



OPEN ACCESS

EDITED BY

Andrew Yates,
Columbia University, United States

REVIEWED BY

Andreas Suhrbier,
The University of Queensland, Australia
Melanie Prague,
Inria Bordeaux - Sud-Ouest Research Centre,
France

*CORRESPONDENCE

Joshua T. Schiffer

✉ jschiffe@fredhutch.org

RECEIVED 30 April 2024

ACCEPTED 22 October 2024

PUBLISHED 07 November 2024

CITATION

Byrne C and Schiffer JT (2024) Ensemble modeling of SARS-CoV-2 immune dynamics in immunologically naïve rhesus macaques predicts that potent, early innate immune responses drive viral elimination. *Front. Immunol.* 15:1426016. doi: 10.3389/fimmu.2024.1426016

COPYRIGHT

© 2024 Byrne and Schiffer. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Ensemble modeling of SARS-CoV-2 immune dynamics in immunologically naïve rhesus macaques predicts that potent, early innate immune responses drive viral elimination

Catherine Byrne and Joshua T. Schiffer*

Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, United States

Introduction: An unprecedented breadth of longitudinal viral and multi-scale immunological data has been gathered during SARS-CoV-2 infection. However, due to the high complexity, non-linearity, multi-dimensionality, mixed anatomic sampling, and possible autocorrelation of available immune data, it is challenging to identify the components of the innate and adaptive immune response that drive viral elimination. Novel mathematical models and analytical approaches are required to synthesize contemporaneously gathered cytokine, transcriptomic, flow cytometry, antibody response, and viral load data into a coherent story of viral control, and ultimately to discriminate drivers of mild versus severe infection.

Methods: We investigated a dataset describing innate, SARS-CoV-2 specific T cell, and antibody responses in the lung during early and late stages of infection in immunologically naïve rhesus macaques. We used multi-model inference and ensemble modeling approaches from ecology and weather forecasting to compare and combine various competing models.

Results and discussion: Model outputs suggest that the innate immune response plays a crucial role in controlling early infection, while SARS-CoV-2 specific CD4+ T cells correspond to later viral elimination, and anti-spike IgG antibodies do not impact viral dynamics. Among the numerous genes potentially contributing to the innate response, we identified IFI27 as most closely linked to viral load decline. A 90% knockdown of the innate response from our validated model resulted in a ~10-fold increase in peak viral load during infection. Our approach provides a novel methodological framework for future analyses of similar complex, non-linear multi-component immunologic data sets.

KEYWORDS

SARS-CoV-2, mathematical modeling, ensemble model, systems immunology, innate immunity, rhesus macaques, within-host infection dynamics

1 Introduction

The COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, spurred an extraordinary global effort to comprehensively understand the pathophysiology of this infection. Longitudinal viral and multi-scale immunological datasets were amassed at an unparalleled scale (1–15), offering a unique opportunity to identify the intricacies of immune defense mechanisms during SARS-CoV-2 infection.

Animal models have proven especially useful for studying SARS-CoV-2 viral and immune kinetics (2, 3, 16–23). In contrast to the limitations faced in human studies, where sampling is often confined to saliva or nasal specimens collected after symptom onset, animal model studies enable sampling from the lung and other various tissue sites throughout the course of infection, including critical early pre-symptomatic time points. In addition, crucial variables such as time of infection, size of viral inoculum, viral variant, vaccination history, prior infection, and frequency of sampling are all experimentally controlled (24). Further, unlike in human studies where viral load is often the sole measure that is sampled longitudinally, animal models permit comprehensive assessments of changes in innate and adaptive immune responses over time (2, 3, 10, 22, 25, 26). SARS-CoV-2 infection in rhesus macaques presents relatively similarly to non-severe infection in humans, with comparable symptoms and duration of infection (10). As such, rhesus macaque data has proven particularly useful to better understand viral infection dynamics following vaccination, treatment, or reinfection with SARS-CoV-2 (2, 3, 10, 22, 23, 27, 28).

Despite the substantial volume of available data characterizing the SARS-CoV-2 immune response in rhesus macaques, there are no standardized methods to discern the relative importance and timing of mechanisms driving viral clearance. Immune responses demonstrate substantial redundancy (29). Moreover, separating and characterizing the dozens of gene products, antibodies, and immune cell populations which may be essential to infection clearance, presents computational challenges. Mathematical models of within-host infection dynamics fit to data offer a methodical way to test competing hypotheses of how a spreading infection and intensifying immune response may interact (30–33). Although numerous mathematical models of the within-host dynamics of SARS-CoV-2 infection have been developed to recapitulate viral load (11, 13, 33–38), only a few have concurrently integrated viral and immune data, and these have typically focused on a single arm of the immune system (35, 39–41). To our knowledge, no model has been fit concurrently to detailed longitudinal innate, cell-mediated, and humoral response data, with these responses variously coupled to SARS-CoV-2 elimination (42–44), and no model has focused specifically on lung immune responses.

When testing different mathematical models, multiple models may adequately fit observed, complex non-linear data such that several competing hypotheses to explain the data remain viable (45–47). This issue is compounded by the fact that even in carefully planned experiments, the sampling frequency may be too low during critical intervals to discriminate models with slightly differing assumptions. Consequently, testing, comparing, and amalgamating the outcomes of multiple potential models provides a more comprehensive and thorough result, allowing the weighting

of multiple hypotheses and projecting necessary uncertainty into subsequent model predictions (48–50).

In the fields of climatology, ecology, and epidemiological modeling (51–57), there are methods for testing many different models and synthesizing predictions into ensemble models. This approach has yet to be adopted in within-host models of infectious disease but will be necessary to account for the rapid emergence of multi-component immune data. Whereas traditional viral dynamic models were fit to viral load alone, future models will be tasked to recapitulate viral load, gene signature data (11), T cell subset, B cell, innate immune cell, cytokine (58), and antibody levels (35) over time. Additional challenges will be weighing the strength of each data type for fitting and assessing for different degrees of misclassification across multiple assays. Akin to weather forecasts which attempt to predict precipitation, wind, and temperature, such detailed models can be utilized to predict viral and immune kinetics in individuals with different immune starting conditions due to prior vaccination or infection.

Here, using detailed rhesus macaque infection data, we develop and test 160 different mathematical models describing the virologic and immunologic dynamics of SARS-CoV-2 infection in the lung. Using the concepts of multi-model inference (50), we predict the importance, timing, and contribution of each tested immune control mechanism in clearing infected cells and free virus. Further, we develop ensemble model predictions that combine top model results and examine how these predictions diverge from those made by individual models. Together, our results provide further insight into the immune mechanisms of SARS-CoV-2 infection clearance, provide new methods for the analysis of similar data sets, but also highlight limitations of current methods to predict infection outcomes.

2 Results

2.1 Analysis of bronchoalveolar lavage fluid reveals dynamic immune and viral kinetics within the lung during SARS-CoV-2 infection

To obtain a representation of infection kinetics within the lung, Nelson et al. performed bronchoalveolar lavage (BAL) on SARS-CoV-2-infected rhesus macaques and analyzed BAL fluid (BALF) for viral genomic RNA concentrations with qPCR, anti-spike IgG titers with ELISA, SARS-CoV-2-specific CD4+ and CD8+ concentrations with cell staining and flow cytometry, and the expression of interferon (IFN) and IFN-stimulating genes (ISGs) with scRNAseq (2). We selected this dataset for modeling based on the longitudinal, multi-model measurements of immune response from the most relevant site of infection, the lung. Such data is not available from human studies. The data was notable for a decline in SARS-CoV-2 RNA starting on day 1 (Figure 1A), relatively stable anti-spike IgG levels with a slight increase in some animals (Figure 1B), late increases in virus-specific CD4+ and CD8+ T cells (Figure 1C), and early increases and abrupt decreases in multiple innate gene signatures (Figure 1D).

While wanting to describe as much of these data as possible within our mathematical model, it was unfeasible to mechanistically

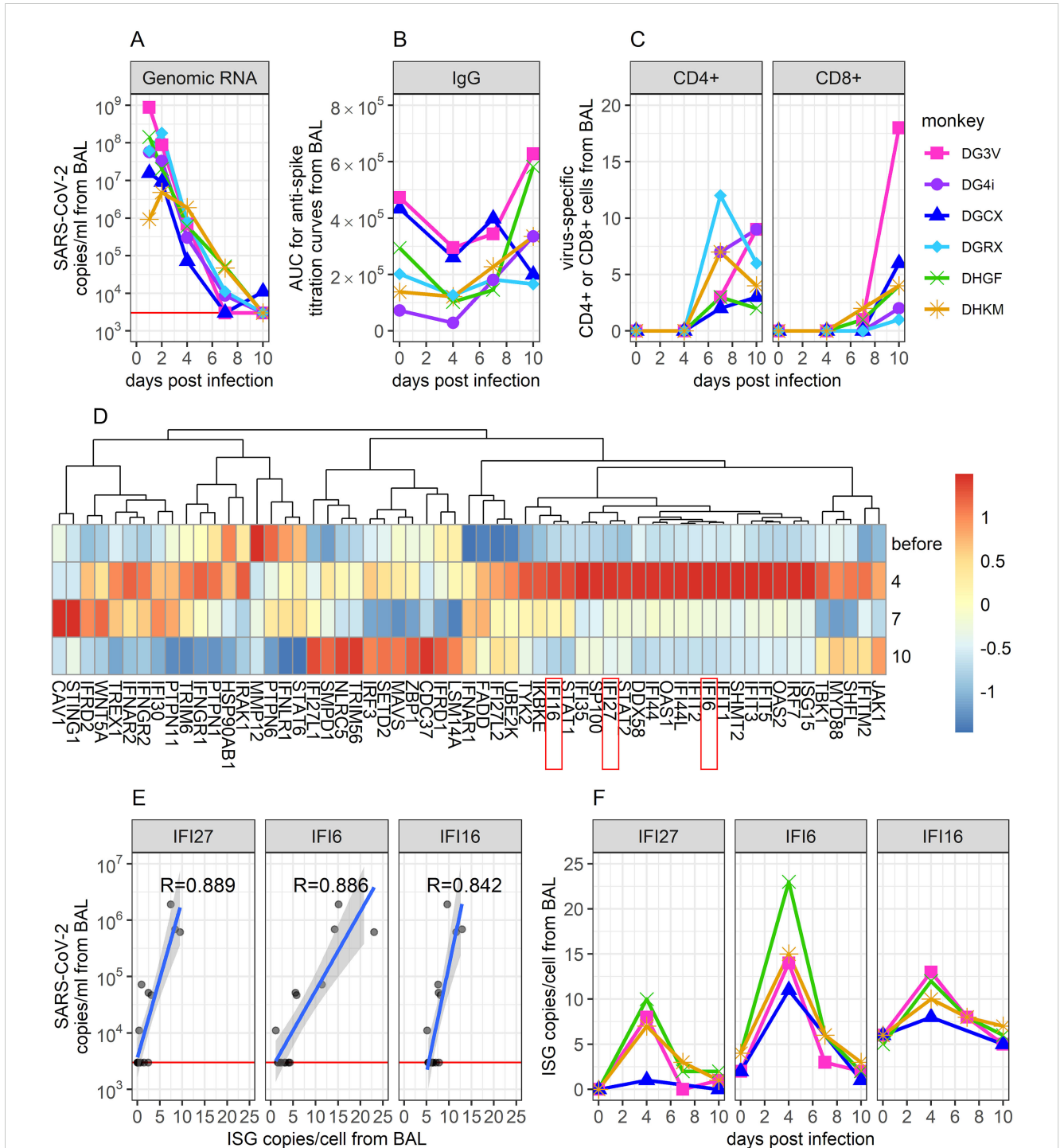


FIGURE 1

Description of viral and immunological data collected from BAL during SARS-CoV-2 infection in immunologically naive Rhesus Macaques. **(A)** SARS-CoV-2 genomic (g)RNA copies/ml of BALF were measured using qPCR. The red horizontal line shows the threshold of detection (3000 copies/ml). Data points on days 1 and 2 are extrapolated from corresponding viral loads measured in the throat and the nose on these days (see Methods). **(B)** Area under the curve (AUC) from ELISA titration curves of SARS-CoV-2 anti-spike proteins in collected BALF. **(C)** Percentage of virus-specific CD4+ and CD8+ T cells in BALF, measured via flow cytometry. CD4+ and CD8+ cells were considered virus-specific if they stain positively for IFN γ or TNF following exposure to a megapool of SARS-CoV-2 antigens. **(D)** Normalized average expression values of IFN genes and ISGs, measured via scRNAseq. **(E)** Correlation between ISG expression and same-day gRNA levels in BALF. The three genes with the most significant correlation are shown (correlation coefficient indicated). These genes also showed the greatest change over time, according to feature selection. **(F)** Time-series data of these top three genes. Data comes from (2).

model the dynamics of all genes examined through scRNAseq without making the model unworkably complex. To identify genes of interest, we performed feature selection (59) on IFN and ISG time series data and examined how gene expression correlated with same-day viral loads. Results indicated that the IFI27, IFI6, and IFI16 genes showed the strongest correlations with viral load (Figure 1E) and the greatest variability over time (Figure 1F). In addition, these genes were selected based on mechanistic importance and high dynamics during human SARS-CoV-2 infection (60–67). Thus, in addition to viral genomic RNA concentrations, anti-spike IgG titers, and virus-specific CD4+ and CD8+ concentrations, we included the dynamics of these three genes in our mathematical model of infection. As ISGs are known to mediate the innate immune response (8, 68), these three genes were assumed to capture the dynamics of the innate immune response and describe its impact on viral dynamics within our model.

2.2 Competing mathematical models of immune containment SARS-CoV-2 infection

Many models of within-host dynamics during SARS-CoV-2 infections exist (11, 13, 33–37), but few incorporate detailed data on how different arms of the immune response might impact infection

(35, 39–41). To test the role of both the innate and adaptive immune response on SARS-CoV-2 elimination, we developed the following model (Figure 2) describing how the kinetics of infection within the lung change over time, t .

$$\frac{dS}{dt} = d_S S_0 - \eta \frac{S}{S_0} e^{-\gamma_A A} V - d_S S$$

$$\frac{dI}{dt} = \eta \frac{S}{S_0} e^{-\gamma_A A} V - \sum_i b_{Fi} I F_i - nIT - mIE - \delta I$$

$$\frac{dV}{dt} = p e^{-\sum_i \gamma_{Fi} F_i} I - cV$$

$$\frac{dF_i}{dt} = d_{Fi} F_{i0} + w_{Fi} I - d_{Fi} F_i$$

$$\frac{dT}{dt} = d_T T_0 + w_T I - d_T T$$

$$\frac{dE}{dt} = d_E E_0 + w_E I - d_E E$$

$$\frac{dA}{dt} = d_A A_0 + w_A I - d_A A$$

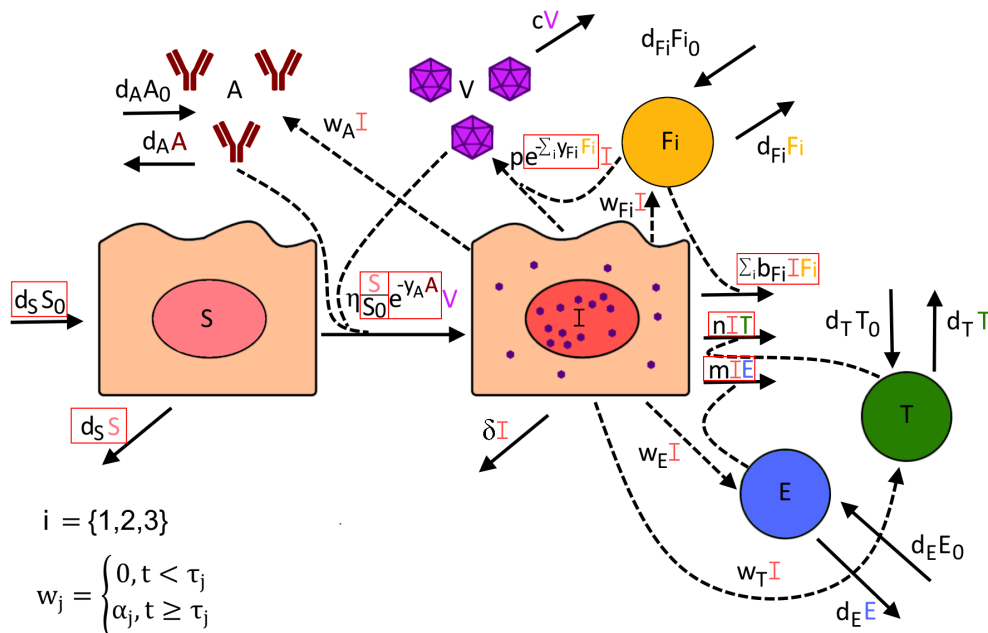


FIGURE 2

Visual description of all potential terms included within our mathematical model of SARS-CoV-2 infection. Red boxes indicate which model terms were alternatively included or excluded to determine how well each version of this model fits the data. The rate at which susceptible cells (S) become infected (I) is dependent on the number of susceptible cells, the amount of virus (V) present, and the presence of anti-spike IgG (A), which may dampen infection rates through neutralization of virus. Infected cells can potentially be cleared by interacting with virus-specific CD8+ T cells (T), virus-specific CD4+ T cells (E), or the innate immune response (Fi). The rate of viral production is dependent on the number of infected cells but can be dampened by the innate immune response inducing an antiviral state in infected cells. The rate of proliferation for anti-spike IgG antibody, virus-specific CD8+ T cells, virus-specific CD4+ T cells, and innate immune cells is proportional to the number of infected cells but is not turned on until time τ_j , where j is specific to the type of immune response. For the innate immune response (Fi), we test its dynamics being represented by three possible ISGs: IFN27 ($i=1$), IFI6 ($i=2$), and IFI16 ($i=3$).

where

$$i = \{1, 2, 3\}$$

and

$$w_j = \begin{cases} 0, & t < \tau_j \\ \alpha_j, & t \geq \tau_j \end{cases}$$

In this model, susceptible cells (S) are produced at a constant rate $d_S S_0$, where S_0 is the number of susceptible cells in the lung before infection ($t = 0$), and naturally die at a rate $d_S S$. These cells become infected by SARS-CoV-2 virions (V) at rate $\eta \frac{S}{S_0} e^{-\gamma_A A} V$ and move into the infected compartment (I). This rate incorporates the potential role of anti-spike IgG (A) in dampening infection rates through virus neutralization, with the scaling factor $e^{-\gamma_A A}$ taking on a value between 0 and 1. Thus, higher IgG antibody levels may lower infection rates according to their concentration.

Virus-specific CD4+ T cells (T), virus-specific CD8+ T cells (E), the innate immune response captured by IFI27, IFI6, and IFI16 expression ($F1$, $F2$, $F3$, respectively), and anti-spike IgG each are assumed to have constant rates of production and per-capita natural death rates that maintain an equilibrium in the absence of infection. Infection also stimulates the production of each compartment, with the rate of proliferation proportional to the number of infected cells. We, however, assume that infection-induced proliferation for each compartment remains zero until time τ_j at which point onwards it is equal to $\alpha_j I$, where j is specific to the immune compartment.

Infected cells may be cleared by the innate immune response, virus-specific CD4+ T cells, or virus-specific CD8+ T cells, at rates $\sum_i b_{Fi} F_i$, nIT , and mIE , respectively. Infected cells may also naturally die at a per-capita rate δ . Virus is produced by infected cells at a per-capita rate p which may be dampened by the innate immune response by a factor $e^{-\sum_i \gamma_{Fi} F_i}$. Virus is naturally cleared at a per-capita rate c .

To test different versions of the model, we alternated setting parameters b_{F1} , b_{F2} , b_{F3} , n , m , γ_{F1} , γ_{F2} , γ_{F3} and γ_A to zero as well as including or excluding potential target cell limitation, where when excluded susceptible cells were not modeled and the infectivity term $\eta \frac{S}{S_0} e^{-\gamma_A A} V$ was set to $\eta e^{-\gamma_A A} V$. As the impact of the separate ISGs is likely hard to distinguish, we only allowed one ISG at a time to have an impact on clearing infection and/or dampening viral production. With these combinations, a total of 160 versions of this model were developed and fit to available data (Figure 1), with appropriate transformations to make each set of data representative of counts within BALF (see Supplementary Material for further details).

2.3 Ensemble modeling to synthesize multiple well-fitting mathematical model versions for multi-model inference

Upon fitting each version of our mathematical model (described in Methods), we calculated Akaike Information Criterion (AIC) scores to determine how well each model performed. Here, the AIC

score for model j is defined as

$$AIC_j = 2k_j - 2\ln(L_j)$$

where k_j is the number of model-estimated parameters in model j and L_j is model j 's maximized likelihood value.

Based on AIC scores, each model was assigned a rank, with lower AIC scores corresponding to a better model. In examining the dynamics of all top-ranked models, we noted that they all predicted the size of the anti-spike IgG population to remain unchanged over the 10-day study (Figure 3C). This comes from the data being relatively flat with only a slight increase noted after viral elimination, as has been observed in infection of immunologically naïve humans (35). As such, the $e^{-\gamma_A A}$ component within the model's infection term ($\eta \frac{S}{S_0} e^{-\gamma_A A} V$) remained constant over time and model results including this term were indistinguishable from models where this constant was incorporated into the value of η . As such, we chose to exclude all models that did not set γ_A to zero from our analysis, reducing the total number of models to 80. The ranking of all 80 models appears in the Supplementary Material (Supplementary Figure S2).

We compared models through several quantifiers. We calculated AIC score differences (ΔAIC_j), capturing the difference between the AIC score of model j and the minimum AIC score across all models, the evidence ratios (ER_j), capturing how much more likely the model with the lowest AIC score is to model j , and AIC weights (w_j), capturing the probability that model j best captured the data. Definitions of these terms are provided in the Methods.

Values for the top-ranked models are displayed in Table 1. The AIC weights of the top 12 ranked models (Figure 3A and Table 1) summed to 0.95, thus giving us the 95% confidence set that we used for further analysis. Among these 12 models, likelihood and AIC scores were relatively similar, with no single model being overwhelmingly the best. However, the models' Akaike weights indicate large variability in how probable each was at best capturing the data. For example, while the top-ranked model has a probability of 0.30 for best-capturing data, the probability is 0.012 for the 10th-ranked model, making the first model 25 times more likely.

To capture the combined predictions of our different models, we developed an ensemble model. While ensemble modelling can encompass a host of different ways of combining model predictions (51–57), here we define ours as the Akaike-weighted median behaviors of the 95% confidence set over time. As such, this model tracks the median and confidence intervals of all viral and immune variables exhibiting dynamics within the observation window. By choosing the 95% confidence set of models, we are 95% confident that one of the models we have included is the best at approximating the data (50). The ensemble model's predictions of each model compartment's dynamics over time are displayed in Figures 3B–E, along with projections of each of the top 12 models. Our ensemble model predicted there to be a period of very slight viral rebound occurring around days 4 and 5 post-infection. Viral loads were predicted to rebound from a local minimum of 3.3×10^5 copies/ml of BALF at 3.6 days post-infection to a local maximum of 1.3×10^6 copies/ml of BALF at 5.2 days post-infection, before

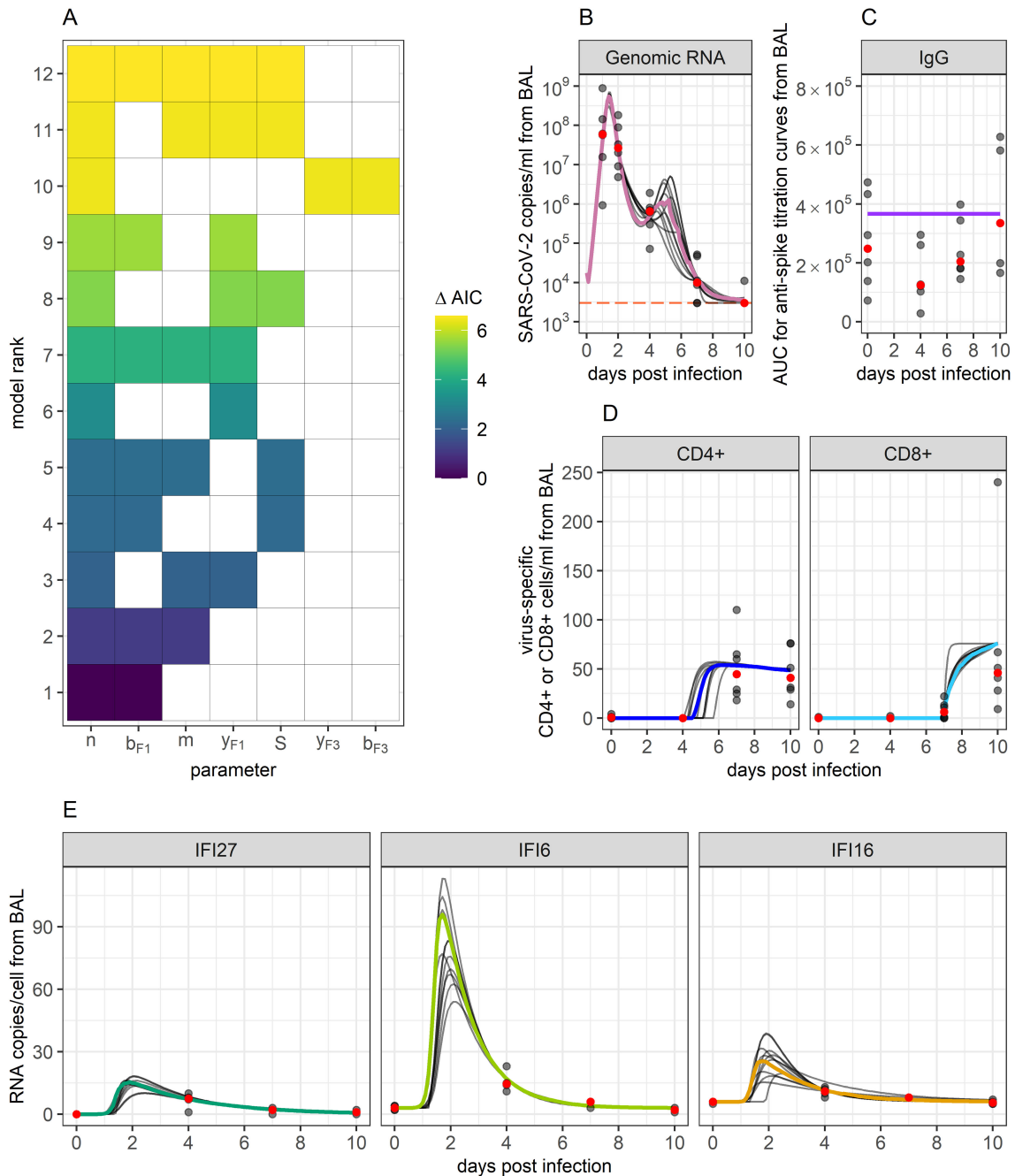


FIGURE 3
 Ensemble model fits to SARS-CoV-2 viral and immune data. **(A)** Ranking of the top 12 models that best fit biological data and whose AIC scores add up to a summed weight of 0.95. The X-axis displays which of the tested parameters were included within a particular model, while the Y-axis ranks models based on their AIC scores. Term “S” represents the inclusion of target cell limitation. AIC scores are indicated by the color of the filled-in boxes. These top 12 models were used to create an ensemble model to capture the combined results. **(B–E)** Ensemble weighted median (colored lines) and the individual top 12 model fits (grey lines) to data. Grey dots indicate data points. Red dots indicate the median of the data points. Orange horizontal line indicates the threshold of detection for qPCR.

continuing back downwards. This resurgence in virus corresponds with a decreasing innate response, while eventual re-control corresponds to rising T cell responses. The model also predicted a surge in innate responses peaking slightly after maximal viral load at comparable orders of magnitude for IFI27, IFI16 and IFI6 relative to baseline to those in humans (11, 69).

2.4 Importance and timing of the innate and adaptive immune response in clearing SARS-CoV-2 infection

Using Akaike weights from the models within the 95% confidence set, we calculated how important each tested model

TABLE 1 95% confidence set of best-ranked models describing SARS-CoV-2 infection in Rhesus Macaques.

Rank	Included Parameters/ Compartments	$\ln(L_j)$	k_j	AIC_j	ΔAIC_j	ER_j	w_j	Summed Weight
1	n, b_{F1}	-741.273	23	1528.547	0.000	1.000	0.305	0.305
2	n, b_{F1}, m	-740.845	24	1529.691	1.144	1.771	0.172	0.477
3	n, m, y_{F1}	-741.292	24	1530.584	2.038	2.770	0.110	0.587
4	n, b_{F1}, S	-741.392	24	1530.784	2.237	3.061	0.100	0.686
5	n, b_{F1}, m, S	-740.433	25	1530.865	2.318	3.187	0.096	0.782
6	n, y_{F1}	-742.860	23	1531.720	3.173	4.887	0.062	0.844
7	n, b_{F1}, m, y_{F1}	-741.381	25	1532.763	4.216	8.231	0.037	0.882
8	n, y_{F1}, S	-742.958	24	1533.916	5.369	14.648	0.021	0.902
9	n, b_{F1}, y_{F1}	-743.086	24	1534.172	5.625	16.650	0.018	0.921
10	n, y_{F3}, b_{F3}	-743.486	24	1534.972	6.425	24.847	0.012	0.933
11	n, m, y_{F1}, S	-742.507	25	1535.014	6.467	25.370	0.012	0.945
12	n, b_{F1}, m, y_{F1}, S	-741.568	26	1535.136	6.589	26.964	0.011	0.956

Here, for each model j within the confidence set, the \ln of the likelihood (L_j), the number of free parameters (k_j), the Akaike information criterion score (AIC_j), the evidence ratio (ER_j), the weight (w_j) and the summed weight is displayed. The tested model parameters that were included in each model are listed. Models were ranked in order of their AIC scores.

term was to obtaining optimal fit to viral and immune dynamic measures during SARS-CoV-2 infection (Table 2). To calculate importance, we summed the Akaike weights of the models in which each term appeared and normalized by the total accumulated weights of the 95% confidence set. We repeated this analysis assuming a lower cut off for late viral loads and detected no significant differences in model selection (not shown).

Our results indicate that parameters describing the impact of the innate immune response on infection were essential to describe observed viral and immune data, appearing in every model within the 95% confidence set, and thus had an importance of 1. Of these innate immune parameters, infected cell clearance by IFI27 appeared most frequently, having an importance of 0.74. Infected cell clearance by virus-specific CD4+ T cells also ranked highly having an importance of 0.95, while infected cell clearance by virus-specific CD8+ T cells had an importance of 0.44. The inclusion of target cell limitation within the model had an importance of 0.24.

Parameter values from the 95% confidence set were relatively preserved across models (Figure 4). As our model predicts when the proliferation of each immune compartment begins, we developed a timeline of how infection is eliminated. Our model suggests that interferon expression, characterizing the innate immune response, expands early during infection. IFI27, IFI6, and IFI16 expansion were predicted to initiate a median of 0.27, 0.72 and 0.52 days post-infection, respectively. This was followed by increases in virus-specific CD4+ T cells beginning at 4.5 days post-infection. Virus-specific CD8+ T cell proliferation was predicted to begin 6.9 days post-infection. Anti-spike IgG antibodies were predicted to increase a median of 9.7 days post-infection; however, IgG expansion rates were very low with the AUC under titration curves expected to increase by a median of 1.9×10^{-7} per infected cell per day. By this time point, BALF viral loads in 5 out of the 6 monkeys were below

the threshold of detection, providing further support that these antibodies had little role in controlling primary infection in this immunologically naïve cohort.

When comparing the models that include infection clearance by the different ISGs, we identified that the per-capita clearance rate of infected cells is a median of 2.5 times faster per IFI27 gene expressed than it is per IFI6 gene expressed (b_{F1}/b_{F3}). Further, we found that the coefficient y_{F1} , allowing for dampening of viral production, is a median of 5.8 times larger per IFI27 gene expressed than it is per IFI6 gene expressed (y_{F1}/y_{F3}). None of the models in the 95% confidence set included the role of infection clearance or dampening of viral production by IFI16. When performing a similar comparison for the models that included infection clearance by T cells, we found that the per-capita clearance rate of infected cells was a median of 3.7 times faster per virus-specific CD4+ T cell/ml than it was per virus-specific CD8+ T cell/ml in the BALF (m/n).

The values of the remaining parameters in our model appear in the Supplementary Material (Supplementary Figure S3), as does an analysis of the correlation between fitted parameter values (Supplementary Figure S4). Notably, some parameter values are positively and negatively correlated indicating a potential lack of complete mathematical identifiability. Alternatively, these correlations may have a mechanistic underpinning. For instance, the three cytokines often proliferate at correlated rates which could reflect an equivalent cellular source; cytokine proliferation often correlates with enhanced cell killing by cytokines as well as CD4+ T cells which may suggest coupled innate and acquired immune responses.

We next used our 95% confidence set of models to examine how the rates of infected cell clearance (Figures 5A, B) and virus production (Figure 5C) fluctuated throughout the 10-day study

TABLE 2 The relative importance of tested parameters in capturing SARS-CoV-2 rhesus macaque infection dynamics.

Parameters/ Compartment	Description	Importance
$b_{F1}, b_{F3}, \gamma_{F1}, \gamma_{F3}$	Impact of the innate response on infection	1
n	Infected cell clearance by CD4+ T cells	0.95
b_{F1}	Infected cell clearance by IFI27	0.74
m	Infected cell clearance by CD8s	0.44
γ_{F1}	Hindrance of viral production by IFI27	0.27
S	Target cell limitation	0.24
γ_{F3}	Hindrance of viral production by IFI6	0.012
b_{F3}	Infected cell clearance by IFI6	0.012
b_{F2}	Infected cell clearance by IFI16	0
γ_{F2}	Hindrance of viral production by IFI16	0
γ_A	Hindrance of infection by anti-spike IgG	0

Importance is determined by summing the Akaike weights of the 95% confidence set of best-ranked models for which a particular parameter appears.

period as a function of the different arms of the immune response. We identified that the rate of infected cell clearance is predicted to peak around day 2 post-infection. Clearance of infected cells was predicted to depend solely on the innate immune response for the first 4 days post-infection. Between days 4 and 5 post-infection, the dominating immune response switched, with virus-specific CD4+ T cells increasing and appearing most closely linked with infected cell lowered viral load. Around 7 days post-infection, some remaining clearance was predicted to be attributed to the virus-specific CD8+ response, though the role of virus-specific CD4+ T cells continued to predominate.

Figure 5C shows how the rate of virus production per-infected-cell-per-day changes over time as a function of the innate immune response. Each infected cell was predicted to produce the lowest amount of virus per day around day 2 post-infection when ISG expression was highest (Figure 3E), with rates at their lowest being a median of 64% of what they were before the innate immune response began to proliferate.

2.5 Projections of infection assuming weaker innate and T-cell responses

To further predict the importance of including each tested term in our mathematical model, we simulated *in silico* knock-down experiments, scaling down model parameter values and observing how these changes impacted predicted viral loads over time. We identified that predicted viral loads showed the greatest increase when we dampened the IFI27-mediated clearance rate of infected

cells, regulated by parameter b_{F1} , and the virus-specific CD4+ T cell-mediated clearance rate of infected cells, regulated by parameter n (Figure 6).

If parameter b_{F1} was knocked down by 50%, 90%, or 99%, the peak viral load was projected to be 2.0, 9.4, or 89 times higher than in the baseline model, respectively. While viral loads reached higher values in these scenarios, this in turn stimulated more virus-specific CD4+ T cell production due to the greater number of infected cells present (not shown). As a result, there was a faster decline in viral levels once virus-specific CD4+ T cell proliferation began. This is a prediction that should be interpreted with caution because innate regulation also helps induce T cell activation. Similar results occurred when we dampened all innate immune response parameters appearing within the 95% confidence set of models ($b_{F1}, b_{F2}, b_{F3}, \gamma_{F1}, \gamma_{F2}$ and γ_{F3}) (Figure 6). This result suggests that IFI27 appears to be the gene best capturing the innate immune response and that its impact on clearing infected cells, rather than dampening infection rates, best described the data.

If parameter n was knocked down, the early viral dynamics remained unaffected due to virus-specific CD4+ T cell proliferation having not yet initiated. Yet, the resurgence of virus predicted at day 4 post-infection was amplified, as the innate response was waning. This outcome is concordant with various degrees of viral rebound noted in untreated individuals in large human cohorts of SARS-CoV-2 infection, particularly during the pre-vaccine era (33). When parameter n was dampened by 50%, 90%, or 99%, viral loads were predicted to rebound to 1.8×10^6 , 7.7×10^6 , or 6.5×10^7 copies/ml of BALF between days 5 and 6 post-infection, respectively. Dampening of other immune parameters or the removal of target cell limitation led to minimal changes in the predicted viral loads of our ensemble model (Supplementary Material, Supplementary Figure S5), suggesting they have a lesser role in hindering or clearing infection in immunologically naïve animals.

2.6 Model forecasts of viral load trajectories given varying degrees of pre-existing innate and acquired immunity

All model fitting was performed on data from previously uninfected and unvaccinated rhesus macaques. The animals therefore had no pre-existing immune memory to SARS-CoV-2. Using our 95% confidence set of models, we attempted to forecast the viral dynamics of re-infection scenarios in rhesus macaques and to assess whether the ensemble model made similar predictions to individual models.

We assumed the same viral inoculum as in the above data but assumed that CD4+ and CD8+ T cell memory would be present due to prior infection, and thus the response to infection by virus-specific T cells would begin sooner. Given the uncertainty regarding the speed of antigen recognition and presentation during early infection, we simulated different scenarios where the CD4+ and CD8+ T cell response was assumed to begin proliferating between 1 and 5 days post-infection (70, 71). Similarly, given the long estimated half-life of memory CD4+ and CD8+ T cells of 200 days (4), we varied the number of virus-specific T cells present in

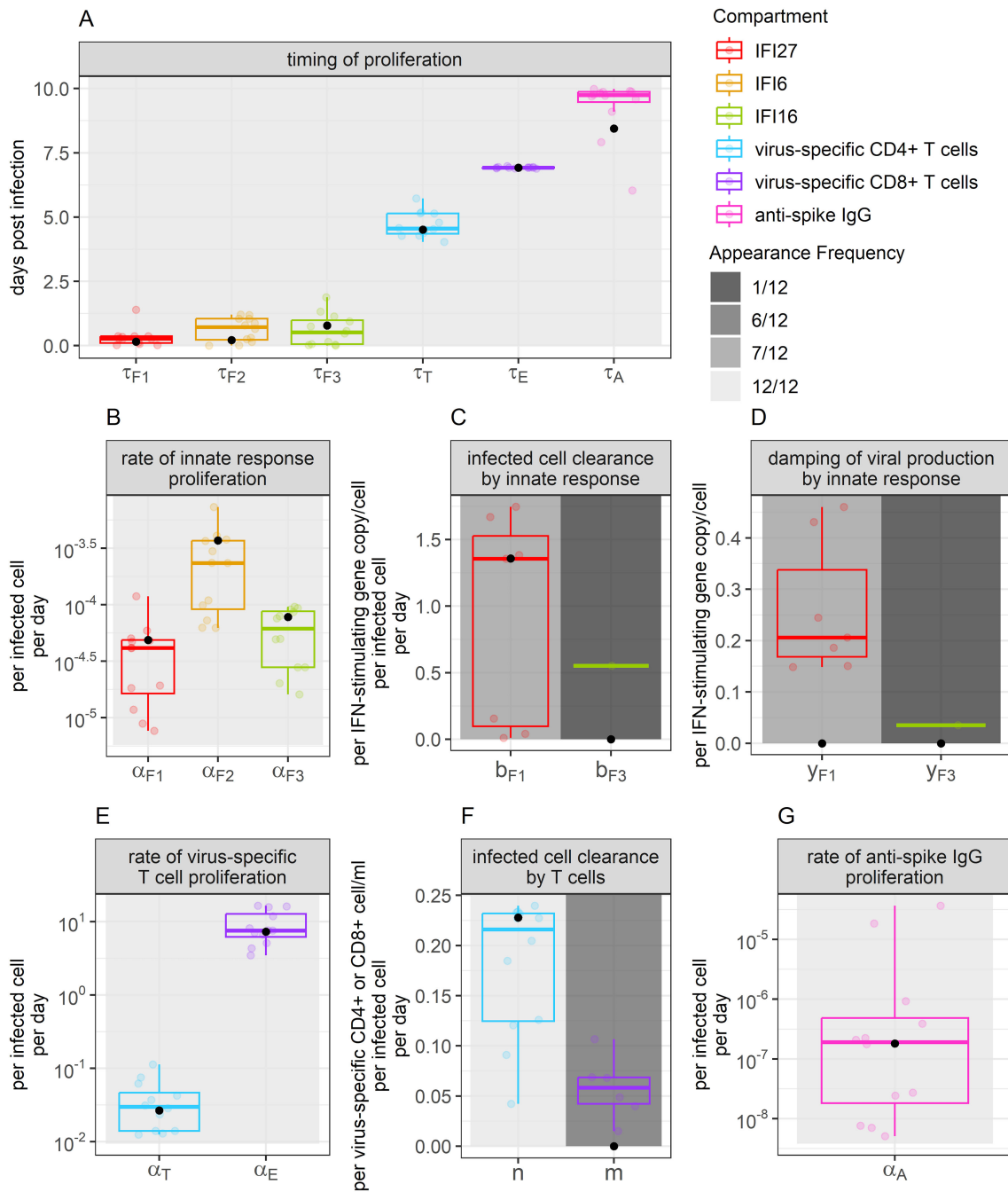


FIGURE 4

Parameter values appearing for the best-ranked models. Boxplot color indicates the immune compartment to which a parameter belongs. Background shade indicates how often a parameter appeared within the 12 best-ranked models. Colored dots show individual parameter values for the models that include each parameter (value not set to 0). Boxplots show the median and interquartile range (IQR), while whiskers indicate 1.5 times the IQR. Black dots indicate the weighted median of each parameter value, determined using model Akaike weights when accounting for all models in the 95% confidence set. Panel (A) displays values related to the timing of proliferation. Panel (B) displays values related to innate immune response proliferation. Panels (C, D) display values related to the clearance of infected cells and damping of viral production by the innate response, respectively. Panels (E, F) display values related to the proliferation of, and clearance of infection by T cells, respectively. Panel (G) displays values describing the rate of anti-spike IgG proliferation.

the BALF from values similar to those observed immediately following infection clearance (≈ 65 cells/ml from BAL) to what might be observed 2 years after initial infection (≈ 5 cells/ml from BAL).

Example results from these simulations are shown in Figure 7. Our ensemble model indicates that the number of virus-specific CD4+ T cells present at the start of re-infection has a larger impact on the predicted maximum viral load than the number of virus-

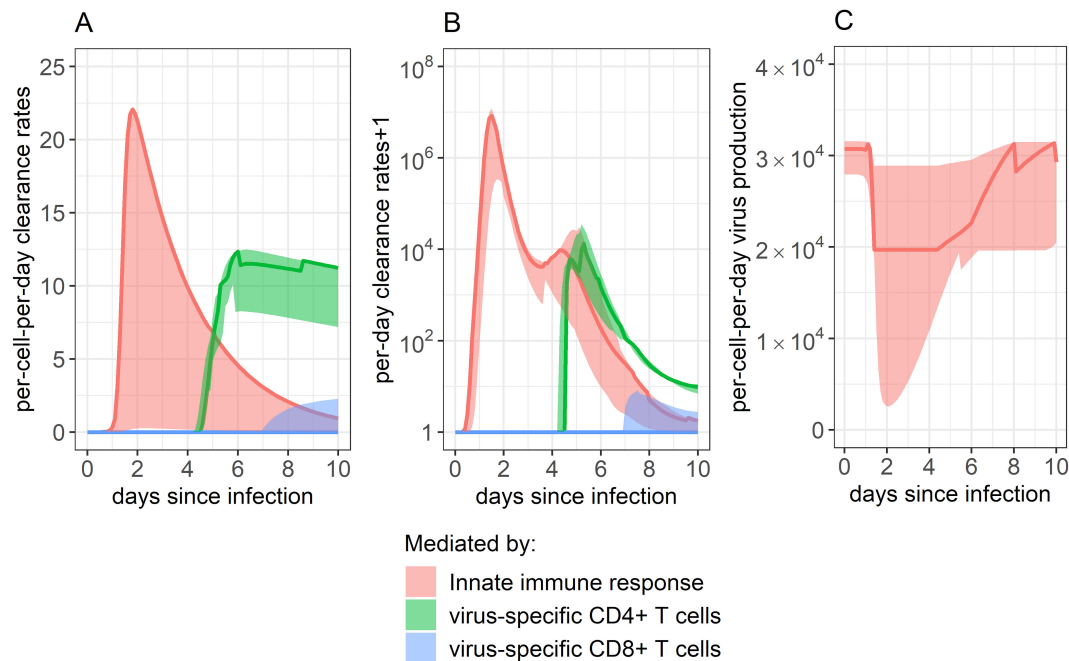


FIGURE 5

Impact of each immune component in eliminating SARS-CoV-2 infected cells. (A, B) Rate of infected cell clearance throughout infection, as mitigated by the innate immune response, CD4+ T cell response, and CD8+ T cell response. Panel A shows the per-infected cell-per-day clearance, while panel B shows the total per-day clearance rate. Note that panel B shows the “per-day clearance rate+1” so that values may be displayed on a log scale. (C) Innate immune system’s impact on dampening the rate of viral production by inducing an antiviral state in infected cells. The innate immune system’s impact was determined by combining the impact of all ISGs examined. As our best-fitting models did not include any impact of anti-spike IgG on infection control, their role is not displayed here. Solid lines show the weighted median, while ribbons show the weighted interquartile range, calculated using the Akaike weights from the 95% confidence set of best-ranked models.

specific CD8+ T cells (Figure 7A). Using general linear models to predict the relationship between peak viral load and virus-specific CD4+ and CD8+ T cells, we found that while 17 additional virus-specific CD4+ T cells/ml of BALF present at the beginning of infection would lower the peak viral load ten-fold, 450 additional virus-specific CD8+ T cells/ml of BALF would be needed at the beginning of infection to have the same effect. The timing of when virus-specific CD4+ and CD8+ T cells were assumed to begin proliferation positively correlated with the predicted maximum viral load, where each additional delay of a day would increase the maximum viral load by approximately 1/3 of a log.

Despite these trends and the similar dynamics predicted by each of the models within the 95% confidence set when describing primary infection (Figure 3), the results of these individual models vary substantially in some of the reinfection scenarios simulated (Figure 7B). While all models in the 95% confidence set include infection clearance by virus-specific CD4+ T cells, only half include infection clearance by virus-specific CD8+ T cells. Thus, while increases in virus-specific CD4+ T cells lead to declines in viral load in all models, increases in virus-specific CD8+ T cells only impact viral loads in the models that include their role in infection clearance. This diversion in model predictions is most apparent when the number of virus-specific CD4+ or CD8+ T cells is high (Figure 7B) and points to the need for greater model specification

against larger and more granular similar datasets, as well as predictive validation against data from animals with previous vaccination and/or infection. Notably, the results from these simulations generate a wide diversity of viral load trajectories with variable peak viral loads, duration of shedding, time to peak, expansion slopes, clearance slopes, and rebound kinetics, all in keeping with results from large human cohorts in the post-vaccination era (33). Overall, these results highlight a potential key impact of pre-infection local immunity on determining the outcome of viral infections.

3 Materials and methods

3.1 Rhesus macaque infection and data collection

Six healthy male rhesus macaques were infected with 2×10^6 TCID₅₀ via a combined intranasal and intratracheal inoculation with the USA-WA1 strain of SARS-CoV-2 as described in (2). Infected rhesus macaques were monitored for ten days, with throat, nasal, BALF, and plasma samples collected throughout. BALF samples were used to perform qPCR, cell staining, flow cytometry, ELISA, and scRNAseq as described in (2). All animal experiments and data collection were performed by Nelson et al. (2)

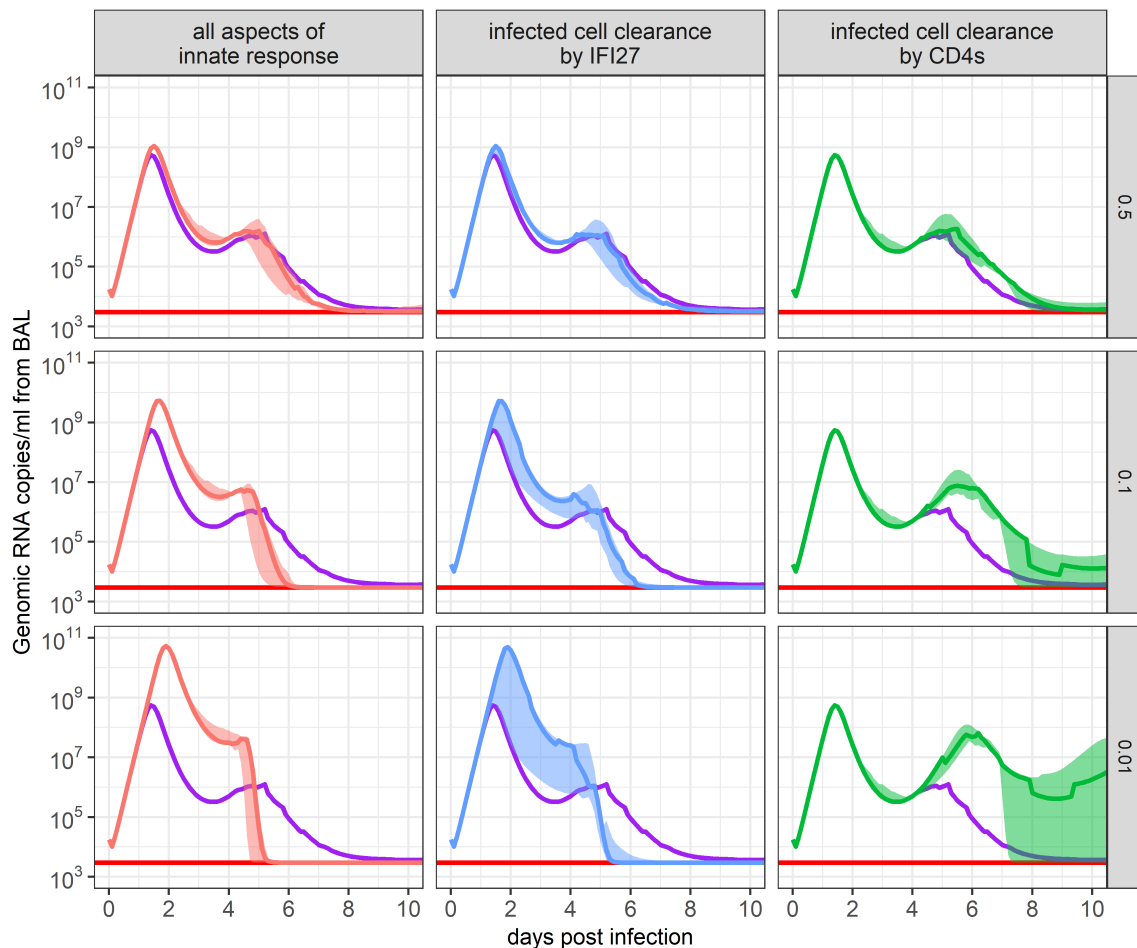


FIGURE 6

Impact of immune response parameters on the ensemble model's SARS-CoV-2 viral load projections. The 95% confidence set of best-ranked models was run where each immune response parameter was multiplied by a scaling factor to maintain or dampen its impact on infection. The resulting ensemble weighted median and IQR of predicted viral loads are shown. The scaled parameters descriptions are in the x-axis strip text while the scaling factor is in the y-axis strip text. Red lines show the threshold of detection and purple lines show the weighted median of the unchanged ensemble model. The panel labeled "all aspects of innate response" indicates the impact of all innate immune response parameters (b_{Fi} and y_{Fi} , where $i=1,2,3$) appearing in the set of best-ranked models.

at the National Institutes of Health in Bethesda, MD. The single cell RNA sequencing read data can be found at NCBI GEO under accession GSE196980. All other source data and detailed methods for data collection can be found at Nelson et al. (2).

3.2 Determining genes of interest based on scRNAseq data

Daily counts/cell for each IFN and ISG examined were averaged across monkeys and then z-score normalized. These values were used in feature selection following the algorithm outlined in (59) and executed using the FSHMM package in R, ranking genes based on their degree of variability over time.

Daily counts/cell were also correlated with the corresponding \log_{10} -transformed genomic RNA measure made on the same day in the same monkey, as shown in Figure 1.

3.3 Simulating and fitting our mathematical model

Our mathematical models were simulated and fit using the POMP package in R (72), with a description of this method appearing in the Supplementary Material. Models were fit to data presented in Figure 1 that were transformed to be representative of counts within BALF. A full explanation of how data was transformed and how initial parameter values were chosen appears in the Supplementary Material. Models were fit by searching parameter space to maximize each model's likelihood using the Nelder-Mead algorithm.

The models were concurrently fit to all rhesus macaque data to give maximum power, with a full description of the fitting process in the Supplementary Material. Code describing how the model was defined and fit can be found on GitHub at <https://github.com/catherinebyrne/SARS-CoV-2-ensemble-model>.

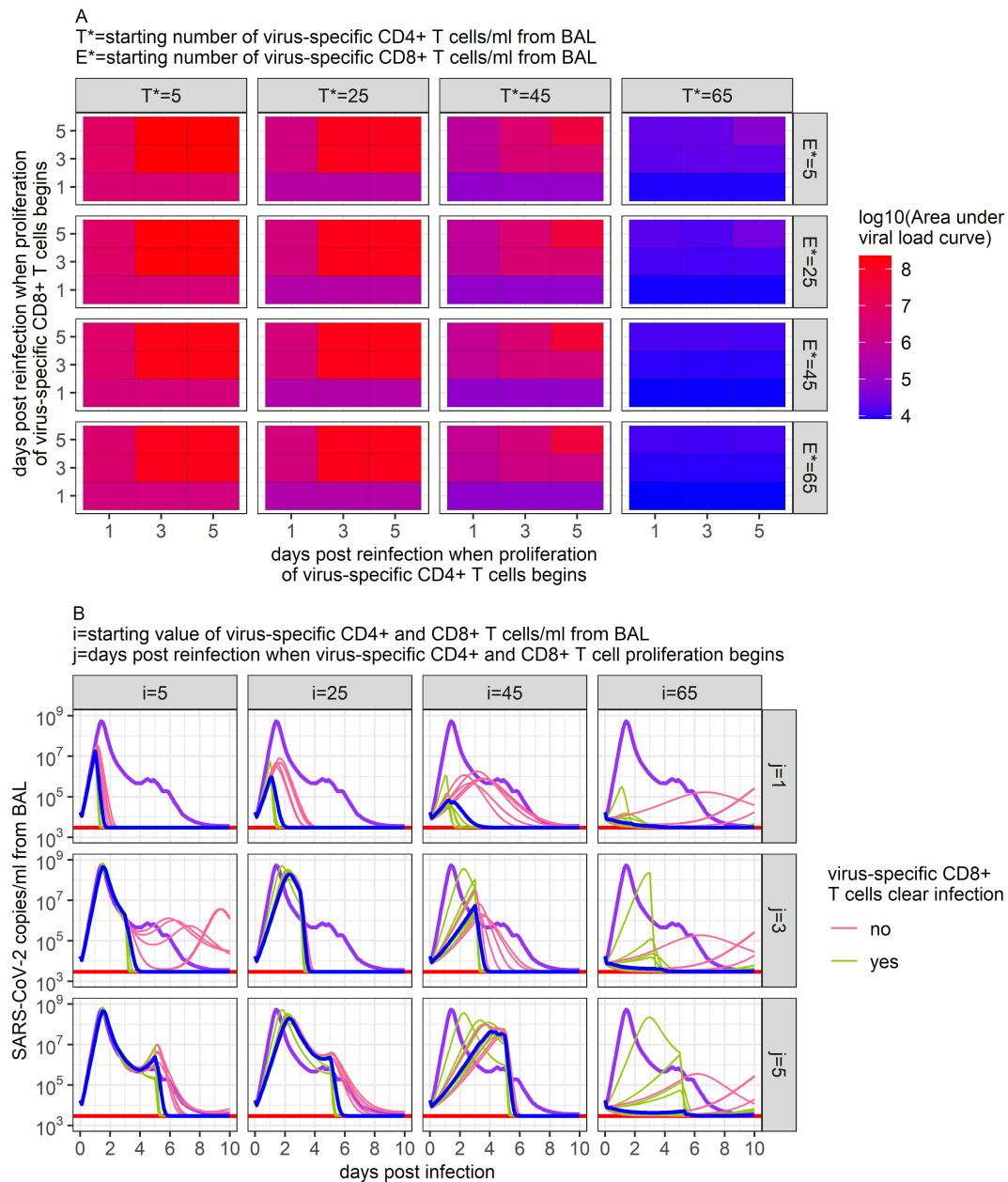


FIGURE 7

Predictions of SARS-CoV-2 infection viral dynamics assuming different memory T cell conditions. Here, we varied the initial number of virus-specific CD4+ T cells present in the BALF (i) and the time at which memory T cell proliferation was assumed to begin (j) to capture potential conditions during a SARS-CoV-2 reinfection. **(A)** Ensemble model's predicted peak SARS-CoV-2 viral load (copies/ml) measured from BAL from all scenarios examined. **(B)** Time series dynamics of predicted viral loads from a subset of the scenarios examined (ratio of virus-specific CD4+ T cells to virus-specific CD8+ T cells is 1:1). Purple lines show the ensemble model's prediction of the viral load during primary infection. Pink and green lines show the individual predictions of the top 12 models within the 95% confidence set. Blue lines show the ensemble model's predictions of the viral load under each reinfection scenario.

3.4 Multi-model inference calculations

We compared our different mathematical models with metrics based on AIC scores to better quantify the probability that each model best represented biological data. We calculated the ΔAIC_j for each model i , where

$$\Delta AIC_j = AIC_j - AIC_{min}$$

and AIC_{min} is the minimum AIC score across all models. From these, we calculated the evidence ratio (ER_j) of each model which quantifies how much more likely the model with the lowest AIC score is to model j , and is defined as

$$ER_j = \frac{e^{-\frac{1}{2}\Delta AIC_{min}}}{e^{-\frac{1}{2}\Delta AIC_j}}$$

Lastly, we calculated AIC weights, where the weight of model j (w_j) is defined as

$$w_j = \frac{e^{-\frac{1}{2}\Delta AIC_j}}{\sum_{g=1}^G e^{-\frac{1}{2}\Delta AIC_g}}$$

and G is the total number of models. These weights represent the probability that a particular model is the best-fitting model of a set, each taking a value between 0 and 1 and summing up to 1.

4 Discussion

The ever-growing volume and complexity of non-linear immunologic data from animal model infection experiments and from human research protocols underscores the necessity of developing new computational methods for transforming these data from being descriptive to more mechanistic. Many SARS-CoV-2 datasets would benefit from such methods, as the data's multidimensionality poses substantial challenges in discerning the precise components of the innate and adaptive immune response driving viral elimination, and immunopathogenesis.

While mechanistic mathematical modeling can be highly beneficial to gain insight into the non-linear dynamics and interplay between pathogens and the immune response (30–33), how to appropriately scale up these models to incorporate multiple immune data types is largely unestablished. In many fields, like climate/weather modeling, ecology, and epidemiology, many models are often created, compared for likelihood, and ultimately combined into an ensemble model, allowing for a more realistic and unbiased approach to forming predictions and understanding a system (51–57). However, this technique has yet to be widely applied to within-host mathematical models of infectious disease. Through this approach, we attempted to fit 160 different mathematical models to data describing SARS-CoV-2 infection within the lungs of rhesus macaques, with each model encoding a specific set of hypotheses describing how the immune response may target and clear SARS-CoV-2 infection.

Through feature selection and correlation of IFN and ISG expression with viral load, we identified IFI27, IFI6, and IFI16 as innate genes likely linked to SARS-CoV-2 clearance and ones whose dynamics may be representative of the entire innate immune response. These genes were also selected based on experimental evidence from human infection linking their expression to viral clearance, though conflicting data from mouse models emphasizes the need to test our model's conclusions further (11, 60, 69, 73). Upon fitting our models to these data, IFI27 emerged as the gene best capturing the innate immune response against SARS-CoV-2. Previous studies have identified the IFI27 gene to be one that is upregulated in the lung tissue of SARS-CoV-2 patients (74), and where expression temporally aligns with SARS-CoV-2 viral loads (11). IFI27 is suggested to regulate the innate immune response during SARS-CoV-2 infection by interfering with the RIG-I

pathway, which is critical for recognizing viral infection and preventing hyper-inflammation and excessive innate immune responses (60, 73). Through analysis of our top-ranked models, the proliferation of the innate immune response was suggested to largely limit infection through direct clearance of infected cells.

Analysis of our different mathematical models also indicated that virus-specific CD4+ T cell-mediated clearance of infection had high importance in capturing infection dynamics, appearing in all top-ranked models. Similarly, clearance of infection by virus-specific CD8+ T cells appeared in half of top-ranked models. Previous research has highlighted the importance of the T cell response in clearing and mitigating the severity of SARS-CoV-2 infection, with the cytotoxic role of CD8+ T cells well-established (42, 75–77). This feature of infection may have been underestimated in our model based on the inclusion of only immunologically naïve animals and highlights the importance of future work fitting models concurrently to infected immune experienced animals as well. While CD4+ T cells are usually thought to help clear infection indirectly through activating and recruiting other immune cells, cytokine production, and regulating the immune response, some studies of SARS-CoV-2 infection have reported significant expansion of cytotoxic CD4+ T cells within infected patients, particularly in those with severe COVID-19 (78–80). Thus, the importance of virus-specific CD4+ T cell-mediated clearance of infection within our model may be a result of its help to other immune cell populations, or through direct cytotoxic activity.

Our model suggested that anti-spike IgG antibodies do not play a discernable role in infection clearance within the first 10 days of infection in immune-naïve animals. Anti-spike IgG titration curves showed no significant change over the 10 days of study (2), and previous studies suggest that titers do not peak until 3 to 7 weeks post-infection (81). While much literature describes the importance of the antibody response against SARS-CoV-2 (22, 81, 82), these responses may instead play more of a role in preventing and controlling re-infection. Indeed, previous mathematical analysis has also shown that plasma SARS-CoV-2 antibodies likely do not impact primary infection, with increases in titers not occurring until later post-infection (35).

Through simulating our top-ranked models when assuming the presence of T cell memory, our results suggest how memory T cells may limit the severity of reinfection, with virus-specific CD4+ T cells having a larger impact than virus-specific CD8+ T cells. While all top-ranked models performed similarly when capturing available primary infection data, they notably diverged in their forecasts of infection when large populations of memory T cells were assumed to be present. By using an ensemble model approach, our predictions consider these differences and describe the uncertainty associated with projections in a way that a single-model analysis would not. This result underscores the importance of ensemble model techniques, especially when forming predictions beyond what available data describes.

The uncertainty of our predictions could be narrowed if our ensemble model reflected the input of fewer models, if top-ranked models differed less substantially in terms of assumptions, and if the top-ranked models had a higher relative likelihood of explaining the

data. To achieve this aim would require fitting models to a wider array of strategically gathered data. In current work, this could be achieved by increasing the amount and type of viral and immune data with a possible focus on gathering these measures at least daily. A common problem for model fitting is that longitudinal sampling is too infrequent. However, in our case, more data time points describing early primary infection may or may not have narrowed the scope of possible best models, as all our top-ranked models predicted similar dynamics during the first 10 days of primary infection. Instead, further testing our models with data collected from other infection scenarios, such as reinfection, or infection following vaccination might effectively rule out the likelihood of multiple models within our current 95% confidence set. These different scenarios may also better reveal the influence of cellular and humoral immunity, which may play a more critical role in repeated infections than our top models indicate.

Another limitation of our models is that they do not include the inherent spatial, multi-compartment dynamics of infection. Our mathematical models are meant to be representative of the infection kinetics within the lung, with the data used to fit our models measured from BALF. While BALF does provide some representation of the immune and viral changes within the secretions of the alveoli of the lower respiratory system, its sampling does not directly capture what is occurring within the lung's mucosa, where infected cells would be present, and tissue-resident immunity may reside. Further, our model treats the lung as a homogeneous region; instead, patches of infected cells and a heterogeneous immune response across different tissue regions are likely present.

Another possible pitfall is subjectivity in terms of model design such that the most likely model may not be included in the list of considered models. Similarly, if key immune mediators of viral clearance are not measured and used for model fit, then these mechanisms will inherently be underrepresented in the ensemble model. These problems are not unique to ensemble models and are relevant to more traditional viral dynamic models as well as immunologic experiments in general. Nevertheless, future work will need to strive to achieve a balance between comprehensive sampling of all arms of the immune response to avoid biasing model conclusions.

Our study demonstrates the potential of integrating diverse mathematical models and ensemble techniques to dissect the dynamics of viral clearance and immune response. These methods will become increasingly necessary due to the breadth and complexity of immune data being collected during longitudinal infection studies. By identifying key genes and predicting the timing and way in which the innate and adaptive immune systems each respond to SARS-CoV-2 infection, our results provide early insights into the intricate interplay between SARS-CoV-2 and the host immune system, while highlighting the need for careful experimental design to best inform these models.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Author contributions

CB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing, Software. JS: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research presented here was financially supported by the National Institute of Allergy and Infectious Diseases (1R01AI177512-01).

Acknowledgments

We would like to thank Dr. Daniel Barber, Dr. Christine Nelson, and Dr. Sivaranjani Namasivayam at Laboratory of Parasitic Diseases, National Institutes of Allergy and Infectious Disease, National Institutes of Health, Bethesda, MD, USA for their assistance in acquiring and comprehending their data on SARS-CoV-2 infections in rhesus macaques.

Conflict of interest

JS has done previous consultation for GSK and Pfizer.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1426016/full#supplementary-material>

References

- Hay JA, Kissler SM, Fauver JR, Mack C, Tai CG, Samant RM, et al. Quantifying the impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound: A retrospective cohort study. *eLife*. (2022) 11:e81849. doi: 10.7554/eLife.81849
- Nelson CE, Namasivayam S, Foreman TW, Kauffman KD, Sakai S, Dorosky DE, et al. NIAID/DIR Tuberculosis Imaging Program, Brooks K, Potter EL, et al. Mild SARS-CoV-2 infection in rhesus macaques is associated with viral control prior to antigen-specific T cell responses in tissues. *Sci Immunol*. (2022) 7:eabo0535. doi: 10.1126/sciimmunol.abo0535
- Chandrashekar A, Liu J, Yu J, McMahan K, Tostanoski LH, Jacob-Dolan C, et al. Prior infection with SARS-CoV-2 WA1/2020 partially protects rhesus macaques against reinfection with B.1.1.7 and B.1.351 variants. *Sci Transl Med*. (2021) 13:eabj2641. doi: 10.1126/scitranslmed.abj2641
- Cohen KW, Linderman SL, Moodie Z, Czartoski J, Lai L, Mantus G, et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep Med*. (2021) 2:100354. doi: 10.1016/j.xcrm.2021.100354
- Hvidt AK, Guo H, Andersen R, Lende SSF, Vibholm LK, Søgaard OS, et al. Long-term humoral and cellular immunity after primary SARS-CoV-2 infection: a 20-month longitudinal study. *BMC Immunol*. (2023) 24:45. doi: 10.1186/s12865-023-00583-y
- He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med*. (2020) 26:672–5. doi: 10.1038/s41591-020-0869-5
- Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N Engl J Med*. (2020) 382:1177–9. doi: 10.1056/NEJMc2001737
- Cheemarla NR, Watkins TA, Mihaylova VT, Wang B, Zhao D, Wang G, et al. Dynamic innate immune response determines susceptibility to SARS-CoV-2 infection and early replication kinetics. *J Exp Med*. (2021) 218:e20210583. doi: 10.1084/jem.20210583
- Borremans B, Gamble A, Prager K, Helman SK, McClain AM, Cox C, et al. Quantifying antibody kinetics and RNA detection during early-phase SARS-CoV-2 infection by time since symptom onset. *eLife*. (2020) 9:e60122. doi: 10.7554/eLife.60122
- Salguero FJ, White AD, Slack GS, Fotheringham SA, Bewley KR, Gooch KE, et al. Comparison of rhesus and cynomolgus macaques as an infection model for COVID-19. *Nat Commun*. (2021) 12:1260. doi: 10.1038/s41467-021-21389-9
- Lim FY, Kim S-Y, Kulkarni KN, Blazevic RL, Kimball LE, Lea HG, et al. High-frequency home self-collection of capillary blood correlates IFI27 expression kinetics with SARS-CoV-2 viral clearance. *J Clin Invest*. (2023) 133:e173715. doi: 10.1172/JCI173715
- Elie B, Roquebert B, Sofonea MT, Trombert-Paolantoni S, Foulongne V, Guedj J, et al. Variant-specific SARS-CoV-2 within-host kinetics. *J Med Virol*. (2022) 94:3625–33. doi: 10.1002/jmv.27757
- Ke R, Martinez PP, Smith RL, Gibson LL, Mirza A, Conte M, et al. Daily longitudinal sampling of SARS-CoV-2 infection reveals substantial heterogeneity in infectiousness. *Nat Microbiol*. (2022) 7:640–52. doi: 10.1038/s41564-022-01105-z
- To KK-W, Tsang OT-Y, Leung W-S, Tam AR, Wu T-C, Lung DC, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis*. (2020) 20:565–74. doi: 10.1016/S1473-3099(20)30196-1
- Zhang B, Yue D, Wang Y, Wang F, Wu S, Hou H. The dynamics of immune response in COVID-19 patients with different illness severity. *J Med Virol*. (2021) 93:1070–7. doi: 10.1002/jmv.26504
- Chan JF-W, Zhang AJ, Yuan S, Poon VK-M, Chan CC-S, Lee AC-Y, et al. Simulation of the clinical and pathological manifestations of coronavirus disease 2019 (COVID-19) in a golden Syrian hamster model: Implications for disease pathogenesis and transmissibility. *Clin Infect Dis*. (2020) 71:2428–46. doi: 10.1093/cid/ciaa325
- Sia SF, Yan L-M, Chin AWH, Fung K, Choy K-T, Wong AYL, et al. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature*. (2020) 583:834–8. doi: 10.1038/s41586-020-2342-5
- Dong W, Mead H, Tian L, Park J-G, Garcia JJ, Jaramillo S, et al. The K18-human ACE2 transgenic mouse model recapitulates non-severe and Severe COVID-19 in response to an infectious dose of the SARS-CoV-2 virus. *J Virol*. (2022) 96:e00964–21. doi: 10.1128/JVI.00964-21
- Dinnon KH, Leist SR, Schäfer A, Edwards CE, Martinez DR, Montgomery SA, et al. A mouse-adapted model of SARS-CoV-2 to test COVID-19 countermeasures. *Nature*. (2020) 586:560–6. doi: 10.1038/s41586-020-2708-8
- Ryan KA, Bewley KR, Fotheringham SA, Slack GS, Brown P, Hall Y, et al. Dose-dependent response to infection with SARS-CoV-2 in the ferret model and evidence of protective immunity. *Nat Commun*. (2021) 12:81. doi: 10.1038/s41467-020-20439-y
- Chu H, Chan JF-W, Yuen K-Y. Animal models in SARS-CoV-2 research. *Nat Methods*. (2022) 19:392–4. doi: 10.1038/s41592-022-01447-w
- McMahan K, Yu J, Mercado NB, Loos C, Tostanoski LH, Chandrashekar A, et al. Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature*. (2021) 590:630–4. doi: 10.1038/s41586-020-03041-6
- Deng W, Bao L, Liu J, Xiao C, Liu J, Xue J, et al. Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. *Science*. (2020) 369:818–23. doi: 10.1126/science.abc5343
- Brady C, Tipton T, Longet S, Carroll MW. Pre-clinical models to define correlates of protection for SARS-CoV-2. *Front Immunol*. (2023) 14:1166664. doi: 10.3389/fimmu.2023.1166664
- Yu J, Collins ND, Mercado NB, McMahan K, Chandrashekar A, Liu J, et al. Protective efficacy of gastrointestinal SARS-CoV-2 delivery against intranasal and intratracheal SARS-CoV-2 challenge in rhesus macaques. *J Virol*. (2022) 96:e0159921. doi: 10.1128/JVI.01599-21
- Hasenkrug KJ, Feldmann F, Myers L, Santiago ML, Guo K, Barrett BS, et al. Recovery from acute SARS-CoV-2 infection and development of anamnestic immune responses in T cell-depleted rhesus macaques. *mBio*. (2021) 12:10.1128. doi: 10.1128/mbio.01503-21
- Yu J, Tostanoski LH, Peter L, Mercado NB, McMahan K, Mahrokhian SH, et al. DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science*. (2020) 369:806–11. doi: 10.1126/science.abc6284
- Vogel AB, Kanevsky I, Che Y, Swanson KA, Muik A, Vormehr M, et al. BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature*. (2021) 592:283–9. doi: 10.1038/s41586-021-03275-y
- Nish S, Medzhitov R. Host defense pathways: Role of redundancy and compensation in infectious disease phenotypes. *Immunity*. (2011) 34:629–36. doi: 10.1016/j.immuni.2011.05.009
- Byrne CM, Johnston C, Orem J, Okuku F, Huang M-L, Rahman H, et al. Examining the dynamics of Epstein-Barr virus shedding in the tonsils and the impact of HIV-1 coinfection on daily saliva viral loads. *PLoS Comput Biol*. (2021) 17:e1009072. doi: 10.1371/journal.pcbi.1009072
- Schiffer JT, Swan DA, Roychoudhury P, Lund JM, Prlc M, Zhu J, et al. A fixed spatial structure of CD8(+) T cells in tissue during chronic HSV-2 infection. *J Immunol*. (2018) 201:1522–35. doi: 10.4049/jimmunol.1800471
- Perelson AS, Ke R. Mechanistic modeling of SARS-CoV-2 and other infectious diseases and the effects of therapeutics. *Clin Pharmacol Ther*. (2021) 109:829–40. doi: 10.1002/cpt.2160
- Owens K, Esmaili-Wellman S, Schiffer JT. Heterogeneous SARS-CoV-2 kinetics due to variable timing and intensity of immune responses. *JCI Insight*. (2024) 9:e176286. doi: 10.1172/jci.insight.176286
- Goyal A, Cardozo-Ojeda EF, Schiffer JT. Potency and timing of antiviral therapy as determinants of duration of SARS-CoV-2 shedding and intensity of inflammatory response. *Sci Adv*. (2020) 6:eabc7112. doi: 10.1126/sciadv.abc7112
- Yang S, Jerome KR, Greninger AL, Schiffer JT, Goyal A. Endogenously produced SARS-CoV-2 specific IgG antibodies may have a limited impact on clearing nasal shedding of virus during primary infection in humans. *Viruses*. (2021) 13:516. doi: 10.3390/v13030516
- Ejima K, Kim KS, Bento AI, Iwanami S, Fujita Y, Aihara K, et al. Estimation of timing of infection from longitudinal SARS-CoV-2 viral load data: Mathematical modelling study. *BMC Infect Dis*. (2022) 22:656. doi: 10.1186/s12879-022-07646-2
- Vaidya NK, Bloomquist A, Perelson AS. Modeling within-host dynamics of SARS-CoV-2 infection: A case study in ferrets. *Viruses*. (2021) 13:1635. doi: 10.3390/v13081635
- Jenner AL, Aogo RA, Alfonso S, Crowe V, Deng X, Smith AP, et al. COVID-19 virtual patient cohort suggests immune mechanisms driving disease outcomes. *PLoS Pathog*. (2021) 17:e1009753. doi: 10.1371/journal.ppat.1009753
- Clairon Q, Prague M, Planas D, Bruel T, Hocqueloux L, Prazuck T, et al. Modeling the kinetics of the neutralizing antibody response against SARS-CoV-2 variants after several administrations of Bnt162b2. *PLoS Comput Biol*. (2023) 19:e1011282. doi: 10.1371/journal.pcbi.1011282
- Uwamino Y, Nagashima K, Yoshifuji A, Suga S, Nagao M, Fujisawa T, et al. Estimating immunity with mathematical models for SARS-CoV-2 after COVID-19 vaccination. *NPI Vaccines*. (2023) 8:1–7. doi: 10.1038/s41541-023-00626-w
- Blanco-Rodríguez R, Du X, Hernández-Vargas E. Computational simulations to dissect the cell immune response dynamics for severe and critical cases of SARS-CoV-2 infection. *Comput Methods Programs BioMed*. (2021) 211:106412. doi: 10.1016/j.cmpb.2021.106412
- Moss P. The T cell immune response against SARS-CoV-2. *Nat Immunol*. (2022) 23:186–93. doi: 10.1038/s41590-021-01122-w
- Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell*. (2020) 181:1489–501.e15. doi: 10.1016/j.cell.2020.05.015

44. Qi H, Liu B, Wang X, Zhang L. The humoral response and antibodies against SARS-CoV-2 infection. *Nat Immunol.* (2022) 23:1008–20. doi: 10.1038/s41590-022-01248-5
45. James LP, Salomon JA, Buckee CO, Menzies NA. The use and misuse of mathematical modeling for infectious disease policymaking: Lessons for the COVID-19 pandemic. *Med Decis Making.* (2021) 41:379–85. doi: 10.1177/0272989X211990391
46. Holmdahl I, Buckee C. Wrong but useful — What Covid-19 epidemiologic models can and cannot tell us. *N Engl J Med.* (2020) 383:303–5. doi: 10.1056/NEJMp2016822
47. Stoica P, Selén Y, Li J. Multi-model approach to model selection. *Digit Signal Process.* (2004) 14:399–412. doi: 10.1016/j.dsp.2004.03.002
48. Oidtman RJ, Omodei E, Kraemer MUG, Castañeda-Orjuela CA, Cruz-Rivera E, Misnaza-Castrillón S, et al. Trade-offs between individual and ensemble forecasts of an emerging infectious disease. *Nat Commun.* (2021) 12:5379. doi: 10.1038/s41467-021-25695-0
49. Beik SP, Harris LA, Kochen MA, Sage J, Quaranta V, Lopez CF. Unified tumor growth mechanisms from multimodel inference and dataset integration. *PLoS Comput Biol.* (2023) 19:e1011215. doi: 10.1371/journal.pcbi.1011215
50. Symonds MRE, Moussalli A. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behav Ecol Sociobiol.* (2011) 65:13–21. doi: 10.1007/s00265-010-1037-6
51. Parker WS. Ensemble modeling, uncertainty and robust predictions. *WIREs Clim Change.* (2013) 4:213–23. doi: 10.1002/wcc.220
52. Li Y, Gao X, An Q, Sun Z, Wang H. Ecological niche modeling based on ensemble algorithms to predicting current and future potential distribution of African swine fever virus in China. *Sci Rep.* (2022) 12:15614. doi: 10.1038/s41598-022-20008-x
53. Mahony CR, Wang T, Hamann A, Cannon AJ. A global climate model ensemble for downscaled monthly climate normals over North America. *Int J Climatol.* (2022) 42:5871–91. doi: 10.1002/joc.7566
54. Bai H, Xiao D, Wang B, Liu DL, Feng P, Tang J. Multi-model ensemble of CMIP6 projections for future extreme climate stress on wheat in the North China plain. *Int J Climatol.* (2021) 41:E171–86. doi: 10.1002/joc.6674
55. Geary WL, Bode M, Doherty TS, Fulton EA, Nimmo DG, Tulloch AIT, et al. A guide to ecosystem models and their environmental applications. *Nat Ecol Evol.* (2020) 4:1459–71. doi: 10.1038/s41559-020-01298-8
56. den Boon S, Jit M, Brisson M, Medley G, Beutels P, White R, et al. Guidelines for multi-model comparisons of the impact of infectious disease interventions. *BMC Med.* (2019) 17:163. doi: 10.1186/s12916-019-1403-9
57. Tebaldi C, Knutti R. The use of the multi-model ensemble in probabilistic climate projections. *Philos Trans R Soc Math Phys Eng Sci.* (2007) 365:2053–75. doi: 10.1098/rsta.2007.2076
58. Waghmare A, Krantz EM, Baral S, Vasquez E, Loeffelholz T, Chung EL, et al. Reliability of self-sampling for accurate assessment of respiratory virus viral and immunologic kinetics. *J Infect Dis.* (2022) 226:278–86. doi: 10.1093/infdis/jiaa451
59. Cárdenas-Ovando RA, Fernández-Figueroa EA, Rueda-Zárate HA, Noguez J, Rangel-Escareño C. A feature selection strategy for gene expression time series experiments with hidden Markov models. *PLoS One.* (2019) 14:e0223183. doi: 10.1371/journal.pone.0223183
60. Villamayor L, Rivero V, López-García D, Topham DJ, Martínez-Sobrido L, Nogales A, et al. Interferon alpha inducible protein 6 is a negative regulator of innate immune responses by modulating RIG-I activation. *Front Immunol.* (2023) 14:1105309. doi: 10.3389/fimmu.2023.1105309
61. Gajate-Arenas M, García-Pérez O, Chao-Pellicer J, Domínguez-De-Barros A, Dorta-Guerra R, Lorenzo-Morales J, et al. Differential expression of antiviral and immune-related genes in individuals with COVID-19 asymptomatic or with mild symptoms. *Front Cell Infect Microbiol.* (2023) 13:1173213. doi: 10.3389/fcimb.2023.1173213
62. Maher AK, Burnham KL, Jones EM, Tan MMH, Saputit RC, Baillon L, et al. Transcriptional reprogramming from innate immune functions to a pro-thrombotic signature by monocytes in COVID-19. *Nat Commun.* (2022) 13:7947. doi: 10.1038/s41467-022-35638-y
63. Persson J, Andersson B, van Veen S, Haks MC, Obudulu O, Torkzadeh S, et al. Stratification of COVID-19 patients based on quantitative immune-related gene expression in whole blood. *Mol Immunol.* (2022) 145:17–26. doi: 10.1016/j.molimm.2022.03.004
64. Shojaei M, Shamshirian A, Monkman J, Grice L, Tran M, Tan CW, et al. IFI27 transcription is an early predictor for COVID-19 outcomes, a multi-cohort observational study. *Front Immunol.* (2023) 13:1060438. doi: 10.3389/fimmu.2022.1060438
65. Huang L, Shi Y, Gong B, Jiang L, Zhang Z, Liu X, et al. Dynamic blood single-cell immune responses in patients with COVID-19. *Signal Transduct Target Ther.* (2021) 6:110. doi: 10.1038/s41392-021-00526-2
66. Ziegler CGK, Miao VN, Owings AH, Navia AW, Tang Y, Bromley JD, et al. Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19. *Cell.* (2021) 184:4713–33.e22. doi: 10.1016/j.cell.2021.07.023
67. Singh DK, Aladyeva E, Das S, Singh B, Esaulova E, Swain A, et al. Myeloid cell interferon responses correlate with clearance of SARS-CoV-2. *Nat Commun.* (2022) 13:679. doi: 10.1038/s41467-022-28315-7
68. Crosse KM, Monson EA, Beard MR, Helbig KJ. Interferon-stimulated genes as enhancers of antiviral innate immune signaling. *J Innate Immun.* (2017) 10:85–93. doi: 10.1159/000484258
69. Lim FY, Lea HG, Dostie A, van Neel T, Hassan G, Takezawa MG, et al. homeRNA self-blood collection enables high-frequency temporal profiling of pre-symptomatic host immune kinetics to respiratory viral infection: a prospective cohort study. *MedRxiv Prepr Serv Health Sci.* (2024) 2023:23296835. doi: 10.1101/2023.10.12.23296835
70. Koutsakos M, Reynaldi A, Lee WS, Nguyen J, Amarasena T, Taiaroa G, et al. SARS-CoV-2 breakthrough infection induces rapid memory and *de novo* T cell responses. *Immunity.* (2023) 56:879–92.e4. doi: 10.1016/j.immuni.2023.02.017
71. Whitmire JK, Eam B, Whitton JL. Tentative T cells: Memory cells are quick to respond, but slow to divide. *PLoS Pathog.* (2008) 4:e1000041. doi: 10.1371/journal.ppat.1000041
72. King AA, Nguyen D, Ionides EL. Statistical inference for partially observed Markov processes via the R package pomp. *J Stat Softw.* (2016) 69:1–43. doi: 10.18637/jss.v069.i12
73. Villamayor L, López-García D, Rivero V, Martínez-Sobrido L, Nogales A, DeDiego ML. The IFN-stimulated gene IFI27 counteracts innate immune responses after viral infections by interfering with RIG-I signaling. *Front Microbiol.* (2023) 14:1176177. doi: 10.3389/fmicb.2023.1176177
74. Kulasinghe A, Tan CW, Miggiolaro AFR dos S, Monkman J, SadeghiRad H, Bhuvu DD, et al. Profiling of lung SARS-CoV-2 and influenza virus infection dissects virus-specific host responses and gene signatures. *Eur Respir J.* (2022) 59:2101881. doi: 10.1183/13993003.01881-2021
75. Tan AT, Linster M, Tan CW, Le Bert N, Chia WN, Kunasegaran K, et al. Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *Cell Rep.* (2021) 34:108728. doi: 10.1016/j.celrep.2021.108728
76. Luo M, Liu J, Jiang W, Yue S, Liu H, Wei S. IL-6 and CD8+ T cell counts combined are an early predictor of in-hospital mortality of patients with COVID-19. *JCI Insight.* (2020) 5:e139024. doi: 10.1172/jci.insight.139024
77. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3
78. Baird S, Ashley CL, Marsh-Wakefield F, Alca S, Ashhurst TM, Ferguson AL, et al. A unique cytotoxic CD4+ T cell-signature defines critical COVID-19. *Clin Transl Immunol.* (2023) 12:e1463. doi: 10.1002/cti2.1463
79. Bieberich F, Vazquez-Lombardi R, Yermanos A, Ehling RA, Mason DM, Wagner B, et al. A single-cell atlas of lymphocyte adaptive immune repertoires and transcriptomes reveals age-related differences in convalescent COVID-19 patients. *Front Immunol.* (2021) 12:701085. doi: 10.3389/fimmu.2021.701085
80. Onofrio LI, Marin C, Dutto J, Brugo MB, Baigorri RE, Bossio SN, et al. COVID-19 patients display changes in lymphocyte subsets with a higher frequency of dysfunctional CD8 α T cells associated with disease severity. *Front Immunol.* (2023) 14:1223730. doi: 10.3389/fimmu.2023.1223730
81. Post N, Eddy D, Huntley C, van Schalkwyk MCI, Shrotri M, Leeman D, et al. Antibody response to SARS-CoV-2 infection in humans: A systematic review. *PLoS One.* (2020) 15:e0244126. doi: 10.1371/journal.pone.0244126
82. Cromer D, Steain M, Reynaldi A, Schlub TE, Wheatley AK, Juno JA, et al. Neutralising antibody titres as predictors of protection against SARS-CoV-2 variants and the impact of boosting: a meta-analysis. *Lancet Microbe.* (2022) 3:e52–61. doi: 10.1016/S2666-5247(21)00267-6