

1 **Title:** Viral and host factors associated with SARS-CoV-2 disease severity in Georgia, USA

2 **Authors:** Ludy R. Carmola^{1*}, Allison Dorothy Roebing^{2,3,4}, Dara Khosravi¹, Rose M.

3 Langsjoen¹, Andrei Bombin^{1,4}, Bri Bixler⁵, Alex Reid¹, Cara Chen¹, Ethan Wang¹, Yang Lu¹,

4 Ziduo Zheng⁶, Rebecca Zhang⁶, Phuong-Vi Nguyen⁴, Robert A. Arthur⁷, Eric Fitts¹, Dalia Arafat

5 Gulick⁸, Dustin Higginbotham⁸, Azmain Taz¹, Alaa Ahmed^{1,9}, John Hunter Crumpler¹, Colleen

6 Kraft^{1,4}, Wilbur A. Lam^{10,11,12,13}, Ahmed Babiker^{1,4}, Jesse J. Waggoner⁴, Kyle P. Openo^{2,3,4},

7 Laura M. Johnson¹⁴, Adrianna Westbrook¹⁴, Anne Piantadosi^{1,4*}

8 **Affiliations:**

9 ¹Department of Pathology and Laboratory Medicine; Emory University School of Medicine;
10 Atlanta, GA, 30322; USA.

11 ²Georgia Emerging Infections Program; Georgia Department of Health; Atlanta, GA, 30303;
12 USA.

13 ³Atlanta Veterans Affairs Medical Center; Decatur, GA, 30033; USA.

14 ⁴Division of Infectious Diseases; Department of Medicine, Emory University School of
15 Medicine; Atlanta, GA, 30322; USA.

16 ⁵Graduate Program in Genetics and Molecular Biology, Emory University; Atlanta, GA, 30322;
17 USA.

18 ⁶Department of Biostatistics and Bioinformatics; Rollins School of Public Health, Emory
19 University; Atlanta, GA, 30322; USA

20 ⁷Emory Integrated Computational Core; Emory University School of Medicine; Atlanta, GA,
21 30322; USA.

22 ⁸Georgia Clinical & Translational Science Alliance; Emory University School of Medicine;
23 Atlanta, GA, 30322; USA.

24 ⁹Emory Integrated Genomics Core; Emory University School of Medicine; Atlanta, GA, 30322;
25 USA.

26 ¹⁰The Atlanta Center for Microsystems-Engineered Point-of-Care Technologies; Atlanta, GA,
27 30322; USA.

28 ¹¹Department of Pediatrics, Emory University School of Medicine; Atlanta, GA, 30322; USA.

29 ¹²Aflac Cancer and Blood Disorders Center at Children's Healthcare of Atlanta; Atlanta, GA,
30 30322; USA

31 ¹³Wallace H. Coulter Department of Biomedical Engineering, Emory University and Georgia
32 Institute of Technology, Atlanta, GA, USA.

33 ¹⁴Pediatric Biostatistics Core; Department of Pediatrics; School of Medicine; Emory University;
34 Atlanta, GA, 30322; USA

35 *Corresponding authors. Email: anne.piantadosi@emory.edu, ludy.registre@emory.edu

36 **Keywords:** SARS-CoV-2, post-vaccine infection, disease severity

37

38 **ABSTRACT**

39 While SARS-CoV-2 vaccines have shown strong efficacy, their suboptimal uptake
40 combined with the continued emergence of new viral variants raises concerns about the ongoing
41 and future public health impact of COVID-19. We investigated viral and host factors, including
42 vaccination status, that were associated with SARS-CoV-2 disease severity in a setting with low
43 vaccination rates. We analyzed clinical and demographic data from 1,957 individuals in the state
44 of Georgia, USA, coupled with viral genome sequencing from 1,185 samples. We found no
45 difference in disease severity between individuals infected with Delta and Omicron variants
46 among the participants in this study, after controlling for other factors, and we found no specific
47 mutations associated with disease severity. Compared to those who were unvaccinated,
48 vaccinated individuals experienced less severe SARS-CoV-2 disease, and the effect was similar
49 for both variants. Vaccination within 270 days before infection was associated with decreased
50 odds of moderate and severe outcomes, with the strongest association observed at 91-270 days
51 post-vaccination. Older age and underlying health conditions, especially immunosuppression and
52 renal disease, were associated with increased disease severity. Overall, this study provides
53 insights into the impact of vaccination status, variants/mutations, and clinical factors on disease
54 severity in SARS-CoV-2 infection when vaccination rates are low. Understanding these
55 associations will help refine and reinforce messaging around the crucial importance of
56 vaccination in mitigating the severity of SARS-CoV-2 disease.

57

58 INTRODUCTION

59 Vaccinations against SARS-CoV-2 have undeniably demonstrated high efficacy in
60 preventing COVID-19 infections and improving disease outcomes¹⁻³, but their impact is
61 challenged by the emergence of new variants carrying immune evasion mutations and the
62 waning of immune responses over time⁴. Partly due to these issues, public trust in COVID-19
63 vaccines has diminished with each round of booster recommendations, especially in the
64 southeast United States (US)⁵. Throughout the SARS-CoV-2 pandemic, the state of Georgia has
65 consistently held one of the lowest rates of vaccine coverage in the nation⁶. One important
66 approach to addressing vaccine hesitancy is to conduct studies – and disseminate their results –
67 to populations with low vaccine uptake.

68 Understanding the impact of vaccination on SARS-CoV-2 disease outcomes requires
69 consideration of other host factors such as underlying health conditions, sex, race, and
70 socioeconomic factors, which impact SARS-CoV-2 disease severity⁷⁻¹⁷. Viral factors are also
71 important, especially as variants emerge with distinctive properties affecting pathogenesis and
72 immune evasion. For example, the Delta variant has been associated with high risk of ICU
73 admission and mortality^{11,18,19}, while Omicron shows low neutralization sensitivity to vaccine
74 induced immunity²⁰.

75 In order to elucidate viral and host factors that contribute to disease outcomes in the
76 context of low vaccination rates, we analyzed demographic and clinical data from 1,957
77 individuals in the state of Georgia. We also analyzed full viral genome sequences from 1,185 of
78 these individuals to examine the influence of variants and mutations on disease severity. We
79 leveraged a large study population and extensive demographic, clinical, and sequence data to

80 define factors associated with SARS-CoV-2 disease severity in a region marked by low
81 vaccination rates.

82

83 **RESULTS**

84 **Clinical and demographic factors differ by SARS-CoV-2 vaccine status**

85 Between May 2021 and May 2022, we identified 1,957 individuals who tested positive
86 for SARS-CoV-2 within the Emory Healthcare system in Atlanta, Georgia. The majority of
87 participants (66%) were residents of the Metro Atlanta area (Table 1, Figure 1A). The median
88 age was 51 years (Interquartile range [IQR]=36,65), and 56% of individuals were female. The
89 racial distribution of participants was predominantly Black (58%) or White (31%). Individuals
90 experienced a range of clinical presentations and outcomes, from asymptomatic infection (15%)
91 to death (2.8%) (Table 2). Among the 967 individuals in this study who were hospitalized, 625
92 (65%) of the hospitalizations were due to COVID-19.

93 During the period of this study, the state of Georgia had the 7th lowest vaccination rate in
94 all 50 United States and the District of Columbia, with 55.1% of the population vaccinated
95 (Figure 1B). In our study, a slightly lower proportion (48%) of individuals were vaccinated, in
96 part based our study was designed to ensure inclusion of unvaccinated individuals. Vaccinated
97 individuals were significantly older than unvaccinated individuals (median age of 58 years vs 43
98 years, $p<0.001$) (Table 1). We observed a large disparity in vaccine status by race; among the
99 unvaccinated participants, 70% were Black and 21% White, while among the vaccinated
100 participants, 45% were Black and 43% were White (Table 1). Vaccination status was not
101 significantly associated with any other demographic variable evaluated.

102 We found that vaccinated individuals were more likely to have underlying medical
103 comorbidities than unvaccinated individuals. Notably, 28% of vaccinated individuals were
104 immunocompromised, compared to 17% of unvaccinated individuals ($p < 0.001$, Table 2).
105 Vaccinated individuals were also more likely to have hypertension (56% vs. 39%, $p < 0.001$),
106 cardiovascular disease (36% vs. 23%, $p < 0.001$), diabetes (28% vs. 19%, $p < 0.001$), renal disease
107 (24% vs. 11%, $p < 0.001$), and autoimmune disease (7% vs. 4%, $p = 0.03$) (Table S1).
108 Vaccinated individuals were less likely to be pregnant (1% vs. 6%, $p < 0.001$) (Table 2, Table
109 S1), however, at the time of the study, vaccines were not yet approved for pregnant individuals.
110 These findings are consistent with higher rates of vaccination in individuals with medical
111 comorbidities.

112 We also found differences in clinical symptoms by vaccination status. Compared to
113 unvaccinated individuals, those who were vaccinated were less likely to have fever (52% vs.
114 42%, $p < 0.001$), chills (38% vs. 32%, $p = 0.02$), nausea/vomiting (35% vs. 25%, $p < 0.001$), and
115 shortness of breath or difficulty breathing (49% vs. 39%, $p < 0.001$); they were more likely to
116 have sore throat (17% vs. 22%, $p = 0.02$) and runny nose/nasal congestion (27% vs. 45%, p
117 < 0.001) (Table S1). These findings are consistent with milder disease in vaccinated individuals.

118 SARS-CoV-2 C_T value by qRT-PCR is often used as a rough proxy for viral load, and we
119 found several key factors associated with SARS-CoV-2 C_T . Because multiple qRT-PCR assays
120 were used within the Emory Healthcare system during this time, we controlled for assay
121 variability (Table 3). Vaccinated individuals had a slightly lower C_T than those who were
122 unvaccinated (-0.70 , $SE=0.26$, $p=0.01$). Though this finding seems counterintuitive, we attribute
123 it to vaccinated individuals presenting for testing earlier after symptom onset, compared to
124 unvaccinated individuals (Table 2). Symptom duration was significantly inversely associated

125 with C_T (Table 3). Specifically, after controlling for other variables, and compared to
126 asymptomatic individuals, those who were tested within 0-3 days after symptom onset had a
127 lower C_T value by 3.0 cycles (Standard Error [SE]=0.55, $p < 0.001$) and those who were tested
128 within 4-7 days had a lower C_T value by 1.62 cycles (SE=0.61, $p = 0.01$). Those who were tested
129 8 or more days from symptom onset did not have a significantly different C_T than those who
130 were asymptomatic ($p = 0.77$) (Table 3). No significant differences in C_T were observed
131 between variants, after adjusting for other factors (Table 3).

132 **Variant frequency differs between vaccinated and unvaccinated individuals**

133 We sequenced full SARS-CoV-2 genomes from residual nasopharyngeal swab samples
134 from 1,185 individuals. The minimum genome coverage of the samples was 76% and the median
135 sequencing depth was 1,804 (Supplementary Data File). Sequences were primarily Delta (68%)
136 and the BA.1 sublineage of the Omicron variant (23%), followed by other Omicron sublineages
137 (5.1%), Alpha (2.5%) and less than 1% each of Beta, Gamma, Mu, A.2.5, and B.1 (Table 4). In
138 Georgia, during the time of this study (May 2021- May 2022), Delta accounted for 55% of
139 infections, Omicron for 40%, and Alpha for 2.5% (Figure 1C, Table S2). Therefore, our study
140 included a somewhat higher proportion of Delta and a lower proportion of Omicron than was
141 circulating in the state. The distribution of all other variants aligned with the overall variant
142 distribution observed in Georgia (Table S2, Table 4).

143 The distribution of SARS-CoV-2 variants in our study was different for vaccinated and
144 unvaccinated individuals ($p < 0.001$) (Table 4). Omicron had a higher frequency in vaccinated
145 individuals (32%) than unvaccinated (25%), whereas Alpha and Delta were more common in
146 unvaccinated compared to vaccinated individuals (Table 4). These differences correspond with

147 the timing of each variant’s circulation compared to vaccine rollout, though decreased vaccine
148 effectiveness against Omicron may also contribute²¹.

149 In addition to the frequency of VOCs among vaccinated and unvaccinated individuals,
150 we investigated the frequency of individual mutations. Within each VOC – Alpha, Delta, and
151 Omicron – no sequence characteristics, including mutations, deletions, and insertions, were
152 different between viruses infecting vaccinated and unvaccinated individuals (Figure 2, Figure
153 S1). However, the number of non-lineage defining mutations was lower in vaccinated compared
154 to unvaccinated individuals (Figure 1D) suggesting that vaccination may have an impact on viral
155 diversity within a host. Phylogenetic analysis demonstrated that sequences from vaccinated
156 individuals were intermixed with sequences from unvaccinated individuals, further confirming
157 no distinct features of post vaccination infections (Figure S2).

158 In-depth metagenomic analysis of 513 samples did not reveal any viral co-infections
159 (Supplementary Data File).

160 **Age, underlying conditions, and vaccination status are associated with disease severity**

161 We evaluated associations between disease severity and demographic characteristics,
162 underlying health conditions, vaccination status, and SARS-CoV-2 variant. Disease severity was
163 defined according to the WHO clinical progression scale²². Mild disease included asymptomatic
164 infection and symptomatic infection without hospitalization. Moderate disease included
165 hospitalized individuals without oxygen therapy or oxygen by mask or nasal prongs. Severe
166 disease included the use of oxygen by noninvasive or high flow, intubation and mechanical
167 ventilation, vasopressors, dialysis, or extracorporeal membrane oxygenation. Death included in-
168 hospital deaths directly linked to COVID.

169 Using an adjusted multinomial logistic regression model that included demographic
170 characteristics, underlying conditions, vaccination status and variant, we found that age, certain
171 underlying medical conditions, and time since most recent vaccination were significantly
172 associated with disease severity (Table 5). Underlying health conditions significantly associated
173 with disease severity included chronic lung disease, renal disease, and the use of systemic
174 immunosuppressive therapy prior to hospitalization. Other conditions showed weak or non-
175 significant associations with disease severity, including pregnancy, diabetes, liver disease, and
176 autoimmune disease. The Omicron variant did not show a significant association with disease
177 severity compared to the Delta variant in the adjusted model although an association was
178 observed in an unadjusted model (Table 5, Table S3). Finally, compared to unvaccinated
179 individuals, vaccination within 91-180 days (about 3 – 6 months) was associated with lower odds
180 of moderate disease, severe disease, and death (Table 5). There were similar effects across all
181 disease severity outcomes when vaccination occurred between 181 – 270 days (about 6 to 9
182 months). When vaccination occurred more than 270 days ago, however, there were no significant
183 differences in disease severity when compared to the unvaccinated group. Additionally,
184 vaccination that had occurred within the past 90 days (about 3 months) was associated with
185 lower odds of moderate disease relative to mild disease but was not associated with lower odds
186 of severe disease or death. In short, vaccinations were most protective against moderate disease,
187 severe disease, and death when they occurred within the prior 3-9 months. Overall, after
188 controlling for multiple host and viral factors, we found that age, chronic lung disease, renal
189 disease, and immunosuppressive therapy increased the odds of progressively more severe
190 COVID-19 disease, while vaccination decreased the odds and SARS-CoV-2 variant had no
191 effect.

192

193 **DISCUSSION**

194 Our comprehensive analysis of host and viral factors associated with SARS-CoV-2
195 disease severity in a setting with low vaccination rates led to several key findings. First, age and
196 underlying health conditions – especially chronic lung disease, renal disease, and the use of
197 immunosuppressive therapy – were associated with more severe disease and death. Second,
198 SARS-CoV-2 variant and viral mutations were not associated with disease severity in this study
199 population, which was comprised of individuals who sought medical care. And third, vaccination
200 was protective against severe outcomes for both Delta and Omicron variants to a similar degree.
201 Unique features of our study included the analysis of a large number of SARS-CoV-2 full viral
202 genome sequences linked to extensive clinical and demographic data, and our focus on a
203 relatively under-studied region of the U.S.

204 Georgia is an important proxy for the southeastern U.S. and other populations with high
205 numbers of vaccine refusals, inequitable access to healthcare, and low insurance coverage²³⁻²⁵.
206 Emphasizing the positive impact of SARS-CoV-2 vaccination among this population, similar to
207 others in the U.S.^{26,27}, is critical as new variants emerge. It is also important to note that among
208 the individuals in this study who contracted SARS-CoV-2 after being vaccinated, a greater
209 proportion reported milder upper respiratory symptoms like sore throat and runny nose, while a
210 lower percentage experienced more severe symptoms such as nausea/vomiting, fever, and
211 shortness of breath/difficulty breathing, in comparison to unvaccinated individuals. Vaccinations
212 were most protective against severe COVID outcomes when they occurred within the prior 3-9
213 months. This finding has timely implications on a national level, given persistently low vaccine
214 uptake, especially of the bivalent SARS-CoV-2 vaccine⁵, and the need for ongoing vaccine

215 updates targeting emerging variants⁴. Results from our study will help emphasize the benefits of
216 vaccination to the public as a means of safeguarding against severe COVID outcomes.

217 Our results also indicated the importance of demographic and clinical factors associated
218 with SARS-CoV-2 disease severity, despite vaccination. Age was a key risk factor; after
219 accounting for vaccination status, demographic factors, health conditions, and SARS-CoV-2
220 variant, our analysis revealed that for each additional year of age, the odds of experiencing more
221 severe outcomes compared to mild disease increased by 5%. The association between age and
222 disease severity has been consistently observed, particularly among individuals aged 65 and
223 above⁷⁻¹⁰. Thus, relying solely on vaccination may be insufficient for reducing disease severity
224 and mortality among older individuals. U.S. Census Bureau data indicates that the population is
225 aging, with Georgians aging at an even faster rate²⁸, underscoring the need for additional
226 preventative and treatment measures.

227 In addition to age, we found that chronic lung disease, renal disease, and the use of
228 immunosuppressive therapy also increased the odds of experiencing moderate and/or severe
229 infection and/or death. In previous studies, cardiovascular disease, diabetes, chronic respiratory
230 conditions, obesity, and compromised immune systems have also been found to increase the risk
231 of severe illness¹⁰⁻¹⁷. Interestingly, we found that cardiovascular disease was only associated with
232 moderate disease (without any association with severe disease or mortality), while diabetes and
233 being overweight were not associated with disease severity in our final multivariate model. One
234 explanation for this discrepancy could be that these two conditions serve as indicators for factors
235 that we controlled in our study. It is noteworthy that when we did not adjust for any factors, these
236 conditions were significantly associated with increased odds of severe disease.

237 It is surprising that we found no difference in disease severity between Delta and
238 Omicron variants, since multiple prior studies have found that the risk of hospitalization, ICU
239 admission, and mortality vary by variant^{11,18,19,29}. This discrepancy may be due to the fact that all
240 individuals in our study sought medical care, so we did not include individuals with minimal
241 symptoms. Our observation that vaccination was similarly protective for individuals with Delta
242 and Omicron is consistent with results from another recent study among hospitalized patients³⁰.

243 We did not find viral factors associated with post-vaccine infection. Most post-vaccine
244 infections were caused by the predominant lineage of the time. Within each variant (Delta and
245 Omicron), no SARS-CoV-2 SNPs, deletions, or insertions were more common in vaccinated
246 individuals than unvaccinated individuals. These results are different from a prior study of
247 similar size, which found more resistance mutations (e.g. L452* and E484*) in vaccinated
248 compared to unvaccinated individuals in the pre-Omicron era¹⁶. Our negative finding likely
249 reflects the challenge of identifying the effect of an individual virus mutation in an increasingly
250 complex immune landscape. Interestingly, we found that vaccinated individuals had fewer non-
251 lineage-defining SNPs than unvaccinated individuals, suggesting less diversity and potentially
252 less viral evolution within vaccinated individuals. This is consistent with a recent study
253 investigating within-host genetic diversity of SARS-CoV-2 in unvaccinated and vaccinated
254 individuals.³¹

255 Our study had several limitations: We only included individuals who presented to care,
256 thus skewing our study population towards individuals with more severe disease than the general
257 population. In addition, due to our study design, we could not collect reliable data regarding prior
258 SARS-CoV-2 infection(s) thus, we were unable to account for natural immunity. Given the
259 retrospective study design, there may be residual confounding, however we adjusted for

260 important variables such age, demographics, pre-existing health conditions and vaccination
261 status.

262 In summary, our findings underscore the critical role of vaccination status, age, and
263 medical comorbidities – especially immunosuppression, chronic kidney disease, and chronic
264 lung disease – in determining disease severity and outcomes among SARS-CoV-2 infected
265 individuals, regardless of virus variant. We contribute valuable insights into the nuanced
266 relationship between these factors, highlighting the importance of considering demographic,
267 clinical, and genetic variables when evaluating disease severity. Ultimately, our results will be
268 valuable in strengthening and reinforcing messaging around SARS-CoV-2 vaccination,
269 especially in settings of low vaccine uptake.

270

271 **METHODS**

272 **Clinical and demographic data**

273 This study was approved by the institutional review board at Emory University under
274 protocol STUDY00000260, with a waiver of consent. All positive SARS-CoV-2 samples from
275 Emory University Hospital Molecular and Microbiology Laboratories collected between 5/3/21
276 and 5/31/22 were reviewed for inclusion in this study. In addition to symptomatic testing, SARS-
277 CoV-2 tests were administered before admission or outpatient procedure as part of Emory
278 Healthcare System’s universal SARS-CoV-2 screening. Individuals were considered vaccinated
279 if they had received a complete vaccine series (2 doses of the Pfizer-BioNTech or Moderna
280 vaccines or 1 dose of the Janssen vaccine) at least 14 days before their first positive test result.
281 Individuals were excluded from the study if they were partially vaccinated or reported out-of-

282 state residency. From May 2021- September 2021, individuals were included in the study on a
283 case-match basis; for each post-vaccine case identified, 2-3 non-vaccinated control cases were
284 selected at random from the positive samples tested in the same calendar week. From October
285 2021- May 2022, all SARS-CoV-2 positive individuals identified who met inclusion criteria
286 were included.

287 The Centers for Disease Control and Prevention (CDC)-funded Georgia Emerging
288 Infections Program (GA EIP) performs active, population- and laboratory- based surveillance for
289 hospitalized cases of SARS-CoV-2 in metropolitan Atlanta, GA (population ~4 million). Patient
290 vaccination status was retrieved by GA EIP from the Georgia Registry of Immunization
291 Transactions and Services (GRITS) database. State of residency was retrieved from Georgia’s
292 State Electronic Notifiable Disease Surveillance System (SENDSS). Patient demographics,
293 SARS-CoV-2 RT-PCR results, and C_T value, underlying medical conditions, symptomatic
294 illness, hospitalization, and disease outcome were obtained from the electronic medical record
295 (EMR). In constructing symptom categories, systemic symptoms were defined as fatigue, fever,
296 chills, rigors, myalgia, and headache. Gastrointestinal symptoms were defined as
297 nausea/vomiting and diarrhea. Upper respiratory symptoms were defined as sore throat, runny
298 nose, and nasal congestion. Lower respiratory symptoms were defined as cough, and shortness of
299 breath/difficulty breathing. Immunosuppressed was defined as HIV, active cancer, autoimmune
300 disease, or immunosuppressive therapy. “Other” underlying conditions were defined as
301 overweight, diabetes, renal, cardiovascular, pregnant, and liver disease.

302 Disease severity was defined according to the WHO clinical progression scale²². The
303 scale includes mild disease- asymptomatic and symptomatic SARS-CoV-2 infection without
304 hospitalization, moderate disease- hospitalization without oxygen therapy or hospitalization with

305 oxygen by mask or nasal prongs, severe disease- hospitalization with use of oxygen by
306 noninvasive or high flow, intubation and mechanical ventilation, vasopressors, dialysis, or
307 extracorporeal membrane oxygenation, and death.

308 Data management and cleaning were conducted in Excel v16.73 and SAS studio v3.81.

309 **SARS-CoV-2 sequencing and analysis**

310 Residual nasopharyngeal (NP) swab samples were obtained from the Emory University
311 Hospital Molecular and Microbiology Laboratories. NP samples underwent RNA extraction,
312 DNase treatment, and cDNA synthesis followed by metagenomic or amplicon-based library
313 construction. For metagenomic sequencing, Nextera XT (Illumina) and Illumina sequencing
314 were performed as previously described³². Amplicon-based sequencing was performed using the
315 xGEN SARS-CoV-2 kit (IDT) as previously described³³.

316 Reference-based SARS-CoV-2 genome assembly was performed using viral-ngs
317 v2.1.12.0³⁴ or Viralrecon³⁵ for metagenomic and amplicon sequencing, respectively, with
318 reference strain NC_045512. SARS-CoV-2 lineages were determined using Pangolin³⁶.
319 Sequences were aligned and visualized in Geneious Prime (<https://www.geneious.com>).
320 Consensus-level single nucleotide polymorphisms (SNPs) and insertions/deletions were
321 identified using the Nextstrain web-based mutation calling tool³⁷.

322 For phylogenetic analysis, 411,634 reference sequences, collected between May 1, 2021,
323 and May 31, 2022, were downloaded from NCBI and were aligned with our study sequences to
324 reference strains Wuhan/Hu-1/2019 and Wuhan/WHO/2019 using Nextalign within the
325 Nextstrain v3.2.4 pipeline 7. This dataset was subsampled in Nextstrain using a custom scheme,
326 in which crowd penalty was set to 0.0 to select 1000 sequences most genetically similar to our

327 sequence dataset. Maximum likelihood phylogenetic trees were constructed using default
328 settings of the Nextstrain SARS-CoV-2 Workflow with TreeTime v0.8.6³⁸.

329 **Viral metagenomic analysis**

330 To assess the presence of viral co-infections in 513 samples that underwent metagenomic
331 sequencing, reads were first passed through a pre-processing pipeline including deduplication
332 with Clumpify.sh in the BMap tools (<https://sourceforge.net/projects/bbmap/>). Deduplicated
333 reads were trimmed with Trimmomatic Version 0.40 and filtered for quality, with flags
334 leading:3, trailing:3, slidingwindow:4:15, minlen:36
335 (<https://github.com/usadellab/Trimmomatic>). Pre-processed reads were run through kraken2
336 v2.1.3 against the k2_pluspf_20210127 database to assign each read to a taxonomic group, then
337 adjusted for significance with Bracken. Within the Kraken Tools packages, the
338 extract_kraken_reads.py script was used to separate reads by taxonomic ID for human
339 taxID_hg="9606", bacteria taxID_bac="2", fungus taxID_fungus="4751", viruses
340 taxID_virus="10239", and COVID-19 taxID_COVID="2697049". Custom shell and R scripts
341 were used to determine if the following viruses were found in each sample:
342 Human mastadenovirus C taxID=129951, Coronavirus HKU1 taxID=443239, Coronavirus
343 NL63 taxID=277944, Coronavirus 299E taxID=11137, Coronavirus OC43 taxID=31631, SARS-
344 CoV-2 taxID=2697049, Paramyxoviridae taxID=11158, Human metapneumovirus
345 taxID=162145, Parainfluenza virus taxID=2905673, Respiratory syncytial virus taxID=12814,
346 Picornaviridae taxID=12058, Rhinovirus taxID=31708, Enterovirus taxID=12059,
347 Orthomyxoviridae taxID=11308, Influenza A taxID=382835, and Influenza B taxID=11520.
348

349 **Statistical Analysis**

350 Demographics, symptoms, underlying conditions, and outcomes were described using
351 frequency distributions for categorical variables and medians and interquartile ranges for
352 continuous variables. Subgroup differences were evaluated to compare individuals who were
353 vaccinated with individuals who were not vaccinated, using chi-square tests and Fisher's exact
354 tests for categorical variables and Wilcoxon rank-sum test for continuous variables.

355 Prior to testing, the association between clinical factors and mean C_T was compared
356 between qRT-PCR testing platforms using ANOVA. There were significant differences between
357 the platforms indicating that platform is an important covariate to control for while modeling
358 factors associated with C_T .

359 Missing values were imputed using the chained equations algorithm in the MICE R
360 package³⁹ (R Version 4.1.3) to create 10 imputed data sets. Predictive mean matching, logistic
361 regression imputation, and polytomous regression imputation were used for numerical, binary,
362 and multcategory variables, respectively. Results from the 10 linear regression models ran on
363 the imputed datasets were then pooled according to Rubin's rule to provide an overall estimate
364 for the variables.

365 Multinomial logistic regressions were used to test the association between demographic
366 characteristics, underlying conditions, vaccine status, and SARS-CoV-2 variant with disease
367 severity (1=Mild [Reference Category], 2=Moderate, 3=Severe, 4=Death). First, unadjusted
368 models tested each variable's association with disease severity separately. Next, multivariable
369 models were constructed in a step-wise fashion, adding variables in blocks of demographic
370 variables, including age (continuous), sex (0=Female[Reference], 1=Male), and
371 race(0=Black[Reference], 1=White), followed by a block of underlying condition variables, all
372 of which were binary (0=No, 1=Yes; pregnant, chronic lung disease, hypertension, overweight,

373 cardiovascular disease, diabetes, renal disease, liver disease, autoimmune disease,
374 immunocompromised, systemic immunosuppressive therapy/medications), and the final model
375 added a block of SARS-CoV-2-related characteristics including vaccination status
376 (1=Unvaccinated [Reference], 2=Vaccinated, 3=Vaccinated and Boosted), Days Since Most
377 Recent Vaccination/Booster (1=Unvaccinated [Reference], 2=Within past 90 Days, 3=91 - 180
378 Days Ago, 4=181 - 270 Days Ago, 5=More than 270 Days Ago) and variant
379 (1=Delta[Reference], 2=Omicron). The multivariable models were tested for multicollinearity
380 using variance inflation factor (VIF), though collinearity was not present, so all variables were
381 retained.

382 All analyses were conducted in R Version 4.1.3 (R Foundation for Statistical Computing,
383 Vienna, Austria).

384 **ACKNOWLEDGEMENTS:**

385 The study was supported by Centers for Disease Control and Preventions contract
386 75D30121C10084 under BAA ERR 20-15-2997 (to A.P) and by the Emory WHSC COVID-19
387 Urgent Research Engagement (CURE) Center, made possible by generous philanthropic support
388 from the O. Wayne Rollins Foundation and the William Randolph Hearst Foundation (to A.P.).
389 L.R.C. was supported by Award Number T32AI074492 from the National Institute of Allergy
390 and Infectious Diseases. EIP Surveillance of COVID-19 was funded through the Centers for
391 Disease Control and Preventions Emerging Infections Program [U50CK000485]. The study was
392 supported by the National Institute of Biomedical Imaging and Bioengineering at the National
393 Institutes of Health under award U54 EB027690 (to W.A.L.) and the National Center for
394 Advancing Translational Sciences of the National Institutes of Health under Award Number
395 UL1TR002378 (to W.A.L.). Additional support was provided by the Georgia Clinical &

396 Translational Science Alliance of the National Institutes of Health under Award Number
397 UL1TR002378. The study was supported in part by the Emory Integrated Genomics Core
398 (EIGC) and Emory Integrated Computational Core (EICC), which are subsidized by the Emory
399 University School of Medicine and are one of the Emory Integrated Core Facilities. The study
400 was supported by the Data Analytics and Pediatric Biostatistics Core, Department of Pediatrics,
401 Emory University School of Medicine. The content is solely the responsibility of the authors and
402 does not necessarily represent the official views of the National Institute of Allergy and
403 Infectious Diseases or the National Institutes of Health. The funders played no role in study
404 design, data collection and analysis, decision to publish, or preparation of the manuscript.

405

406 **Data Availability**

407 All sequence data are available in NCBI under BioProject PRJNA634356. The GISAID
408 accession number for each sequence is listed in the Supplementary Data File.

409

410 **Author contributions:**

411 Conceptualization: LRC, ADR, AB, AP

412 Methodology: LRC, KPO, JJW, AP, RZ, ZZ, RA, LJ, AW

413 Investigation: LRC, ADR, JJW, DHK, ECF, AP, PN, AT

414 Validation: ADR, DHK, AP

415 Visualization: LRC, EW, YL, RML

416 Data Curation: LRC, ADR, DHK, AB, AA, JHC

417 Funding acquisition: KPO, WAL, AP

418 Project administration: LRC, DHK, AP

419 Supervision: LRC, DAG, AP, RZ

420 Writing – original draft: LRC, RML, ZZ

421 Writing – review & editing: DHK, ECF, AB, AP, RML, RZ, RA, JJW, LJ, AW, WAL

422 Resources: DHK, ECF, DAG, AP, CK, DH

423 Statistical analyses: RZ, AA, ZZ, LJ, AW

424 Unrestricted access to all data: KPO, AP

425 First draft of the manuscript, reviewed it and edited it: LRC, AP

426 All authors agreed to submit the manuscript, read and approved the final draft and take full

427 responsibility of its content, including the accuracy of the data and the fidelity of the trial to the

428 registered protocol and its statistical analysis.

429 **Competing interests:**

430 The authors have declared that no competing interests exist.

431

432

433 **REFERENCES:**

- 434 1. Polack FP, Thomas SJ, Kitchin N, et al. Safety and Efficacy of the BNT162b2 mRNA
435 Covid-19 Vaccine. *N Engl J Med*. Dec 31 2020;383(27):2603-2615.
436 doi:10.1056/NEJMoa2034577
- 437 2. Baden LR, El Sahly HM, Essink B, et al. Efficacy and Safety of the mRNA-1273 SARS-
438 CoV-2 Vaccine. *New England Journal of Medicine*. 2020;384(5):403-416.
439 doi:10.1056/NEJMoa2035389
- 440 3. Sadoff J, Gray G, Vandebosch A, et al. Safety and Efficacy of Single-Dose
441 Ad26.COV2.S Vaccine against Covid-19. *New England Journal of Medicine*.
442 2021;384(23):2187-2201. doi:10.1056/NEJMoa2101544
- 443 4. Food and Drug Administration. Updated COVID-19 Vaccines for Use in the United
444 States Beginning in Fall 2023. Updated June 16, 2023. Accessed August 10, 2023,
445 [https://www.fda.gov/vaccines-blood-biologics/updated-covid-19-vaccines-use-united-states-](https://www.fda.gov/vaccines-blood-biologics/updated-covid-19-vaccines-use-united-states-beginning-fall-2023)
446 [beginning-fall-2023](https://www.fda.gov/vaccines-blood-biologics/updated-covid-19-vaccines-use-united-states-beginning-fall-2023)
- 447 5. Centers for Disease Control and Prevention. COVID-19 Vaccinations in the United
448 States. Updated May 11, 2023. Accessed August 10, 2023, [https://covid.cdc.gov/covid-data-](https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-people-booster-percent-pop5)
449 [tracker/#vaccinations_vacc-people-booster-percent-pop5](https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-people-booster-percent-pop5)
- 450 6. Murthy BP, Sterrett N, Weller D, et al. Disparities in COVID-19 Vaccination Coverage
451 Between Urban and Rural Counties - United States, December 14, 2020-April 10, 2021. *MMWR*
452 *Morb Mortal Wkly Rep*. May 21 2021;70(20):759-764. doi:10.15585/mmwr.mm7020e3
- 453 7. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult
454 inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*.
455 2020;395(10229):1054-1062. doi:10.1016/S0140-6736(20)30566-3
- 456 8. Antonelli M, Penfold RS, Merino J, et al. Risk factors and disease profile of post-
457 vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: a
458 prospective, community-based, nested, case-control study. *Lancet Infect Dis*. Sep 1
459 2021;doi:10.1016/S1473-3099(21)00460-6
- 460 9. Chemaitelly H, Tang P, Hasan MR, et al. Waning of BNT162b2 Vaccine Protection
461 against SARS-CoV-2 Infection in Qatar. *N Engl J Med*. Dec 9 2021;385(24):e83.
462 doi:10.1056/NEJMoa2114114
- 463 10. Levin EG, Lustig Y, Cohen C, et al. Waning Immune Humoral Response to BNT162b2
464 Covid-19 Vaccine over 6 Months. *N Engl J Med*. Dec 9 2021;385(24):e84.
465 doi:10.1056/NEJMoa2114583
- 466 11. Fisman DN, Tuite AR. Evaluation of the relative virulence of novel SARS-CoV-2
467 variants: a retrospective cohort study in Ontario, Canada. *Cmaj*. Oct 25 2021;193(42):E1619-
468 e1625. doi:10.1503/cmaj.211248
- 469 12. Lustig Y, Sapir E, Regev-Yochay G, et al. BNT162b2 COVID-19 vaccine and correlates
470 of humoral immune responses and dynamics: a prospective, single-centre, longitudinal cohort
471 study in health-care workers. *Lancet Respir Med*. Sep 2021;9(9):999-1009. doi:10.1016/S2213-
472 2600(21)00220-4
- 473 13. Russell CD, Lone NI, Baillie JK. Comorbidities, multimorbidity and COVID-19. *Nature*
474 *Medicine*. 2023/02/01 2023;29(2):334-343. doi:10.1038/s41591-022-02156-9
- 475 14. Di Fusco M, Moran MM, Cane A, et al. Evaluation of COVID-19 vaccine breakthrough
476 infections among immunocompromised patients fully vaccinated with BNT162b2. *J Med Econ*.
477 Jan-Dec 2021;24(1):1248-1260. doi:10.1080/13696998.2021.2002063

- 478 15. Johnson AG, Amin AB, Ali AR, et al. COVID-19 Incidence and Death Rates Among
479 Unvaccinated and Fully Vaccinated Adults with and Without Booster Doses During Periods of
480 Delta and Omicron Variant Emergence - 25 U.S. Jurisdictions, April 4-December 25, 2021.
481 *MMWR Morb Mortal Wkly Rep.* Jan 28 2022;71(4):132-138. doi:10.15585/mmwr.mm7104e2
482 16. Servellita V, Morris MK, Sotomayor-Gonzalez A, et al. Predominance of antibody-
483 resistant SARS-CoV-2 variants in vaccine breakthrough cases from the San Francisco Bay Area,
484 California. *Nat Microbiol.* Feb 2022;7(2):277-288. doi:10.1038/s41564-021-01041-4
485 17. Tenforde MW, Self WH, Adams K, et al. Association Between mRNA Vaccination and
486 COVID-19 Hospitalization and Disease Severity. *JAMA.* 2021;326(20):2043-2054.
487 doi:10.1001/jama.2021.19499
488 18. Lin L, Liu Y, Tang X, He D. The Disease Severity and Clinical Outcomes of the SARS-
489 CoV-2 Variants of Concern. Systematic Review. *Frontiers in Public Health.* 2021-November-30
490 2021;9doi:10.3389/fpubh.2021.775224
491 19. Funk T, Pharris A, Spiteri G, et al. Characteristics of SARS-CoV-2 variants of concern
492 B.1.1.7, B.1.351 or P.1: data from seven EU/EEA countries, weeks 38/2020 to 10/2021. *Euro*
493 *Surveill.* Apr 2021;26(16)doi:10.2807/1560-7917.Es.2021.26.16.2100348
494 20. Willett BJ, Grove J, MacLean OA, et al. SARS-CoV-2 Omicron is an immune escape
495 variant with an altered cell entry pathway. *Nature Microbiology.* 2022/08/01 2022;7(8):1161-
496 1179. doi:10.1038/s41564-022-01143-7
497 21. Chaguza C, Coppi A, Earnest R, et al. Rapid emergence of SARS-CoV-2 Omicron
498 variant is associated with an infection advantage over Delta in vaccinated persons. *Med.* May 13
499 2022;3(5):325-334 e4. doi:10.1016/j.medj.2022.03.010
500 22. Marshall JC, Murthy S, Diaz J, et al. A minimal common outcome measure set for
501 COVID-19 clinical research. *The Lancet Infectious Diseases.* 2020;20(8):e192-e197.
502 doi:10.1016/S1473-3099(20)30483-7
503 23. Medcalfe SK, Slade CP. Racial residential segregation and COVID-19 vaccine uptake: an
504 analysis of Georgia USA county-level data. *BMC Public Health.* Jul 20 2023;23(1):1392.
505 doi:10.1186/s12889-023-16235-0
506 24. Nguyen TH, Shah GH, Schwind JS, Richmond HL. Community Characteristics and
507 COVID-19 Outcomes: A Study of 159 Counties in Georgia, United States. *J Public Health*
508 *Manag Pract.* May-Jun 01 2021;27(3):251-257. doi:10.1097/phh.0000000000001330
509 25. Berman AE, Miller DD, Rahn DW, et al. A County-Level Analysis of Socioeconomic
510 and Clinical Predictors of COVID-19 Incidence and Case-Fatality Rates in Georgia, March-
511 September 2020. *Public Health Reports.* 2021/09/01 2021;136(5):626-635.
512 doi:10.1177/00333549211023267
513 26. Skarbinski J, Wood MS, Chervo TC, et al. Risk of severe clinical outcomes among
514 persons with SARS-CoV-2 infection with differing levels of vaccination during widespread
515 Omicron (B.1.1.529) and Delta (B.1.617.2) variant circulation in Northern California: A
516 retrospective cohort study. *Lancet Reg Health Am.* Aug 2022;12:100297.
517 doi:10.1016/j.lana.2022.100297
518 27. Lewnard JA, Hong VX, Patel MM, Kahn R, Lipsitch M, Tartof SY. Clinical outcomes
519 associated with SARS-CoV-2 Omicron (B.1.1.529) variant and BA.1/BA.1.1 or BA.2 subvariant
520 infection in Southern California. *Nat Med.* Sep 2022;28(9):1933-1943. doi:10.1038/s41591-022-
521 01887-z
522 28. United States Census Bureau. U.S. Older Population Grew From 2010 to 2020 at Fastest
523 Rate Since 1880 to 1890. Updated May 25,2023. Accessed August 10, 2023,

- 524 [https://www.census.gov/library/stories/2023/05/2020-census-united-states-older-population-](https://www.census.gov/library/stories/2023/05/2020-census-united-states-older-population-grew.html#:~:text=The%20older%20population%20increased%20by,total%20population%20(ab)
525 [grew.html#:~:text=The%20older%20population%20increased%20by,total%20population%20\(ab](https://www.census.gov/library/stories/2023/05/2020-census-united-states-older-population-grew.html#:~:text=The%20older%20population%20increased%20by,total%20population%20(ab)
526 [out%20200%25\)](https://www.census.gov/library/stories/2023/05/2020-census-united-states-older-population-grew.html#:~:text=The%20older%20population%20increased%20by,total%20population%20(ab).
- 527 29. Hyams C, Challen R, Marlow R, et al. Severity of Omicron (B.1.1.529) and Delta
528 (B.1.617.2) SARS-CoV-2 infection among hospitalised adults: a prospective cohort study in
529 Bristol, United Kingdom. *The Lancet Regional Health – Europe*.
530 2023;25doi:10.1016/j.lanpe.2022.100556
- 531 30. Robinson ML, Morris CP, Betz JF, et al. Impact of Severe Acute Respiratory Syndrome
532 Coronavirus 2 Variants on Inpatient Clinical Outcome. *Clinical Infectious Diseases*.
533 2022;76(9):1539-1549. doi:10.1093/cid/ciac957
- 534 31. Gu H, Quadeer AA, Krishnan P, et al. Within-host genetic diversity of SARS-CoV-2
535 lineages in unvaccinated and vaccinated individuals. *Nature Communications*. 2023/03/31
536 2023;14(1):1793. doi:10.1038/s41467-023-37468-y
- 537 32. Babiker A, Bradley HL, Stittleburg VD, et al. Metagenomic Sequencing To Detect
538 Respiratory Viruses in Persons under Investigation for COVID-19. *J Clin Microbiol*. Dec 17
539 2020;59(1)doi:10.1128/JCM.02142-20
- 540 33. Sexton ME, Waggoner JJ, Carmola LR, et al. Rapid Detection and Characterization of
541 Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Omicron Variant in a
542 Returning Traveler. *Clin Infect Dis*. Aug 24 2022;75(1):e350-e353. doi:10.1093/cid/ciac032
- 543 34. Park DJ, Dudas G, Wohl S, et al. Ebola virus epidemiology, transmission, and evolution
544 during seven months in Sierra Leone. *Cell*. 2015;161(7):1516-1526.
- 545 35. Patel H, Varona S, Monzón S, et al. nf-core/viralrecon: nf-core/viralrecon v2.2 - Tin
546 Turtle. 2021;doi:<https://doi.org/10.5281/zenodo.5146252>
- 547 36. Rambaut A, Holmes EC, O'Toole A, et al. A dynamic nomenclature proposal for SARS-
548 CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. Nov 2020;5(11):1403-1407.
549 doi:10.1038/s41564-020-0770-5
- 550 37. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution.
551 *Bioinformatics*. Dec 1 2018;34(23):4121-4123. doi:10.1093/bioinformatics/bty407
- 552 38. Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic
553 analysis. *Virus Evolution*. 2018;4(1)doi:10.1093/ve/vex042
- 554 39. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained
555 Equations in R. *Journal of Statistical Software*. 12/12 2011;45(3):1 - 67.
556 doi:10.18637/jss.v045.i03

557

558

560 **TABLES:**

Table 1. Demographic Characteristics by Vaccination Status

Variable	Overall, N = 1,957 ¹	Unvaccinated, N = 1,024 ¹	Vaccinated, N = 933 ¹	p ²
Age				
Age (in years)	51 (36, 65)	43 (31, 59)	58 (43, 70)	<0.001
Age (by decade)				<0.001
18-29	275 (14%)	216 (21%)	59 (6.3%)	
30-39	342 (17%)	225 (22%)	117 (13%)	
40-49	325 (17%)	178 (17%)	147 (16%)	
50-59	317 (16%)	152 (15%)	165 (18%)	
60-69	332 (17%)	133 (13%)	199 (21%)	
70-79	228 (12%)	80 (7.8%)	148 (16%)	
80-89	113 (5.8%)	31 (3.0%)	82 (8.8%)	
90+	25 (1.3%)	9 (0.9%)	16 (1.7%)	
Sex				0.10
Female	1,093 (56%)	590 (58%)	503 (54%)	
Male	864 (44%)	434 (42%)	430 (46%)	
Race				<0.001
American Indian/Alaska Native	64 (3.3%)	23 (2.2%)	41 (4.4%)	
Asian/Native Hawaiian/ Pacific Islander	3 (0.2%)	1 (<0.1%)	2 (0.2%)	
Black	1,137 (58%)	721 (70%)	416 (45%)	
White	611 (31%)	211 (21%)	400 (43%)	
Other	6 (0.4%)	5 (0.5%)	1 (0.1%)	
Unknown	134 (6.9%)	62 (6.1%)	72 (7.7%)	
(Missing)	2	1	1	
Ethnicity				0.99
Hispanic/Latino	64 (3.6%)	34 (3.6%)	30 (3.6%)	
Non-Hispanic/Latino	1,726 (96%)	918 (96%)	808 (96%)	
(Missing)	167	72	95	
Residence Region				0.28
Metro Atlanta	1,301 (66%)	692 (68%)	609 (65%)	
Other	656 (34%)	332 (32%)	324 (35%)	

¹ n (%) or median (Interquartile Range [IQR])

² Pearson's Chi-squared test or Fisher's exact test for categorical variables and Wilcoxon rank-sum test for continuous variables

Table 2. Clinical Characteristics by Vaccination Status

Variable	Overall, N = 1,957 ¹	Unvaccinated, N = 1,024 ¹	Vaccinated, N = 933 ¹	p ²
COVID Symptoms				
Any Systemic Symptoms ³	1,301 (82%)	713 (84%)	588 (81%)	0.08
(Missing)	379	175	204	
Any GI Symptoms ⁴	642 (42%)	387 (47%)	255 (37%)	<0.001
(Missing)	440	205	235	
Any Upper Respiratory ⁵ Symptoms	636 (43%)	281 (35%)	355 (53%)	<0.001
(Missing)	474	216	258	
Any Lower Respiratory ⁶ Symptoms	1,275 (80%)	675 (80%)	600 (80%)	0.71
(Missing)	364	177	187	
Symptom Duration at Time of Testing				<0.01
0-3 days	681 (43%)	338 (39%)	343 (48%)	
4-7 days	428 (27%)	255 (29%)	173 (24%)	
8+ days	238 (15%)	147 (17%)	91 (13%)	
Asymptomatic	233 (15%)	125 (14%)	108 (15%)	
(Missing)	377	159	218	
Underlying Conditions				
Any Immunosuppressed Underlying Condition ⁷	407 (22%)	162 (17%)	245 (28%)	<0.001
(Missing)	145	73	72	
Any Other Underlying Condition ⁸	1,164 (65%)	573 (60%)	591 (70%)	<0.001
(Missing)	155	71	84	
Disease Severity				0.70
Mild	1,148 (62%)	609 (62%)	539 (62%)	
Moderate	463 (25%)	251 (26%)	212 (24%)	
Severe	182 (9.9%)	92 (9.4%)	90 (10%)	
Death	51 (2.8%)	24 (2.5%)	27 (3.1%)	
(Missing)	113	48	65	
Vaccination				
Days Since Full Vaccination	198 (141, 272)	—	198 (141, 272)	
(Missing)	1,024	—	0	
Days Since Booster	115 (63, 163)	—	115 (63, 163)	
(Missing)	1,796	—	772	
Days Since Most Recent Vaccination/Booster				
Unvaccinated	1,024 (52%)	1,024 (100%)	—	
Within past 90 Days	136 (6.9%)	—	136 (15%)	
91 - 180 Days Ago	402 (21%)	—	402 (43%)	
181 - 270 Days Ago	283 (14%)	—	283 (30%)	
More than 270 Days Ago	112 (5.7%)	—	112 (12%)	
Viral Load				
C _T Value	23 (19, 30)	24 (19, 31)	22 (18, 27)	<0.001
(Missing)	339	164	175	

¹ n (%) or median (Interquartile Range [IQR])

² Pearson's Chi-squared test or Fisher's exact test for categorical variables and Wilcoxon rank-sum test for continuous variables

³ Any Systemic symptoms included fatigue, fever, chills, rigors, myalgia, and headache

⁴ Any GI symptoms included nausea/vomiting, and diarrhea

⁵ Any Upper Respiratory symptoms included sore throat, runny nose, and nasal congestion

⁶ Any Lower Respiratory symptoms included cough, shortness of breath, and difficulty breathing

⁷ Any Immunosuppressed Underlying Condition included HIV, active cancer, autoimmune disease, immunosuppressed, or immunosuppressive therapy

⁸ Any Other Underlying Condition included overweight, diabetes, renal, cardiovascular, pregnant and liver disease

562

563

Table 3. Association between C_T value and clinical factors

Variable	Beta	SE	p
Vaccination status			
Unvaccinated	Ref	—	—
Vaccinated	-0.70	0.26	0.01
Lineage			
Alpha	Ref	—	—
Delta	-0.23	0.87	0.79
Omicron	1.68	0.92	0.07
Symptom Duration at Time of Testing			
Asymptomatic	Ref	—	—
0-3 days	-2.99	0.55	<0.001
4-7 days	-1.62	0.61	0.01
8+ days	0.20	0.68	0.77
Assay			
Cepheid Gene Xpert	Ref	—	—
BioFire Defense, LLC - BioFire COVID-19 Test	-1.78	2.82	0.54
Cepheid - Xpert Xpress SARS-CoV-2/Flu/RSV	-0.39	2.85	0.89
Roche Cobas	-0.56	0.35	0.11
Other	4.37	7.81	0.59

SE = Standard Error, Ref = Reference Level

564

Table 4. SARS-CoV-2 Variants by Vaccination Status

Variable	Overall, N = 1,957¹	Unvaccinated, N = 1,024¹	Vaccinated, N = 933¹	p²
Variant				<0.001
A.2.5	1 (<0.1%)	0 (0%)	1 (0.2%)	
Alpha	30 (2.5%)	24 (4.0%)	6 (1.0%)	
B.1	1 (<0.1%)	0 (0%)	1 (0.2%)	
Beta	1 (<0.1%)	0 (0%)	1 (0.2%)	
Delta	805 (68%)	415 (70%)	390 (66%)	
Gamma	6 (0.5%)	5 (0.8%)	1 (0.2%)	
Mu	4 (0.3%)	4 (0.7%)	0 (0%)	
Omicron	337 (28%)	146 (25%)	191 (32%)	
(Missing)	772	430	342	

565

Table 5. Adjusted Multinomial Logistic Regression Models Testing the Association between Demographics, Underlying Health Conditions, and COVID-Related Characteristics with Disease Severity

Variable	Moderate ¹			Severe ¹			Death ¹		
	aOR ²	95% CI ²	p	aOR ²	95% CI ²	p	aOR ²	95% CI ²	p
<i>Demographics</i>									
Age	1.05	1.03, 1.06	<0.001	1.05	1.03, 1.07	<0.001	1.09	1.05, 1.13	<0.001
Male	1.28	0.89, 1.85	0.18	1.34	0.82, 2.19	0.24	1.21	0.52, 2.81	0.66
White	0.70	0.46, 1.06	0.09	0.89	0.51, 1.54	0.67	0.73	0.27, 1.96	0.53
<i>Underlying Conditions</i>									
Chronic Lung Disease	2.10	1.38, 3.17	<0.001	1.98	1.15, 3.41	0.01	1.31	0.50, 3.46	0.58
Hypertension	1.02	0.65, 1.59	0.94	0.97	0.53, 1.76	0.91	3.46	0.88, 13.57	0.08
Overweight	1.40	0.95, 2.07	0.09	1.17	0.69, 1.99	0.56	1.61	0.65, 4.01	0.30
Cardiovascular Disease	1.63	1.06, 2.50	0.03	1.45	0.82, 2.55	0.20	1.08	0.44, 2.63	0.87
Diabetes	0.99	0.64, 1.54	0.98	1.45	0.83, 2.54	0.20	1.09	0.43, 2.74	0.86
Renal Disease	2.64	1.61, 4.32	<0.001	1.91	1.01, 3.64	0.05	3.63	1.41, 9.36	0.01
Liver Disease	1.13	0.48, 2.69	0.78	1.84	0.69, 4.92	0.22	1.48	0.25, 8.78	0.67
Autoimmune Disease	1.78	0.87, 3.65	0.11	1.95	0.77, 4.95	0.16	3.96	1.00, 15.65	0.05
Immunocompromised ³	1.08	0.55, 2.11	0.82	0.59	0.25, 1.40	0.23	3.26	0.83, 12.71	0.09
Systemic Immunosuppressive Therapy or Meds.	3.45	1.88, 6.32	<0.001	5.01	2.38, 10.51	<0.001	1.48	0.38, 5.71	0.57
<i>Days Since Most Recent Vaccination/Booster</i>									
Unvaccinated	Ref	—		Ref	—		Ref	—	
Within past 90 Days⁴	0.37	0.16, 0.88	0.02	0.40	0.13, 1.26	0.12	0.72	0.18, 2.86	0.64
91 - 180 Days Ago⁴	0.32	0.19, 0.54	<0.001	0.42	0.21, 0.82	0.01	0.31	0.10, 0.96	0.04
181 - 270 Days Ago⁴	0.32	0.18, 0.56	<0.001	0.38	0.17, 0.80	0.01	0.14	0.04, 0.59	0.01
More than 270 Days ⁴ Ago	0.53	0.23, 1.21	0.13	0.88	0.34, 2.29	0.79	0.34	0.06, 1.95	0.23
<i>Lineage</i>									
Delta	Ref	—		Ref	—		Ref	—	
Omicron ⁵	0.95	0.61, 1.47	0.81	1.42	0.81, 2.50	0.22	2.07	0.84, 5.12	0.12

¹ "Mild" is the Reference Category

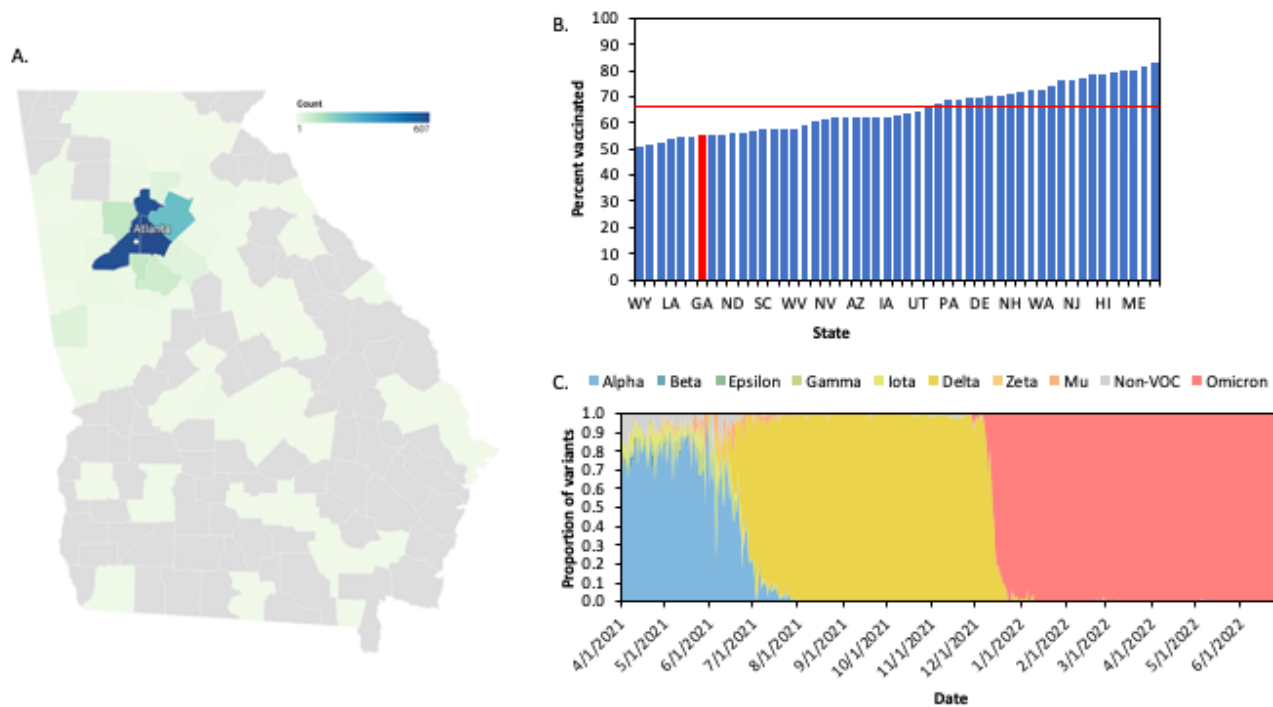
² aOR = Adjusted Odds Ratio, CI = Confidence Interval, Ref = Reference Level

³ HIV infection, active cancer, solid organ transplant, hematopoietic stem cell transplant

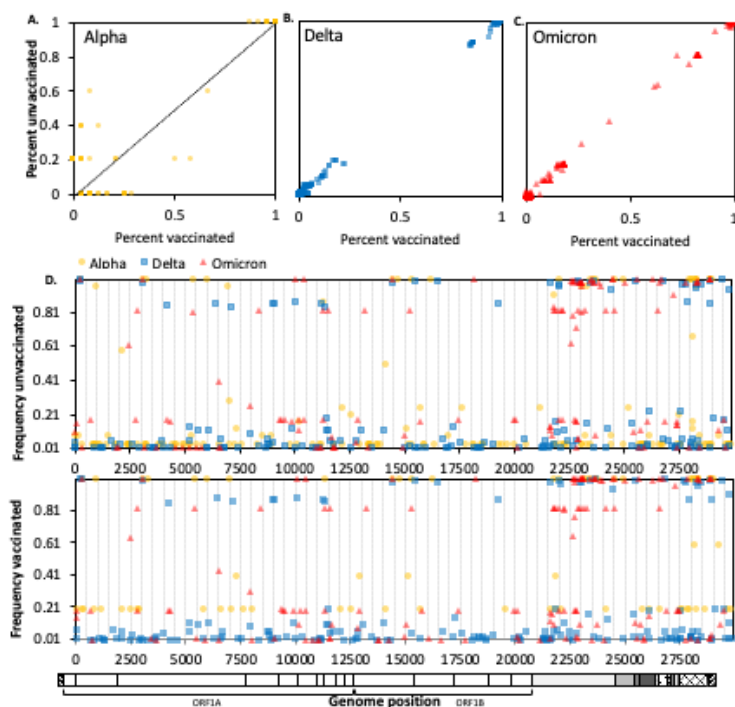
⁴ Compared to unvaccinated individuals

⁵ Compared to Delta

566 **FIGURES:**



567 **Figure 1. COVID-19 cases from May 2021-May 2022 in Georgia, US. A.** 1,957 Emory
568 Healthcare COVID-19 case mapped by counties in and around Atlanta, GA. **B.** Percent of
569 population fully vaccinated against SARS-CoV-2 by May 2022 in each state in the United States.
570 Red line represents national average (66.7%). Red bar represents Georgia. **C.** Proportion of
571 variants circulating in Georgia from April 2021- June 2022. Sequences obtained from GISAID.
572



573

574 **Figure 2. Frequencies of single nucleotide polymorphisms (SNPs) among SARS-COV-2**
575 **genome sequences from vaccinated and unvaccinated individuals.** Each point represents a
576 single SNP plotted by its frequency in sequences from unvaccinated individuals (Y-axis) versus
577 its frequency in sequences from vaccinated individuals (X-axis). Data is divided by WHO variant
578 classifications Alpha (A), Delta (B), and Omicron (C). Mutations observed along the diagonal
579 depict mutations observed equally among vaccinated and unvaccinated individuals. Mutations
580 observed moving away from the diagonal represent mutations observed in either vaccinated (X-
581 axis) or unvaccinated (Y-axis) individuals. (D) Frequency of SNPs in SARS-CoV-2 sequences
582 from unvaccinated (top) and vaccinated (bottom) individuals, by genome position (x-axis). In all
583 panels, Alpha is represented by yellow circles, Delta by blue squares, and Omicron by red
584 triangles.

585 SUPPLEMENTARY TABLES

Table S1. Extended COVID Symptoms and Underlying Conditions by Vaccination Status

Variable	Overall, N = 1,957 ¹	Unvaccinated, N = 1,024 ¹	Vaccinated, N = 933 ¹	p ²
COVID Symptoms				
Fever	720 (47%)	430 (52%)	290 (42%)	<0.001
(Missing)	435	197	238	
Chills	537 (35%)	314 (38%)	223 (32%)	0.02
(Missing)	435	198	237	
Rigors	29 (2.0%)	17 (2.1%)	12 (1.8%)	0.67
(Missing)	491	221	270	
Myalgia	501 (34%)	279 (35%)	222 (33%)	0.59
(Missing)	484	218	266	
Headache	521 (35%)	299 (37%)	222 (33%)	0.12
(Missing)	476	215	261	
Sore Throat	286 (20%)	138 (17%)	148 (22%)	0.02
(Missing)	496	228	268	
Nausea / Vomiting	451 (30%)	282 (35%)	169 (25%)	<0.001
(Missing)	455	209	246	
Diarrhea	369 (25%)	212 (26%)	157 (23%)	0.13
(Missing)	470	220	250	
Fatigue	630 (41%)	337 (41%)	293 (41%)	0.81
(Missing)	418	195	223	
Runny Nose or Nasal Congestion	519 (35%)	217 (27%)	302 (45%)	<0.001
(Missing)	473	215	258	
Cough	1,087 (70%)	576 (69%)	511 (72%)	0.37
(Missing)	414	195	219	
Shortness of Breath or Difficulty Breathing	690 (44%)	408 (49%)	282 (39%)	<0.001
(Missing)	401	188	213	
Loss of Taste	264 (18%)	151 (19%)	113 (17%)	0.33
(Missing)	505	233	272	
Underlying Conditions				
Pregnant	69 (3.6%)	59 (6.0%)	10 (1.1%)	<0.001
(Missing)	66	42	24	
Chronic Lung Disease	387 (21%)	193 (20%)	194 (22%)	0.33
(Missing)	134	75	59	
Hypertension	874 (47%)	377 (39%)	497 (56%)	<0.001
(Missing)	95	50	45	
Overweight	584 (32%)	315 (33%)	269 (32%)	0.50
(Missing)	150	70	80	
Cardiovascular Disease	534 (29%)	217 (23%)	317 (36%)	<0.001
(Missing)	126	65	61	
Diabetes	426 (23%)	180 (19%)	246 (28%)	<0.001
(Missing)	101	53	48	

Renal Disease	316 (17%)	106 (11%)	210 (24%)	<0.001
(Missing)	132	69	63	
Liver Disease	83 (4.5%)	37 (3.8%)	46 (5.2%)	0.16
(Missing)	107	59	48	
Autoimmune Disease	101 (5.5%)	42 (4.4%)	59 (6.8%)	0.03
(Missing)	135	69	66	
Immunocompromised ³	269 (15%)	89 (9.3%)	180 (21%)	<0.001
(Missing)	137	69	68	
Systemic Immunosuppressive Therapy or Medications	293 (16%)	117 (12%)	176 (20%)	<0.001
(Missing)	115	59	56	

¹ n (%) or median (Interquartile Range [IQR])

² Pearson's Chi-squared test or Fisher's exact test for categorical variables and Wilcoxon rank-sum test for continuous variables

³ Immunocompromised (e.g., HIV infection, active cancer, solid organ transplant, hematopoietic stem cell transplant)

586

587

Table S2. Circulation of VOC/VOI in Georgia, USA as measured by available sequences on GISAID

VOC/VOI	Period variant circulated in GA ¹	Date range of sequences obtained from GISAID	Sequences available on GISAID	Total sequences available from May 2021 to May 2022 (%)
Alpha	12/10/2020-08/26/2021	05/01/2021-08/26/2021	1,361	2.66
Beta	02/01/2021-05/25/2021	05/01/2021-05/25/2021	16	0.03
Delta	4/29/2021-03/10/2022	05/01/2021-03/10/2022	28,320	55.5
Gamma	03/05/2021-09/24/2021	05/01/2021-09/24/2021	185	0.36
Lambda	02/07/2021-08/05/2021	05/01/2021-08/05/2021	18	0.04
Mu	04/23/2021-09/14/2021	05/01/2021-09/14/2021	142	0.28
Omicron	11/30/2021-4/18/2023	11/30/2021-04/18/2023	20,759	40.6
Total VOC/VOI			50,801	99.46
Total sequences available on GISAID from May 2021 to May 2022 from the state of Georgia			51,075	---

¹Period of circulation based on sequences available on GISAID including Emory Healthcare samples.

588

Table S3. Unadjusted Multinomial Logistic Regression Models Testing the Association between Demographics, Underlying Health Conditions, and COVID-Related Characteristics with Disease Severity

Variable	Moderate ¹			Severe ¹			Death ¹		
	OR ²	95% CI ²	p	OR ²	95% CI ²	p	OR ²	95% CI ²	p
Unadjusted Models									
Age	1.04	1.03, 1.05	<0.001	1.05	1.04, 1.05	<0.001	1.07	1.05, 1.09	<0.001
Male ³	1.40	1.13, 1.74	<0.01	1.30	0.95, 1.78	0.10	1.26	0.72, 2.22	0.40
White ⁴	0.86	0.67, 1.10	0.20	1.27	0.91, 1.77	0.20	1.04	0.57, 1.91	0.90
Chronic Lung Disease	2.07	1.59, 2.68	<0.001	1.95	1.35, 2.81	<0.001	2.65	1.44, 4.86	<0.01
Hypertension	2.94	2.35, 3.68	<0.001	4.19	2.98, 5.90	<0.001	11.4	5.09, 25.6	<0.001
Overweight	1.45	1.15, 1.83	<0.01	1.58	1.14, 2.19	0.01	2.02	1.13, 3.61	0.02
Cardiovascular Disease	2.89	2.28, 3.67	<0.001	3.46	2.48, 4.81	<0.001	4.80	2.70, 8.54	<0.001
Diabetes	2.04	1.58, 2.63	<0.001	3.21	2.29, 4.50	<0.001	4.35	2.43, 7.80	<0.001
Renal Disease	3.88	2.90, 5.19	<0.001	3.86	2.62, 5.67	<0.001	9.86	5.46, 17.8	<0.001
Liver Disease	1.47	0.87, 2.50	0.20	2.36	1.25, 4.45	<0.01	3.08	1.16, 8.21	0.02
Autoimmune Disease	1.60	1.01, 2.56	0.046	1.58	0.82, 3.03	0.20	2.45	0.93, 6.47	0.07
Immunocompromised	2.65	1.97, 3.58	<0.001	2.44	1.61, 3.69	<0.001	3.69	1.93, 7.09	<0.001
Systemic Immunosuppressive Therapy/Meds.	3.49	2.61, 4.67	<0.001	3.24	2.18, 4.81	<0.001	3.48	1.79, 6.78	<0.001
<u>Vaccination Status</u>									
Unvaccinated	Ref	—		Ref	—		Ref	—	
Vaccinated ⁵	0.80	0.64, 1.01	0.07	0.86	0.61, 1.21	0.40	1.00	0.54, 1.84	>0.90
Vaccinated & boosted ⁵	2.28	1.52, 3.43	<0.001	3.24	1.95, 5.40	<0.001	3.69	1.58, 8.59	<0.01
<u>Days Since Most Recent Vaccination/Booster</u>									
Unvaccinated	—	—		—	—		—	—	
Within past 90 Days ⁵	1.34	0.86, 2.08	0.20	1.68	0.93, 3.03	0.09	2.81	1.17, 6.79	0.02
91 - 180 Days Ago ⁵	0.98	0.74, 1.29	0.90	0.92	0.60, 1.40	0.70	1.14	0.55, 2.37	0.70
181 - 270 Days Ago ⁵	0.72	0.51, 1.01	0.06	0.74	0.45, 1.24	0.30	0.71	0.27, 1.89	0.50
More than 270 Days Ago ⁵	1.18	0.72, 1.94	0.50	2.49	1.43, 4.36	<0.01	1.91	0.64, 5.72	0.20
<u>Lineage</u>									
Delta	Ref	—		Ref	—		Ref	—	
Omicron⁶	1.51	1.09, 2.07	0.01	2.25	1.46, 3.48	<0.001	4.55	2.25, 9.20	<0.001

¹ "Mild" is the Reference Category

² OR = Odds Ratio, CI = Confidence Interval, Ref = Reference Level

³ Compared to "Female"

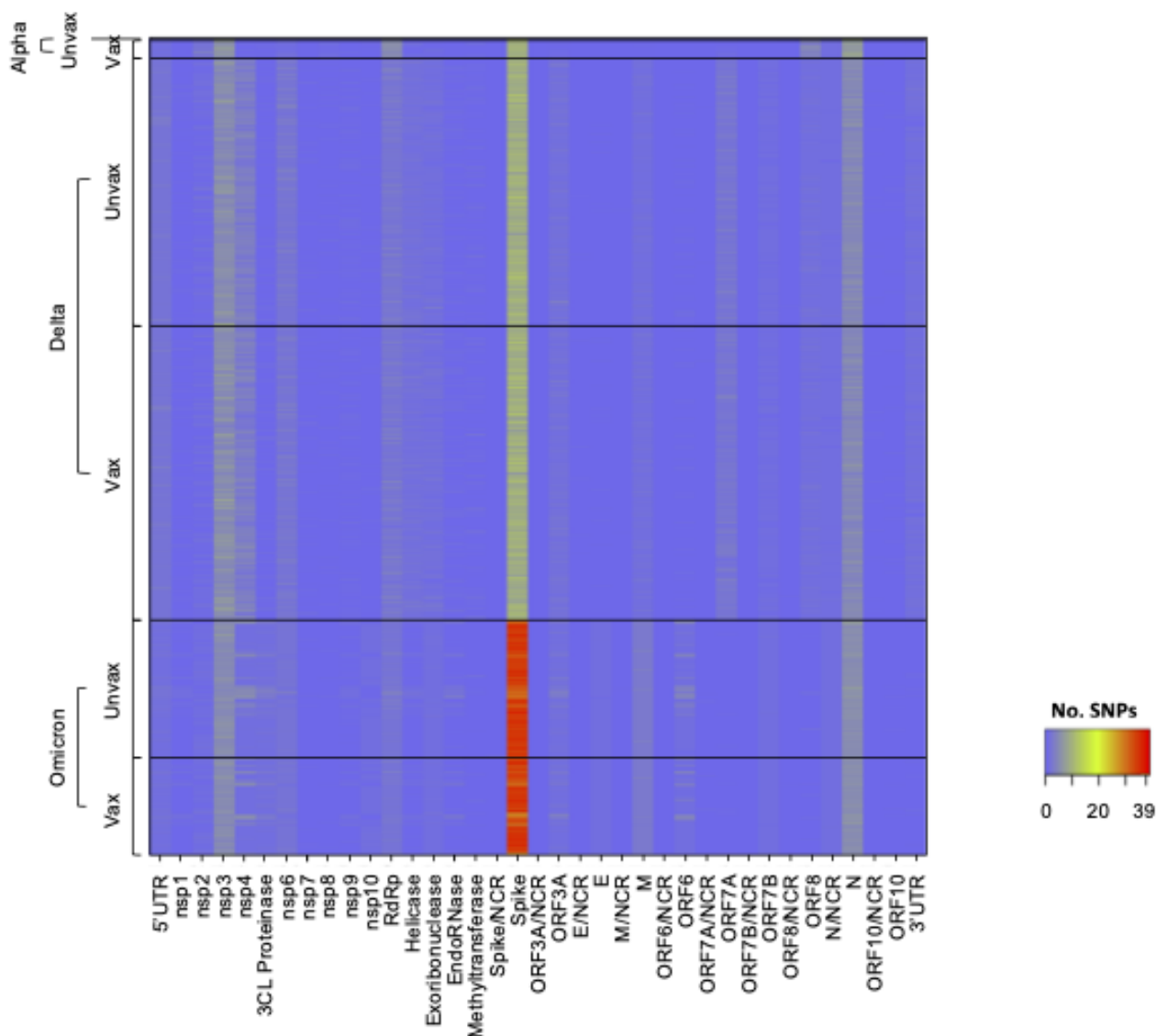
⁴ Compared to "Black"

⁵ Compared to "Unvaccinated individuals"

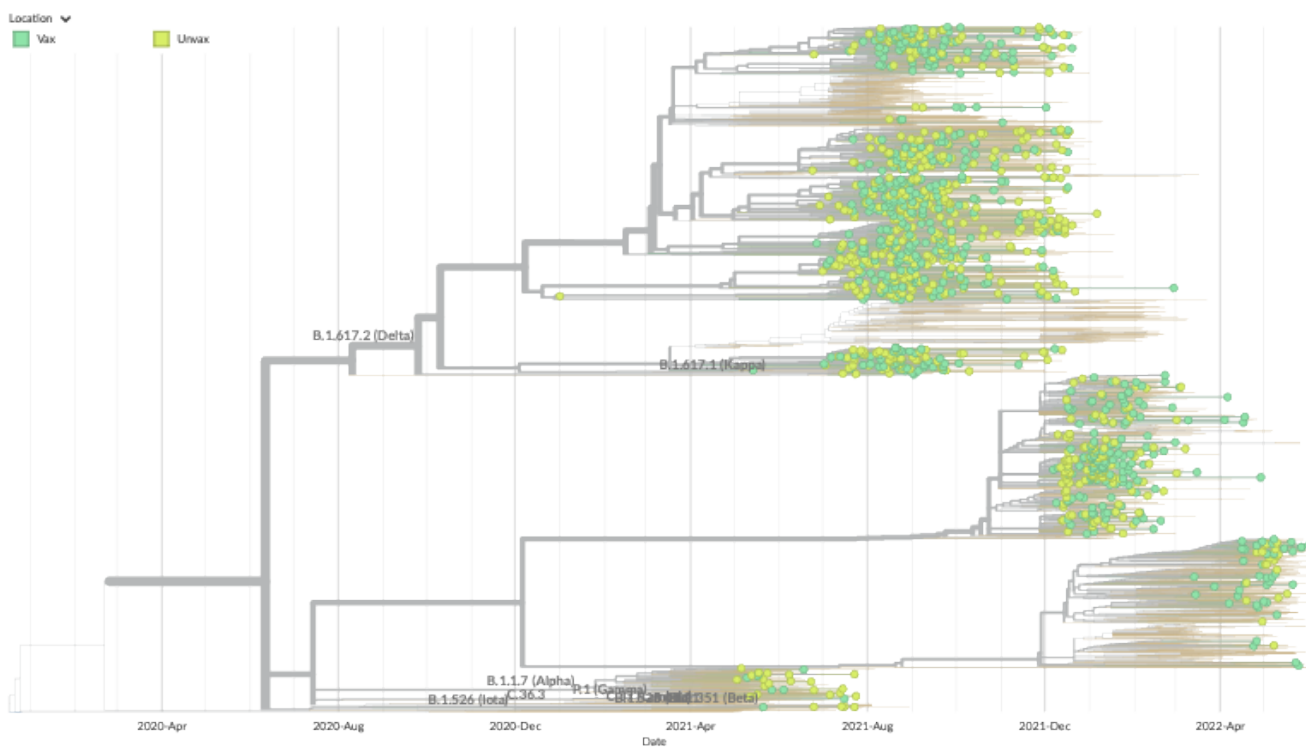
⁶ Compared to "Delta"

590 SUPPLEMENTARY FIGURES

591



592
593 **Figure S1. Mutations across SARS-CoV-2 genes.** Alpha, delta, and omicron sequences were
594 aligned to Wuhan/Hu-1/2019. For each variant, samples from vaccinated (“vax”) and
595 unvaccinated (“unvax”) individuals are labeled. Each square in the heatmap represents the
596 number of single nucleotide polymorphisms (SNPs) in each gene, labeled on the x-axis. Each
597 row represents one sample.



598

599 **Figure S2. Phylogenetic analysis does not reveal differences in SARS-CoV-2 sequences**
600 **from vaccinated and unvaccinated individuals.** Maximum likelihood tree containing
601 sequences from vaccinated (green) and unvaccinated (yellow) individuals in the context of 2000
602 global sequences from GISAID (orange) selected by a custom Nextstrain subsampling scheme
603 and rooted to NC_045512.

604 **KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Residual NP swabs	Emory Microbiology Lab	N/A
Critical commercial assays		
Nextera XT DNA Library Preparation kit	Illumina	FC-131-1096
Illumina MiSeq kit	Illumina	MS-102-3001
Deposited data		
SARS-CoV-2 consequence sequences	This paper	GISAID: PRJNA634356
SARS-CoV-2 reads	This paper	NCBI: PRJNA634356
Software and algorithms		
Geneious		https://www.geneio.us.com
Nextstrain	Hadfield et al., 2018	https://clades.nextstrain.org/
IQtree	Trifinopoulos et al., 2016	http://iqtree.cibiv.univie.ac.at/
Interactive Tree of Life (iTOL)	Letunic et al., 2021	https://itol.embl.de/
ViReMA	Routh et al., 2013	https://sourceforge.net/projects/virema/
Pangolin	Rambaut et al., 2020	https://pangolin.cog-uk.io
Pilon	Walker et al., 2014	https://github.com/broadinstitute/pilon
Bowtie2	Langmead et al., 2012	http://bowtie-bio.sourceforge.net/index.shtml
samtools	Li et al., 2009	http://www.htslib.org/
seqtk		https://github.com/lh3/seqtk
DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
R Studios	RStudio: Integrated Development for R	Rstudio.com
SAS	SAS Institute	https://www.sas.com/en_us/home.html

605

606