



Kinetics of HIV-Specific CTL Responses Plays a Minimal Role in Determining HIV Escape Dynamics

Yiding Yang^{1*} and Vitaly V. Ganusov^{1,2,3}

¹Department of Microbiology, University of Tennessee, Knoxville, TN, United States, ²National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN, United States, ³Department of Mathematics, University of Tennessee, Knoxville, TN, United States

Cytotoxic T lymphocytes (CTLs) have been suggested to play an important role in controlling human immunodeficiency virus (HIV-1 or simply HIV) infection. HIV, due to its high mutation rate, can evade recognition of T cell responses by generating escape variants that cannot be recognized by HIV-specific CTLs. Although HIV escape from CTL responses has been well documented, factors contributing to the timing and the rate of viral escape from T cells have not been fully elucidated. Fitness costs associated with escape and magnitude of the epitope-specific T cell response are generally considered to be the key in determining timing of HIV escape. Several previous analyses generally ignored the kinetics of T cell responses in predicting viral escape by either considering constant or maximal T cell response; several studies also considered escape from different T cell responses to be independent. Here, we focus our analysis on data from two patients from a recent study with relatively frequent measurements of both virus sequences and HIV-specific T cell response to determine impact of CTL kinetics on viral escape. In contrast with our expectation, we found that including temporal dynamics of epitope-specific T cell response did not improve the quality of fit of different models to escape data. We also found that for well-sampled escape data, the estimates of the model parameters including T cell killing efficacy did not strongly depend on the underlying model for escapes: models assuming independent, sequential, or concurrent escapes from multiple CTL responses gave similar estimates for CTL killing efficacy. Interestingly, the model assuming sequential escapes (i.e., escapes occurring along a defined pathway) was unable to accurately describe data on escapes occurring rapidly within a short-time window, suggesting that some of model assumptions must be violated for such escapes. Our results thus suggest that the current sparse measurements of temporal CTL dynamics in blood bear little quantitative information to improve predictions of HIV escape kinetics. More frequent measurements using more sensitive techniques and sampling in secondary lymphoid tissues may allow to better understand whether and how CTL kinetics impacts viral escape.

Keywords: HIV, CTL escape, multiple responses, mathematical model, model fitting, likelihood

Abbreviations: CTL, cytotoxic T lymphocyte; HIV, human immunodeficiency virus; SIV, simian immunodeficiency virus.

OPEN ACCESS

Edited by:

Juarez Antonio Simões Quaresma,
Instituto Evandro Chagas, Brazil

Reviewed by:

Mario M. D'Elia,
University of Florence, Italy

Bin Su,
Capital Medical University, China

*Correspondence:

Yiding Yang
yyang42@utk.edu

Specialty section:

This article was submitted to
Microbial Immunology,
a section of the journal *Frontiers in
Immunology*

Received: 27 September 2017

Accepted: 16 January 2018

Published: 08 February 2018

Citation:

Yang Y and Ganusov VV (2018)
Kinetics of HIV-Specific CTL
Responses Plays a Minimal Role in
Determining HIV Escape Dynamics.
Front. Immunol. 9:140.
doi: 10.3389/fimmu.2018.00140

1. INTRODUCTION

In 2014, the number of people living with human immunodeficiency virus 1 (HIV-1 or simply HIV) was estimated as 36.9 million (1), with roughly 2 million new HIV infections and 1.2 million people dead of HIV-induced diseases (AIDS) (2). Cytotoxic CD8⁺ T lymphocyte (CTL) responses play an important role in control of virus replication (3, 4) by modulating some important predictors of disease progression (e.g., viral set-point and the rate of CD4⁺ T cell loss (5)). Generation of HIV-specific CD8⁺ T cells by vaccination is one of the current approaches in developing HIV vaccines (6, 7). However, HIV is able to generate mutants (termed “CTL escape mutants”) that are not recognized by HIV-specific T cells, which may be one of the reasons for failure of T cell based vaccines (8–10). Better understanding of mechanisms of viral escape and principles governing CD8⁺ T cell responses to HIV may allow us to evaluate *in silico* a potential efficacy of T cell-based HIV vaccines.

Viral escape from CTL responses follows a somewhat predictive pattern with more dominant (larger magnitude) CTL responses leading to earlier viral escape (11, 12). However, not every CTL response elicits an escape and sometimes viral mutations occur in regions predicted to be recognized by CTLs but in the absence of detectable response (13). To understand the timing and kinetics of CTL escape in HIV/SIV infection, mathematical models have been proposed previously on the dynamics of viral escape from a single CTL response (e.g., Ref. (14–20)). These initial models made a strong assumption of independent viral escape—i.e., it was assumed that viruses escaping from different CTL responses do not compete. Recent work, however, suggested presence of clonal interference and genetic hitchhiking among immune escape variants through reconstruction of HIV whole genome haplotypes (21), and similar concurrent CTL escapes were observed in four HIV-infected patients (22). Clonal interference was suggested to impact the estimates of the escape rates (23, 24). Even though several models have been developed to describe the dynamics of escapes from multiple CTL responses (e.g., Ref. (17, 18, 23–26)), many of these studies involved only model simulations and did not use information on the actual kinetics of HIV-specific CTL responses in predicting viral escape.

Here, we explored whether including experimentally measured CTL kinetics improves description of the viral escape data. In doing so, we compared predictions of three alternative models of viral escape from CTL responses such as independent escapes, sequential escapes, and concurrent escapes. In the first model (independent escapes), we assumed that escape from any given CTL response occurs independently of other escapes and directly from the wild-type, i.e., we ignored the effects of clonal interference—in essence assuming high effective population size and/or high recombination rate. Of note, several recent experimental papers also assumed independent escapes (11–13). In the second model (sequential escape), we assumed that escapes from different CTL responses occur along a defined pathway, generally set by the sequences of escape occurrence in the data. This model assumes strong clonal interference, which may arise at low effective population size or when recombination rate is low. Finally, in the third model (concurrent escape), we tracked all escape variants

simultaneously, thus allowing for co-existence of multiple escape variants (i.e., escapes could occur along multiple alternative pathways). Interestingly, we found that for well-sampled data on virus evolution, the estimated CTL killing efficacies were independent of the model for viral escape. Some escape data could not be well described by the sequential escape model for biologically reasonable parameters. Furthermore, explicitly taking CTL kinetics into account did not improve the quality of fit of different models to escape data. Our results suggest that CTL kinetics in the blood as it is currently available may bear limited information relevant to improve description of kinetics of HIV escape from CTL responses.

2. MATERIALS AND METHODS

2.1. Experimental Data

Experimental details of patient enrollment and data collection were described in detail previously (12, 13). In short, data from 17 patients in the Center for HIV/AIDS Vaccine Immunology (CHAVI) infected acutely with HIV-1 (subtypes B or C) were analyzed in great detail. All patients were infected with a single transmitted/founder (T/F) virus as determined by the single genome amplification and sequencing (SGA/S), and there were enough samples to accurately quantify CTL response to the whole viral proteome. In each patient, the kinetics of virus-specific CTL (CD8⁺ T cell) responses were measured using peptide-stimulated IFN- γ ELISPOT assay and/or intracellular cytokine staining (ICS) 6 months after enrollment using peptides matched to the founder virus sequence (12, 13). For CTL responses measured by ELISPOT, the reported magnitude of the response was the number of cells, producing IFN- γ , per 10⁶ peripheral blood mononuclear cells (PBMC). Multiple viruses were sequenced by SGA/S, and all sequences were compared at sites coding for CTL epitopes, and changes in the percentage of transmitted (wild-type) sequences were followed over time (12). The dynamics of the HIV-specific CTL responses and viral escape from epitope-specific CTL responses were measured longitudinally. Escape mutants were identified as viral variants with mutations in regions recognized by patient's CTL responses with a reduced (or fully abrogated) production of IFN- γ following T cell stimulation. In many cases, mutation in a single position was responsible for the escape. In our analysis, all viral variants, which did not have the wild-type amino acid in the epitope region, were considered as escape variants.

Review of the virus evolution and CTL dynamics data in all 17 patients revealed some data limitations. In particular, data for many patients lacked adequate temporal resolution to accurately estimate virus escape rates. In the vast majority of viral escape variants, escapes often occurred rapidly between two sequential time points with the frequency of the escape variant jumping from 0 to 1. While previously it was suggested that such data may be modified to provide an estimate of the escape rate (14, 15, 17), such approaches may lead to biased parameter estimates (25). While development of a method for unbiased estimation of escape rate from sparse data was recently proposed (27), for this analysis, we focused on patients CH131 and CH159 in which viral escape rates could potentially be accurately estimated due to sufficiently frequent sampling. While data from these patients were presented

before (12), linking of escape and CTL response dynamics was not yet performed.

2.2. Model of Viral Escape from a Single CTL Response

Models describing the dynamics of viral escape from a single cytotoxic T lymphocyte (CTL) response have been developed and adopted by different researchers (e.g., Ref. (14–18)). Here, we start with the basic model formulated earlier (18) and extend it to viral escape dynamics from multiple CTL responses. The model of viral escape from a single CTL response can be extended from the basic viral dynamics model (28) in the following way:

$$\begin{aligned} \frac{dT(t)}{dt} &= s(T_0 - T(t)) - \beta_w T(t)V_w(t) - \beta_m T(t)V_m(t), \\ \frac{dI_w(t)}{dt} &= \beta_w(1 - \mu)T(t)V_w(t) - \delta I_w(t) - kI_w(t), \\ \frac{dI_m(t)}{dt} &= \beta_m T(t)V_m(t) + \beta_w \mu T(t)V_w(t) - \delta I_m(t), \\ \frac{dV_w(t)}{dt} &= p_w I_w(t) - c_v V_w(t), \\ \frac{dV_m(t)}{dt} &= p_m I_m(t) - c_v V_m(t), \end{aligned} \quad (1)$$

where $T(t)$ is the density of uninfected target cells; $I_w(t)$ and $I_m(t)$ is the density of target cells infected by the wild-type or escape variant viruses, respectively; $V_w(t)$ and $V_m(t)$ is the density of wild-type or escape variant viruses, respectively; s is the turnover rate of uninfected target cells; T_0 is the preinfection level of uninfected target cells; β_w and β_m is infection rate of wild-type or escape variant viruses, respectively; μ is the probability of mutation from wild-type to escape mutant during reverse transcription of viral RNA into proviral DNA; δ is the death rate of infected cells due to viral pathogenicity; k is the killing rate of wild-type virus infected cell due to CTL response; p_w and p_m is the rate at which cells infected by wild-type or escape mutant viruses produce viruses; and c_v is the clearance rate of free viral particles.

In this model (equation (1)), we assume that target cells infected by wild-type ($V_w(t)$) and escape viruses ($V_m(t)$) differ by two factors: viral infectivity (β_w and β_m) and the rate of virus production (p_w and p_m). Given that *in vivo* viral particles are short-lived (29, 30), to a good approximation, we may assume a quasi steady state for the virus particle concentration leading to $V_w^*(t) = \frac{p_w}{c_v} I_w(t)$ and $V_m^*(t) = \frac{p_m}{c_v} I_m(t)$. We define a fitness cost $c = 1 - \frac{\beta_m p_m}{\beta_w p_w}$, where c can be positive or negative. Positive c means true fitness cost of escape mutations, which is escape variant and has a lower replication rate ($\beta_m p_m \leq \beta_w p_w$) (31), and negative c implies fitness advantage of escape virus (31, 32). By straightforward calculation, the system (equation (1)) can be written as

$$\begin{aligned} \frac{dV_w^*(t)}{dt} &= [(1 - \mu)r(t) - \delta - k]V_w^*(t), \\ \frac{dV_m^*(t)}{dt} &= [(1 - c)r(t) - \delta]V_m^*(t) + \mu r(t)V_w^*(t)\frac{p_m}{p_w}. \end{aligned} \quad (2)$$

For convenience, we replace $V_w^*(t)$ and $V_m^*(t)$ by $w(t)$ or $m(t)$, respectively, and assume that the wild-type and escape viruses

differ only in the rate of infectivity (that is $\beta_w \geq \beta_m$ and $p_w = p_m$) (13), the system (2) can be simplified as

$$\begin{aligned} \frac{dw(t)}{dt} &= [(1 - \mu)r(t) - \delta - k]w(t), \\ \frac{dm(t)}{dt} &= [(1 - c)r(t) - \delta]m(t) + \mu r(t)w(t), \end{aligned} \quad (3)$$

where $r(t) = \frac{\beta_w p_w}{c_v} T(t)$ is the replication rate of cells infected by wild-type virus, and $c = 1 - \frac{\beta_m}{\beta_w}$ is the cost of the escape mutation defined as a selection coefficient. The frequency of the escape variant in the whole population is given by $f(t) = \frac{m(t)}{w(t)+m(t)}$. This is perhaps the simplest model for a viral escape from a single CTL response. This is denoted as model 1 in the paper.

2.3. Models of Viral Escapes from Multiple CTL Responses

Mathematical model given in equation (3) tracks changes in densities of wild-type virus and a single variant that has escaped recognition from a single epitope-specific CTL response. In acute HIV infection, the virus can escape from recognition of multiple CTL responses, which are specific to several viral epitopes (13, 33). Several models have been developed to describe the dynamics of escapes from multiple CTL responses (e.g., Ref., (17, 18, 26)). Our model is an extension of previous models (17, 18) incorporating mutations from wild-type virus to different viral escapes. In contrast with previous studies, in our analyses, here, we used experimentally measured time courses of different CTL responses (12).

To track the dynamics of viral escape from multiple responses, we assume that there are in total n CTL responses that control viral growth, and virus can potentially escape from all n responses. We use m_i to denote the density of variants where \mathbf{i} is a vector $\mathbf{i} = (i_1, i_2, \dots, i_n)$ denoting the positions of n epitopes, and we define $i_j = 0$ if there is no mutation in the j th CTL epitope and $i_j = 1$ if there is a mutation leading to an escape from the j th ($1 \leq j \leq n$) CTL response. We denote the set of escape variant as I , which is $\mathbf{i} \in I$. The wild-type variant is then denoted as $(0, 0, \dots, 0)$.

For our analysis, we neglect recombination and backward mutation from mutant to wild-type. We use k_i , c_i , and μ_i to denote killing rate due to i th CTL response, cost of escape mutation from the i th CTL response and mutation rate for the i th epitope, respectively. Due to a small rate of double mutation (34), we assume that escape virus is generated with only one mutation in a single generation. That is, for two escape variants $m_i = m_{(i_1, i_2, \dots, i_n)}$ and $m_j = m_{(j_1, j_2, \dots, j_n)}$, we define the mutation rate $M_{i,j}$ from m_i to m_j as μ_k , if and only if m_j has only one more mutation at position k than m_i and all other positions are exactly same. For example, when there are 3 CTL responses, the mutation rate from $m_{(1,0,0)}$ to $m_{(1,1,0)}$ is μ_2 , and the mutation rate from $m_{(0,0,0)}$ to $m_{(0,0,1)}$ is 0. Assuming multiplicative fitness (detailed deviation is given in Section S2 in Supplementary Material), that is, the fitness cost of a variant $\mathbf{i} = (i_1, i_2, \dots, i_n)$ is $C_i = 1 - \prod_{j=1}^n (1 - c_j i_j)$. The death rate of the escape variant $\mathbf{i} = (i_1, i_2, \dots, i_n)$ due to remaining CTL responses is given by $K_i = \sum_{j=1}^n k_j (1 - i_j)$, where we assume that killing of infected cells by different CTL responses is additive.

Similar to equation (3), the dynamics of the wild-type and escape variants are given by

$$\frac{dm_i(t)}{dt} = \left[r(1 - C_i) \left(1 - \sum_{j \in I} M_{i,j} \right) - K_i - \delta \right] m_i(t) + \sum_{j \in I} r(1 - C_j) M_{j,i} m_j(t), \quad \mathbf{i} \in I. \quad (4)$$

We define $M(t) = \sum_{i \in I} m_i$ as the total density of all variants in the population, and $f_j(t)$ ($j = 1, \dots, n$) is the fraction of viral variants that have escaped recognition from the j th CTL response. The frequency of a viral variant escaping from the j th response is given by

$$f_j(t) = \sum_{i \in J} m_i(t)/M(t), \quad J = (i_1, \dots, i_j, \dots, i_n) \text{ with } i_j = 1. \quad (5)$$

Based on previous work (22, 25, 35), we assume that there are two alternative ways to generate escape mutants (**Figure 1**). The first way can be called “sequential” escape (model 2), that is escape mutants are generated sequentially along a defined path from wild-type viruses. This is likely to happen when the effective population size of HIV is small and when the rate of recombination is negligible. The second way can be described as “concurrent” escape (model 3), in which the virus can escape from n CTL responses simultaneously along multiple different pathways. This is likely to happen when the HIV effective population size is large. With n CTL responses, there are n escape variants for “sequential” escape and $2^n - 1$ escape variants for “concurrent” escape in addition to the wild-type variant. For example, with $n = 3$ CTL responses, for “sequential” escape, there are 3 escape variants: $m_{(1,0,0)}$, $m_{(1,1,0)}$, and $m_{(1,1,1)}$ with $m_{(0,0,0)}$ being the wild-type virus. For “concurrent” escape, there are 7 escape variants: $m_{(1,0,0)}$, $m_{(0,1,0)}$, $m_{(0,0,1)}$, $m_{(1,1,0)}$, $m_{(1,0,1)}$, $m_{(0,1,1)}$, and $m_{(1,1,1)}$ with $m_{(0,0,0)}$ being the wild-type virus (**Figure 1**). Detailed equations for both models with $n = 3$ CTL responses can be found in Supplement (Section S2 in Supplementary Material). It is interesting to note that “sequential” escape is a simplification of “concurrent” escape when the effective population size is small. Previous work did not fully resolve whether CTL escapes in HIV infection occur sequentially or concurrently (22, 25); most likely the type of escape varies by patient.

2.4. Models for CTL Response

The killing rate k_i of the CTL response specific to the i th epitope in all three models is composed of two parts: the per-cell killing efficacy of CTLs (k'_i) and the number of epitope-specific CTLs (E_i) (16). Previously the killing rates k_i were often set to a constant (e.g., Ref. (16, 18)), or were set to a certain form $k'_i g(E_i(t))$ where $g(E_i(t))$ is a function of epitope-specific CTL responses $E_i(t)$ (e.g., Ref. (24, 36)). With the measured epitope-specific CTL response dynamics (13), we adopted two forms of killing rate: constant k_i (termed as “constant response”) or time-dependent killing rate $k'_i E_i(t)$ (termed as “interpolated/fitted response”). We used the “mass-action” killing term to describe effect of CTLs on virus dynamics because it is the simplest form, it involves minimum parameters, and it is supported by some experimental data (37).

Based on the available time course information of epitope-specific T cell response $E_i(t)$, we used the first-order interpolation function (termed as “interpolated response”) or the fitted response function (termed as “fitted response”) by the $T_{on}-T_{off}$ model (38) to quantify the kinetics of HIV-specific CTL responses. The $T_{on}-T_{off}$ model assumes that the response starts with E_0 epitope-specific CD8⁺ T cells that become activated at time T_{on} . Activated T cells start proliferating at a rate ρ and reach the peak at time T_{off} . After the peak, epitopes-specific CD8⁺ T cells decline at a rate α . The dynamics of the CD8⁺ T cell response $E(t)$ is given thus by the following differential equation:

$$\frac{dE}{dt} = \begin{cases} 0, & \text{if } t < T_{on}, \\ \rho E, & \text{if } T_{on} \leq t \leq T_{off}, \\ -\alpha E, & \text{if } t > T_{off} \end{cases} \quad (6)$$

with $E(0) = E_0$. Here the “precursor frequency” E_0 is a generalized recruitment parameter, which combines the true precursor frequency and the recruitment rate/time (38, 39). Our recent work showed that this model (equation (6)) reasonably well describes kinetics of HIV-specific CTL responses in acute HIV infection (40). When fitting the model (equation (6)) to experimental data of CTL dynamics, we changed all initial undetected response values from 0 to 1; the latter was the detection limit in the data.

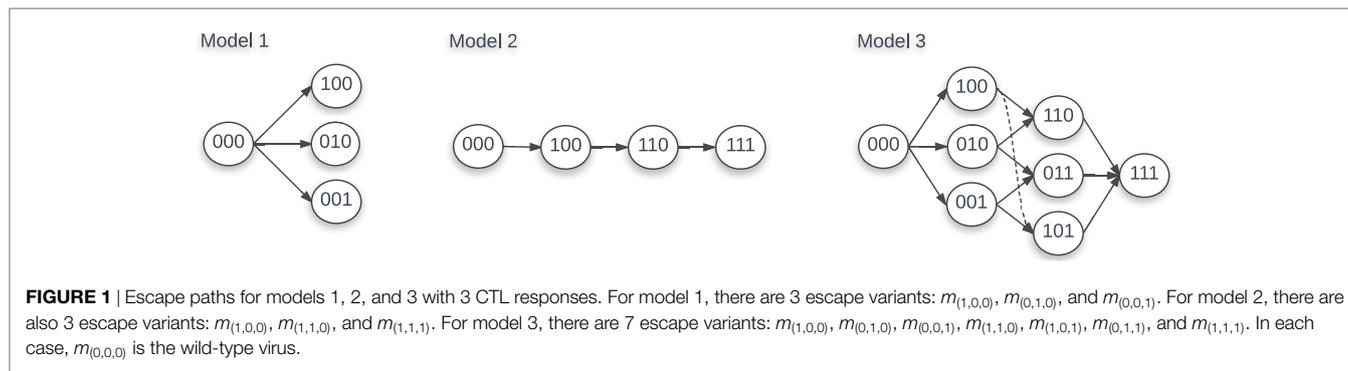
2.5. Statistics

Previously, under the assumption that some mutants are present initially, researchers (e.g., Ref. (16, 36)) fit a logistic model to data on viral escape kinetics by the method of nonlinear least squares (41). In essence, this is a maximum likelihood method, which assumes normally distributed residuals. While this standard statistical method provides reasonable parameter estimates, it assumes equal weights to different data points independently of how many viral sequences were measured at every time point, which is likely to be unrealistic for most experimental studies. Here, we follow the method proposed recently (18) to use binomial distribution (and thus different weights for different measurements/time points) in the likelihood of the model given the escape data. For HIV escape from a single CTL response, the log-likelihood function is given by

$$\mathcal{L} = \sum_{j=1}^{T_i} [a_j \ln(f(t_j)) + (N_j - a_j) \ln(1 - f(t_j))], \quad (7)$$

where a_j is the number of escape variant sequences in a sample of N_j sequences at the sample time t_j , T_i is the number of measured time points for a i th specific viral escape trajectory, and $f(t_j)$ is the predicted frequency of a specific viral escape variant at time t_j . Model parameters were thus found by maximizing the log-likelihood function (equation (7)).

To discriminate between alternative models under different parameter constrains, we used corrected Akaike information criterion (AIC) scores (42). The model fit with the minimum AIC score among tested models was treated as the best model; however, a difference of less than 3 AIC units is generally viewed as not



significant (42). To test the statistical significance of the differences between parameters found by fitting different models, we used a bootstrap approach (43). In this approach, we resampled the data 1,000 times using the Random routine in Mathematica assuming beta distribution for sequencing data (44), fitted models to bootstrap samples, and recorded all estimated parameters. For the same parameter, we use either paired and unpaired *t*-test to compare the parameter averages for different models.

Both fitness costs of escape mutations and the killing efficacy of the CTL response determine the kinetics of viral escape from T cells (14–16), and that viral escape (sequence) data in most cases are not sufficient to estimate both rates (16). Therefore, in our analyses, to avoid overfitting, we set fitness cost of escape to 0 $c_i = 0$. In all fits, we assumed that the rate of virus replication $r = 1.5/\text{day}$ (28).

While multiple models may be able to describe accurately experimental data, some models may do so at biologically unreasonable parameters. For example, estimated rate of mutation at different epitopes may be unrealistically large. Thus, in our analysis, we assume that mutation rates, which are above 10^{-3} are likely to be unrealistic given that currently estimated HIV mutation rate is about 3.2×10^{-5} per bp per generation (34) and size of a CTL epitope is 8–10 amino acids ($3 \times 10 \times 3.2 \times 10^{-5} \approx 10^{-3}$).

To fit the $T_{\text{on}}-T_{\text{off}}$ model [equation (6)] to experimental data using non-linear least squares, we log-transformed the model predictions and the data.

When interpolating CTL response kinetics, there was often not enough information on the starting point (day 0). In such situations, we set the initial CTL density as 1 (the detection level for this data set) for simplicity. Other starting points (e.g., intersection point of the CTL response axis and the reverse extension line of the interpolation function) were also tested and led to similar results (not shown). This was largely due to the fact that, in our models, CTLs at low densities are not expected to exert large selective pressure on the virus population due to assumed mass-action killing term.

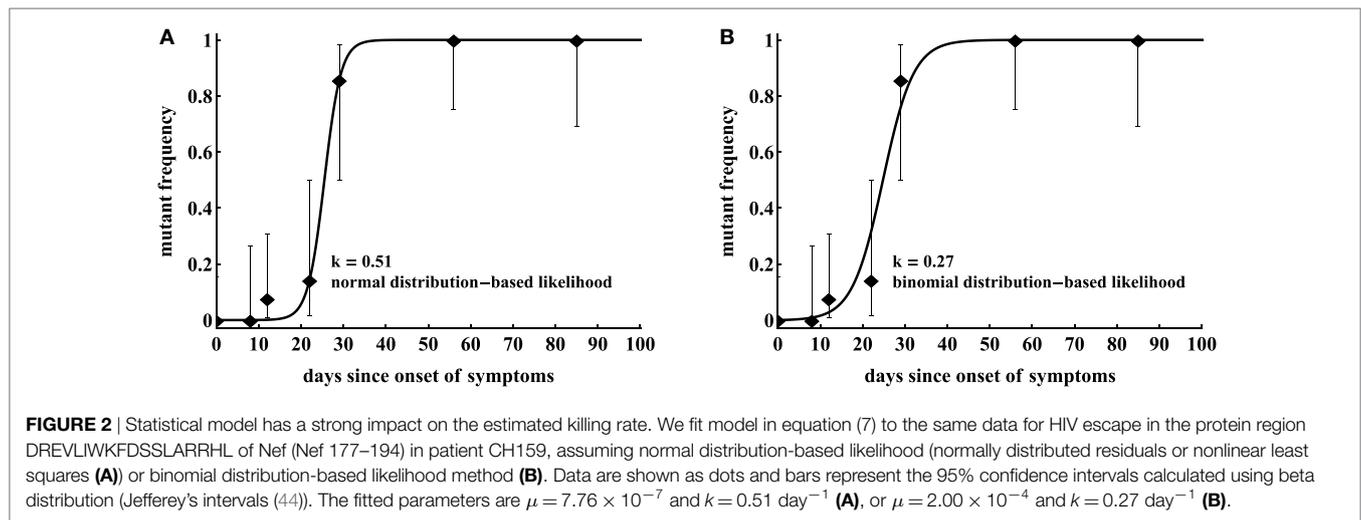
3. RESULTS

3.1. Statistical Model Impacts Estimation of the Escape (Killing) Rate

Given virus evolution data, we may be often interested in quantifying selecting pressures driving specific changes in the virus

population. Following HIV-1 infection, the virus escapes from several cytotoxic T lymphocyte (CTL) responses (45), and multiple studies used mathematical models of various levels of complexity to estimate the predicted efficacy at which CTLs recognize and eliminate cells, infected with the wild-type (unescaped) virus (14–18, 25). Many of these previous studies estimated the rate of HIV escape from immunity using nonlinear least squares, which explicitly assumes normal distribution of the deviations between model predictions and data (14–17). However, the assumption of normally distributed residuals is likely to be violated for data when only a handful of viral genomes are sequenced—which is common in many studies involving single genome amplification and sequencing techniques (SGA/S). We have recently proposed to use a likelihood approach, which assumes virus genome sampling to follow a binomial distribution (18). This binomial distribution-based likelihood approach showed to impact the estimates of the CTL killing rate (escape rate can be proportional to the killing rate under an assumption of constant CTL response) when compared to normal distribution-based likelihood approach (least squares) (18). However, this previous comparison was done on data, which were fairly sparse and comparison involved modifications of data to allow for non-zero and non-one frequencies of the escape variant (14, 15), and thus, it remained unclear if estimates of escape rates are truly dependent on the statistical model for better sampled data.

Unfortunately, in our cohort of 17 patients (12), very few patients were sampled frequently enough to observe gradual accumulation of escape variants in the population (i.e., data with two sequential time points with mutant frequency in the range $0 < f < 1$ were rare). For the analysis, we, therefore, used the escape data from two patients, CH131 and CH159, where CTL and HIV sequence measurements were sufficiently frequent to address our modeling questions. We fitted a simple mathematical model describing escape of the virus from a single constant (non-changing) CTL response (equation (3)) to the data from one patient CH159 (Figure 2) assuming two different statistical models: with normally distributed residuals (least squares) or binomial distribution-based likelihood (equation (7)). Consistent with our previous observation, we found that the type of statistical model impacts the estimate of the escape rate (k in Figure 2) with difference being nearly twofold ($k = 0.27/\text{day}$ vs. $k = 0.51/\text{day}$). It is interesting to note that, visually, the least squares method appear to describe the data better by accurately fitting the points with intermediate frequency of the escape variant in 20–30 days



after the symptoms (but missing the another intermediate data point (12, 0.08)). However, this visually better fit is not supported by the statistics: likelihood of the model for these data is -12.64 or -10.53 for normal (**Figure 2A**) or binomial (**Figure 2B**) distribution, respectively (and AIC scores being 31.0 vs. 26.8, respectively). Interestingly, the main difference in the estimated escape rates was driven by just one data point ($(t, f) = (12, 0.08)$); removing this data point from the data led to identical estimates of the escape rate, $k = 0.51/\text{day}$, from two statistical models (results not shown). This is not surprising because with this data point removed, the information on escape rate is only coming from two data points when the frequency of the escape variant is intermediate ($0 < f < 1$).

As discussed before, least squares may not allow to estimate escape rates, e.g., in cases when mutant frequency jumps from 0 to 1 between two subsequent time points unless data are modified (14, 15). Similarly, models assuming normally distributed residuals may not be able to fit other types of data, in which frequency of the mutant has an intermediate value ($0 < f < 1$) at one time point only. In particular, in our analysis of another escape in patient CH159 (Rev GRPTEPVFPQLPPLRLC, see **Figure 3**), we could not obtain finite estimates of the escape rate using normally distributed residuals (results not shown). Rather, the model fits tended to describe accurately two data points ($t = 22$ days and $t = 29$ days) and ignore another data point ($t = 56$ days) leading to extremely high predicted escape rates (results not shown). Interestingly, using binomial distribution-based likelihood allowed for an accurate fit of the model to data and the fit compromised between describing early and late data points (**Figure 4A**). The reason for the compromise is that a fit predicting fast escape and nearly 100% escape variant by 56 days since symptoms is highly disfavored by the binomial distribution-based likelihood because some wild-type variants were still present at day 56 (thus, the weight for missing this point by the model fit was very high in binomial distribution-based likelihood but not in the normal distribution-based likelihood). Taken together, these results suggest that the type of the statistical model used to estimate HIV escape rates influences the final estimates. Therefore, many previous studies on HIV escape assuming normally distributed

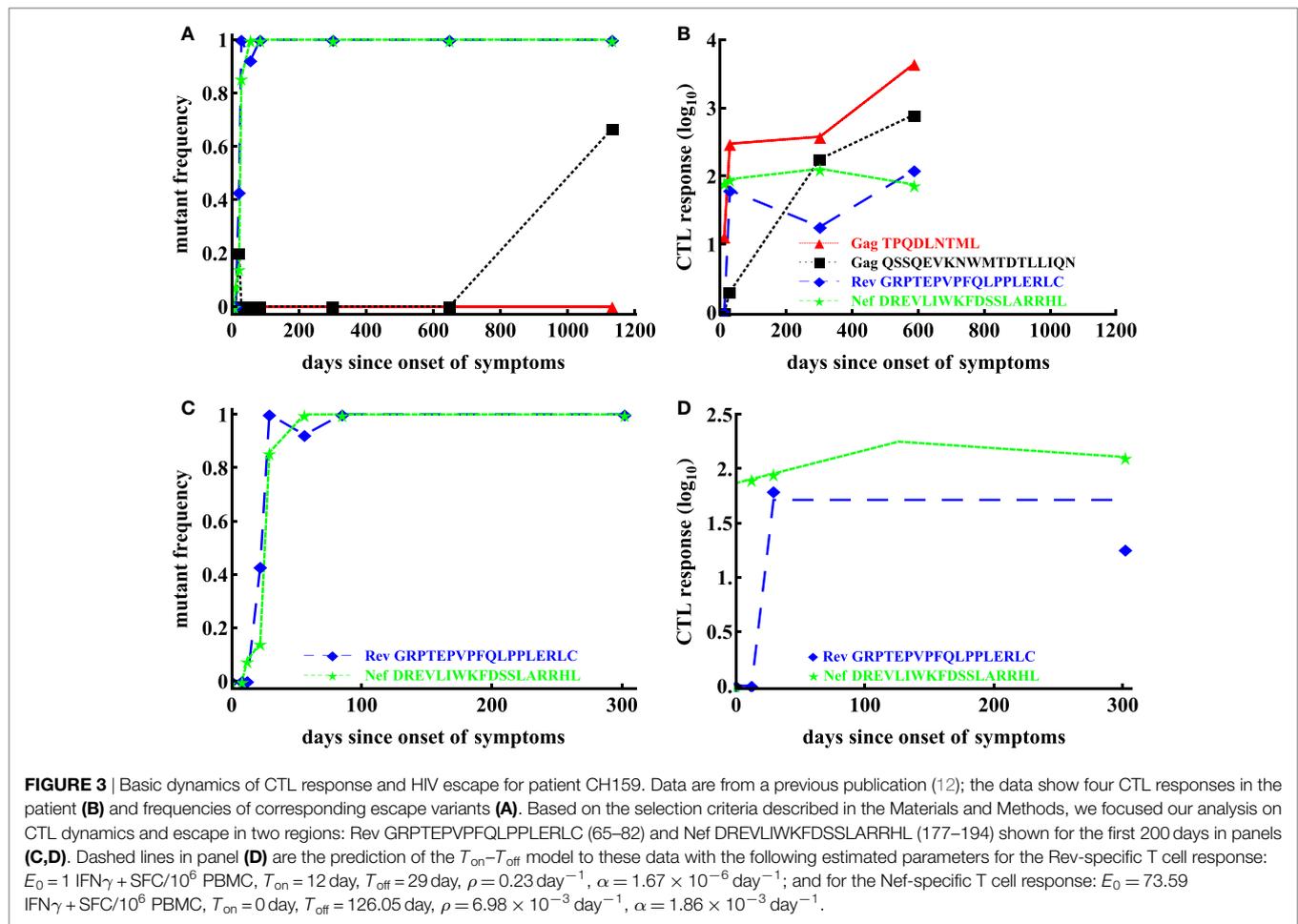
residuals may need to be re-evaluated for the robustness of their conclusions.

3.2. CTL Response Kinetics Do Not Improve Description of the Escape Data

As CTL responses drive HIV escape from epitope-specific T cells, it is expected that the magnitude of the CTL response should naturally impact escape kinetics. Previous studies provided some evidence that the relative magnitude of a given CTL response in the total HIV-specific CTL response early in infection (% immunodominance) predicts the timing of viral escape (11, 12). Immune response was also shown to impact escape of simian immunodeficiency virus (SIV) from T cell responses (19, 46, 47). Immune response magnitude, and as a consequence, the overall CTL killing efficacy is important in determining both timing and speed of viral escape with the rate of viral escape being directly related to the immune response efficacy (16, 17). In contrast, both initial mutant frequency, virus mutation rate, and CTL killing efficacy determine timing of viral escape (17). Whether inclusion of the experimentally measured CTL dynamics impacts ability of mathematical models to accurately describe viral escape data has not been tested.

To test the benefits of using longitudinally measured CTL responses in describing viral escape data, we considered several alternative models for the CTL dynamics and viral escape. Our model 1 describes the dynamics of viral escape from each CTL response independently. Models 2 and 3 describe escape from multiple CTL response that occurs sequentially or concurrently, respectively (see Materials and Methods for more details). CTL dynamics was either considered to be unimportant (i.e., killing rate k_i was set constant over time), or when killing rate was proportional to the experimentally measured CTL frequency ($k_i E_i(t)$), respectively. To describe CTL dynamics, we either used the first order interpolation function or the $T_{\text{on}}-T_{\text{off}}$ model (equation (6) and see Materials and Methods for more detail).

In patient CH159, four CTL responses were detected (**Figure 3B**), and three of these responses were escaped within nearly 4 years of infection. Interestingly, the response specific



to Gag TPQDLNTML was dominant (Figure 3B), but the corresponding escape mutant Gag TPQDLNTMLNTVGGHQAA did not appear up to 1,132 days since onset of symptoms (Figure 3A).

Patient CH159 had two escape mutants in regions Rev GRPTEPVPFQLPPLERLC (Rev 65–82) and Nef DREVLWKFSSSLARRHL (Nef 177–194) satisfying our selection criteria (Figure 3C). Despite a relative small magnitude of CTL responses specific to Rev65 and Nef177 early in infection (up to 29 days since onset of symptoms), escape mutants appeared early and their frequencies arose rapidly.

We fitted three alternative mathematical models for viral escape and three alternative models for the CTL dynamics to the data on viral escape (Figure 3C) using binomial distribution-based likelihood method (see Materials and Methods for more detail). Surprisingly, we found that the models 1 and 3 with a constant immune response described the data with best quality as judged by the AIC (or likelihood). Parameter estimates in the model 1, which assumes independent escape were nearly identical to the parameters in the model 3, which assumed concurrent escape (Figure 4; Table 1). Importantly, adding experimentally measured CTL response dynamics (as interpolated function or by using parameterized $T_{on}-T_{off}$ model) did not improve the quality of the model fit to escape data (Table 1). Even worse, for models 1 and

3, the fits with a fitted response were of lower quality as judged by the large increase in AIC (Table 1). Models that included an interpolated CTL response provided better fits than models with a fitted response (Table 1).

The exact reasons of why including experimentally measured CTL response dynamics led to worse fits of the escape data are unclear but perhaps rapid change in magnitude of CTL responses in this patient—if response directly impacts killing of infected cells—was simply not reflected in the kinetics of viral escape (Figures 4D,G). Specifically, CTL kinetics-driven escape would predict non-monotonic rise in the escape variant frequency, which was not observed in the data, thus, favoring a model with a constant killing rate by CTLs.

Interestingly, the model 2 fits of the data resulted in unphysiologically large estimates for the mutation rate μ_2 (Table 1). As we elaborate later (see below), this failure of the model to describe these data stems from the fact that escapes in the data occur nearly at the same time and assuming that escapes are sequential led to an unrealistic mutation rate in the second epitope. This suggests that the observed dynamics of viral escape in patient CH159 is not consistent with sequential escape.

Models 1 and 3 also predicted slightly higher than expected mutation rate μ_1 (bigger than 10^{-3}) for the peptide Rev 65–82. Constraining this parameter to remain $\mu_1 \leq 10^{-3}$ led to fits of

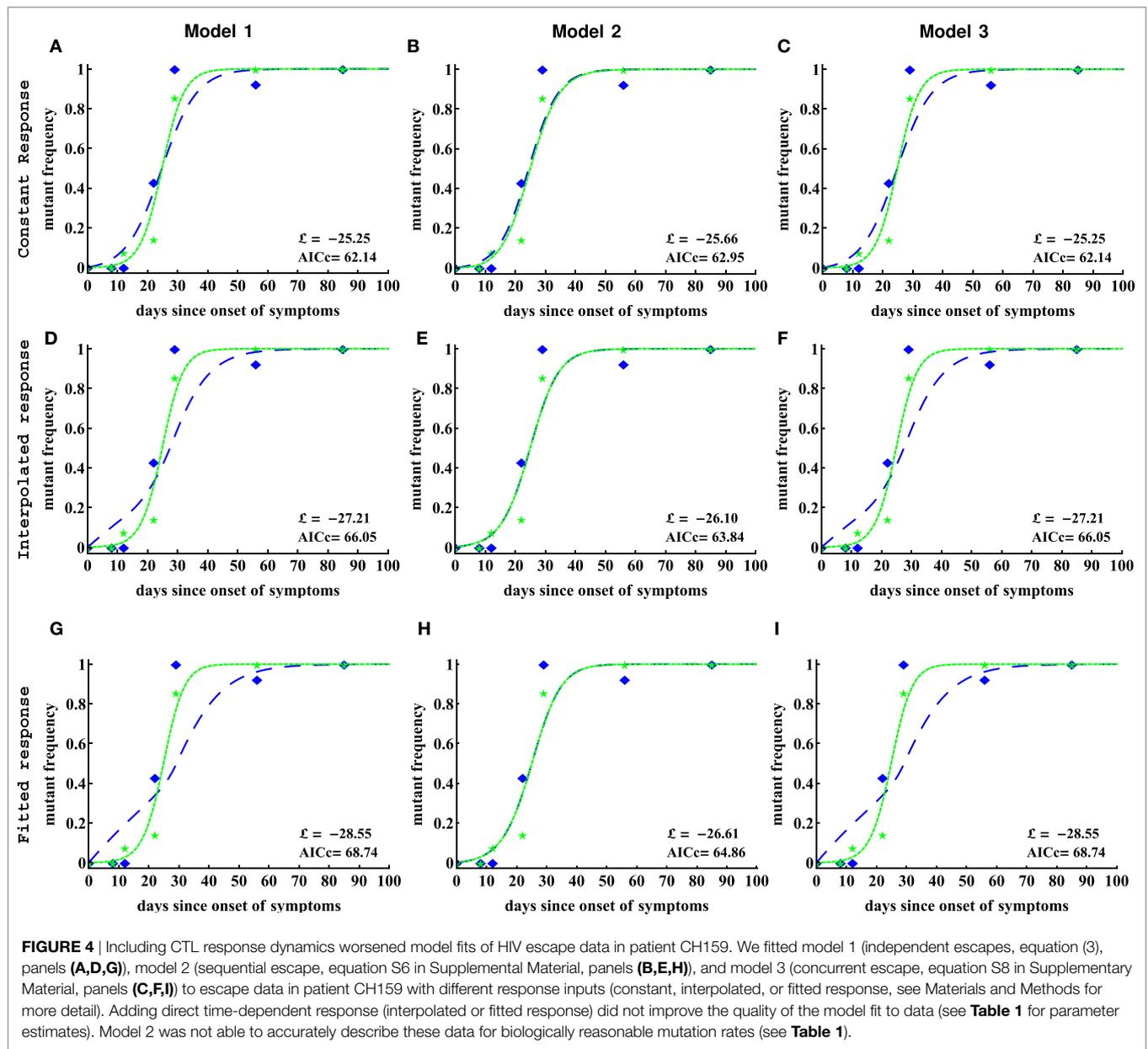


FIGURE 4 | Including CTL response dynamics worsened model fits of HIV escape data in patient CH159. We fitted model 1 (independent escapes, equation (3), panels (A,D,G)), model 2 (sequential escape, equation S6 in Supplemental Material, panels (B,E,H)), and model 3 (concurrent escape, equation S8 in Supplementary Material, panels (C,F,I)) to escape data in patient CH159 with different response inputs (constant, interpolated, or fitted response, see Materials and Methods for more detail). Adding direct time-dependent response (interpolated or fitted response) did not improve the quality of the model fit to data (see Table 1 for parameter estimates). Model 2 was not able to accurately describe these data for biologically reasonable mutation rates (see Table 1).

significantly lower quality (likelihood ratio test, $p < 0.05$). Due to large length of the peptide, the overall mutation rate in this region could indeed be slightly higher than our calculated high bound for the mutation rate (see Materials and Methods for more detail). Furthermore, since peptide Rev 65–82 is the epitope in which first escape occurred, it was possible that the high estimate of the mutation rate could be due to late sampling of viral sequences. In these, data sampling was done after patients were diagnosed with infection; however, viral escape could have started earlier and for escapes starting earlier, it may be possible to describe the data with a lower mutation rate (18, 48).

Therefore, to test whether the timing of the start of the escape influences the estimate of the mutation rate we did the following. We shifted the data for two escapes forward by adding some initial zeroes to data and reverse extended the predicted CTL response

curves. Then we refitted models 1 and 3 to the data under the constrain $\mu \leq 10^{-3}$. We found shifting the data did not improve the quality of the model fits as compared to unmodified data when CTL dynamics is explicitly taken into account as interpolated or fitted response (results not shown). However, assuming a constant response allowed to obtain lower, more physiological estimates of the mutation rate. These results suggest that inability of the models, which explicitly incorporate CTL dynamics to explain kinetics of first escape with physiologically reasonable mutation rate is due to late appearance of the CTL response. Indeed, escape can only accumulate when CTL response is present and extending the time window for virus evolution but not having CTL response active will not significantly impact estimates of the mutation rate.

Given our results for one patient, we next sought to investigate whether our conclusions will remain robust when looking at

TABLE 1 | Parameters for the three models fitted to escape data from patient CH159.

Peptide		Model 1		Model 2		Model 3	
Constant response		Mutation rate	Killing rate	Mutation rate	Killing rate	Mutation rate	Killing rate
		$(\mu_i, i = 1, 2)$	$(k_i, i = 1, 2)$	$(\mu_i, i = 1, 2)$	$(k_i, i = 1, 2)$	$(\mu_i, i = 1, 2)$	$(k_i, i = 1, 2)$
	Rev 65–82	<i>1.68×10^{-3}</i>	0.17	<i>9.71×10^{-4}</i>	0.20	<i>1.68×10^{-3}</i>	0.17
Nef 177–194	2.02×10^{-4}	0.27	0.11	6.29×10^{-12}	2.0×10^{-4}	0.27	
		$\mathcal{L} = -25.25, AICc = 62.14$		$\mathcal{L} = -25.66, AICc = 62.95$		$\mathcal{L} = -25.25, AICc = 62.14$	
Interpolated response		Mutation rate	Killing rate	Mutation rate	Killing rate	Mutation rate	Killing rate
		$(\mu_i, i = 1, 2)$	$(k_i', i = 1, 2)$	$(\mu_i, i = 1, 2)$	$(k_i', i = 1, 2)$	$(\mu_i, i = 1, 2)$	$(k_i', i = 1, 2)$
	Rev 65–82	<i>8.88×10^{-3}</i>	2.12×10^{-3}	<i>1.64×10^{-3}</i>	2.03×10^{-10}	<i>8.88×10^{-3}</i>	2.12×10^{-3}
Nef 177–194	4.94×10^{-4}	3.23×10^{-3}	697.77	2.32×10^{-3}	4.93×10^{-4}	3.23×10^{-3}	
		$\mathcal{L} = -27.21, AICc = 66.05$		$\mathcal{L} = -26.10, AICc = 63.84$		$\mathcal{L} = -27.21, AICc = 66.05$	
Fitted response		Mutation rate	Killing rate	Mutation rate	Killing rate	Mutation rate	Killing rate
		$(\mu_i, i = 1, 2)$	$(k_i', i = 1, 2)$	$(\mu_i, i = 1, 2)$	$(k_i', i = 1, 2)$	$(\mu_i, i = 1, 2)$	$(k_i', i = 1, 2)$
	Rev 65–82	<i>1.43×10^{-2}</i>	1.39×10^{-3}	<i>1.73×10^{-3}</i>	8.50×10^{-18}	<i>1.43×10^{-2}</i>	1.39×10^{-3}
Nef 177–194	2.46×10^{-4}	3.25×10^{-3}	13,004.84	2.29×10^{-3}	2.47×10^{-4}	3.25×10^{-3}	
		$\mathcal{L} = -29.68, AICc = 70.99$		$\mathcal{L} = -26.61, AICc = 64.86$		$\mathcal{L} = -29.68, AICc = 70.99$	

Fits of the model to data are shown in Figure 4. \mathcal{L} and AICc are the log-likelihood and the corrected Akaike information criterion value, respectively. In bold, we show maximum \mathcal{L} and minimum AICc reached by the models 1 and 3 with constant response. There are some unrealistic mutation rates given by model 2 (much bigger than 10^{-3} , highlighted as italic), and models 1 and 3 also led to slightly unrealistic mutation rates at the peptide Rev 65–82 (slightly bigger than 10^{-3}). Units for k_i and k_i' are day⁻¹ and μ_i is dimensionless (same for all tables below).

data from another patient. Patient CH131 had 6 CTL responses, and there was escape from at least 5 of these responses in 2 years since symptoms (Figure 5). One escape, Nef EEVGF-PVKPQV (Nef 64–74), occurred very early in infection, and two escapes, Env RQGYSPLSFQTLIPNPRG (Env 709–726) and Gag VKVIEEKAFSPEVIMFT (Gag 156–173), occurred late (Figure 5). In this patient, the pattern of escape followed the ranking of immunodominance of CTL responses (12): Nef64-specific CTLs were dominant at symptoms and drove earlier escape, while Env 709- and Gag156-specific CTLs arose later with escapes occurring later in infection (Figures 5A,B). However, there were apparently discrepancies such as two escapes in Tat epitopes (Tat DPWNHPGSQPKTACNNCY, that is Tat 9–26 and Tat FQKKGLGISY, that is Tat 38–47) occurred at the same time while CTL responses specific to these different epitopes were of different sizes (Figures 5A,B). Because escapes in these two Tat epitopes occurred rapidly and did not have two intermediate measurements of the mutant frequency, our following analysis was only restricted to escapes in three CTL epitopes: Nef64, Env709, Gag156 (Figures 5C,D).

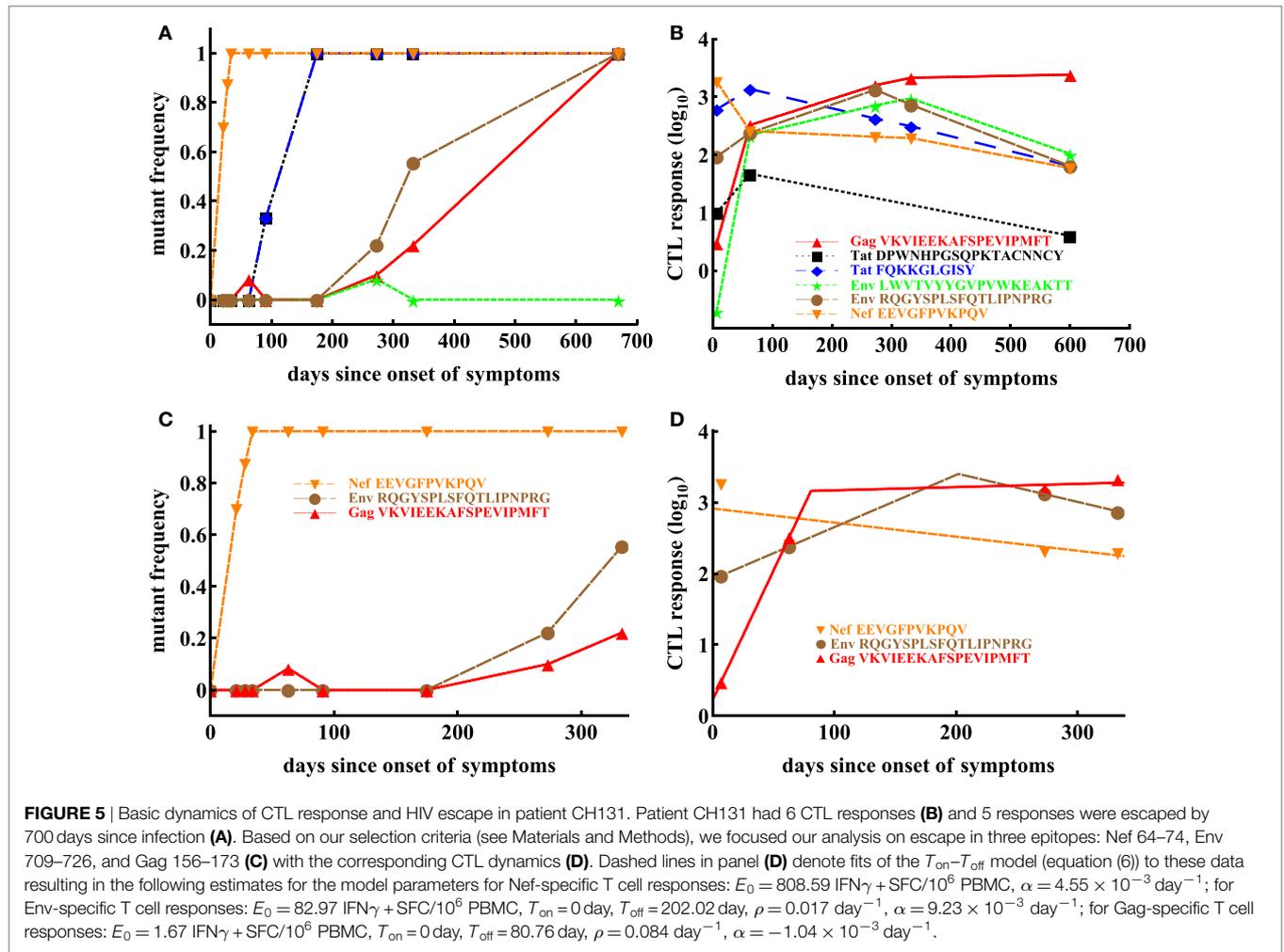
We thus fitted 3 different models of viral escape combined with 3 different models for the CTL dynamics to the data on viral escape (Figure 6). Importantly, as with the analysis of data from patient CH159, we found that including the data-driven CTL dynamics in the escape models did not improve the quality of the model fit to the escape data (Table 2). In contrast with the previous results, though, the assumption of the constant and time-variable killing efficacy (i.e., due to variation in the immune response magnitude) did not strongly impact the quality of the model fit as judged by the AIC or likelihood (Table 2). Importantly, however, models 1 and 3 gave nearly identical estimates of the CTL killing efficacy, suggesting that for data with good temporal resolution model estimates of the CTL killing efficacy (or by inference, escape rates) are not strongly dependent on the specific mechanisms used to describe escape (independent vs. concurrent escape). This observation also suggests that exclusion of the data on escape occurring at intermediate times after symptoms in Tat should not

influence the accuracy of estimation of the killing rates of CTLs specific to other epitopes in CH131.

Extending the observation made with the patient CH159 data, we found that model assuming sequential escape (model 2) could not accurately describe the dynamics of viral escape for biologically reasonable parameter values specifically for the third escape in Gag156 although this inability was significant only for a constant killing efficacy (Table 2). Allowing time-dependent killing efficacy resulted in small yet larger values for the mutation rate than that expected from basic calculations. Forcing the mutation rate μ_3 to be constrained ($\mu_3 \leq 10^{-3}$) significantly reduced the quality of the model fit to data (likelihood ratio test, $p \ll 0.001$). Furthermore, estimates for the CTL killing efficacy differed between model 2 and models 1 and 3 suggesting that model choice (sequential vs. concurrent) may indeed influence estimates of the killing efficacy.

3.3. No Difference in Predicted Killing Efficacy of CTLs, Specific to Different Epitopes

Our analyses, so far, demonstrated that several different mathematical models were capable of accurately describing the escape data, but this ability was dependent on the specific pathway of how escape mutants were generated and the assumption on whether data-driven CTL dynamics was included in the model. In cases, when a model was able to accurately describe the data, we generally observed different estimates for the parameters for HIV escape in different epitopes; for example, for the data in patient CH131 estimated CTL killing rate in the model 1 (independent escapes) with interpolated response differed nearly 100-fold between k_1' and k_3' (Table 2). Knowing which immune responses may be more efficient on a per cell basis in killing virus-infected cells may be beneficial for inducing such responses by vaccination. We, therefore, investigated how robust these differences in estimated per capita killing rates are. For that, we fitted mathematical models assuming equal killing efficacies to the data on escape.



As expected, reducing the number of fitted parameters led to fits of lower quality (as judged by the log-likelihood); however, this reduction in complexity of the model was favored by the AIC and in most cases by the likelihood ratio test (Tables S2 and S4 in Supplementary Material). Visually, the reduction in the quality of the model fit to data was also relatively small (Figures S2 and S4 in Supplementary Material). Thus, for these data, we found no strong evidence in the difference in the estimated per capita killing efficacy of the CTL response specific to different viral epitopes.

3.4. Identifying Conditions When the Model 2 (Sequential Escapes) Fails

In analysis of data from both patients, we found that model 2, describing sequential escape from CTL responses, was not able to accurately describe experimental data for biologically reasonable parameter values; these model fits predicted extremely high mutation rates (e.g., see Tables 1 and 2). Additional analyses demonstrated that fitting the models with constrained mutation rates, $\mu_i \leq 10^{-3}$ led to fits of significantly lower quality (based on increased AIC, results not shown).

A closer look at the experimental data for which model 2 provided unreasonably high mutation rates revealed that the trajectories of two subsequent escapes in the model 2 were too

close to each other, which naturally required a high mutation rate from one variant to another. Therefore, only when trajectories are separated in time mutation rate μ_2 is expected to be biologically reasonable. Indeed, by simulating virus dynamics using model for sequential escapes by varying model parameters, we found that CTL killing rate has the major impact on the time delay between two escapes (Figure 7). This analysis thus suggested that for the model 2 (sequential escape) to be consistent with the data, escapes from 2 responses must be separated in time by about 20–50 days.

4. DISCUSSION

CTL responses play a major role in HIV within-host evolution (45, 49). Recent studies suggested that a relative magnitude of the CTL response (relative immunodominance) plays an important role in determining the time of viral escape from T cell responses (11, 12). These previous studies, however, only utilized a maximum value of the CTL response early in infection, in general, within 50 days since the onset of symptoms, and thus impact of the kinetics of CTL response on the rate of virus escape remained undetermined. Furthermore, the pathways of HIV escape from CTL responses were not fully resolved as escapes occurring sequentially and

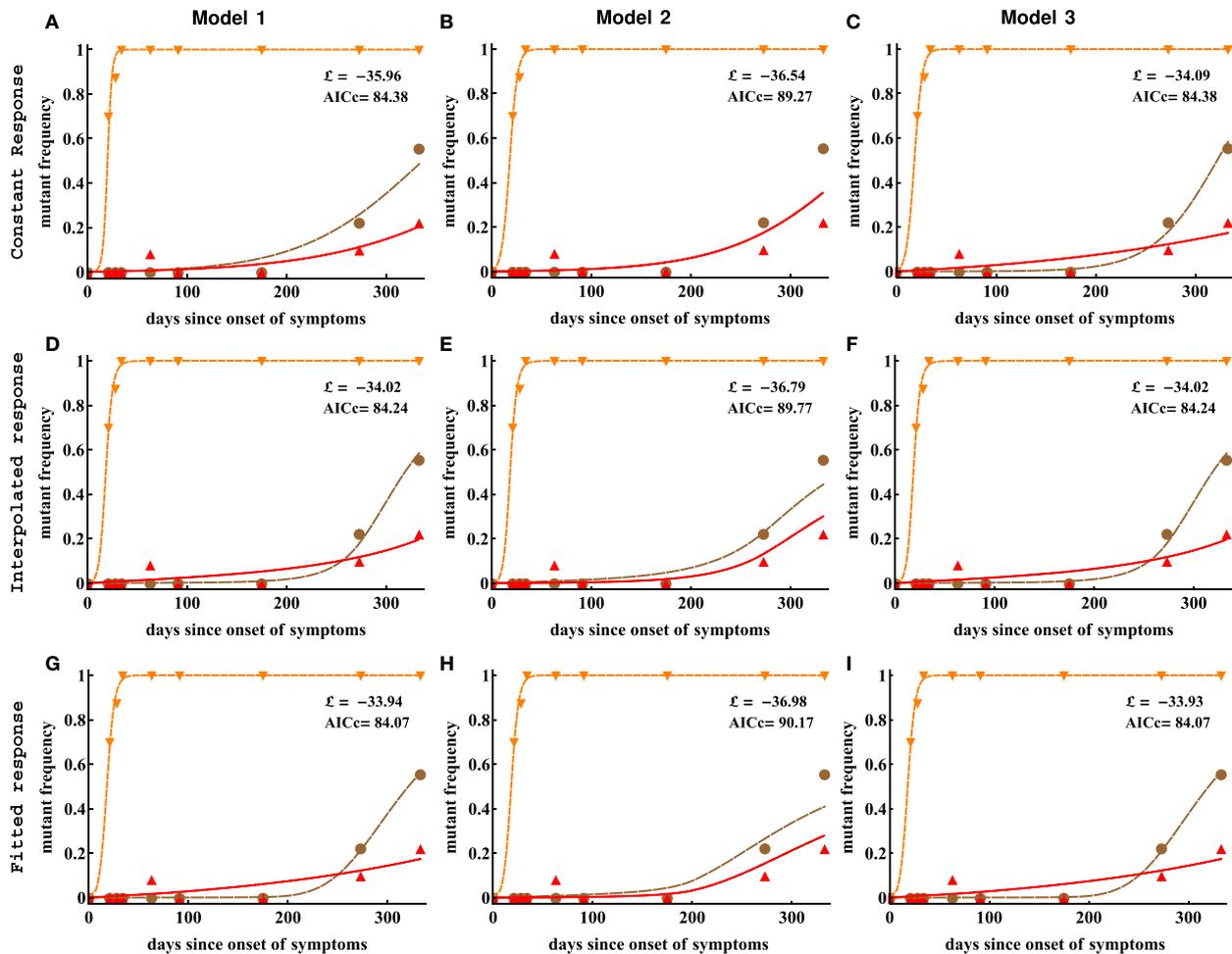


FIGURE 6 | Including CTL response dynamics did not improve model fits of HIV escape data in patient CH131. We fitted model 1 (independent escapes, panels (A,D,G)), model 2 (sequential escape, panels (B,E,H)), and model 3 (concurrent escape, panels (C,F,I)) to escape data in patient CH131 with different CTL response inputs (constant, interpolated, or fitted response). Adding data-derived time-dependent CTL response (interpolated or fitted response) does not improve the fitting results in most cases (Table 2). Notably, model 2 was unable to accurately describe late escape for biologically reasonable mutation rate μ_3 . Model parameters providing the best fit are given in Table 2.

concurrently have been proposed (21, 22, 25), and several previous studies assumed that escapes occur independently from each other (14, 15, 17). Here, by using experimental data on evolution of HIV sequences from acute infection into chronic phase and temporally resolved dynamics of HIV-specific CTL responses, we tested the hypothesis that CTL dynamics plays an important role in virus escape.

Perhaps, in contrast with our initial expectations (e.g., due to Ref. (11, 50)), we found that including experimentally measured dynamics of epitope-specific CTL responses did not lead to a better description of the kinetics of viral escape from T cells (e.g., in patient CH131, Table 2), or even reduced the quality of the model for viral escape fit to data (e.g., in patient CH159, Table 1). This was not because we assumed that killing of virus-infected cells was dependent on the absolute magnitude of epitope-specific CTL responses; assuming frequency-dependent killing, that is, when killing of infected cells expressing i th epitope was given by $k_i E_i(t) / \sum_{j=1}^n E_j(t)$ ($1 \leq i \leq n$), led to similar conclusions (results not shown). Because previous work suggested that kinetics of

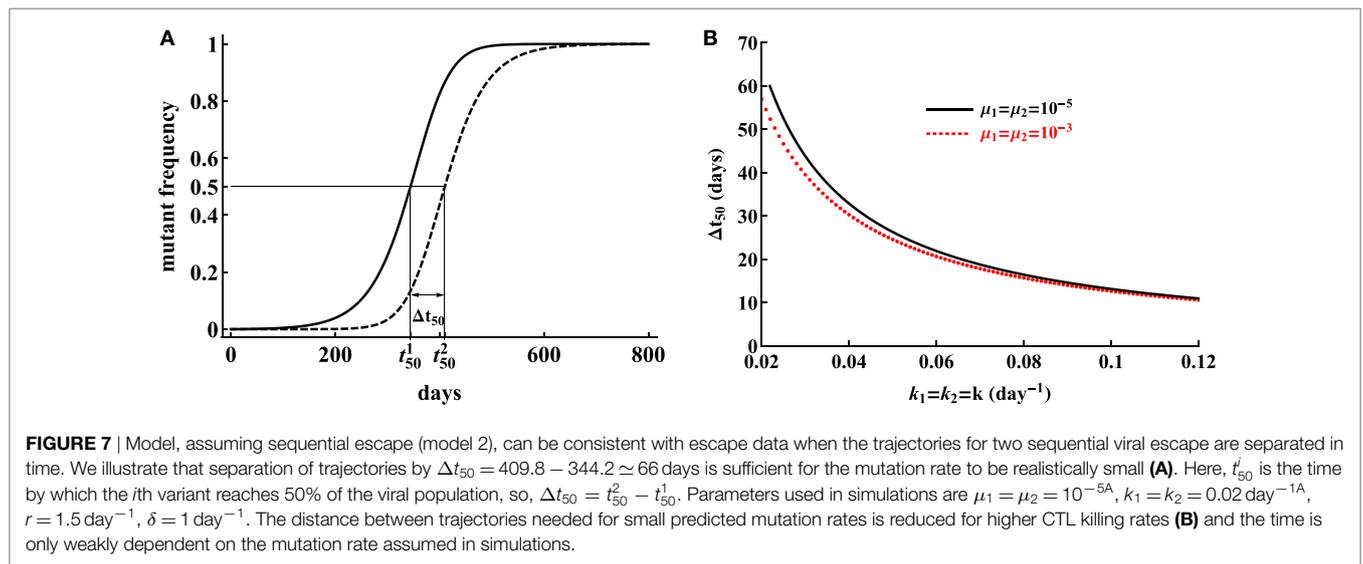
escape was independent of the specific mechanism of how CTLs suppress wild-type virus (e.g., killing of infected cells or virus production by infected cells) (16), we did not investigate non-lytic control of HIV by T cells. It is interesting that the lack of correlation between the rate of viral escape and CTL response magnitude was highlighted previously (17).

Reasons of why a model with time-variable CTL response did not describe experimental data better than a model with a constant response remain unclear but several hypotheses could be generated. First, frequency of sampling of the viral sequences may not be high enough to detect change in the speed at which mutant viruses accumulate in the population. Indeed, in mathematical models, CTL dynamics has a direct impact on the rate of escape (e.g., see equation (3)), and the observed changes in CTL densities may not be reflected in escape data if measurements are infrequent. Second, virus sequence data could simply be noisy. Because only handful of viral sequences were analyzed by the SGA/S, measurements of frequencies of viral variants have in general large expected error (e.g., Figure 2). Third, CTL dynamics

TABLE 2 | Parameters estimated by fitting different models of viral escape to escape data in patient CH131 assuming constant killing rates k_i (panels A–C), or time-varying killing rates due to interpolated CTL response (panels D–E) or CTL response in the $T_{on}-T_{off}$ model (panels G–I).

Peptide		Model 1		Model 2		Model 3	
Constant response	Nef 64–74	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k_i, i = 1, 2, 3$)	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k_i, i = 1, 2, 3$)	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k_i, i = 1, 2, 3$)
	Env 709–726	1.75×10^{-3}	0.25	1.72×10^{-3}	0.25	1.78×10^{-3}	0.25
	Gag 156–173	1.03×10^{-7}	0.031	3.18×10^{-5}	5.45×10^{-3}	9.91×10^{-7}	0.031
		$\mathcal{L} = -34.09, AICc = 84.38$		$\mathcal{L} = -36.54, AICc = 89.27$		$\mathcal{L} = -34.09, AICc = 84.38$	
Interpolated response	Nef 64–74	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k'_i, i = 1, 2, 3$)	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k'_i, i = 1, 2, 3$)	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k'_i, i = 1, 2, 3$)
	Env 709–726	4.33×10^{-4}	1.97×10^{-4}	3.95×10^{-4}	2.00×10^{-4}	4.30×10^{-4}	1.96×10^{-4}
	Gag 156–173	7.07×10^{-6}	3.01×10^{-5}	8.76×10^{-5}	1.56×10^{-5}	7.17×10^{-6}	3.00×10^{-5}
		$\mathcal{L} = -34.02, AICc = 84.24$		$\mathcal{L} = -36.79, AICc = 89.77$		$\mathcal{L} = -34.02, AICc = 84.24$	
Fitted response	Nef 64–74	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k'_i, i = 1, 2, 3$)	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k'_i, i = 1, 2, 3$)	Mutation rate ($\mu_i, i = 1, 2, 3$)	Killing rate ($k'_i, i = 1, 2, 3$)
	Env 709–726	3.25×10^{-3}	3.38×10^{-4}	2.99×10^{-3}	3.46×10^{-4}	3.16×10^{-3}	3.41×10^{-4}
	Gag 156–173	1.38×10^{-6}	2.59×10^{-5}	8.90×10^{-5}	1.05×10^{-5}	1.12×10^{-6}	2.66×10^{-5}
		$\mathcal{L} = -33.94, AICc = 84.07$		$\mathcal{L} = -36.98, AICc = 90.17$		$\mathcal{L} = -33.94, AICc = 84.07$	

Alternative models assume independent escape (model 1, panels A, D, and G), sequential escape (model 2, panels B, E, and H), or concurrent escape (model 3, panels C, F, and I). Fits of models 1 and 3 gave very close parameter values, but there were some unrealistic parameter values (italicized in the table) from fits of the model 2. \mathcal{L} and AICc give the log-likelihood score and the correlated Akaike information criterion value, respectively. Models 1 and 3 fit almost equally with three types of response inputs and the lowest \mathcal{L} and AICc are shown in bold.



in the blood may not reflect CTL dynamics in tissues such as secondary lymphoid organs (lymph nodes and spleen). While it is well known that T cells recirculate in the body (51), how quickly CTLs in the tissues migrate into the blood and then back to the tissues during HIV infection is not known. Finally, it is possible that the measured CTL responses were not the drivers of escape. While the ability of CTLs to recognize the wild-type virus and inability of the same CTLs to recognize mutant viruses is generally interpreted as evidence that these CTLs drove viral escape, such observations are correlational in nature, and thus cannot fully establish the causality of escape, at least in humans.

Our results may be interpreted as contradictory to several previous studies that found a strong correlation between the time

of viral escape (time when an escape variant reaches frequency of 50% in the viral population) and a relative magnitude of CTL response (relative or “vertical” immunodominance) (11, 12). However, our studies are not directly compatible because this previous work focused on the timing of escape while we primarily focused on the rate of viral escape. These two parameters are differently impacted by the CTL response (17) and may have different clinical importance. In our simple mathematical model (e.g., equation (3)), CTL response magnitude is expected to directly impact the rate at which an escape mutant accumulates in the population, independently of when this escape may occur. In contrast, timing of viral escape also depends on the mutation rate. Biologically, however, timing of escape may be more important

than the rate because it may be more beneficial to the patient if viral escape occurs 5 years after infection but rapidly as compared to slow escape in just 1 year. This conjecture clearly depends on the premise that HIV escapes from CTL responses are detrimental to patients.

In our analysis, we generally found that for well sampled data, the pathway of generation of escape mutants played a minor role in predicting overall CTL killing efficacy; assuming escapes that occur independently (model 1) or concurrently (model 3) gave nearly identical estimates of the CTL killing efficacy (e.g., **Tables 1** and **2**). In contrast, the model assuming sequential escape (model 2) often failed to accurately explain experimental data; this was due to some escapes co-occurring at nearly the same time, which obviously violated the model assumption of sequential escape. This inability of the sequential escape model to describe the data may be the result of the way we compared models to data: by using deterministic model approach and by ignoring recombination. Using deterministic model may be justified because, in acute infection, the effective population size of HIV may be sufficiently large and ignoring recombination may again be appropriate because very few cells in HIV infection are generally infected by 2 or more viruses (52, 53). However, further work is needed to demonstrate whether our conclusions regarding inability of sequential escape model to accurately explain some escape data is due to some of the assumptions made in the model by running stochastic simulations and by allowing some degree of recombination.

Many of our model fits predicted a high mutation rate for the first epitope to be escaped by the virus (e.g., **Table 2**). This model prediction could not be changed by shifting the experimental data to allow for more time to generate escape mutant; in part, this test failed because in the absence of epitope-specific T cells escape variants accumulate rather slowly mainly driven by mutations. It may indicate that immune pressure on the virus population starts much earlier than it is reflected in the blood, echoing our concerns of whether CTL dynamics in the blood is an accurate reflection of T cell response in lymphoid tissues. Currently, it is believed that lymphoid tissues and not the blood are the major places of interactions between the virus and CTLs (50, 54).

Our analysis further highlights the importance of choosing the appropriate statistical model for the analysis of the escape data—assuming normally distributed residuals, and therefore, using least squares approach, may not be appropriate for some escape data with very few sequences analyzed. Importantly, we confirm that the type of statistical model has an impact on the estimate of the escape rate (18).

We found that experimental data on HIV escape can be explained well if we assume identical per capita killing efficacy of CTLs, specific to different viral epitopes. This suggests that individual per capita killing rates not accurately estimated from these data. While it is possible that this result was the consequence of assuming additive killing of virus-infected cells by different

CTL responses, we currently do not have any *in vivo* data to support more complex killing terms.

Overall, analyses of data from two patients suggested that models assuming independent escape of HIV from different CTL responses (model 1) or models assuming concurrent escape from multiple CTL responses (model 3) fit the data well and provide very similar (often nearly identical) estimates for the killing efficacy of CTL response. Thus, for well sampled data, assumption of independent escapes may be sufficient to accurately estimate HIV escape rates. Also, the model with data-driven time-dependent CTL response (interpolated or fitted response input) did not improve the quality of the model fit to data, so, at present, it appears to be unnecessary to incorporate the experimentally measured CTL response dynamics in the model describing viral escapes. Yet, because our results were found only for two patients, whether similar conclusions will be reached in other studies/patients remains to be determined. Our analysis nevertheless demonstrates how mathematical modeling may help to quantify HIV evolution in presence of CTL responses and to highlight potential limitations with experimental measurements.

ETHICS STATEMENT

This paper uses experimental data obtained previously, and no new observations requiring patient consent or institutional review board approval have been performed.

AUTHOR CONTRIBUTIONS

YY and VG designed the study, contributed to analysis, interpolation of data, and simulation results, and wrote the paper. YY performed the simulations.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Nilu Goonetilleke and the Center for HIV/AIDS Vaccine Immunology (CHAVI) for data access.

FUNDING

This work was supported by the American Heart Association (AHA: 13SDG16960053) grant to VG and in part by the National Institutes of Health (NIH: R01-GM118553) grant to VG. Funding for open access to this research was provided by University of Tennessee's Open Publishing Support Fund.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://www.frontiersin.org/articles/10.3389/fimmu.2018.00140/full#supplementary-material>.

REFERENCES

1. WHO. *Global Health Observatory (GHO) Data*. Geneva: WHO (2016).
2. WHO. *Global HIV and AIDS Statistics*. Geneva: WHO (2016).
3. Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB. Virus-specific CD8⁺ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *J Virol* (1994) 68(9):6103–10.

4. Ndhlovu ZM, Kanya P, Mewalal N, Kloverpris HN, Nkosi T, Pretorius K, et al. Magnitude and kinetics of CD8⁺ T cell activation during hyperacute HIV infection impact viral set point. *Immunity* (2015) 43(3):591–604. doi:10.1016/j.immuni.2015.08.012
5. Streeck H, Jolin JS, Qi Y, Yassine-Diab B, Johnson RC, Kwon DS, et al. Human immunodeficiency virus type 1-specific CD8⁺ T-cell responses during primary infection are major determinants of the viral set point and loss of CD4⁺ T cells. *J Virol* (2009) 83(15):7641–8. doi:10.1128/jvi.00182-09
6. Hansen SG, Ford JC, Lewis MS, Ventura AB, Hughes CM, Coyne-Johnson L, et al. Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature* (2011) 473(7348):523–7. doi:10.1038/nature10003
7. Walker BD, Ahmed R, Plotkin S. Moving ahead an HIV vaccine: use both arms to beat HIV. *Nat Med* (2011) 17(10):1194–5. doi:10.1038/nm.2529
8. Barouch DH, Kunstman J, Kuroda MJ, Schmitz JE, Santra S, Peyrel FW, et al. Eventual AIDS vaccine failure in a rhesus monkey by viral escape from cytotoxic T lymphocytes. *Nature* (2002) 415(6869):335–9. doi:10.1038/415335a
9. Goulder PJR, Watkins DI. HIV and SIV CTL escape: implications for vaccine design. *Nat Rev Immunol* (2004) 4(8):630–40. doi:10.1038/nri1417
10. Rolland M, Tovanabutra S, Decamp AC, Frahm N, Gilbert PB, Sanders-Buell E, et al. Genetic impact of vaccination on breakthrough HIV-1 sequences from the STEP trial. *Nat Med* (2011) 17(3):366–71. doi:10.1038/nm.2316
11. Barton JP, Goonetilleke N, Butler TC, Walker BD, McMichael AJ, Chakraborty AK. Relative rate and location of intra-host HIV evolution to evade cellular immunity are predictable. *Nat Commun* (2016) 7:11660. doi:10.1038/ncomms11660
12. Liu MK, Hawkins N, Ritchie AJ, Ganusov VV, Whale V, Brackenridge S, et al. Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. *J Clin Invest* (2013) 123(1):380–93. doi:10.1172/JCI65330
13. Goonetilleke N, Liu MK, Salazar-Gonzalez JF, Ferrari G, Giorgi E, Ganusov VV, et al. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *J Exp Med* (2009) 206(6):1253–72. doi:10.1084/jem.20090365
14. Asquith B, Edwards CTT, Lipsitch M, McLean AR. Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells in vivo. *PLoS Biol* (2006) 4(4):e90. doi:10.1371/journal.pbio.0040090
15. Fernandez CS, Stratov I, Rose RD, Walsh K, Dale CJ, Smith MZ, et al. Rapid viral escape at an immunodominant simian-human immunodeficiency virus cytotoxic T-lymphocyte epitope exacts a dramatic fitness cost. *J Virol* (2005) 79(9):5721–31. doi:10.1128/jvi.79.9.5721-5731.2005
16. Ganusov VV, De Boer RJ. Estimating costs and benefits of CTL escape mutations in SIV/HIV infection. *PLoS Comput Biol* (2006) 2(3):e24. doi:10.1371/journal.pcbi.0020024
17. Ganusov VV, Goonetilleke N, Liu MKP, Ferrari G, Shaw GM, McMichael AJ, et al. Fitness costs and diversity of the cytotoxic T lymphocyte (CTL) response determine the rate of CTL escape during acute and chronic phases of HIV infection. *J Virol* (2011) 85(20):10518–28. doi:10.1128/jvi.00655-11
18. Ganusov VV, Neher RA, Perelson AS. Mathematical modeling of escape of HIV from cytotoxic T lymphocyte responses. *J Stat Mech* (2013) 2013(01):01010. doi:10.1088/1742-5468/2013/01/p01010
19. Mandl JN, Regoes RR, Garber DA, Feinberg MB. Estimating the effectiveness of simian immunodeficiency virus-specific CD8⁺ T cells from the dynamics of viral immune escape. *J Virol* (2007) 81(21):11982–91. doi:10.1128/jvi.00946-07
20. Petravic J, Loh L, Kent SJ, Davenport MP. Cd4⁺ target cell availability determines the dynamics of immune escape and reversion in vivo. *J Virol* (2008) 82(8):4091–101. doi:10.1128/jvi.02552-07
21. Pandit A, De Boer RJ. Reliable reconstruction of HIV-1 whole genome haplotypes reveals clonal interference and genetic hitchhiking among immune escape variants. *Retrovirology* (2014) 11(1):56. doi:10.1186/1742-4690-11-56
22. Levisyang S, Ganusov VV. Broad CTL response in early HIV infection drives multiple concurrent CTL escapes. *PLoS Comput Biol* (2015) 11(10):e1004492. doi:10.1371/journal.pcbi.1004492
23. Garcia V, Feldman MW, Regoes RR. Investigating the consequences of interference between multiple CD8⁺ T cell escape mutations in early HIV infection. *PLoS Comput Biol* (2016) 12(2):e1004721. doi:10.1371/journal.pcbi.1004721
24. Garcia V, Regoes RR. The effect of interference on the CD8⁺ T cell escape rates in HIV. *Front Immunol* (2015) 5:661. doi:10.3389/fimmu.2014.00661
25. Kessinger TA, Perelson AS, Neher RA. Inferring HIV escape rates from multi-locus genotype data. *Front Immunol* (2013) 4:252. doi:10.3389/fimmu.2013.00252
26. Van Deutekom HWM, Wijnker G, De Boer RJ. The rate of immune escape vanishes when multiple immune responses control an HIV infection. *J Immunol* (2013) 191(6):3277–86. doi:10.4049/jimmunol.1300962
27. Ganusov VV. Time intervals in sequence sampling, not data modifications, have a major impact on estimates of HIV escape rates. *bioRxiv* (2017). doi:10.1101/221812
28. Perelson AS. Modelling viral and immune system dynamics. *Nat Rev Immunol* (2002) 2(1):28–36. doi:10.1038/nri700
29. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* (1996) 271(5255):1582–6. doi:10.1126/science.271.5255.1582
30. Ramratnam B, Bonhoeffer S, Binley J, Hurley A, Zhang L, Mittler JE, et al. Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet* (1999) 354(9192):1782–5. doi:10.1016/s0140-6736(99)02035-8
31. Song H, Pavlicek JW, Cai F, Bhattacharya T, Li H, Iyer SS, et al. Impact of immune escape mutations on HIV-1 fitness in the context of the cognate transmitted/founder genome. *Retrovirology* (2012) 9(1):89. doi:10.1186/1742-4690-9-89
32. Wright JK, Brumme ZL, Carlson JM, Heckerman D, Kadie CM, Brumme CJ, et al. Gag-protease-mediated replication capacity in HIV-1 subtype c chronic infection: associations with HLA type and clinical parameters. *J Virol* (2010) 84(20):10820–31. doi:10.1128/jvi.01084-10
33. Turnbull EL, Wong M, Wang S, Wei X, Jones NA, Conrod KE, et al. Kinetics of expansion of epitope-specific T cell responses during primary HIV-1 infection. *J Immunol* (2009) 182(11):7131–45. doi:10.4049/jimmunol.0803658
34. Mansky LM, Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol* (1995) 69(8):5087–94.
35. Levisyang S. Computational inference methods for selective sweeps arising in acute HIV infection. *Genetics* (2013) 194(3):737–52. doi:10.1534/genetics.113.150862
36. Althaus CL, De Boer RJ. Dynamics of immune escape during HIV/SIV infection. *PLoS Comput Biol* (2008) 4(7):e1000103. doi:10.1371/journal.pcbi.1000103
37. Ganusov VV, Barber DL, De Boer RJ. Killing of targets by CD8⁺ T cells in the mouse spleen follows the law of mass action. *PLoS One* (2011) 6(1):e15959. doi:10.1371/journal.pone.0015959
38. De Boer RJ, Oprea M, Antia R, Murali-Krishna K, Ahmed R, Perelson AS. Recruitment times, proliferation, and apoptosis rates during the CD8⁺ T-cell response to lymphocytic choriomeningitis virus. *J Virol* (2001) 75(22):10663–9. doi:10.1128/jvi.75.22.10663-10669.2001
39. De Boer RJ, Homann D, Perelson AS. Different dynamics of CD4⁺ and CD8⁺ T cell responses during and after acute lymphocytic choriomeningitis virus infection. *J Immunol* (2003) 171:3928–35. doi:10.4049/jimmunol.171.8.3928
40. Yang YD, Ganusov VV. Defining kinetic properties of HIV-specific CD8⁺ T-cell responses in acute infection. *bioRxiv* (2017). doi:10.1101/158683
41. Bates DM, Watts DG. *Nonlinear Regression Analysis and Its Applications*. New York: Wiley (1988).
42. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York: Springer (2002).
43. Efron B, Tibshirani R. *An Introduction to the Bootstrap*. New York: Chapman & Hall (1993).
44. Brown L, Cai T, DasGupta A, Agresti A, Coull B, Casella G, et al. Interval estimation for a binomial proportion. *Stat Sci* (2001) 16(2):101–33. doi:10.1214/ss/1009213285
45. McMichael AJ, Borrow P, Tomaras GD, Goonetilleke N, Haynes BF. The immune response during acute HIV-1 infection: clues for vaccine development. *Nat Rev Immunol* (2010) 10(1):11–23. doi:10.1038/nri2674
46. Love TMT, Thurston SW, Keefer MC, Dewhurst S, Lee HY. Mathematical modeling of ultradeep sequencing data reveals that acute CD8⁺ T-lymphocyte responses exert strong selective pressure in simian immunodeficiency virus-infected macaques but still fail to clear founder epitope sequences. *J Virol* (2010) 84(11):5802–14. doi:10.1128/JVI.00117-10
47. Martyushev AP, Petravic J, Grimm AJ, Alinejad-Rokny H, Gooneratne SL, Reece JC, et al. Epitope-specific CD8⁺ T cell kinetics rather than viral variability determine the timing of immune escape in simian immunodeficiency virus infection. *J Immunol* (2015) 194(9):4112–21. doi:10.4049/jimmunol.1400793

48. Kijak GH, Sanders-Buell E, Chenine A-L, Eller MA, Goonetilleke N, Thomas R, et al. Rare HIV-1 transmitted/founder lineages identified by deep viral sequencing contribute to rapid shifts in dominant quasispecies during acute and early infection. *PLoS Pathog* (2017) 13:e1006510. doi:10.1371/journal.ppat.1006510
49. McMichael AJ, Phillips RE. Escape of human immunodeficiency virus from immune control. *Annu Rev Immunol* (1997) 15:271–96. doi:10.1146/annurev.immunol.15.1.271
50. Li Q, Skinner PJ, Ha SJ, Duan L, Mattila TL, Hage A, et al. Visualizing antigen-specific and infected cells in situ predicts outcomes in early viral infection. *Science* (2009) 323(5922):1726–9. doi:10.1126/science.1168676
51. Ganusov VV, Auerbach J. Mathematical modeling reveals kinetics of lymphocyte recirculation in the whole organism. *PLoS Comput Biol* (2014) 10:e1003586. doi:10.1371/journal.pcbi.1003586
52. Josefsson L, King MS, Makitalo B, Brännström J, Shao W, Maldarelli F, et al. Majority of CD4⁺ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. *Proc Natl Acad Sci U S A* (2011) 108(27):11199–204. doi:10.1073/pnas.1107729108
53. Josefsson L, Palmer S, Faria NR, Lemey P, Casazza J, Ambrozak D, et al. Single cell analysis of lymph node tissue from HIV-1 infected patients reveals that the majority of CD4⁺ T-cells contain one HIV-1 DNA molecule. *PLoS Pathog* (2013) 9:e1003432. doi:10.1371/journal.ppat.1003432
54. Haase AT. Population biology of HIV-1 infection: viral and CD4⁺ T cell demographics and dynamics in lymphatic tissues. *Annu Rev Immunol* (1999) 17:625–56. doi:10.1146/annurev.immunol.17.1.625

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Yang and Ganusov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner 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.