Epidemiological analysis of the SARS-CoV-2 generation time for the Alpha and Delta variants

WS Hart, E Miller, NJ Andrews, P Waight, PK Maini, S Funk, RN Thompson

Appendix 2

Contents

Supplementary Methods

Details of household study protocol

PCR positive index cases were identified through community testing in England (Pillar 2 testing).[1] Anybody in England can access a Pillar 2 test if they have symptoms of COVID-19 or if they are part of a local or national mass testing programme.[1] Index cases were contacted by study nurses to request their participation, and the participation of household contacts, in the study (in our analysis, the index case was not assumed to necessarily be the primary case that began the household transmission cluster). Nurses interviewed the index case and all consenting household members at recruitment (Day 1) and Day 21 to document symptom onset dates (for individuals who developed one of a set of solicited symptoms), dates of any previous positive test result (where available) and vaccination history (number of vaccine doses received, the vaccine type, and the date(s) of vaccination). Only households with at least one contact who had not had a previous positive PCR test were included. Households in which the index case was under 18 years of age were excluded to ensure comparability in age between unvaccinated and vaccinated index cases, as only adults were eligible for vaccination during the study period. Swabs in virus transport medium (∑-Virocult® from mwe Medical Wire) were couriered to households for the index case and household contacts to take combined nose and throat swabs on recruitment (Day 1), and again on Days 3 and 7. Completed tests were analysed by dual target RT-PCR (ORF and E genes) at the Virus Reference Department, Colindale. Whole genome sequencing was performed on RT-PCR positive samples[2] to identify the variant responsible for the infection as part of the COVID-19 Genomics UK Consortium (COG-UK) initiative.[3] Viral amplicons were sequenced using Illumina library preparation kits (Nextera) on Illumina short-read sequencing machines (Nextseq or Hiseq), as described by Jeffery-Smith and colleagues.[4]

We assumed individuals were infected during the household transmission cluster if they returned a positive PCR test (including tests taken up to 28 days before the index case first tested positive) and/or developed at least one of the solicited symptoms. Otherwise, we assumed individuals remained uninfected if they returned only negative tests. Household members who did not participate in the study, or withdrew before either taking a test or developing symptoms, were excluded from our analysis (but were included in the household size).

General mathematical modelling framework

We assumed the expected force of infection, $\beta(\tau)$, exerted by an infected host onto each susceptible member of their household at time τ since infection, to be given by

$$\beta(\tau) = \frac{\beta_0}{(n-1)} f(\tau),$$

for a host who develops symptoms, and

$$\beta(\tau) = \frac{\alpha_A \beta_0}{(n-1)} f(\tau),$$

for a host who remains asymptomatic throughout infection. Here:

- β_0 represents the overall transmissibility for a host who develops symptoms.
- n is the household size (we assumed frequency-dependent household transmission).
- $f(\tau)$ is the intrinsic generation time distribution (i.e., the generation time distribution in the absence of susceptible depletion during the course of infection).
- α_A is the relative infectiousness of infected hosts who remain asymptomatic throughout infection (compared to infected hosts who develop symptoms). We assumed a value of $\alpha_A = 0{\cdot}35$ in our main analysis[5] (however, since other studies have estimated different values of this quantity,[6,7] we considered the sensitivity of our results to this value in Figure S7).

The (expected) instantaneous probability density of the infected host under consideration infecting a given susceptible household member (denoted j) is given by $\eta_j \beta(\tau)$, where η_j describes the relative susceptibility of individual j. This susceptibility was assumed to depend on vaccination status according to previous estimates of vaccine efficacy against infection by the Alpha and Delta variants (see Table S3).[8]

In our approach for estimating the generation time, infectiousness is explicitly linked to symptoms, so that the infectiousness profile of a given infector depends on exactly when they develop symptoms. Throughout the Supplementary Material, we denote the expected force of infection exerted by an infected host onto each susceptible member of their household at time τ since infection, conditional on incubation period τ_{inc}, by $\beta(\tau \mid \tau_{inc})$. Therefore, we have

$$\beta(\tau) = \int_0^\infty \beta(\tau \mid \tau_{inc}) f_{inc}(\tau_{inc}) \mathrm{d}\tau_{inc},$$

where $f_{inc}(\tau_{inc})$ denotes the probability density function of the incubation period (i.e., $\beta(\tau)$ is the average of $\beta(\tau \mid \tau_{inc})$ over the incubation period distribution).

Individual infectiousness model

We considered a mathematical model in which each infected host (who develops symptoms) progresses through independent latent (*E*), presymptomatic infectious (*P*) and symptomatic infectious (*I*) stages of infection. The transmission rates of the host during the *P* and *I* stages are denoted by β_P and β_I, respectively, and we denote their ratio $\alpha_P = \beta_P/\beta_I$. We assumed the duration of each stage, denoted $y_{E/P/I}$, to be gamma distributed:

$$y_E \sim \mathrm{Gamma}(k_E, 1/(k_{inc}\gamma)),$$
$$y_P \sim \mathrm{Gamma}(k_P, 1/(k_{inc}\gamma)),$$
$$y_I \sim \mathrm{Gamma}(k_I, 1/(k_I\mu)),$$

where we write $X \sim \mathrm{Gamma}(a, b)$ for a gamma distributed random variable with shape parameter a and scale parameter b. We assumed that $k_E + k_P = k_{inc}$, so that the incubation period, $\tau_{inc} = y_E + y_P$, is gamma distributed, with

$$\tau_{inc} \sim \mathrm{Gamma}(k_{inc}, 1/(k_{inc}\gamma)).$$

Hosts who remain asymptomatic throughout infection were assumed to follow the same *E*/*P*/*I* stages, although in this case the distinction between the *P* and *I* stages has no epidemiological meaning. Stage durations, as well as the value of α_P, were assumed to be identical for entirely asymptomatic hosts and those who develop symptoms. Similarly, vaccination was not assumed to affect model parameters (other than the relative susceptibility of vaccinated individuals), although we explored the effect of vaccination on realised household generation times in Figure 2. We also considered different values of the relative infectiousness of vaccinated infected individuals in Figure S8.

In our main analysis, we fixed the values of the parameters k_{inc} and $1/\gamma$ (which represent the shape parameter of the incubation period distribution and the reciprocal of the mean incubation period, respectively) in order to obtain an incubation period distribution of mean 5·8 days and standard deviation 3·1 days[9] (we considered the sensitivity of our results to the exact incubation period distribution in Figure S6). The values of k_{inc} and $1/\gamma$ are given in Table S3. We assumed that $k_I = 1$, so the symptomatic infectious period is exponentially distributed. The following quantities were then estimated for the Alpha and Delta variants when we fitted the model to the household transmission data:

- The ratio between the mean latent (*E*) period and the mean incubation (combined *E* and *P*) period, k_E/k_{inc}.
- The mean symptomatic infectious (*I*) period, $1/\mu$.
- The ratio between the transmission rates when potential infectors are in the *P* and *I* stages, α_P.
- The overall transmissibility parameter, β_0.

Conditional infectiousness

For a host who develops symptoms, conditional on incubation period τ_{inc}, their expected infectiousness at time since infection τ is[10]

$$\beta(\tau \mid \tau_{inc}) = \begin{cases} \dfrac{\alpha_P C \beta_0}{(n-1)}(1 - F_{Beta}(1 - \tau/\tau_{inc}; k_P, k_E)), & 0 < \tau < \tau_{inc}, \\ \dfrac{C\beta_0}{(n-1)}(1 - F_I(\tau - \tau_{inc})), & \tau > \tau_{inc}. \end{cases}$$

Here, β_0 is the overall transmissibility parameter, n is the household size, $F_I(y)$ is the cumulative distribution of the duration of the *I* stage, $F_{Beta}(x; a, b)$ is the cumulative distribution of a beta distributed random variable with shape parameters a and b, and the constant

$$C = \frac{k_{inc}\gamma\mu}{\alpha_P k_P \mu + k_{inc}\gamma}$$

is chosen to ensure that the overall infectiousness (i.e., the integral of $\beta(\tau \mid \tau_{inc})$ over all times since infection, τ), averaged over the incubation period distribution, is given by $\beta_0/(n-1)$ (see the next paragraph). The term

$$(1 - F_{Beta}(1 - \tau/\tau_{inc}; k_P, k_E))$$

is the probability that the duration of the *P* stage of infection exceeds $(\tau_{inc} - \tau)$, conditional on incubation period τ_{inc} (i.e., the probability that the host has become infectious by time since infection $\tau < \tau_{inc}$), while

$$(1 - F_I(\tau - \tau_{inc}))$$

is the probability that the duration of the *I* stage exceeds $(\tau - \tau_{inc})$ (i.e., the probability that the host remains infectious at time since infection $\tau > \tau_{inc}$).

The cumulative conditional infectiousness up to time τ can be calculated to be

$$B(\tau \mid \tau_{inc}) = \int_0^{\tau} \beta(\tilde{\tau} \mid \tau_{inc}) \mathrm{d}\tilde{\tau}$$

$$= \begin{cases} (\tau - \tau_{inc})\beta(\tau \mid \tau_{inc}) + \dfrac{\alpha_P C \beta_0}{(n-1)}\left[\dfrac{k_p \tau_{inc}}{k_{inc}}\left(1 - F_{Beta}(1 - \tau/\tau_{inc}; k_P + 1, k_E)\right)\right], & 0 \le \tau < \tau_{inc}, \\ (\tau - \tau_{inc})\beta(\tau \mid \tau_{inc}) + \dfrac{C\beta_0}{(n-1)}\left[\dfrac{\alpha k_p \tau_{inc}}{k_{inc}} + \dfrac{1}{\mu} F_{Gamma}\left(\tau - \tau_{inc}; k_I + 1, \dfrac{1}{k_I \mu}\right)\right], & \tau \ge \tau_{inc}, \end{cases}$$

where $F_{Gamma}(x; a, b)$ is the cumulative distribution of a gamma distributed random variable with shape parameter a and scale parameter b. The total force of infection on each household member (over the infector's course of infection) is then

$$B(\infty \mid \tau_{inc}) = \frac{C\beta_0}{(n-1)}\left(\alpha_P \times \frac{k_P \tau_{inc}}{k_{inc}} + \frac{1}{\mu}\right) = \frac{\beta_0}{(n-1)}\left(\frac{\alpha_P k_P \gamma \mu \tau_{inc} + k_{inc}\gamma}{\alpha_P k_P \mu + k_{inc}\gamma}\right).$$

Here,

$$\alpha_P \times \frac{C\beta_0}{(n-1)} \times \frac{k_P \tau_{inc}}{k_{inc}}$$

represents the total force of infection exerted by the host on each susceptible household member during the *P* stage of infection (where $k_P \tau_{inc}/k_{inc}$ is the mean duration of the *P* stage, conditional on incubation period τ_{inc}), and

$$\frac{C\beta_0}{(n-1)} \times \frac{1}{\mu}$$

represents the analogous total force of infection exerted during the *I* stage (where $1/\mu$ is the mean duration of the *I* stage. The mean of the expression for $B(\infty \mid \tau_{inc})$ over the incubation period distribution is then $\beta_0/(n-1)$.

For a host who remains asymptomatic throughout infection, conditional on the combined duration of the *E* and *P* stages ($\tau_{inc} = y_E + y_P$), the infectiousness ($\beta(\tau \mid \tau_{inc})$) is given by the product of α_A and the corresponding expression for a host who develops symptoms. We note that in this case, τ_{inc} has no epidemiological interpretation, but this formulation was convenient when fitting the model to data (see "Parameter fitting" below).

Intrinsic generation time distribution

Here, we consider the intrinsic generation time distribution, $f(\tau)$ – i.e., the generation time distribution assuming that a constant supply of susceptible individuals is available to each infected host throughout their infection. This distribution describes the relative expected infectiousness profile of an infected individual at each time since infection, normalised so that it represents a valid probability density function.

The generation time for a single transmission, τ_{gen}, can be written as

$$\tau_{gen} = y_E + y^*,$$

where y_E is the length of the latent (*E*) stage, and y^* is the interval from the start of the presymptomatic infectious (*P*) stage to the transmission occurring (i.e., the time from becoming infectious to the transmission occurring). As shown in our previous work,[10] the intrinsic probability density function of y^* (neglecting the effect of susceptible depletion during infection) is

$$f^*(y^*) = C\left(\alpha_P\left(1 - F_P(y^*)\right) + \int_0^{y^*} \left(1 - F_I(y^* - y_P)\right) f_P(y_P) \mathrm{d}y_P\right).$$

This density represents the (normalised) expected infectiousness profile of an infected host at time y^* since entering the *P* stage. The term

$$\big(1 - F_P(y^*)\big)$$

represents the probability that the host remains in the *P* stage for a duration of at least y^*, while

$$\int_0^{y^*} \big(1 - F_I(y^* - y_P)\big) f_P(y_P) \mathrm{d}y_P$$

gives the probability that the infected host has entered, but not yet left, the *I* stage after time y^* since entering the *P* stage (the integral is obtained by conditioning on the duration of the *P* stage, y_P, where $0 < y_P < y^*$).

Using the expression for $f^*(y^*)$, it can be shown that the moments of this distribution are

$$E[(y^*)^m] = \frac{C}{m+1}(\alpha_P E[{y_P}^{m+1}] + E[(y_P + y_I)^{m+1} - {y_P}^{m+1}]).$$

In particular,

$$E[y^*] = \frac{C}{2}(\alpha_P E[{y_P}^2] + 2E[y_P]E[y_I] + E[{y_I}^2]),$$

and

$$\mathrm{Var}[y^*] = \frac{C}{3}(\alpha_P E[{y_P}^3] + 3E[{y_P}^2]E[y_I] + 3E[y_P]E[{y_I}^2] + E[{y_I}^3]) - (E[y^*])^2.$$

Note that for a gamma distributed random variable, $X \sim \mathrm{Gamma}(a, b)$, we have

$$E[X^m] = \frac{\Gamma(a+m)}{\Gamma(a)} b^m = a(a+1) \ldots \big(a + (m-1)\big) b^m.$$

Therefore, for gamma distributed stage durations, explicit expressions can be obtained for the mean and variance of the intrinsic generation time distribution,

$$E\big[\tau_{gen}\big] = E[y_E] + E[y^*],$$
$$\mathrm{Var}\big[\tau_{gen}\big] = \mathrm{Var}[y_E] + \mathrm{Var}[y^*],$$

where the latter expression holds because y_E and y^* are assumed to be independent.

Likelihood function

Here, we consider a household of size n, in which n_I household members become infected (of whom n_S develop symptoms and n_A remain asymptomatic throughout infection) and $n_U = n - n_I$ remain uninfected. We derive an expression for the likelihood of the vector of unknown model parameters,

$$\theta = (k_E/k_{inc}, 1/\mu, \alpha_P, \beta_0),$$

(where θ was assumed to be different for the Alpha and Delta variants), given:

i. The entire sequence of infection times of individuals in the household ($t_1 < \cdots < t_{n_I}$).
ii. The precise symptom onset time ($t_{s,j}$) of each host, j, who develops symptoms.
iii. The times at which entirely asymptomatic infected hosts enter the *I* stage of infection (also denoted by $t_{s,j}$).

Since exact infection and symptom onset times were not available within study households, we used data augmentation MCMC to fit the two models to the household transmission data using this likelihood function (see further details below).

When deriving the likelihood, we made the following simplifying assumptions:

- The virus is introduced once into the household (i.e., no subsequent infections from the community occur following the infection of the primary case; results excluding data from households in which long gaps between symptom onset dates suggested the possibility of multiple introductions of the virus into the household are shown in Figure S10).
- No co-primary cases (we relaxed this assumption to account for the possibility of co-primary cases in Figure S9 – see also the "Extension of framework to account for co-primary cases" section below).
- Potential bias towards more recent infection of the primary host if community prevalence is increasing, or less recent if prevalence is decreasing,[11–13] was neglected.

We denote the conditional infectiousness of household member j, at time τ since infection, by $\beta_j(\tau \mid t_{s,j} - t_j)$, where $(t_{s,j} - t_j)$ corresponds to the incubation period for a host who develops symptoms. The total (instantaneous) force of infection exerted at time t on each susceptible household member is then

$$\lambda(t) = \sum_{j=1}^{n_I} \beta_j (t - t_j \mid t_{s,j} - t_j),$$

where $\beta_j(t - t_j \mid t_{s,j} - t_j) = 0$ for $t \leq t_j$ (i.e., individual j does not exert any force of infection at or before their time of infection), and the cumulative force of infection is

$$\Lambda(t) = \int_{-\infty}^{t} \lambda(s)\mathrm{d}s = \sum_{j=1}^{n_I} B_j(t - t_j \mid t_{s,j} - t_j),$$

where $B_j(\tau \mid t_{s,j} - t_j)$ denotes the cumulative conditional infectiousness of individual j.

We accounted for the fact that individuals with a higher relative susceptibility (i.e., unvaccinated individuals) were more likely to be the primary case (here denoted as individual 1) by including a likelihood contribution of

$$\frac{\eta_1}{\sum_{j=1}^{n} \eta_j},$$

where η_j denotes the relative susceptibility of host j. Note however that the identity of the primary case within each study household was estimated during the data augmentation MCMC parameter fitting procedure (i.e., the individual with the earliest imputed infection time was taken to be the primary in each step of the chain). For $k = 2, \dots, n_I$, conditional on the sequence of infection times up to time t_k, the probability of host k becoming infected at time t_k is given by

$$\eta_k \lambda(t_k) \exp\big(-\eta_k \Lambda(t_k)\big),$$

where $\exp(-\eta_k \Lambda(t_k))$ represents the probability of host k avoiding infection up to time t.[14,15] For $k = n_I + 1, \dots, n$, conditional on the entire sequence of infection times, $t_1, \dots, t_{n_I}$, the probability of host k remaining uninfected is given by $\exp(-\eta_k \Lambda(\infty))$.

The likelihood, $L(\theta)$, can therefore be written as

$$L(\theta) = \prod_{k=1}^{n} L_{k,1}(\theta) L_{k,2}(\theta).$$

Here, $L_{k,1}(\theta)$ is the contribution to the likelihood from the transmission, or absence of transmission, to host k, i.e.,

$$L_{k,1}(\theta) = \begin{cases} \dfrac{\eta_1}{\sum_{j=1}^{n} \eta_j}, & \text{for } k = 1; \\ \eta_k \lambda(t_k) \exp\left(-\eta_k \Lambda(t_k)\right), & \text{for } k = 2, \dots, n_I; \\ \exp(-\eta_k \Lambda(\infty)), & \text{for } k = n_I + 1, \dots, n. \end{cases}$$

For an infected host k, $L_{k,2}(\theta)$ is the likelihood contribution from their incubation period (or for an entirely asymptomatic infected host, the corresponding combined duration of the *E* and *P* stages of infection), i.e.,

$$L_{k,2}(\theta) = \begin{cases} f_{inc}(t_{s,k} - t_k), & \text{for } k = 1, \dots, n_I; \\ 1, & \text{for } k = n_I + 1, \dots, n; \end{cases}$$

where f_{inc} is the probability density function of the incubation period.

Parameter estimation

Unknown model parameters were estimated for each variant using data augmentation MCMC. The observed transmission data comprised information about whether or not individuals were ever infected and/or displayed symptoms, symptom onset dates, and for some individuals an upper bound on their infection time (corresponding to the date of a positive PCR test). These data were augmented with:

i. The infection time, t_j, of each infected host.
ii. The time, $t_{s,j}$, at which each infected host transitioned from the *P* stage to the *I* stage of infection (this corresponds to the symptom onset time for a host who developed symptoms).

No prior assumptions were made about the order of transmissions within the household (instead, the imputed infection times were used to estimate the order of transmissions).

For each variant, we assumed gamma distributed priors for $1/\mu$ (the mean duration of the symptomatic infectious period), α_P (the ratio of the transmission rates during the presymptomatic infectious and symptomatic infectious stages of infection) and β_0 (the overall transmissibility). We assumed a prior with mean 2.0 for α_P –

this choice was informed by our previous work indicating this quantity to be above one for previously circulating SARS-CoV-2 variants.[10,16] A beta prior was used for k_E/k_{inc} (the ratio of the mean durations of the latent and incubation periods, which was constrained to lie between 0 and 1). All prior distributions were chosen to limit the prior probabilities of extreme parameter values, and were taken to be independent (between different parameters, and between variants). The exact priors we used are given in Table S4.

In the description of the parameter fitting procedure below, we denote the augmented data by

$$\boldsymbol{t} = \left(\boldsymbol{t}^{(1)}, \ldots, \boldsymbol{t}^{(M)}\right),$$

where $\boldsymbol{t}^{(m)}$ represents the augmented data from household $m = 1, \ldots, M$, and M is the total number of households. We denote the vector of fitted parameters by θ, where

$$\theta = (\theta^{(Alpha)}, \theta^{(Delta)})$$

now includes the parameter values for both variants. We write the (overall) likelihood as

$$L(\theta; \boldsymbol{t}) = \prod_{m=1}^{M} L^{(m)}\left(\theta^{(V_m)}; \boldsymbol{t}^{(m)}\right),$$

where V_m denotes the variant responsible for infections in household m, and the likelihood contributions, $L^{(m)}\left(\theta^{(V_m)}; \boldsymbol{t}^{(m)}\right)$, were computed as described in the previous section (i.e., all households in the study were assumed to be independent). Finally, we denote the prior density of θ by $\pi(\theta)$.

In each step of the chain, we carried out (in turn) one of the following:

1. Propose new values for each entry of the vector of model parameters, θ, using a multivariate normal proposal distribution (around the value of θ in the previous step of the chain) with covariance matrix Σ_1. Accept the proposed parameters, θ_{prop}, with probability

$$\min\left(\frac{L\left(\theta_{prop}; \boldsymbol{t}\right)\pi(\theta_{prop})}{L(\theta_{old}; \boldsymbol{t})\pi(\theta_{old})}, 1\right),$$

   where θ_{old} denotes the vector of parameter values from the previous step of the chain, and where the augmented data, $\boldsymbol{t}$, remain unchanged in this step (we defined the likelihood to be zero for negative parameter values, so that if any entry of the proposed parameter vector, θ_{prop}, is negative, then the proposal is rejected).

2. Propose new values for the precise symptom onset times of each symptomatic infected host, using independent uniform proposal distributions (within the day of symptom of onset for each host). For each household, m, accept the proposed augmented data, $\boldsymbol{t}_{prop}^{(m)}$, from that household with probability

$$\min\left(\frac{L^{(m)}\left(\theta^{(V_m)}; \boldsymbol{t}_{prop}^{(m)}\right)}{L^{(m)}\left(\theta^{(V_m)}; \boldsymbol{t}_{old}^{(m)}\right)}, 1\right),$$

   where $\boldsymbol{t}_{old}^{(m)}$ denotes the corresponding augmented data from the previous step of the chain, and where the model parameters, θ, remain unchanged in this step (i.e., proposed times are accepted/rejected independently for each household, according to the likelihood contribution from that household).

3. Propose new values for the infection time of one randomly chosen infected host in each household (either symptomatic or asymptomatic), using independent normal proposal distributions (around the equivalent times in the previous step of the chain) with standard deviation σ_3. For each household, m, accept the proposed augmented data, $\boldsymbol{t}_{prop}^{(m)}$, from that household with probability

$$\min\left(\frac{L^{(m)}\left(\theta^{(V_m)}; \boldsymbol{t}_{prop}^{(m)}\right)}{L^{(m)}\left(\theta^{(V_m)}; \boldsymbol{t}_{old}^{(m)}\right)}, 1\right).$$

4. Propose new values for both the infection time, t, and the time of the start of the *I* stage, t_s, holding $t_s - t$ constant, for one randomly chosen asymptomatic infected host in each household (in households where there was at least one; i.e., the timing of infection is updated, holding the combined duration of the *E* and *P* stages constant – this is different to step 3 above, in which this combined stage duration is updated), using independent normal proposal distributions (around the equivalent times in the previous step of the chain) with standard

deviation σ_4. For each household, m, accept the proposed augmented data, $\boldsymbol{t}_{prop}^{(m)}$, from that household with probability

$$\min\left(\frac{L^{(m)}\left(\theta^{(V_m)};\boldsymbol{t}_{prop}^{(m)}\right)}{L^{(m)}\left(\theta^{(V_m)};\boldsymbol{t}_{old}^{(m)}\right)},1\right).$$

The covariance matrix Σ_1 was chosen to ensure an acceptance rate of approximately 30% in step 1 above (a correlation of 0·5 was used between the proposal distributions of k_E/k_{inc} and α_P, and between those of $1/\mu$ and α_P, for each variant; all other off-diagonal entries of Σ_1 were set to zero). Similarly, the tuning parameters σ_3 and σ_4 were chosen to ensure an overall acceptance rate (averaged over the acceptance rates for each household) of approximately 30% in steps 3 and 4, respectively.

The chain was run for 10,000,000 iterations; the first 2,000,000 iterations were discarded as burn-in. Posteriors were obtained by recording only every 100 iterations of the chain.

Sampling of household generation times

Realised household generation times are shorter than predicted by the intrinsic generation time distribution, $f(\tau)$, due to the depletion of susceptible household members before longer generation times can be attained.[16–18] For example, if infected hosts are (on average) equally infectious at two times since infection, $\tau_1 < \tau_2$, then $f(\tau_1) = f(\tau_2)$. However, because the number of susceptible household members may decrease between these two times (i.e., either the host under consideration, or another infected household member, may transmit the virus within the household in the intervening time), then transmission is in fact more likely to occur in a household at the earlier time, τ_1, when more susceptibles are available.

Therefore, we also estimated the realised generation times within the study households. These were sampled during the parameter fitting procedure as follows: in a given household, for $k > 1$, the instantaneous probability density of individual k (the k^{th} household member to be infected) being infected at time t_k is

$$\eta_k \lambda(t_k) = \eta_k \sum_{j=1}^{k-1} \beta_j(t_k - t_j \mid t_{s,j} - t_j),$$

where:

- $t_1 < t_2 < \cdots t_{(k-1)}$ are the infection times of the first $(k-1)$ individuals to be infected in the household.
- $t_{s,j}$ is the symptom onset time of individual j (or the entry time into the I stage of infection for an entirely asymptomatic infected host).
- $\beta_j(\tau \mid t_{s,j} - t_j)$ is the (conditional) infectiousness of individual j at time since infection τ.
- η_k is the relative susceptibility of individual k.

Conditional on this transmission occurring at time t_k, the probability that host j is responsible for the transmission (for $j = 1, \ldots, (k-1)$) is given by

$$p_j = \frac{\beta_j\left(t_k - t_j \mid t_{s,j} - t_j\right)}{\lambda(t_k)},$$

i.e., the generation time corresponding to the k^{th} transmission is $(t_k - t_j)$ with probability p_j.

During the parameter fitting procedure, we calculated the probabilities p_j corresponding to each transmission given the augmented data. We used these probabilities to sample the infector responsible for each transmission and therefore obtain samples of estimated household generation times. Household generation times could then be compared by variant (Figure 1C-D), as well as other factors such as vaccination status, age and infection month (Figure 2; note that these three factors were not directly accounted for when fitting model parameters, other than our assumption of a dependence of relative susceptibility on vaccination status).

Extension of framework to account for co-primary cases

In our main analysis, we assumed that each household transmission chain was initiated by a single primary case, so that all other infected household members were infected from within the household. However, we also relaxed this assumption by extending our modelling framework to account for the possibility of co-primary cases (Figure S9). Rather than assuming that all co-primary cases were infected at exactly the same time, we instead assumed that each household member could be infected at any time during a primary infection event that

was taken to last one day (the choice of one day was arbitrary but in principle any duration could be used). This enabled us to easily incorporate the possibility of co-primary cases into our data augmentation MCMC approach by adapting the likelihood function as described below.

As in the "Likelihood function" section above, we again here consider a household of size n, in which n_I household members become infected (of whom n_S develop symptoms and n_A remain asymptomatic throughout infection) and n_U remain uninfected. We now denote the total force of infection exerted on each susceptible member of the household by other household members at time t by $\lambda_h(t)$, and the cumulative force of infection by $\Lambda_h(t)$ (i.e., these are equal to the quantities denoted by $\lambda(t)$ and $\Lambda(t)$, respectively, in the "Likelihood function" section above). Assuming each (susceptible) household member is also subject to a constant force of infection, β_p, during a primary transmission event taking place between times $t_{p,start}$ and $t_{p,end}$, the total force of infection exerted on each susceptible household member at time t is

$$\lambda(t) = \lambda_p(t) + \lambda_h(t),$$

where

$$\lambda_p(t) = \begin{cases} \beta_p, & t_{p,start} \le t \le t_{p,end}; \\ 0, & \text{otherwise.} \end{cases}$$

The cumulative force of infection is

$$\Lambda(t) = \Lambda_p(t) + \Lambda_h(t),$$

where

$$\Lambda_p(t) = \int_{-\infty}^{t} \lambda_p(s)\mathrm{d}s = \frac{\beta_p}{2}\left(t_{p,end} - t_{p,start} + \left|t - t_{p,start}\right| - \left|t_{p,end} - t\right|\right).$$

We took $t_{p,start}$ and $t_{p,end}$ to be the start and end of the day of the first household member becoming infected, respectively.

The likelihood contribution from the household, $L(\theta)$, where θ is the vector of unknown model parameters, is then given by

$$L(\theta) = \frac{1}{1 - \exp\left(-\beta_p\left(\sum_{j=1}^{n}\eta_j\right)(t_{p,end} - t_{p,start})\right)} \prod_{k=1}^{n} L_{k,1}(\theta)L_{k,2}(\theta).$$

Here,

$$L_{k,1}(\theta) = \begin{cases} \eta_k\lambda(t_k)\exp\left(-\eta_k\Lambda(t_k)\right), & \text{for } k = 1, \dots, n_I; \\ \exp(-\eta_k\Lambda(\infty)), & \text{for } k = n_I + 1, \dots, n; \end{cases}$$

and

$$L_{k,2}(\theta) = \begin{cases} f_{inc}(t_{s,k} - t_k), & \text{for } k = 1, \dots, n_I; \\ 1, & \text{for } k = n_I + 1, \dots, n; \end{cases}$$

where f_{inc} is the probability density function of the incubation period. The factor

$$\frac{1}{1 - \exp\left(-\beta_p\left(\sum_{j=1}^{n}\eta_j\right)(t_{p,end} - t_{p,start})\right)}$$

is included to condition on at least one household member becoming infected during the primary transmission event.

Using this likelihood function, we used the same parameter fitting procedure described in the "Parameter estimation" section above to fit the household transmission model to data. For simplicity, we assumed a value of β_p = 0·22, which corresponds to a probability of infection during the primary transmission event (for an unvaccinated individual) of

$$1 - \exp\left(-\beta_p\left(t_{p,end} - t_{p,start}\right)\right) = 1 - \exp\left(-\beta_p \times (1\text{ day})\right) = 0{\cdot}2.$$

We note however that we estimated whether each specific individual in the household dataset was a co-primary in each step of the parameter fitting procedure, depending on the augmented data. The probability that specific individuals were determined to be co-primaries therefore depended on the testing and symptom data as well as the value of β_p.

Supplementary Figures

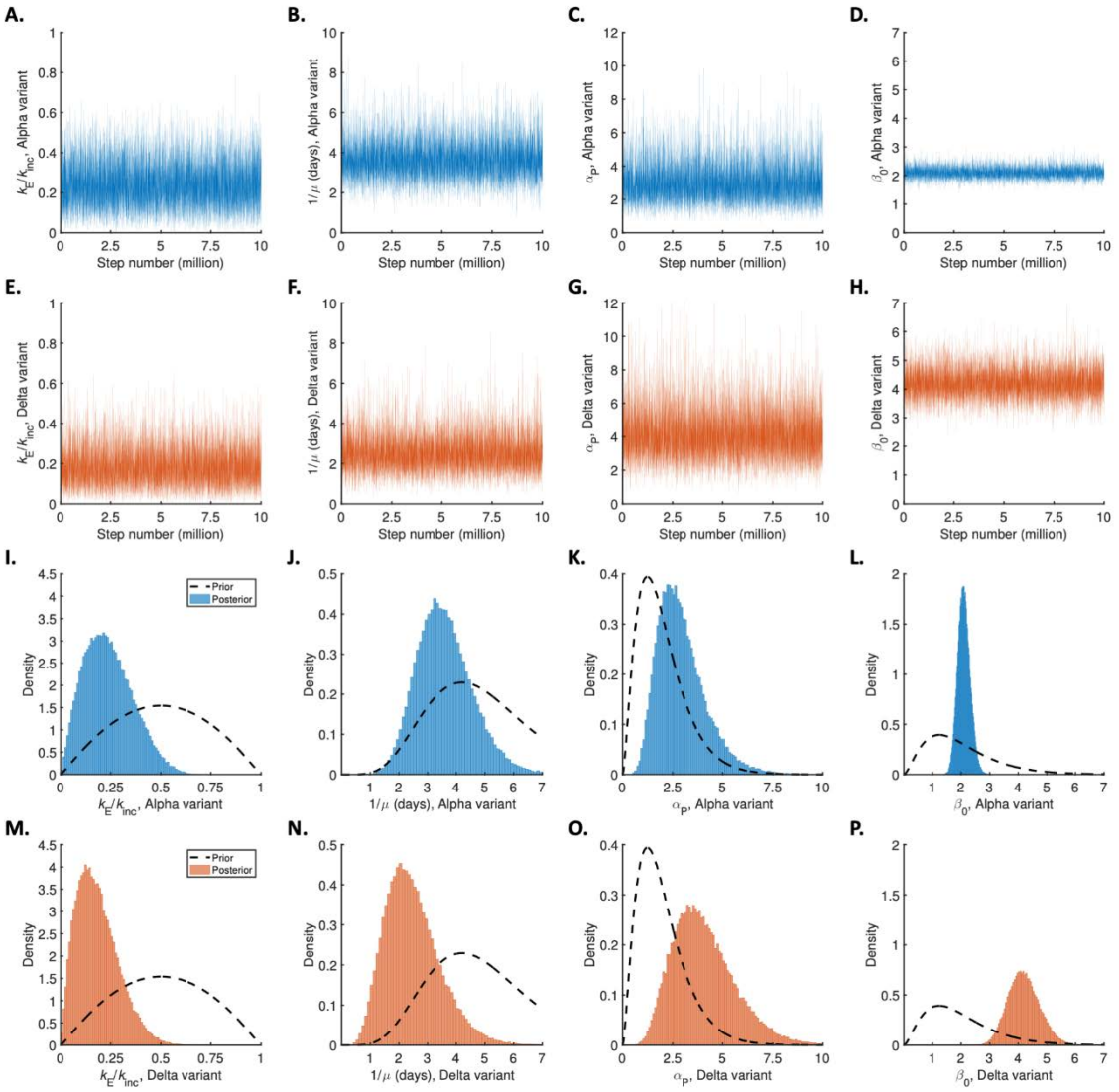

Figure S1. Trace plots and posterior distributions of fitted model parameters. Panels A-H show trace plots of estimated model parameters for the Alpha (A-D, blue lines) and Delta variants (E-H, red lines) obtained in the data augmentation MCMC procedure that we used to fit our mathematical transmission model to the UK household data: A,E. The ratio of the mean durations of the latent (*E*) and incubation (combined *E* and *P*) periods, k_E/k_{inc}; B,F. The mean duration of the symptomatic infectious (*I*) period, $1/\mu$; C,G. The ratio of the transmission rates in the presymptomatic infectious (*P*) and symptomatic infectious (*I*) stages of infection, α_P; D,H. The overall transmissibility parameter, β_0. In each panel, only the output of every 1,000 MCMC iterations is shown, including during the initial burn-in period of 2,000,000 total iterations (whereas one in every 100 iterations after the burn-in period were used to calculate posteriors). In panels I-P, histograms indicate posterior distributions of model parameters for the Alpha (I-L, blue bars) and Delta (M-P, red bars) variants, while black dotted curves indicate prior distributions. The exact priors that we used, as well as posterior medians and 95% credible intervals, are given in Table S4.

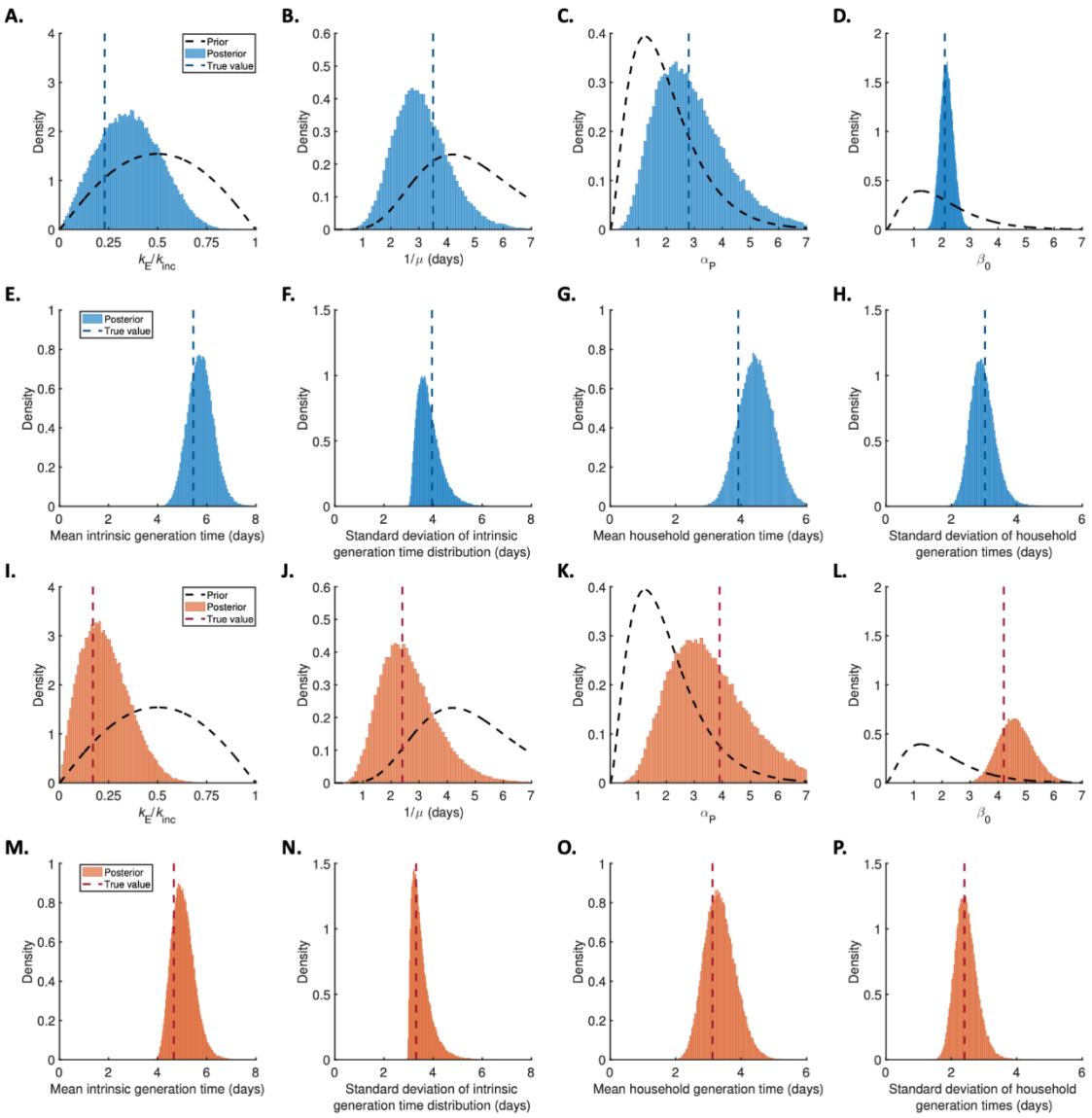

Figure S2. Testing the parameter fitting procedure using simulated data. In addition to our main analysis in which we applied our approach to real-world data, we conducted a simulation study to verify that the posteriors obtained during the parameter fitting procedure included the true values of model parameters. To do this, we generated a synthetic household dataset by simulating the household transmission model using the parameter values shown in Tables S1 and S2 (specifically, we used the median parameter estimates obtained in our main analysis). These are referred to as the "true" underlying parameter values. We used the same household structure and individual vaccination statuses as in the real-world study, and assumed that the households were infected by the same variants as assumed in our main analysis. We assumed that 7% of infections were entirely asymptomatic (which was consistent with the real household data). The synthetic data consisted of (i) whether or not each individual became infected, (ii) whether or not infected individuals developed symptoms, and (iii) symptom onset dates (for simplicity, we did not generate testing data). In panels A-H, histograms (blue bars) indicate posterior estimates for the Alpha variant of: A. The ratio of the mean durations of the latent (*E*) and incubation (combined *E* and *P*) periods, k_E/k_{inc}; B. The mean duration of the symptomatic infectious (*I*) period, $1/\mu$; C. The ratio of the transmission rates in the presymptomatic infectious (*P*) and symptomatic infectious (*I*) stages of infection, α_P; D. The overall transmissibility parameter, β_0; E. The mean intrinsic generation time; F. The standard deviation of the intrinsic generation time distribution; G. The mean household generation time; H. The standard deviation of household generation times. In each panel, vertical dotted lines indicate the "true" value of each quantity (in panels G and H, "true" values were calculated from the synthetic data), while black dotted curves in panels A-D indicate the prior distributions of the fitted model parameters. Panels I-P are equivalent to A-H, but for the Delta variant (using red instead of blue).

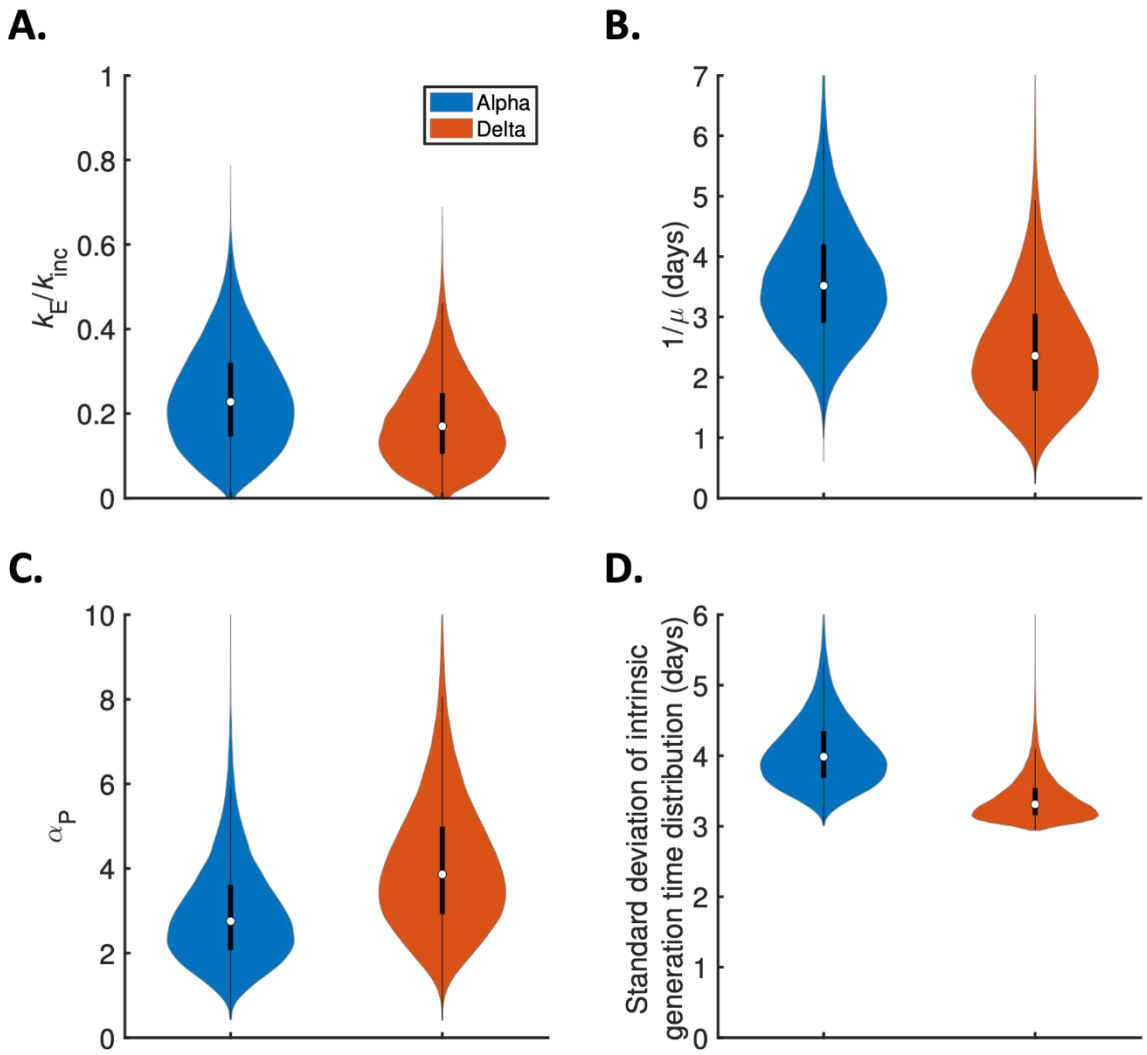

Figure S3. Comparison of posterior distributions of fitted model parameters, and the standard deviation of the intrinsic generation time distribution, between variants. Violin plots indicate posterior estimates for the Alpha (blue) and Delta (red) variants of: A. The ratio of the mean durations of the latent (*E*) and incubation (combined *E* and *P*) periods, k_E/k_{inc}; B. The mean duration of the symptomatic infectious (*I*) period, $1/\mu$; C. The ratio of the transmission rates in the presymptomatic infectious (*P*) and symptomatic infectious (*I*) stages of infection, α_P; D. The standard deviation of the intrinsic generation time distribution. Posterior medians and 95% credible intervals for these quantities are given in Table S4 and Table S5.

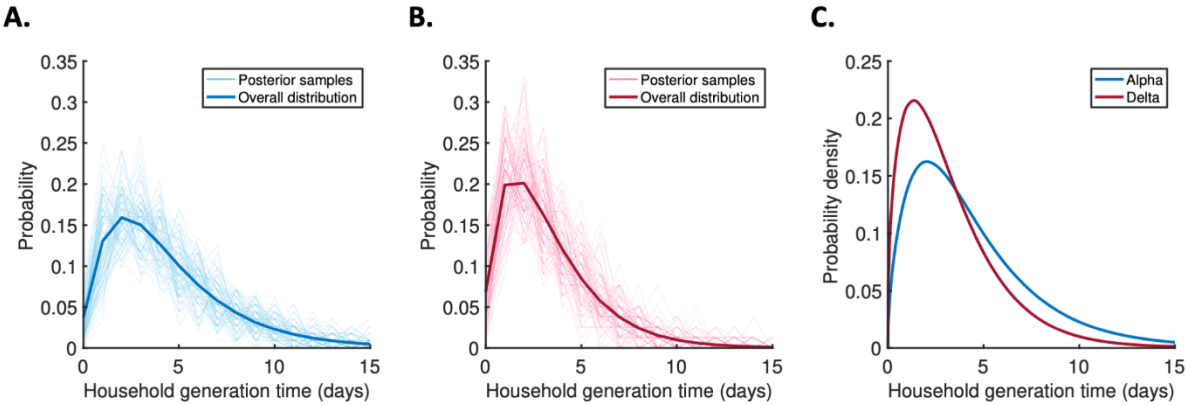

Figure S4. Distributions of household generation times. A-B. The discretised household generation time distribution (i.e., the number of days between the dates of individuals becoming infected and transmitting the virus) for the Alpha (panel A) and Delta (panel B) variants. In both panels, the thin lines show the predicted distribution from 100 randomly selected steps of the MCMC procedure that we used to fit the mathematical transmission model to the UK household data, and the thick, darker, line shows the overall distribution when combining the output of all MCMC steps (after burn-in and thinning). C. Comparison of the overall (continuous-time) household generation time distribution for the Alpha (blue) and Delta (red) variants.

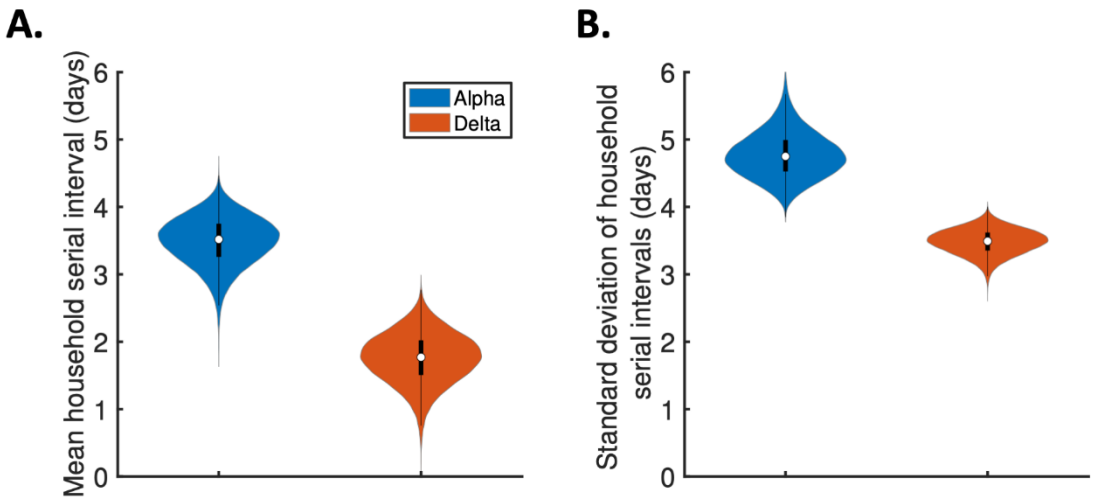

Figure S5. The effect of variant on the household serial interval. Violin plots indicate posterior estimates for the Alpha (blue) and Delta (red) variants of: A. The mean household serial interval (i.e., the mean interval between the symptom onset times infectors and infectees (who both developed symptoms) within study households); B. The standard deviation of household serial intervals.

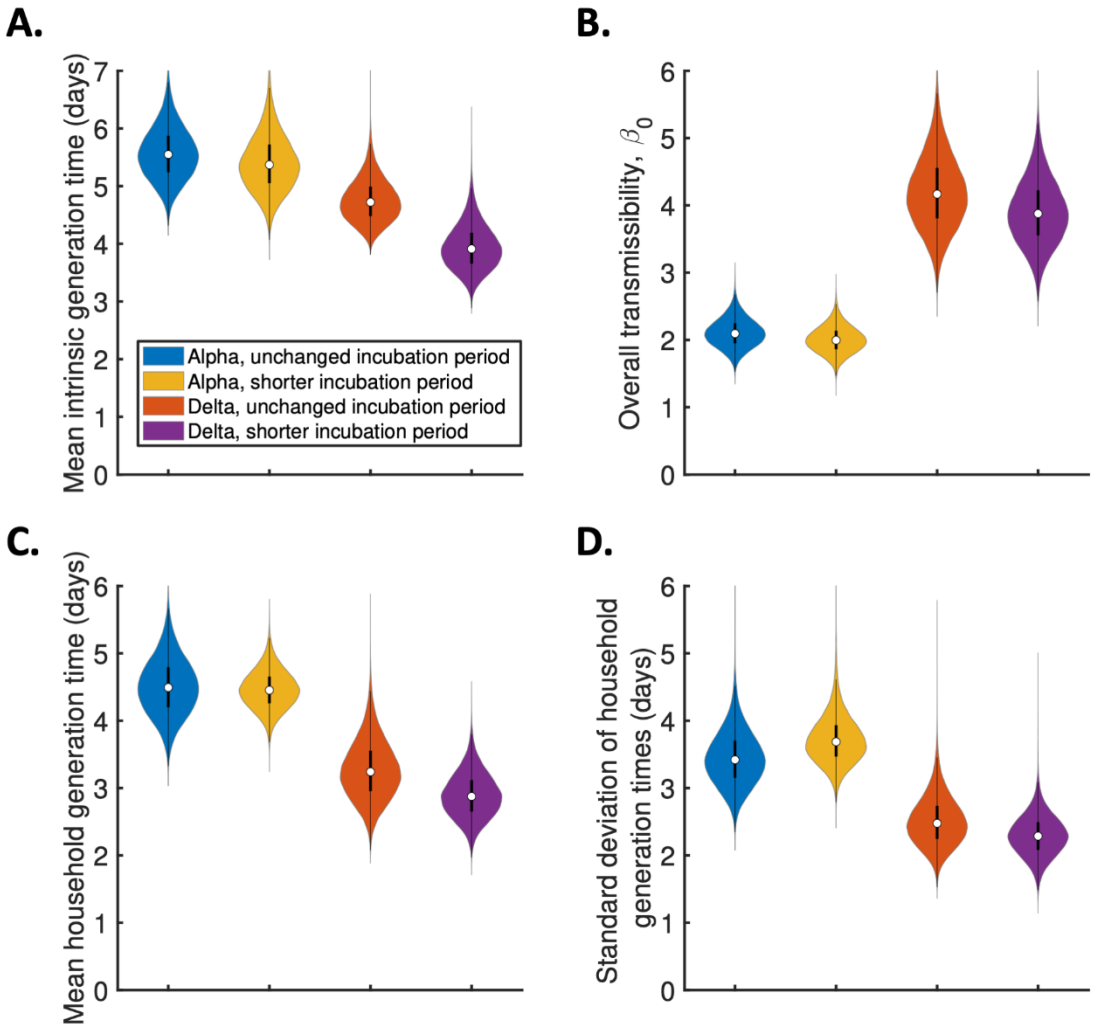

Figure S6. Sensitivity of the results to the assumed incubation period. In our main analysis, we assumed that the incubation period distribution was identical for the Alpha and Delta variants, and used an estimate for SARS-CoV-2 obtained in a meta-analysis carried out before the Alpha and Delta variants emerged (mean 5·8 days and standard deviation 3·1 days[9]). Here, we compare the results shown in Figure 1 with results assuming one or both variants have a shorter incubation period (mean 4·4 days and standard deviation 1·9 days, as estimated for the Delta variant in a study in China using data from 47 individuals[19]). Violin plots indicate posterior estimates for the Alpha variant with unchanged incubation period (blue), the Alpha variant with shorter incubation period (orange), the Delta variant with unchanged incubation period (red), and the Delta variant with shorter incubation period (purple), of: A. The mean intrinsic generation time; B. The overall transmissibility parameter, β_0; C. The mean household generation time; D. The standard deviation of household generation times.

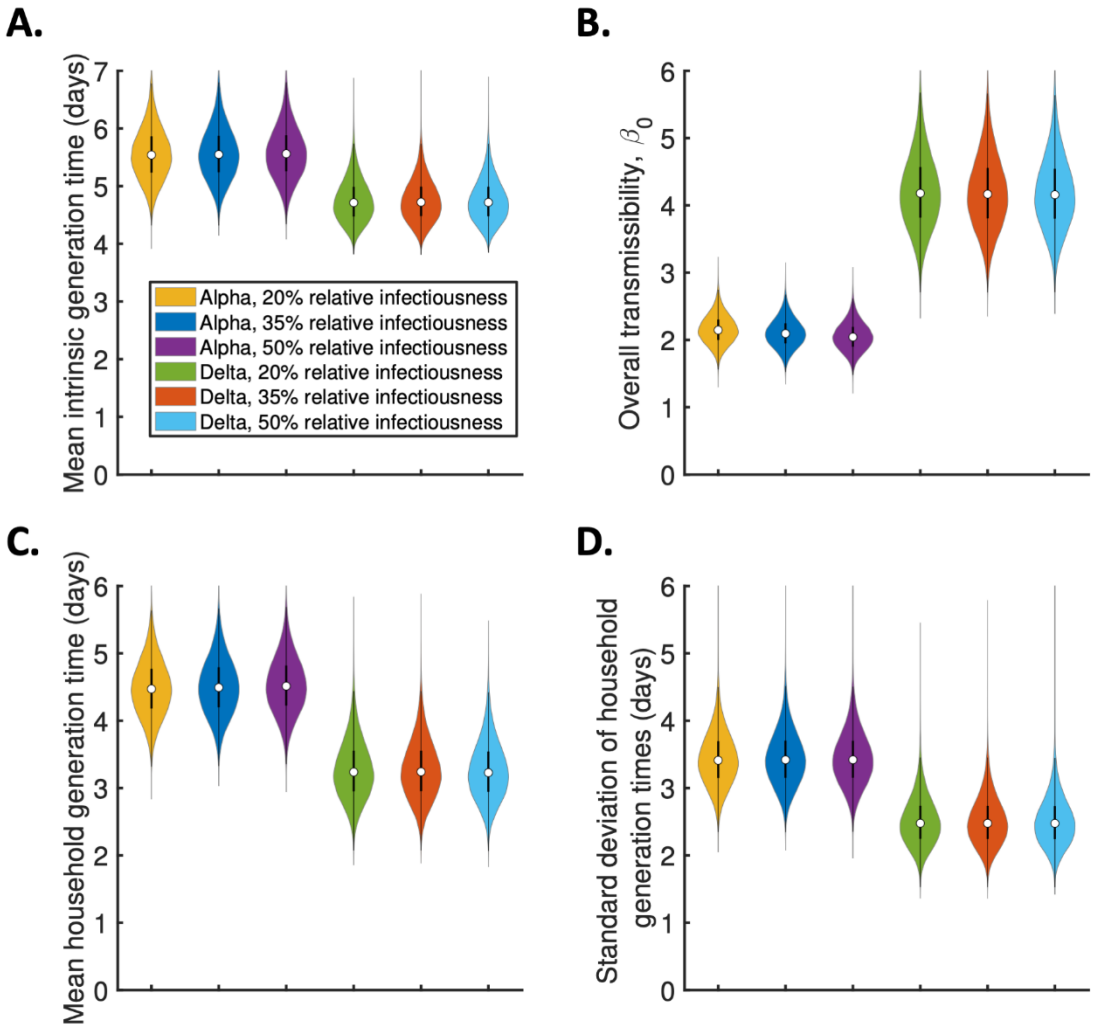

Figure S7. Sensitivity of the results to the assumed relative infectiousness of asymptomatic infected hosts. Violin plots indicate posterior estimates for the Alpha and Delta variants under different values of the relative infectiousness, α_A, of entirely asymptomatic infected hosts (compared to those who developed symptoms), of: A. The mean intrinsic generation time; B. The overall transmissibility parameter, β_0; C. The mean household generation time; D. The standard deviation of household generation times. In each panel, results are shown for values of α_A of 20% (orange and green violins for the Alpha and Delta variants, respectively), 35% (as assumed in the main text; dark blue and red) and 50% (purple and light blue).

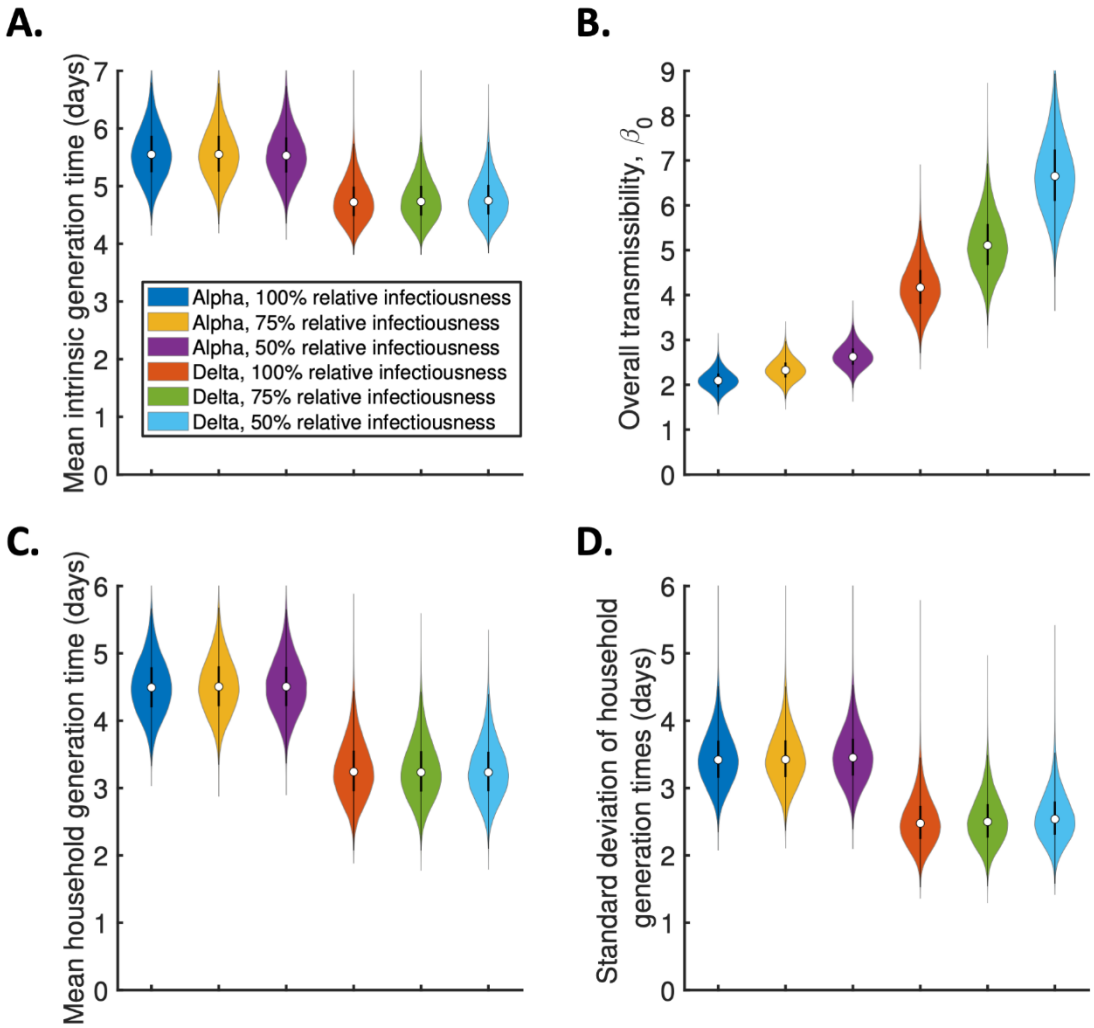

Figure S8. Sensitivity of the results to the assumed relative infectiousness of vaccinated infected hosts. Violin plots indicate posterior estimates for the Alpha and Delta variants under different values of the relative infectiousness of vaccinated infected hosts (compared to those who are unvaccinated) who developed symptoms (both unvaccinated and vaccinated asymptomatic infected hosts were assumed to be 35% as infectious as unvaccinated hosts who developed symptoms, as in our main analysis), of: A. The mean intrinsic generation time; B. The overall transmissibility parameter, β_0; C. The mean household generation time; D. The standard deviation of household generation times. In each panel, results are shown for relative infectiousness values of 100% (as assumed in the main text; dark blue and red violins for the Alpha and Delta variants, respectively), 75% (orange and green) and 50% (purple and light blue). For simplicity, no difference in relative infectiousness was assumed between infected individuals who had received one or two vaccine doses, or between those who received either the Oxford-AstraZeneca or the Pfizer-BioNTech vaccine.

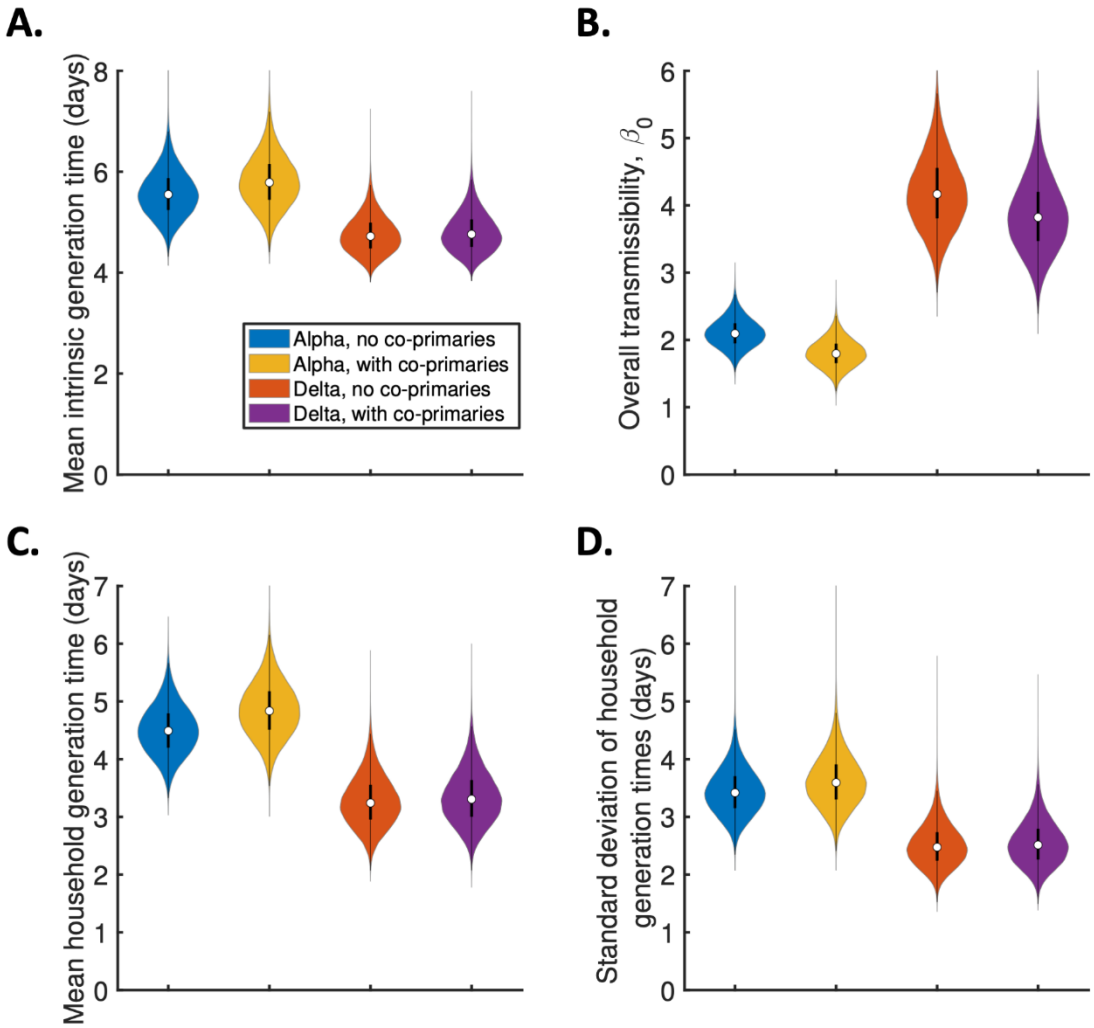

Figure S9. Sensitivity of the results to the assumption of no co-primary cases. In our main analysis, we assumed that each household transmission cluster originated with a single primary case. Here, we compare the results shown in Figure 1 with results accounting for the possibility of co-primary cases (see the "Extension of framework to account for co-primary cases" section of the Supplementary Methods). Violin plots indicate posterior estimates for the Alpha variant with unchanged incubation period (blue), the Alpha variant with shorter incubation period (orange), the Delta variant with unchanged incubation period (red), and the Delta variant with shorter incubation period (purple), of: A. The mean intrinsic generation time; B. The overall transmissibility parameter, β_0; C. The mean household generation time; D. The standard deviation of household generation times.

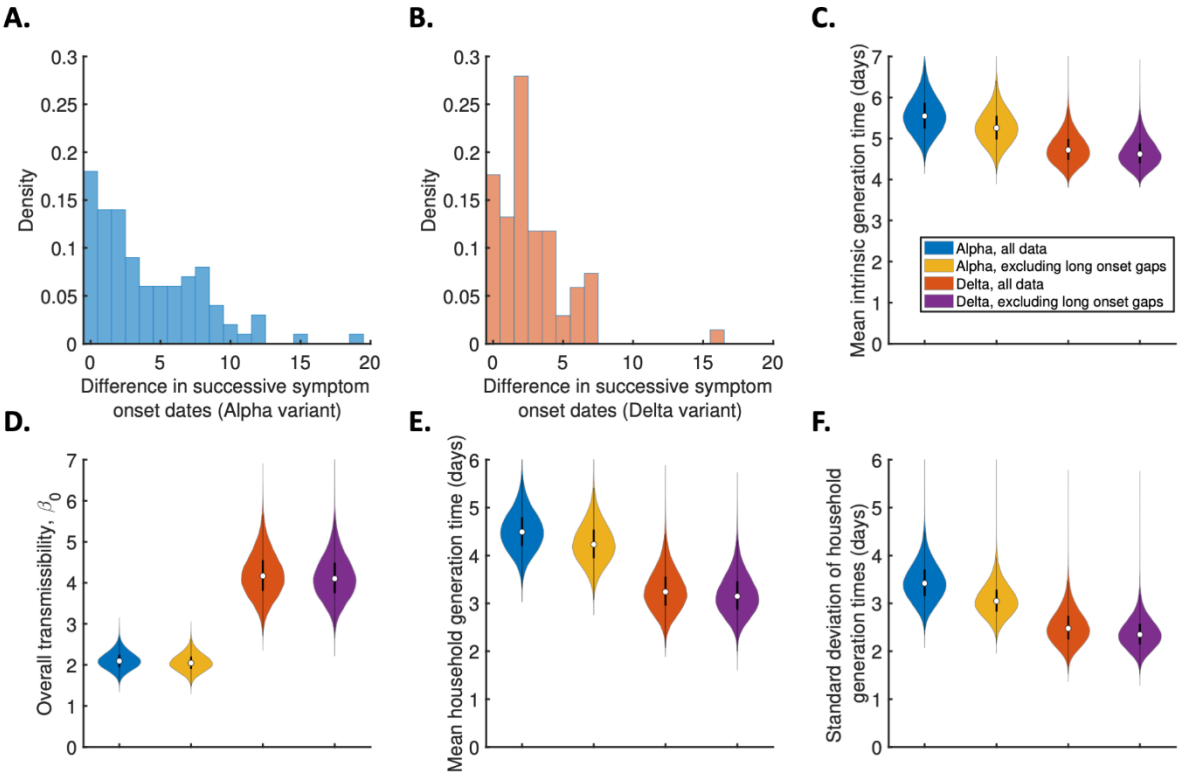

Figure S10. Sensitivity of the results to the exclusion of households with long gaps between successive dates on which household members developed symptoms. Considering the distribution of the number of days between successive symptom onset dates within households (note that this differs from the serial interval because an infected individual may not have been infected by the previous household member to develop symptoms) for the Alpha (panel A) and Delta (B) variants indicated that there were three households in which gaps between household members developing symptoms of over two weeks were observed (two households for the Alpha variant and one for the Delta variant). This could indicate that multiple introductions of the virus into those households occurred. In panels C-F, we compare the results shown in Figure 1 (in which the symptom onset gaps of over two weeks were assumed to result from within-household transmission) with results excluding the data from these three households. Violin plots indicate posterior estimates for the Alpha variant with all data included (blue), the Alpha variant with the three households excluded (orange), the Delta variant with all data included (red), and the Delta variant with the three households excluded (purple), of: C. The mean intrinsic generation time; D. The overall transmissibility parameter, β_0; E. The mean household generation time; F. The standard deviation of household generation times.

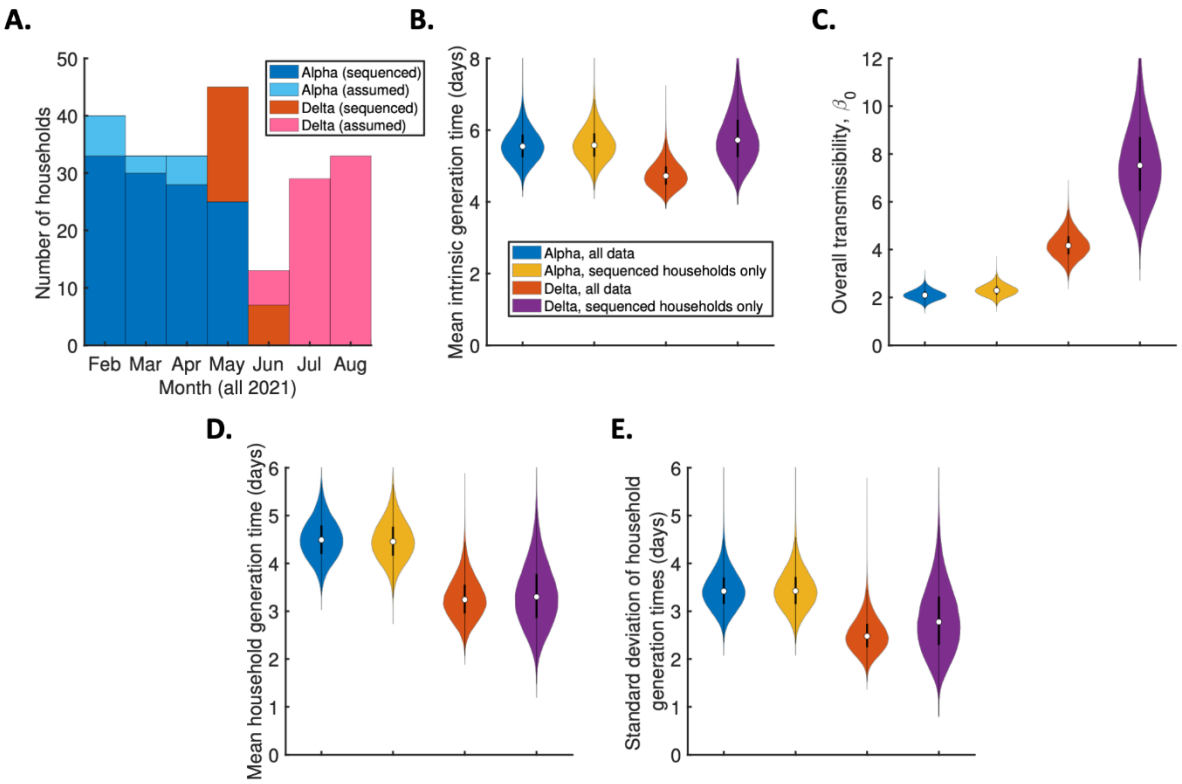

Figure S11. Sensitivity of the results to the identification of variant by index month. In panel A, stacked bars indicate the number of study households in which the index case first tested positive during each month, distinguishing between households in which genomic sequencing indicated infections due to either the Alpha (dark blue bars) or the Delta (red) variant, unsequenced households with an index month of April 2021 or earlier in which the Alpha variant was assumed to be responsible for infections (light blue), and unsequenced households with an index month of June 2021 or later in which the Delta variant was assumed to be responsible for infections (pink). Three unsequenced households recruited during May 2021 were excluded from our analysis and are not shown here. In panels B-F, we compare the results shown in Figure 1 with results using only data from sequenced households. Violin plots indicate posterior estimates for the Alpha variant with all data included (blue), the Alpha variant using only data from sequenced households (orange), the Delta variant with all data included (red), and the Delta variant using only data from sequenced households (purple), of: B. The mean intrinsic generation time; C. The overall transmissibility parameter, β_0; D. The mean household generation time; E. The standard deviation of household generation times.

Supplementary Tables

Variant	Number of vaccine doses received	Number of participants	Number of infections
Alpha	0	188	129
	1	112	90
	2	34	24
Delta	0	52	41
	1	27	22
	2	146	111

Table S1. Number of participants and infections by variant and vaccination status. For the Alpha and Delta variants, and distinguishing between individuals who were either unvaccinated or had received one or two doses of a COVID-19 vaccine, we show the total number of participants and the number of those participants who were infected. Individuals who received their first vaccine dose fewer than 21 days before the first member of their household developed symptoms or returned a positive test were considered to be unvaccinated (and are listed as such here). Additional protection from the second vaccine dose was assumed to take effect immediately. We made these assumptions because no estimates of vaccine protection within three weeks of the first dose are given in[8], whereas the estimates of protection within two weeks of the second dose in that study[8] are similar to the estimates for protection after at least two weeks that we used in our analysis.

Variant	Age group	Number of participants	Number of infections
Alpha	0-10	33	19
	11-18	23	7
	19-54	194	153
	55+	84	64
Delta	0-10	8	5
	11-18	15	9
	19-54	136	114
	55+	66	46

Table S2. Number of participants and infections by variant and age. For the Alpha and Delta variants, and distinguishing between individuals in different age groups, we show the total number of study participants and the number of those participants who were infected.

Parameter	Interpretation	Value	Justification
η_j	Relative susceptibility of individual j (compared to an unvaccinated individual)	Dependent on the vaccination status of individual j: 1 (unvaccinated or within 20 days of first vaccine dose) 0·37 (Alpha variant; one dose of Oxford-AstraZeneca vaccine) 0·21 (Alpha variant; two doses of Oxford-AstraZeneca vaccine) 0·41 (Alpha variant; one dose of Pfizer-BioNTech vaccine) 0·22 (Alpha variant; two doses of Pfizer-BioNTech vaccine) 0·54 (Delta variant; one dose of Oxford-AstraZeneca vaccine) 0·33 (Delta variant; two doses of Oxford-AstraZeneca vaccine) 0·43 (Delta variant; one dose of Pfizer-BioNTech vaccine) 0·20 (Delta variant; two doses of Pfizer-BioNTech vaccine)	Consistent with previous estimates of vaccine efficacy against infection[8]
$1/\gamma$	Mean incubation period	5·8 days	Consistent with previous estimates of incubation period distribution[9]
k_{inc}	Shape parameter of (gamma) incubation period distribution	3·5	Consistent with mean (5·8 days) and standard deviation (3·1 days) of previous estimates of incubation period distribution[9]
k_I	Shape parameter of (gamma) symptomatic infectious period distribution	1	Assumed
α_A	Relative infectiousness of entirely asymptomatic hosts	0·35	Taken from[5]

Table S3. Values of parameters that were not estimated from the household data. All parameters listed were assumed to take the same values for the Alpha and Delta variants except where explicitly stated otherwise. No vaccine protection against infection was assumed for individuals who received their first vaccine dose fewer than 21 days before the first member of their household developed symptoms or returned a positive test, while increased vaccine protection from the second dose was assumed to take effect immediately (this was because no estimates for vaccine protection within three weeks of the first dose are given in[8], whereas the estimates for protection within two weeks of the second dose from that study[8] were similar to the estimates for protection after at least two weeks that we used here). Two individuals in our analysis who received the Moderna vaccine were assumed to have the same relative susceptibility as those who received the Pfizer-BioNTech vaccine. Two individuals who received an unspecified vaccine were assumed to have the same relative susceptibility as those who received the Oxford-AstraZeneca vaccine. The sensitivity of our results to the assumed incubation period distribution (i.e., the values of $1/\gamma$ and k_{inc}) is considered in Figure S6, and sensitivity to the assumed value of α_A is considered in Figure S7.

Parameter	Interpretation	Prior	Posterior median (95% CrI)
k_E/k_{inc}	Ratio of mean durations of the latent (*E*) and incubation (*E*+*P*) periods	Beta(2·1,2·1) [prior mean 0·5, 95% CrI 0·1-0·9]	Alpha: 0·23 (0·04-0·50) Delta: 0·17 (0·03-0·42)
$1/\mu$	Mean symptomatic infectious (*I*) period	Gamma(7,0·7) [prior mean 4·9 days, 95% CrI 2·0-9·1 days]	Alpha: 3·5 days (1·9-5·8 days) Delta: 2·4 days (1·0-4·7 days)
α_P	Ratio of transmission rates in the presymptomatic infectious (*P*) and symptomatic infectious (*I*) stages	Gamma(2·65,0·75) [prior mean 2·0, 95% CrI 0·4-5·0]	Alpha: 2·8 (1·2-5·9) Delta: 3·9 (1·6-7·7)
β_0	Overall transmissibility parameter	Gamma(2·65,0·75) [prior mean 2·0, 95% CrI 0·4-5·0]	Alpha: 2·1 (1·7-2·6) Delta: 4·2 (3·2-5·4)

Table S4. Values of fitted model parameters. Descriptions of fitted model parameters, the prior distributions used (all prior distributions were taken to be independent, both between different parameters and between variants), and the posterior medians and 95% credible intervals obtained for the Alpha and Delta variants.

Quantity	Posterior median (95% CrI) for Alpha variant	Posterior median (95% CrI) for Delta variant
Mean intrinsic generation time	5·5 days (4·7-6·5 days)	4·7 days (4·1-5·6 days)
Standard deviation of intrinsic generation time distribution	4·0 days (3·3-5·3 days)	3·3 days (3·0-4·3 days)
Mean household generation time	4·5 days (3·7-5·4 days)	3·2 days (2·5-4·2 days)
Standard deviation of household generation times	3·4 days (2·7-4·3 days)	2·5 days (1·9-3·3 days)

Table S5. Mean and standard deviation of the intrinsic and household generation time distributions. Posterior medians and 95% credible intervals for the mean and standard deviation of the intrinsic and household generation time distributions for the Alpha and Delta variants.

References

1 UK Government. COVID-19 testing data: methodology note. 2020. https://www.gov.uk/government/publications/coronavirus-covid-19-testing-data-methodology/covid-19-testing-data-methodology-note (accessed Dec 9, 2021).

2 Quick J. nCoV-2019 sequencing protocol V.1. protocols.io. 2020. https://dx.doi.org/10.17504/protocols.io.bbmuik6w (accessed Dec 9, 2021).

3 COVID-19 Genomics UK Consortium. https://www.cogconsortium.uk/ (accessed Nov 29, 2021).

4 Jeffery-Smith A, Dun-Campbell K, Janarthanan R, *et al.* Infection and transmission of SARS-CoV-2 in London care homes reporting no cases or outbreaks of COVID-19: Prospective observational cohort study, England 2020. *Lancet Reg Health – Eur* 2021; **3**: 100038.

5 Buitrago-Garcia D, Egli-Gany D, Counotte MJ, *et al.* Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLOS Med* 2020; **17**: e1003346.

6 Li F, Li Y-Y, Liu M-J, *et al.* Household transmission of SARS-CoV-2 and risk factors for susceptibility and infectivity in Wuhan: a retrospective observational study. *Lancet Infect Dis* 2021; **21**: 617–28.

7 Madewell ZJ, Yang Y, Longini IM Jr, Halloran ME, Dean NE. Factors associated with household transmission of SARS-CoV-2: an updated systematic review and meta-analysis. *JAMA Netw Open* 2021; **4**: e2122240.

8 Pouwels KB, Pritchard E, Matthews PC, *et al.* Impact of Delta on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. *medRxiv* 2021; : 2021.08.18.21262237.

9 McAloon C, Collins Á, Hunt K, *et al.* Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ Open* 2020; **10**: e039652.

10 Hart WS, Maini PK, Thompson RN. High infectiousness immediately before COVID-19 symptom onset highlights the importance of continued contact tracing. *eLife* 2021; **10**: e65534.

11 Ferretti L, Wymant C, Kendall M, *et al.* Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* 2020; **368**: eabb6936.

12 Lehtinen S, Ashcroft P, Bonhoeffer S. On the relationship between serial interval, infectiousness profile and generation time. *J R Soc Interface* 2021; **18**: 20200756.

13 Britton T, Scalia Tomba G. Estimation in emerging epidemics: biases and remedies. *J R Soc Interface* 2019; **16**: 20180670.

14 Cauchemez S, Carrat F, Viboud C, Valleron AJ, Boëlle PY. A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Stat Med* 2004; **23**: 3469–87.

15 Ferguson NM, Cummings DAT, Cauchemez S, *et al.* Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 2005; **437**: 209–14.

16 Hart WS, Abbott S, Endo A, *et al.* Inference of SARS-CoV-2 generation times using UK household data. *medRxiv* 2021; : 2021.05.27.21257936.

17 Fraser C. Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. *PLOS ONE* 2007; **2**: e758.

18 Cauchemez S, Donnelly CA, Reed C, *et al.* Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. *N Engl J Med* 2009; **361**: 2619–27.

19 Zhang M, Xiao J, Deng A, *et al.* Transmission dynamics of an outbreak of the COVID-19 Delta variant B.1.617.2 — Guangdong Province, China, May–June 2021. *China CDC Wkly* 2021; **3**: 584–6.