:https://doi.org/10.1016/j.immuni.2021.06.003
24. West AP, et al. SARS-CoV-2 lineage B.1.526 emerging in the New York region detected by software utility created to query the spike mutational landscape. *bioRxiv*. 2021:2021.02.14.431043. doi:10.1101/2021.02.14.431043
25. Annavajhala MK, et al. A Novel SARS-CoV-2 Variant of Concern, B.1.526, Identified in New York. *medRxiv*. 2021:2021.02.23.21252259. doi:10.1101/2021.02.23.21252259
26. SWIFT Protocol. Swift Normalase™ Amplicon Panels (SNAP). https://swiftbiosci.com/wp-content/uploads/2021/06/PRT-028-Swift-Normalase-Amplicon-Panels-SNAP-SARS-CoV-2-Panels-Rev-9.pdf
27. Bolger AM, et al. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. Aug 1 2014;30(15):2114-20. doi:10.1093/bioinformatics/btu170
28. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. Jul 15 2009;25(14):1754-60. doi:10.1093/bioinformatics/btp324
29. SWIFT Biosciences. Primerclip. https://github.com/swiftbiosciences/primerclip
30. Li H, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. Aug 15 2009;25(16):2078-9. doi:10.1093/bioinformatics/btp352
31. Rambaut A, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology*. 2020/11/01 2020;5(11):1403-1407. doi:10.1038/s41564-020-0770-5
32. CBRC. MAFFT version 7. Accessed June 7th, 2021, https://mafft.cbrc.jp/alignment/software/closelyrelatedviralgenomes.html
33. Miller MA, et al. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. 2010:1-8.

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

34. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic acids research*. Jul 2 2019;47(W1):W256-W259. doi:10.1093/nar/gkz239
35. R Core Team. R: A language and environment for statistical computing. http://www.R-project.org/
36. RStudio Team. RStudio: Integrated Development for R. . RStudio, Inc., Boston, MA http://www.rstudio.com/
37. Los Alamos National Laboratory tools. Highlighter tool. http://www.hiv.lanl.gov/
38. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*. Oct 15 1998;17(19):2265-81. doi:10.1002/(sici)1097-0258(19981015)17:19<2265::aid-sim918>3.0.co;2-b
39. Ho DE, et al. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*. 2011;42(8):1--28.
40. Kidd M, et al. S-variant SARS-CoV-2 lineage B1.1.7 is associated with significantly higher viral loads in samples tested by ThermoFisher TaqPath RT-qPCR. *J Infect Dis*. Feb 13 2021;doi:10.1093/infdis/jiab082
41. Frampton D, et al. Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect Dis*. Apr 12 2021;doi:10.1016/S1473-3099(21)00170-5
42. McCormick KD, et al. The emerging plasticity of SARS-CoV-2. *Science*. Mar 26 2021;371(6536):1306-1308. doi:10.1126/science.abg4493
43. Greaney AJ, et al. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe*. Jan 13 2021;29(1):44-57 e9. doi:10.1016/j.chom.2020.11.007
44. McCarthy KR, et al. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science*. Feb 3 2021;doi:10.1126/science.abf6950
45. Washington NL, et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell*. May 13 2021;184(10):2587-2594 e7. doi:10.1016/j.cell.2021.03.052
46. Davies NG, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*. 2021:eabg3055. doi:10.1126/science.abg3055
47. Volz E, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*. Mar 25 2021;doi:10.1038/s41586-021-03470-x
48. Kissler S, et al. Densely sampled viral trajectories suggest longer duration of acute infection with B.1.1.7 variant relative to non-B.1.1.7 SARS-CoV-2. *Preprint*. 2021;
49. Ratcliff J, et al. Virological and serological characterization of critically ill patients with COVID-19 in the UK: Interactions of viral load, antibody status and B.1.1.7 variant infection. *J Infect Dis*. May 24 2021;doi:10.1093/infdis/jiab283
50. Davies NG, et al. Increased hazard of death in community-tested cases of SARS-CoV-2 Variant of Concern 202012/01. *medRxiv*. 2021:2021.02.01.21250959. doi:10.1101/2021.02.01.21250959
51. Iacobucci G. Covid-19: New UK variant may be linked to increased death rate, early data indicate. *BMJ*. Jan 26 2021;372:n230. doi:10.1136/bmj.n230
52. Challen R, et al. Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. *BMJ*. 2021;372:n579. doi:10.1136/bmj.n579

medRxiv preprint doi: https://doi.org/10.1101/2021.07.05.21259547; this version posted July 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

53. Grint DJ, et al. Case fatality risk of the SARS-CoV-2 variant of concern B.1.1.7 in England, 16 November to 5 February. *Eurosurveillance*. 2021;26(11):2100256. doi:doi:https://doi.org/10.2807/1560-7917.ES.2021.26.11.2100256

54. Collier DA, et al. Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature*. 2021/03/11 2021;doi:10.1038/s41586-021-03412-7

55. Starr TN, et al. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell*. Sep 3 2020;182(5):1295-1310 e20. doi:10.1016/j.cell.2020.08.012

56. Gu H, et al. Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science*. 2020;369(6511):1603-1607. doi:10.1126/science.abc4730

57. Meng B, et al. Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the variant of concern lineage B.1.1.7. *Cell reports*. doi:10.1016/j.celrep.2021.109292

58. Shen X, et al. SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral spike vaccines. *Cell Host & Microbe*. 2021;doi:10.1016/j.chom.2021.03.002

59. Emary KRW, et al. Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): an exploratory analysis of a randomised controlled trial. *The Lancet*. doi:10.1016/S0140-6736(21)00628-0

60. Munitz A, et al. BNT162b2 vaccination effectively prevents the rapid rise of SARS-CoV-2 variant B.1.1.7 in high-risk populations in Israel. *Cell Rep Med*. May 18 2021;2(5):100264. doi:10.1016/j.xcrm.2021.100264

61. Marks M, et al. Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *Lancet Infect Dis*. May 2021;21(5):629-636. doi:10.1016/S1473-3099(20)30985-3