Magnitude and determinants of SARS-CoV-2 household transmission: a longitudinal cohort study

J. Daniel Kelly,[1,2,3,4] Scott Lu,[1,2] Khamal Anglin,[2] Miguel Garcia-Knight,[5] Jesus Pineda-Ramirez,[2] Sarah A. Goldberg,[1] Michel Tassetto,[5] Amethyst Zhang,[5] Kevin Donohue,[6] Michelle C. Davidson,[6] Mariela Romero,[2] Ruth Diaz Sanchez,[2] Manuella Djomaleu,[6] Sujata Mathur, Jessica Y. Chen,[2] Carrie A. Forman,[7] Venice Servellita,[8] Rubi D. Montejano,[6] Joshua R. Shak,[4] George W. Rutherford,[1,2] Steven G. Deeks,[9] Glen R. Abedi (CDC),[10] Melissa A. Rolfes (CDC),[10] Sharon Saydah (CDC),[10] Melissa Briggs-Hagen (CDC),[10] Michael J. Peluso,[9] Charles Chiu,[8] Claire M. Midgley (CDC),[10] Raul Andino,[5] Jeffrey N. Martin[1]

1. Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA;
2. Institute for Global Health Sciences, University of California, San Francisco, CA, USA;
3. F.I. Proctor Foundation, University of California, San Francisco, CA, USA;
4. San Francisco VA Medical Center, San Francisco, CA, USA;
5. Department of Microbiology and Immunology, UCSF
6. School of Medicine, University of California, San Francisco, CA, USA;
7. School of Medicine, Drexel University, Philadelphia, PA, USA;
8. Department of Laboratory Medicine, University of California, San Francisco, CA, USA;
9. Division of HIV, Infectious Diseases and Global Medicine, Zuckerberg San Francisco General Hospital, San Francisco, CA, USA;
10. Respiratory Viruses Branch, Division of Viral Diseases, CDC, Atlanta, GA, USA;

Corresponding author contact information: J. Daniel Kelly, 550 16th Street, Third Floor, P.O. Box 1224, San Francisco, CA 94143-1224, USA; email: dan.kelly@ucsf.edu

Running title: SARS-CoV-2 household transmission

ACCEPTED MANUSCRIPT

© Published by Oxford University Press on behalf of Infectious Diseases Society of America 2022. This work is written by (a) US Government employee(s) and is in the public domain in the US.

Abstract

Background: Households have emerged as important venues for SARS-CoV-2 transmission. Little is known, however, regarding the magnitude and determinants of household transmission in increasingly vaccinated populations.

Methods: From September 2020 to January 2022, symptomatic non-hospitalized individuals with SARS-CoV-2 infection by RNA detection were identified within 5 days of symptom onset; all individuals resided with at least one other SARS-CoV-2-uninfected household member. These infected persons (cases) and their household members (contacts) were subsequently followed with questionnaire-based measurement and serial nasal specimen collection. The primary outcome was SARS-CoV-2 infection among contacts.

Results: We evaluated 42 cases and their 74 household contacts. Among the contacts, 32 (43%) became infected, of whom 5/32 (16%) were asymptomatic; 81% of transmissions occurred by 5 days after the case's symptom onset. From 21 unvaccinated cases, 14-day cumulative incidence of SARS-CoV-2 infection among contacts was 18/40 (45%; 95% CI: 29, 62), most of whom were unvaccinated. From 21 vaccinated cases, 14-day cumulative incidence of SARS-CoV-2 infection was 14/34 (41%; 95% CI: 25, 59) among all contacts and 12/29 (41%; 95% CI: 24, 61) among vaccinated contacts. At least one co-morbid condition among cases and 10 or more days of RNA detection in cases were associated with increased risk of infection among contacts.

Conclusions: Among households including individuals with symptomatic SARS-CoV-2 infection, both vaccinated-to-vaccinated and unvaccinated-to-unvaccinated transmission of SARS-CoV-2 to household contacts was common. Because vaccination alone did not notably reduce risk of infection, household contacts will need to employ additional interventions to avoid infection.

Presentations at meetings: This work has not been presented.

Keywords: SARS-CoV-2; household transmission; epidemiology; infectious viral shedding

ACCEPTED MANUSCRIPT

Introduction

The reported magnitude of SARS-CoV-2 transmission from index cases to household members has varied, depending on multiple factors including viral variant, vaccination status of both the index case and household contacts, and diagnostic procedures.[1] Prior to the introduction of Delta variant, fully vaccinated index cases had a lower proportion of transmission events (cumulative incidence, or secondary attack rates) to household members compared to unvaccinated index cases.[2-4] Studies focused on Delta infections, however, found that there was no difference in the proportion of transmission events between vaccinated and unvaccinated index cases.[5-7] It should be noted that most of these Delta-focused transmission studies did not perform longitudinal specimen collection more than weekly, and as a result, may have missed cases in the household. Very few studies have taken the additional step of generating qualitative and quantitative viral culture data from specimens to assess the effect of infectious viral shedding of cases on infection of contacts.

Without rigorous descriptions of transmission events in households from and between vaccinated and unvaccinated individuals, we have a limited understanding of the causal determinants of SARS-CoV-2 transmission and host susceptibility of infection. We developed a rigorous set of criteria, using viral, epidemiological, and genetic data, to identify primary cases. Then, we sought to assess the host and viral determinants and magnitude of SARS-CoV-2 household transmission, stratifying cases and contacts by vaccination status.

Methods

Overall Design

This was a longitudinal cohort study enrolling vaccinated and unvaccinated individuals (index cases) at the time of SARS-CoV-2 infection and their household contacts. The study was reviewed by the UCSF Institutional Review Board and given a designation of public health

ACCEPTED MANUSCRIPT

surveillance according to federal regulations as summarized in 45 CFR 46.102(d)(1)(2). Written informed consent was obtained from all participants.

Participants

From September 2020 to January 2022, we identified individuals of all ages who were positive for SARS-CoV-2 via provider-ordered molecular testing at UCSF-affiliated testing sites. Individuals were screened for study eligibility by review of available data or by telephone interview. Individuals were eligible for inclusion as index cases if they were non-hospitalized, resided with at least one other individual, and lived in non-congregate settings in the San Francisco Bay Area. Symptomatic individuals were eligible if they could be enrolled within 5 days of symptom onset. To reliably identify asymptomatic index cases early in infection, asymptomatic individuals were only eligible as index cases if they could be enrolled within 10 days of a known high-risk exposure (unprotected exposure within 6 feet for greater than 15 minutes over 24-hour period).

A household contact was defined as any individual who had spent at least one night in the household during the 2 days before illness onset of the index case through to enrollment. If the index case was eligible and at least one household member was willing to participate, then household eligibility was assessed. Households were not eligible for the study if household contacts ever had a history of confirmed or probable SARS-CoV-2 infection or had suspected SARS-CoV-2 infection in the 14 days preceding symptom onset of the index case,[8] unless the first study visit could occur within 5 days of symptom onset of all individuals living in the household.

ACCEPTED MANUSCRIPT

Measurements

Questionnaire-based

Our interviewers administered questionnaires to cases and contacts by telephone to collect data on sociodemographics, exposure and medical history, symptom status and onset, and clinical course of acute COVID-19. Our symptom checklist included 32 symptoms derived from the U.S. Centers for Disease Control (CDC) list of COVID-19 symptoms[9] and the Patient Health Questionnaire Somatic Symptom Scale.[10] We also recorded any other self-reported symptom. Any symptoms were recorded as present if they were new or worsened since the time of SARS-CoV-2 infection. Interviewers reviewed documentation of SARS-CoV-2 infection and vaccination status during study visits. Participants were also assessed for receipt of vaccine boosters. Questionnaires were completed on day (d) of enrollment (dE), and at d9, 14, 21, and 28 after symptom onset of the index case.

Laboratory-based

Study staff also visited participants in their households on the same days as the telephone questionnaires (dE, d9, d14, d21, d28). Detailed operational methods are described in the **Supplemental Material**. Anterior nasal specimens were self-collected daily from dE through d14 and then on d17, d19, d21, and d28. Biospecimens were collected relative to the symptom onset of the index case (or day of first positive test in asymptomatic index cases). These specimens were stored in households at -20 °C for up to one week, until the next in-person study visit by the investigators, and subsequently transported on dry ice to laboratories at UCSF. To provide participants with timely clinical results, additional oropharynx (OP) specimens were collected at enrollment for molecular testing at UCSF clinical laboratories. All specimens tested by clinical and research laboratories were used to measure SARS-CoV-2 infection. We also measured infectious virus and viral lineage. Details of the laboratory assays have been previously reported.[11]

ACCEPTED MANUSCRIPT

RT-PCR

In brief, the research laboratory used all anterior nasal specimens to quantify SARS-CoV-2 RNA through RT-PCR targeting nucleocapsid (N) and envelope (E) genes on a CFX Connect Real-Time PCR detection system (Biorad).

Whole genome sequencing of SARS-CoV-2

Viral sequencing was done using RNA from nasal specimens with the highest RNA level for each SARS-CoV-2-infected individual. The ARTIC Network amplicon-based sequencing protocol for SARS-CoV-2 (using primer versions 3 and 4.1) was followed and sequencing was done on a MinION sequencer (Oxford Nanopore Technologies). The nCoV-2019 novel coronavirus bioinformatics protocol was used to assemble viral genomes and generate consensus sequences.[12] Full consensus genomes were submitted to GISAID and NCBI. Viral lineages were assigned using the PANGOLIN (Phylogenetic Assignment of Named Global Outbreak Lineages) version 3.0.2.

Phylogenetic analyses

A dataset was compiled consisting of all available high-quality whole genome sequences deposited to GISAID from San Francisco and Alameda counties collected between September 2020 and January 2021 (N=5,212) together with 72 genomes generated from our study cohort. Sequence alignment was done using MAFFT v7.388 implemented in CIPRES Science Gateway.[13] Aligned sequences were used as input for the Nextstrain bioinfomatic pipeline Augur version 13.0.2 [14] and maximum likelihood phylogenetic trees were inferred using IQTREE v1.6 and a discreet traits model. Phylogenies were visualized using Auspice.

ACCEPTED MANUSCRIPT

Cytopathic effect assay

Anterior nasal specimens were used to detect infectious virus on Vero-hACE2-TMPRSS2 cells. All specimens up to 14 days after symptom onset were assayed for cytopathic effect (CPE), which allows for a qualitative (yes/no) determination of infectious virus. In cases where CPE was observed within days 11-14, testing was continued until there were three consecutive negative results. Viral cultures with evidence of CPE underwent RT-PCR to confirm the presence of SARS-CoV-2.

Viral plaque assay

Based on qRT-PCR data, the specimen with maximum RNA load for each participant was selected for evaluation of quantitative infectious virus. Conventional plaque assays were performed and plaques were counted to determine infectious viral titers (expressed as plaque forming units/mL).[11]

Analyses

Analytical definitions

The index case for each household was the individual identified by the study team from the overall list of individuals with positive SARS-CoV-2 PCR testing. For analyses, we defined the single primary case as the first SARS-CoV-2-infected individual in the household, based on the first illness onset or, if asymptomatic, the first individual with a positive test. Hereinafter, single primary cases will be referred to as primary cases. Co-primary cases occurred when the first SARS-CoV-2-infected individual was not able to be determined from two or more infected individuals in the household (see next section for more details). A household case was any household contact infected by SARS-CoV-2 after the primary case and not considered a co-primary case. Serial interval was defined as the number of days between the symptom onsets (or date of first positive test in asymptomatic cases) of the primary case and any household

ACCEPTED MANUSCRIPT

case. Shared exposure was broadly defined as having had a high-risk exposure in the same contextual setting (e.g., hospital, indoor dining, family gathering) within 14 days prior to symptom onset of the index case. We created a decision algorithm with serial intervals, shared exposures, and phylogenetic analyses to identify primary versus co-primary cases from index cases. Details describing the decision algorithm can be found in the **Supplemental Material**.

Statistical analyses

We described the entire cohort and whether infected individuals self-reported the presence or absence of any symptom over the infectious period. We restricted analyses to households that had primary cases who were symptomatic. Our primary outcome was SARS-CoV-2 infection among household contacts, defined as the identification of SARS-CoV-2 RNA in any nasal or oropharyngeal specimen longitudinally obtained over the infectious period. All household contacts who developed SARS-CoV-2 infection during the study had positive testing within 10 days of symptom onset of the primary case in the household; we assumed that all infections identified in these households represented transmission from the primary case.

We estimated 14-day cumulative incidence of SARS-CoV-2 infection among household contacts and determined which host and viral factors were associated with household transmission. We stratified cumulative incidence by vaccination status and by variant of the primary case. Participants were considered vaccinated if participants completed a primary series of a COVID-19 vaccine $\geq$14 days prior to enrollment; those who had received booster dose were also included as vaccinated. Partial vaccination was defined as having received the first mRNA dose in a 2-dose series >14 days earlier but were either missing a second dose or <14 days had elapsed since the receipt of the second dose.[15]

We next performed analyses to assess risk factors for infection among household contacts accounting for characteristics of primary cases and household contacts. Host factors included age, sex, race/ethnicity, BMI (<25, 25-30, >30), vaccination status at baseline, and

ACCEPTED MANUSCRIPT

presence of any comorbidities (autoimmunity, cancer, diabetes, HIV/AIDS, heart disease, hypertension, lung disease, kidney disease). Viral factors included variant, maximum RNA load (log copies/mL), maximum infectious viral load (log plaque forming units [log pfu]/mL), and duration of RNA and infectious viral shedding (days after symptom onset until last positive test). If no RNA or live virus was detected, we assigned that individual a duration of zero. In our analyses of maximum RNA level or infectious viral load, we did not include individuals with a value of zero because of the lack of a valid assumption for use of bins.

We used generalized estimating equations (GEE) with modified Poisson regression (log link), clustering by household, to generate marginal estimates of risk ratios (cumulative incidence ratios) and 95% confidence intervals (CIs). We performed a series of unadjusted analyses examining each host and viral factor for primary cases and then each host factor for household contacts. We performed a series of adjusted analyses that included factors for both the primary cases and household contacts. To determine the relevant covariates, we created an adjustment set of covariates in a Directed Acyclic Graph to assess potential confounders, mediators, and colliders (**Suppl. Figure 1**). We considered the literature, expert consultation, and the results of our unadjusted analyses in the determination of final covariate adjustment set. We repeated this process in the assessment of each factor. Details of each model can be found in the footnote of the tables.

Confidence intervals for cumulative incidence were calculated using Agresti-Coull interval. We used GEE to compare cumulative incidence estimates and generated p-values, considering any p-value of less than 0.05 as statistically significant. In the analyses involving statistical modeling, we used confidence intervals to determine statistical significance and considered intervals that did not include the null value to be significant. All analyses were performed using STATA/BE 17.0 (StataCorp, College Station, Texas, USA). As a sensitivity analysis of the cumulative incidence estimate, we repeated the analyses with the entire cohort, inclusive of primary and co-primary cases (**Suppl. Table 1**).

ACCEPTED MANUSCRIPT

Results

From September 2020 to January 2022, we enrolled 65 index cases and their 115 household contacts in the total study cohort (180 participants) (**Figure 1**)**.** Among these 180 participants, 72 (57%) of 126 infected participants had sufficient RNA levels for viral sequencing (**Figure 2**). There were 21 index cases that had no infected household contacts and were classified as primary cases. Among the remaining 44 index cases, we assessed the serial interval and shared exposures and identified 24 additional possible primary cases. From phylogenetic analyses of available genomic data, 24 contacts had viral sequences that could be paired to an index case and assessed as potentially related sequences. Based on these phylogenetic analyses, we reclassified one of the 45 households thought to have a single primary case as having co-primary cases (**Figure 3**). Among the remaining 44 single primary cases, two were asymptomatic. We excluded asymptomatic primary cases to focus on an analysis cohort of 42 symptomatic primary cases and their 74 household contacts (116 participants) (**Figure 4**).

Description of primary cases and household contacts

Primary cases were enrolled a median of 4 days (range: 1 to 5) from symptom onset. Among the 42 primary cases included in this analysis, half (50%) were unvaccinated and half (50%) vaccinated, with 1 having a full primary series and booster vaccine (no partially vaccinated primary cases). Almost all (20/21, 95%) vaccinated primary cases received an mRNA vaccine (3 received Moderna and 17 received Pfizer); one received the viral vector vaccine (Janssen). All vaccinated primary cases were enrolled from June 2021 through January 2022 during the Delta or Omicron periods. Among the 74 household contacts of these 42 primary cases, 40 (54%) were unvaccinated, 4 (5%) were partially vaccinated, and 30 (41%)

ACCEPTED MANUSCRIPT

were vaccinated. Households had a median size of 3 participants (IQR: 2, 4), inclusive of primary cases.

A median of 14 specimens (total of 552; IQR: 12, 15) were collected per primary case. Based on viral sequencing results of the 42 primary cases, 12 (29%) were infected by non-VOI/VOC, 18 (43%) by Delta variant, 3 (7%) by Omicron, 3 (7%) by Epsilon, 1 (2%) by Alpha, and 5 (12%) by unclassified variants. Primary cases had a median maximum RNA viral load of 6.7 log copies/mL (IQR: 5.1, 8.2; n=40), median duration of RNA detection of 9.5 days since symptom onset (IQR: 7, 11; n=42), median maximum infectious viral titers of 8.3 log pfu/mL (IQR: 6.6, 12.3; n=32), and median duration of infectious viral shedding of 5 days since symptom onset (IQR: ≤3, 7; n=39). See **Table 1** for more characteristics of the primary cases and their contacts, and see **Table 2** for virologic characteristics of the infected participants.

Cumulative incidence of SARS-CoV-2 infection

Among the 74 household contacts of the 42 primary cases, a median of 14 specimens (total of 924; IQR 12, 15) were collected per household contact. 32/74 (43%) household contacts were SARS-CoV-2 positive (e.g., household cases). Of 32 household cases, five (16%) were asymptomatic (none were vaccinated), and none were hospitalized. Among household contacts exposed to 21 vaccinated primary cases, cumulative incidence of SARS-CoV-2 infection was 14/34 (41%; 95% CI: 25, 59); among household contacts exposed to 21 unvaccinated primary cases, cumulative incidence was 18/40 (45%; 95% CI: 29, 62; p=0.60).

Most (29/34) household contacts of vaccinated primary cases were also vaccinated, and cumulative incidence among these vaccinated household contacts was 12/29 (41%; 95% CI: 24, 61). Likewise, almost all (34/40) household contacts of unvaccinated primary cases were also unvaccinated, and cumulative incidence among these unvaccinated household contacts was 16/34 (47%; 95% CI: 30, 65). (**Table 3**). Almost all (26/32, 81%) transmission events occurred

ACCEPTED MANUSCRIPT

within 5 days since day of symptom onset of the primary case, with a median serial interval of 3 days (IQR: 2, 4; range: 0, 12) (**Suppl. Table 2**).

Stratifying by viral variant of 42 primary cases, cumulative incidence of SARS-CoV-2 infection was as follows: by Delta variant, 12/29 (41%; 95% CI: 24, 61) contacts from 18 cases; by Omicron variant, 2/5 (40%; 95% CI: 5.3, 84) contacts from 3 cases; by Epsilon, 1/8 (13%; 95% CI 0, 53) contacts from 3 cases; and by non-VOI/VOC, 12/22 (55%; 95% CI: 32, 76) contacts from 12 cases (**Table 4**).

Host and viral determinants of SARS-CoV-2 infection

In unadjusted analyses, household contacts had increased risk of infection when exposed to primary cases who were female, or with at least one comorbidity, and 10 or more days of RNA viral shedding. After adjustment, increased risk of infection among contacts remained associated with primary cases who had at least one comorbidity and RNA viral shedding for ≥10 days. Among these determinants, the greatest risk of infection was seen from primary cases with at least one co-morbidity (adjusted risk ratio [aRR]: 2.1; 95% CI: 1.2, 3.8). Although duration of infectious viral shedding was not statistically significant, household contacts of primary cases with longer duration of infectious viral shedding had 1.05 times the adjusted risk of infection than household contacts of primary cases with shorter duration of shedding (95% CI: 0.96, 1.2) (**Table 5**).

Host susceptibility factors associated with infection were assessed among household contacts, and we found that after adjustment, household contacts of White race had higher risk of infection than other races (**Table 6**). We did not find any statistically significant associations between risk of infection and the following characteristics of household contacts: age, sex, BMI, vaccination status, or comorbidity status.

ACCEPTED MANUSCRIPT

Discussion

In this cohort longitudinally sampled for evidence of SARS-CoV-2, we identified a significant amount of onward transmission from both vaccinated-to-vaccinated and unvaccinated-to-unvaccinated individuals within households. When comparing vaccination status of primary cases, we did not observe a difference in cumulative incidence among household contacts. We used phylogenetic analyses to strengthen evidence of transmission events from and between vaccinated individuals in a similar way as other studies, which have described high-confidence events.[5, 6] Although vaccination prevents severe illness[16], SARS-CoV-2 transmission commonly occurs in vaccinated households, serving as a public health reminder of the ongoing value in masking and other mitigating measures, particularly when community transmission increases.

Although we did not detect associations between infection of household contacts and most household contact characteristics, we did detect associations with host and viral determinants of the primary case, despite small numbers. Increased risk of infection was notably high among household contacts of primary cases with underlying conditions; this finding may inform public health strategies, including targeted case and contact investigations. We also identified an association between household contact infection and duration of viral RNA shedding (10 or more days) in the primary case; we did not detect an association with maximum viral RNA load, controlling for vaccination status, though previous studies have found maximum viral RNA load was associated with the risk of onward transmission[6, 17]. Our study extends the literature with its inclusion of infectious viral data. We did not find that risk of infection was associated with maximum infectious viral titer or maximum viral RNA load, though may have been underpowered to detect such associations. These two virological parameters (maximum loads) have been correlated in previous work.[18] It is possible that duration of RNA and infectious viral shedding are also correlated and may be important virological parameters of transmissibility.

ACCEPTED MANUSCRIPT

Our study has limitations. Although we attempted to reach infected participants as early as possible, some of our primary cases were negative for infectious virus, meaning that we may have missed the presence of viral shedding, which could have possibly left-censored our data. Furthermore, we assessed maximum viral load available for each individual, but this may not have reflected the true peak viral load in most primary cases. Low levels of RNA limited our ability to sequence virus for some primary cases. Among the viruses sequenced, we had a different distribution of variants among the vaccinated versus unvaccinated primary cases, potentially biasing the difference in cumulative incidence among these groups to the null. Our phylogenetic analyses depicted whether participants were from similar clades but were unable to decipher the transmission chain with individual-level resolution. The magnitude of exposure and mitigating factors during the isolation and quarantine periods likely varied and were not considered in our analyses. We also lacked detailed symptom data for the characterization of infected household contacts. This sample may have been enriched for characteristics specific to non-hospitalized index cases who were symptomatic and diagnosed with rapid access to the health system and thus limit the external validity or generalizability of the study. Our sample size was small, however, so we may have been underpowered to detect some associations, such as differences in cumulative incidence by variant or duration of infectious viral shedding. Because of the small sample size, we also were not able to reliably assess the effects of time since vaccination or booster vaccinations.

Although vaccination may reduce severity, it did not significantly reduce transmission from June 2021 to January 2022, which was a period when household transmission was common among vaccinated-to-vaccinated individuals. As new variants emerge and vaccines are updated, we expect that the impact of vaccination on transmission may continue to change and that there will be potential for unrecognized transmission despite widespread COVID-19 vaccination. It will be critical to characterize viral shedding and transmission dynamics through ongoing active surveillance and natural history studies.

ACCEPTED MANUSCRIPT

Acknowledgements

We thank the participants for making this study possible while acutely infected with SARS-CoV-2. We appreciate the input and support of Amy J. Markowitz, Elan L. Guterman, Thomas M. Lietman, Will Brett, Eric Talbert, and others in the CDC COVID-19 response who contributed to this study. Vero TMPRSS2 hAce2 cells were a kind gift from Barney Graham (NIH).

Supplement Sponsorship

This article appears as part of the supplement “Vaccines, Variants, and Vigilance: Strengthening the COVID-19 Public Health Response through Partnerships and Collaborations”, supported by the Infectious Diseases Society of America through Cooperative Agreement NU50CK000574 with the U.S. Centers for Disease Control and Prevention.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the U.S. Centers for Disease Control and Prevention.

Funding

This study was funded by the Centers for Disease Control and Prevention Broad Agency Announcement. The National Institute of Allergy and Infectious Diseases also supported JDK during this study (K23 grant number AI146268). These funding sources had no role in the content of the manuscript nor the decision for publication.

Conflicts of Interest

MBH reports funding support from the Centers for Disease Control and Prevention. All authors submitted ICMJE forms and have no conflicts to report.

ACCEPTED MANUSCRIPT

References

1. Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. Household secondary attack rates of SARS-CoV-2 by variant and vaccination status: an updated systematic review and meta-analysis. medRxiv **2022**.
2. Bender JK, Meyer ED, Sandfort M, Matysiak-Klose D, Bojara G, Hellenbrand W. Low Sensitivity of Rapid Antigen Tests to Detect Severe Acute Respiratory Syndrome Coronavirus 2 Infections Before and on the Day of Symptom Onset in Nursing Home Staff and Residents, Germany, January-March 2021. J Infect Dis **2021**; 224(11): 1987-9.
3. Gazit S, Mizrahi B, Kalkstein N, et al. BNT162b2 mRNA Vaccine Effectiveness Given Confirmed Exposure: Analysis of Household Members of COVID-19 Patients. Clin Infect Dis **2021**.
4. de Gier B, Andeweg S, Joosten R, et al. Vaccine effectiveness against SARS-CoV-2 transmission and infections among household and other close contacts of confirmed cases, the Netherlands, February to May 2021. Euro Surveill **2021**; 26(31).
5. Ng OT, Koh V, Chiew CJ, et al. Impact of Delta Variant and Vaccination on SARS-CoV-2 Secondary Attack Rate Among Household Close Contacts. Lancet Reg Health West Pac **2021**; 17: 100299.
6. Singanayagam A, Hakki S, Dunning J, et al. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. Lancet Infect Dis **2022**; 22(2): 183-95.
7. de Gier B, Andeweg S, Backer JA, et al. Vaccine effectiveness against SARS-CoV-2 transmission to household contacts during dominance of Delta variant (B.1.617.2), the Netherlands, August to September 2021. Euro Surveill **2021**; 26(44).

ACCEPTED MANUSCRIPT

8. Centers for Disease Control and Prevention. National Notifiable Diseases Surveillance System. Coronavirus Disease 2019 (COVID-19) 2021 Case Definition. Available at: https://ndc.services.cdc.gov/case-definitions/coronavirus-disease-2019-2021/.
9. Centers for Disease Control and Prevention. Symptoms of COVID-19. Updated 22 February 2022. Available at: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html.
10. Kroeke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2022; 64:258-266.
11. Garcia-Knight M, Anglin K, Tassetto M, Lu S, Zhang A, Goldberg SA, Catching A, Davidson MC, Shak JR, Romero M, Pineda-Ramirez J, Diaz-Sanchez R, Rugart P, Donohue K, Massachi J, Sans HM, Djomaleu M, Mathur S, Servellita V, McIlwain D, Gaudiliere B, Chen JY, Martinez EO, Tavs JM, Bronstone G, Weiss J, Watson JT, Briggs-Hagen M, Abedi GR, Rutherford GW, Deeks SG, Chiu C, Saydah S, Peluso MJ, Midgley CM, Martin JN, Andino R, Kelly JD. Infectious viral shedding of SARS-CoV-2 Delta following vaccination: a longitudinal cohort study. medRxiv. 2022 May 19; doi.org/10.1101/2202.05.15.22275051. .
12. Tyson JR, James P, Stoddart D, et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. bioRxiv **2020**.
13. Miller MA, Pfeiffer W, and Schwartz T. (2010) "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov 2010, New Orleans, LA pp.1-8.
14. Huddleston J, Hadfield J, Sibley TR, et al. Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens. J Open Source Softw **2021**; 6(57).
15. Danza P, Koo TH, Haddix M, et al. SARS-CoV-2 Infection and Hospitalization Among Adults Aged ≥18 Years, by Vaccination Status, Before and During SARS-CoV-2

ACCEPTED MANUSCRIPT

B.1.1.529 (Omicron) Variant Predominance - Los Angeles County, California, November 7, 2021-January 8, 2022. MMWR Morb Mortal Wkly Rep **2022**; 71(5): 177-81.

16. Maslo C, Friedland R, Toubkin M, Laubscher A, Akaloo T, Kama B. Characteristics and Outcomes of Hospitalized Patients in South Africa During the COVID-19 Omicron Wave Compared With Previous Waves. JAMA **2021**.

17. Marks M, Millat-Martinez P, Ouchi D, et al. Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. Lancet Infect Dis **2021**; 21(5): 629-36.

18. Garcia-Knight M, Anglin K, Tassetto M, Lu S, Zhang A, Goldberg SA, Catching A, Davidson MC, Shak JR, Romero M, Pineda-Ramirez J, Diaz-Sanchez R, Rugart P, Donohue K, Massachi J, Sans HM, Djomaleu M, Mathur S, Servellita V, McIlwain D, Gaudiliere B, Chen JY, Martinez EO, Tavs JM, Bronstone G, Weiss J, Watson JT, Briggs-Hagen M, Abedi GR, Rutherford GW, Deeks SG, Chiu C, Saydah S, Peluso MJ, Midgley CM, Martin JN, Andino R, Kelly JD. Infectious viral shedding of SARS-CoV-2 Delta following vaccination: a longitudinal cohort study. medRxiv. 2022 May 19; doi.org/10.1101/2202.05.15.22275051.

ACCEPTED MANUSCRIPT

Table 1: Description of sociodemographic, epidemiological, clinical, and household characteristics of total cohort (N=180) and analysis cohort (N=116). We stratified the analysis cohort by primary cases (N=42) and their household contacts (N=74).

	Total cohort (N=180)	Household cohort with primary cases who were included in this analysis (N=116)	Primary cases (N=42)	Household contacts of primary cases (N=74)
Age, median (IQR)	33.5 (21.5 to 44)[1]	34 (24.5 to 44.5)	34 (27 to 42)	34 (23 to 46)
Age categories, n (%)				
<18	13 (7.2%)	7 (6.0%)	4 (9.5%)	3 (4.1%)
18-44	113 (62.8%)	79 (68.1%)	33 (78.6%)	46 (62.2%)
≥45	19 (10.6%)	11 (9.5%)	1 (2.4%)	10 (13.5%)
Female sex, n (%)	94 (52.2%)	58 (50.0%)	20 (47.6%)	38 (51.4%)
Race/ethnicity[2]				
Hispanic/Latino	41 (22.8%)	17 (14.7%)	6 (14.3%)	11 (14.9%)
White	97 (53.9%)	68 (58.6%)	24 (57.1%)	44 (59.5%)
Black/African American	6 (3.3%)	6 (5.2%)	2 (4.8%)	4 (5.4%)
Asian	24 (13.3%)	18 (15.5%)	8 (19.0%)	10 (13.5%)
Pacific Islander/Native Hawaiian	3 (1.7%)	2 (1.7%)	0 (0.0%)	2 (2.7%)
American Indian or Alaska Native	2 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
BMI - 3 Categories[2, 3]				
<25	90 (50.0%)	61 (52.6%)	22 (52.4%)	39 (52.7%)
25 to 30	44 (24.4%)	29 (25.0%)	10 (23.8%)	19 (35.7%)
> 30	34 (18.9%)	20 (17.2%)	9 (21.4%)	11 (14.9%)
Education[2, 3]				
At least some HS	32 (17.8%)	18 (15.5%)	7 (16.7%)	11 (14.9%)
At least some college	71 (39.4%)	49 (42.2%)	17 (40.5%)	32 (43.2%)
At least some graduate school	40 (22.2%)	28 (24.1%)	13 (31.0%)	15 (20.3%)
Annual household income[2, 3, 4]				
$50,000 or less	7 (12.1%)	4 (11%)		
$50,000 to $100,000	10 (17.2%)	9 (24%)		
$100,000 to $300,000	21 (36.2%)	11 (30%)		
More than $300,000	5 (8.6%)	4 (11%)		
Any comorbidity	36 (20.0%)	24 (20.7%)	11 (26.2%)	13 (17.6%)
Autoimmune disease	4 (2.2%)	3 (2.6%)	0 (0%)	3 (4.1%)
Cancer treated within past 2 years	5 (2.8%)	3 (2.6%)	2 (4.8%)	1 (1.4%)
Diabetes	5 (2.8%)	4 (3.4%)	4 (9.5%)	0 (0.0%)
Heart attack or heart failure	1 (0.6%)	1 (0.9%)	1 (2.4%)	0 (0.0%)
Hypertension or high blood pressure	11 (6.1%)	5 (4.3%)	3 (7.1%)	2 (2.7%)

ACCEPTED MANUSCRIPT

Lung disease	22 (12.2%)	14 (12.1%)	5 (11.9%)	9 (12.2%)
Vaccination status				
No vaccination	107 (59.4)	60 (51.7%)	21 (50%)	39 (52.7%)
Partially vaccinated	5 (2.8%)	4 (3.4%)	0 (0%)	4 (5.4%)
Fully vaccinated	56 (31.1%)	44 (37.9%)	20 (48%)	24 (32.4)
Booster	12 (6.7%)	8 (6.9%)	1 (2.4%)	7 (9.5%)
Household size				
Participating members per household[5]	3 (2 to 4)	3 (2 to 4)		
2	33 (52.4%)	23 (54.8%)		
3	15 (23.8%)	9 (21.4%)		
4	11 (17.5%)	7 (16.7%)		
5	4 (6.35%)	3 (7.14%)		
Total members per household (including unenrolled) [5]	3 (2 to 4)	3 (2 to 4)		
2	25 (39.7%)	17 (40.5%)		
3	16 (25.4%)	11 (26.2%)		
4	14 (22.2%)	8 (19.1%)		
5	7 (11.1%)	5 (11.9%)		
10	1 (1.6%)	1 (2.38%)		

[1]Median (Interquartile range) unless otherwise specified.
[2]Missing and nonresponse. Race/ethnicity: 7 missing; BMI: 8 missing; Education: 35 missing, 2 prefer not to answer; Annual household income: 7 missing, 24 prefer not to answer.
[3]Categories limited to adult respondents,
[4]Annual household income reported from N=65 total index cases and N=42 primary cases.
[5]Among the 42 households with primary cases, 34 (81%) had full enrollment. Household size was inclusive of cases.

ACCEPTED MANUSCRIPT

Table 2: Description of virologic characteristics of total infected cohort (N=126) and analysis cohort (N=74). We stratified the analysis cohort by primary cases (N=42) and their household contacts (N=32).

	Total cohort (N=126)	Household cohort with primary cases who were included in this analysis (N=74)	Primary cases (N=42)	Household contacts of primary cases (N=32)
Maximum RNA viral load (log copies/mL)[1]	6.03 (3.60 to 8.22)	5.44 (3.26 to 8.14)	6.73 (5.07 to 8.21)	3.81 (3.02 to 8.13)
Duration of RNA detection (days post-symptom onset)[1]	8 (5 to 11) Range: 0 to 28	8 (5 to 11) Range: 0 to 28	9.5 (7 to 11) Range: 0 to 19	7.5 (4 to 9.5) Range: 0 to 28
Maximum infectious viral load (log plaque forming units/mL * 10^3)[1]	12.00 (7.537 to 14.00)	11.29 (6.824 to 13.85)	8.33 (6.633 to 12.28)	13.77 (9.378 to 14.44)
Duration of infectious viral shedding (days post-symptom onset)[1]	5 (0 to 7) Range: 0 to 13	5 (0 to 7) Range: 0 to 13	5 (0 to 7) Range: 0 to 13	2.5 (0 to 7) Range: 0 to 10
Variant				
Alpha	1 (0.8%)	1 (1.4.0%)	1 (24%)	0 (0.0%)
Delta	43 (34.1%)	30 (40.5%)	18 (42.9%)	12 (37.5%)
Epsilon	4 (3.2%)	4 (5.4%)	3 (7.1%)	1 (3.1%)
Omicron	18 (14.3%)	5 (6.8%)	3 (7.1%)	2 (66.%)
Non-VOI/VOC	42 (33.3%)	24 (32.4%)	12 (28.6%)	12 (37.5%)
Unknown	18 (14.3%)	10 (13.5%)	5 (11.9%)	5 (15.6.0%)

[1]Laboratory values: maximum RNA load, 147 participants (40 primary cases and their 51 household contacts); duration of RNA and infectious shedding, 180 participants (all cases and contacts); infectious viral titers, 56 participants (32 primary cases and their 24 household contacts).

ACCEPTED MANUSCRIPT

Table 3: SARS-CoV-2 infection among household contacts, stratified by vaccination status (Primary cases: N=42; household contacts: N=74). Note: p-values were estimated from unadjusted models using generalized estimating equations, which accounted for clustering by household.

	Vaccinated Primary Cases (n = 21)	Unvaccinated Primary Cases (n = 21)	P-values
Proportion of household contacts who were SARS-CoV-2 positive	14/34 (41%, 25 to 59%)*	18/40 (45%, 29 to 62%)	0.60
Proportion of vaccinated household contacts who were SARS-CoV-2 positive	12/29 (41%, 24 to 61%)	0/2 (0%, 0 to 84%)	NA
Proportion of partially vaccinated household contacts who were SARS-CoV-2 positive	NA	2/4 (50%, 7 to 93%)	NA
Proportion of unvaccinated household contacts who were SARS-CoV-2 positive	2/5 (40%, 5.3 to 85%)	16/34(47%, 30 to 65%)	0.88

*proportion (95% CI)

NA: not applicable due to the inability of a statistical model to converge with a zero-value cell of a strata.

ACCEPTED MANUSCRIPT

Table 4: SARS-CoV-2 infection among all household contacts and vaccinated household contacts, stratified by viral variant of the primary case (Primary cases: N=42; household contacts: N=74). Unvaccinated and partially vaccinated primary cases and household contacts are not presented. Missing were unknown or Alpha variant (not shown because only one primary case).

	Number of primary cases	Proportion of household contacts who were SARS-CoV-2 positive	Number of vaccinated primary index cases	Proportion of vaccinated household contacts who were SARS-CoV-2 positive
Non-VOI/VOC	12	12/22 (55%, 32 to 76%)*	0	NA
Delta	18	12/29 (41%, 15 to 51%)	18	10/24 (42%, 22 to 63%)
Epsilon	3	1/8 (13%, 0.0 to 53%)	0	NA
Omicron	3	2/5 (40%, 5.3 to 85%)	3	2/5 (40%, 5.3 to 85%)
Alpha	1	1/2 (50%, 1.3 to 99%)	0	NA
Unknown	5	5/15 (33%, 12 to 62%)	0	NA

*proportion (95% CI)

ACCEPTED MANUSCRIPT

Table 5: Host and viral determinants of cases associated with infection among household contacts* (Primary cases: N=42; household contacts: N=74). Note: risk ratios indicate cumulative incidence ratios. Confidence interval = CI.

Characteristics of the primary case	Household cumulative incidence risk	Unadjusted		Adjusted#*	
	n/N (%)	Risk Ratio	95% CI	Risk Ratio	95% CI
Age					
<18	2/12 (17)	Ref	Ref	Ref	Ref
18-44	19/43 (44)	2.15	0.58, 8.04	1.67	0.41, 6.72
≥45	11/19 (58)	3.05	0.80, 11.6	1.91	0.46, 7.97
Female	24/41 (59)	2.21	1.07, 4.55	1.57	0.71, 3.46
Race/ethnicity					
Hispanic/Latino	6/11 (55)	Ref	Ref	Ref	Ref
White	20/46 (44)	0.81	0.38, 1.75	0.47	0.19, 1.16
Black/African American	1/2 (50)	0.92	0.19, 4.36	0.33	0.13, 0.89
Asian	2/12 (17)	0.36	0.08, 1.59	0.34	0.09, 1.35
BMI					
<25	11/35 (31)	Ref	Ref	Ref	Ref
25-30	11/18 (61)	1.75	0.90, 3.42	1.60	0.67, 3.86
>30	7/17 (41)	1.19	0.51, 2.81	0.93	0.37, 2.35
Vaccinated cases at baseline	14/34 (41)	0.85	0.47, 1.55	1.29	0.52, 3.21
At least one comorbidity	15/18 (83)	2.71	1.66, 4.43	2.11	1.16, 3.84
Greater maximum RNA viral load (log copies/mL)**	39/60 (65)	1.10	0.97,1.25	1.08	0.96, 1.23
Longer duration of RNA detection (days) **	18/41 (44)	1.02	0.94,1.10	1.01	0.96, 1.07
10 or more days of RNA detection	18/41 (44)	1.05	1.03, 1.06	1.05	1.04, 1.08
Greater maximum infectious viral load (log pfu/mL)**	10/20 (50)	0.96	0.87,1.06	0.99	0.87, 1.11
Longer duration of infectious virus detection (days)**	25/46 (54)	1.09	0.99,1.20	1.05	0.96, 1.16
Variant					

ACCEPTED MANUSCRIPT

Delta	11/22 (50)	Ref	Ref	Ref	Ref
Omicron	2/5 (40)	0.84	0.32, 2.21	0.98	0.33, 2.92
Epsilon	1/8 (13)	0.36	0.05, 2.81	0.56	0.08, 4.13
Non-VOI/VOC	12/22 (55)	1.08	0.56, 2.10	0.93	0.45, 1.93
Alpha	1/2 (50)	1.01	0.64, 1.60	1.28	0.75, 2.18
Unknown	5/15 (33)	0.69	0.27, 1.77	0.87	0.34, 2.22

*Each cell in this table represents the output of a generalized estimating equation with modified Poisson regression, accounting for clustering by household.

#Each adjusted model includes factors of case and contact and was developed with a Directed Acyclic Graph (DAG) of potential confounders, mediators, and colliders (age, sex, race/ethnicity, etc.). Please see the **Supplemental Material** for a listing of the variables included in each model.

**Relative values for greater viral load and longer duration of shedding determined as upper 50th percentile of observed values from primary cases

ACCEPTED MANUSCRIPT

Table 6: Host susceptibility factors associated with infection among household contacts* (Primary cases: N=42; household contacts: N=74). Note: risk ratios indicate cumulative incidence ratios. Confidence interval = CI.

Characteristics of the household contact	Household cumulative incidence risk	Unadjusted		Adjusted#*	
	n/N (%)	Risk Ratio	95% CI	Risk Ratio	95% CI
Age					
<18	7/15 (47)	Ref	Ref	Ref	Ref
18-44	17/38 (45)	0.92	0.57, 1.50	1.02	0.59, 1.78
≥45	8/21 (38)	0.79	0.43, 1.46	0.86	0.45, 1.63
Female	17/38 (45)	0.91	0.56, 1.49	1.05	0.65,1.71
Race/ethnicity					
Hispanic/Latino	5/11 (45)	Ref	Ref	Ref	Ref
White	21/44 (48)	1.18	0.59, 2.37	1.15	0.59, 2.23
Black/African American	2/4 (50)	0.84	0.26, 2.69	0.83	0.28, 2.43
Asian	2/10 (20)	0.47	0.13, 1.76	0.46	0.13, 1.60
BMI					
<25	15/39 (39)	Ref	Ref	Ref	Ref
25-30	9/19 (47)	0.96	0.53, 1.75	0.88	0.46, 1.69
>30	5/11 (45)	0.95	0.42, 2.18	0.98	0.40, 2.36
Vaccinated contacts at baseline	12/31 (39)	0.77	0.44, 1.36	0.90	0.37, 2.19
At least one comorbidity	5/13 (38)	0.92	0.48, 1.79	0.72	0.36, 1.48

*Each cell in this table represents the output of a generalized estimating equation with modified Poisson regression, accounting for clustering by household.

#Each adjusted model includes factors of case and contact and was developed with a Directed Acyclic Graph (DAG) of potential confounders, mediators, and colliders (age, sex, race/ethnicity, etc.). Please see the **Supplemental Material** for a listing of the variables included in each model.

ACCEPTED MANUSCRIPT

Figure 1: Flow diagram of household cohort. Households from the initially enrolled cohort were excluded if index cases were determined to have co-primary cases criteria or if single primary cases were asymptomatic. Note: We identified 42 households with single primary cases, also referred to as primary cases. For analysis, we excluded 21 households with co-primary cases and 2 households with asymptomatic primary cases.

Figure 2: Phylogenetic tree situating SARS-CoV-2 associated genomes (colored dots) identified over time from household clusters in the local pandemic of San Francisco and Alameda Counties, California (N=72 participants). Inset, colors represent Pango lineages reflected in the tree.

Figure 3: Phylogenetic sub-tree of a household cluster determined to be co-primary cases. The index case and household contact reported symptoms two days apart (serial interval = 2) as well as a shared exposure, but genomic epidemiology showed the SARS-CoV-2 sequences belonged to distinct monophyletic clades. These participants were classified as having had unrelated sequences.

Figure 4: Flow diagram of determination of single primary cases from the decision algorithm. Determination of single primary or co-primary cases from 65 households (65 index cases). Note: We identified 42 households with single primary cases, also referred to as primary cases.

ACCEPTED MANUSCRIPT

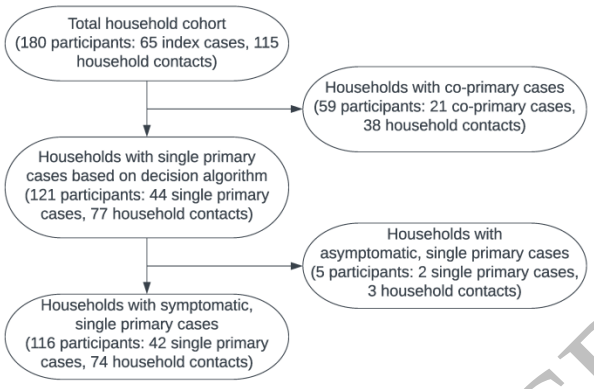
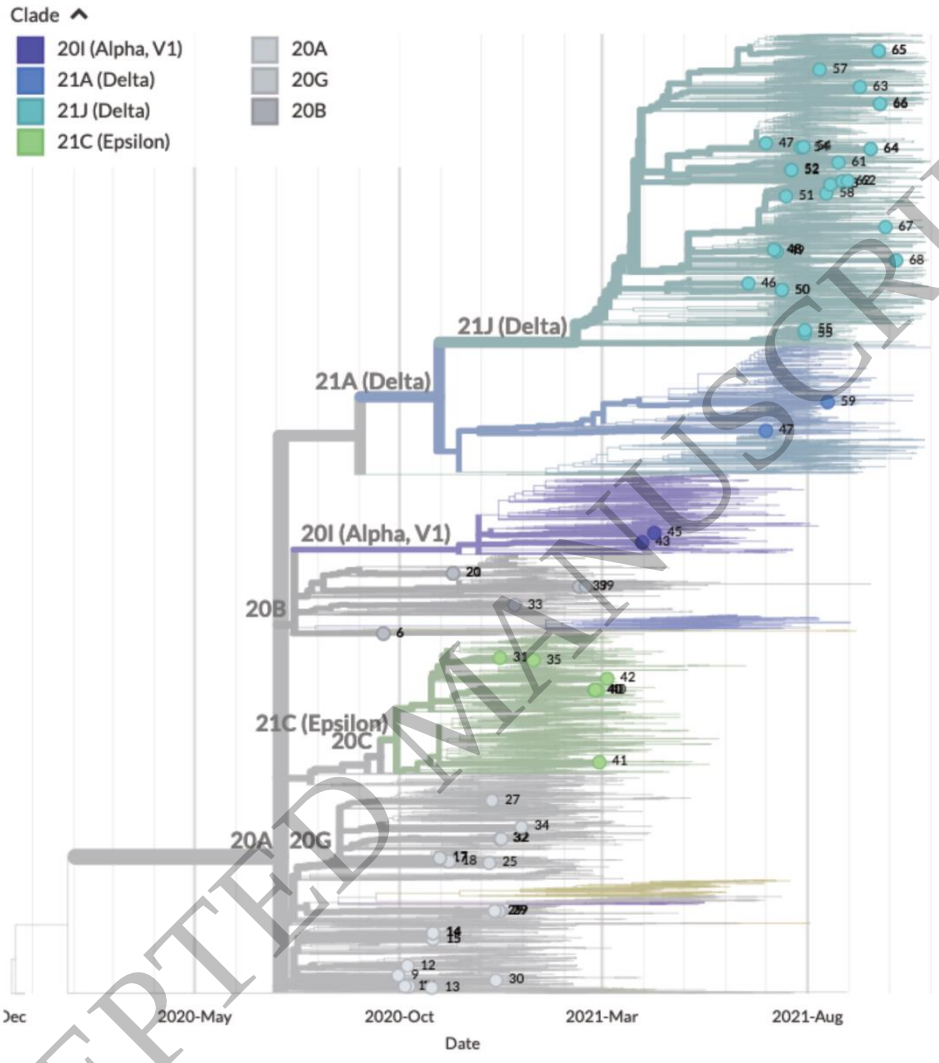


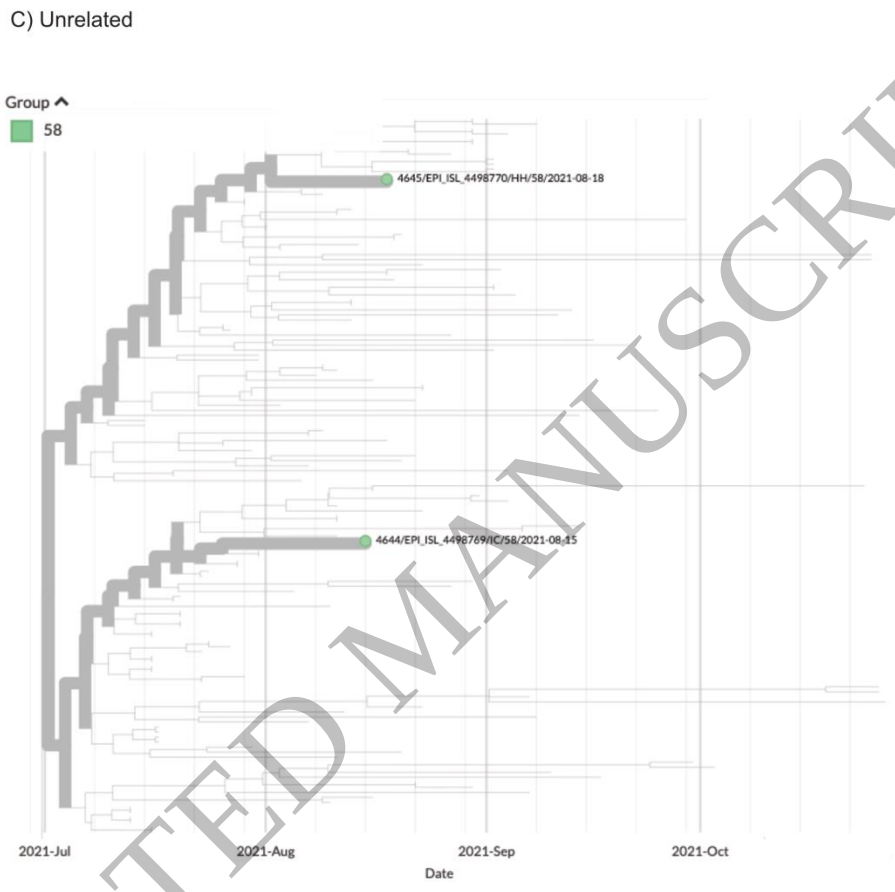
Figure 1
279x216 mm (x DPI)

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

Figure 2
395x559 mm (x DPI)



ACCEPTED MANUSCRIPT

Figure 3
395x559 mm (x DPI)

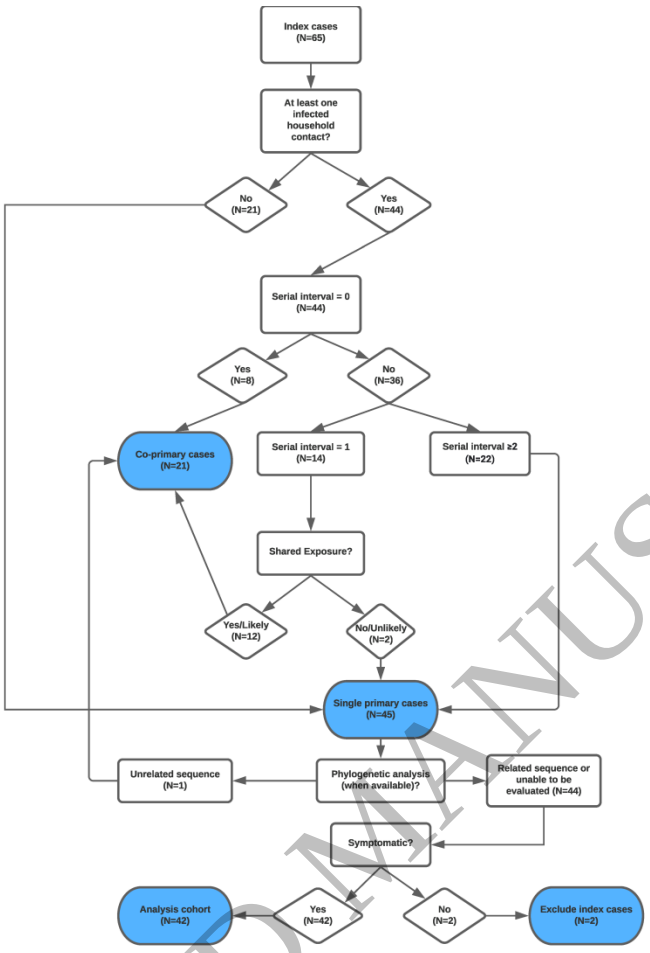


Figure 4
279x216 mm (x DPI)

ACCEPTED MANUSCRIPT