

# Increased oral Epstein-Barr virus shedding with HIV-1 co-infection is due to a combination of B cell activation and impaired cellular immune control

Catherine M. Byrne<sup>1,2,3\*</sup>, Christine Johnston<sup>4,5</sup>, Jackson Orem<sup>5,6</sup>, Fred Okuku<sup>6</sup>, Meei-Li Huang<sup>5,7</sup>, Stacy Selke<sup>7</sup>, Anna Wald<sup>4,5,7,8</sup>, Lawrence Corey<sup>4,5,7</sup>, Joshua T. Schiffer<sup>4,5,9</sup>, Corey Casper<sup>4,5,10</sup>, Daniel Coombs<sup>3,11‡</sup>, Soren Gantt<sup>1,3,12‡</sup>

**1** Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

**2** Institute of Applied Mathematics, University of British Columbia, Vancouver, British Columbia, Canada

**3** British Columbia Children's Hospital Research Institute, Vancouver, British Columbia, Canada

**4** Department of Medicine, University of Washington, Seattle, Washington, United States

**5** Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States

**6** Uganda Cancer Institute, Kampala, Uganda

**7** Department of Laboratory Medicine, University of Washington, Seattle, Washington, United States

**8** Department of Epidemiology, University of Washington, Seattle, Washington, United States

**9** Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States

**10** Infectious Disease research Institute, Seattle, Washington, United States

**11** Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada

**12** Department of Pediatrics, University of British Columbia, Vancouver, British Columbia, Canada

‡These authors contributed equally to this work.

\* cbyrne@bccr.ubc.ca

## Abstract

Epstein-Barr virus (EBV) infection is transmitted by saliva and is a major cause of cancer in people living with HIV/AIDS as well as in the general population. To better understand the determinants of oral EBV shedding we evaluated the frequency and quantity of detectable EBV in the saliva in a prospective cohort study of 85 adults in Uganda, half of whom were co-infected with HIV-1. Participants were not receiving antiviral medications, and those with HIV-1 co-infection had a CD4+ T cell count >200 cells/mm<sup>3</sup>. Daily, self-collected oral swabs were collected over a 4-week period. Compared with HIV-1 uninfected participants, co-infected participants had an increased frequency of oral EBV shedding (IRR=1.27, 95% CI=1.10-1.47). To explain why EBV oral shedding is greater in HIV-1 co-infected participants, we developed a stochastic, mechanistic mathematical model that describes the dynamics of EBV, infected cells, and antiviral cellular immune responses within the tonsillar epithelium, and examined

parameter-specific differences between individuals of different HIV-1 infection statuses. We fit the model to our observational data using Approximate Bayesian Computation. After fitting, model simulations showed high fidelity to daily oral shedding time-courses and matched key summary statistics. Examination of the model revealed that higher EBV loads in saliva are driven by B cell activation causing EBV lytic replication in the tonsils, in combination with a less effective EBV-specific cellular immune response. Thus, both these factors contribute to higher and more frequent EBV shedding in HIV-1 co-infected individuals compared to HIV-1 uninfected individuals. These conclusions were further validated by modelling daily oral EBV shedding in a 26-participant North American cohort. Our results provide insights into the determinants of EBV shedding and implicate B cell activation to be a potential therapeutic target to reduce EBV replication in HIV-1 co-infected individuals at high risk for EBV-related malignancies.

## Author summary

Epstein-Barr virus (EBV) is a ubiquitous infection worldwide. Infection with EBV is associated with the development of several kinds of cancer, including B cell lymphoma and nasopharyngeal carcinoma. Rates of EBV replication and disease are higher in individuals who are also infected with HIV-1. HIV-1 infection is associated with increased B cell activation, which is known to induce EBV reactivation, as well as immunodeficiency resulting from loss of T cells. However, whether these factors contribute to higher rates of EBV replication during co-infection, and by how much, was unknown. We analysed oral EBV shedding data in a cohort of adults from Uganda that were chronically infected with EBV. We found that participants that were HIV-1 infected were much more likely to have detectable quantities of EBV in their saliva. Also, when detected, the quantity of EBV present in the saliva was usually higher in HIV-1 infected participants. To better understand these findings, we developed a mathematical model to describe the dynamics of EBV, EBV-infected cells, and the cellular immune response within the tonsils. By rigorously matching our model to our participant data, we determined that high EBV loads in saliva are caused by high rates of infected B cell activation, as well as worse cellular immune control of EBV infection. These results provide an explanation of the impact of HIV-1 on EBV infection. Further, they suggest that strategies that suppress B cell activation may prevent EBV-related malignancy in people who are also infected with HIV-1.

## Introduction

Epstein-Barr virus (EBV) infection is associated with the development of approximately 200,000 malignancies per year, including B cell lymphomas and nasopharyngeal carcinoma [1]. The risk of EBV-associated malignancies is significantly higher in people co-infected with HIV-1. For example, risks of non-Hodgkin lymphoma in the U.S., an AIDS-defining cancer, are 10-fold higher in HIV-1 infected individuals than the general population [2]. Individuals with EBV/HIV-1 co-infection tend to have higher EBV viral loads in saliva and blood [3–5]. Uncovering the mechanisms by which HIV-1 impairs control of EBV infection may provide clues relevant to prevention of EBV-related disease as well as insights into basic EBV pathobiology.

EBV is primarily transmitted via saliva, and infection is nearly universal, occurring during early childhood in developing countries and by young adulthood in highly developed countries [6–8]. During primary infection, it is thought that EBV first infects oral epithelial cells overlying the lymphoid tissue known as Waldeyer’s ring [9]. This area consists of the tubal, adenoid, palatine and lingual tonsils [9]. Infected epithelial

cells produce large numbers of infectious virions [10], facilitating latent infection of naïve B cells in the underlying lymphoid tissue. EBV drives these naïve B cells to mature into resting memory B cells and circulate throughout the body, through the expression of only a small number of latent gene products [11,12]. Viral shedding is highest during primary EBV infection but remains frequent throughout chronic infection [5]. During chronic infection, B cells latently infected with EBV can return to Waldeyer's ring, encounter cognate antigen, and become activated to mature into plasma cells, triggering lytic reactivation and production of infectious virions [13–15]. This process initiates a new round of epithelial infection in the tonsils and viral shedding in the saliva. Increased EBV shedding with HIV-1 co-infection may be due to more frequent reactivation of EBV-infected B cells, and thus increased viral seeding of oral tissue, and/or impaired T cell-mediated immune control of EBV replication, prolonging or inhibiting the clearance of infected epithelial cells [16–18].

The dynamics of chronic herpesvirus infections in humans have been elucidated by longitudinally sampling mucosal surfaces and blood, revealing the patterns of latency, reactivation, and dissemination, as well as giving insight into viral pathogenesis and host-pathogen interactions [19,20]. While several studies have examined EBV mucosal shedding patterns in HIV-1-infected persons [5,21–25], the majority have been in the setting of advanced HIV-1 infection or in persons receiving highly active antiretroviral therapy (HAART) [23–25]. The mechanisms by which co-infection with HIV-1 significantly increases viral loads in saliva have not been explained in part due to the requirement for frequent quantitative measures of shedding amenable to mathematical modelling [23–25]. Previous mathematical models have examined the within-host dynamics of EBV infection, developing quantitative descriptions of how infected cells, B cells, antibodies, virus, and T cells interact, and determining biologically relevant parameter values for the rates at which infection dynamics occur in the tonsillar epithelium [10,26–30]. However, none have examined the differences between EBV shedding in HIV-1 co-infected and HIV-1 uninfected individuals. In this study, we analysed rich oral EBV shedding data from a cohort of Ugandan adults with or without HIV-1 co-infection and constructed a mechanistic mathematical model to identify the determinants of EBV replication.

## Results

### HIV-1 infection is associated with increased frequency and quantity of oral EBV shedding

Among this cohort of 85 participants, a total of 2264 daily oral swabs were collected, with a median of 29 swabs per participant (range 1-32). 43 (51%) participants were HIV-1 seropositive. Additional details of the cohort have been previously published [31], and other data from the cohort, including EBV loads in genital swabs and plasma samples, can be found in the Supporting Information. We began our analysis by first examining whether HIV-1 infection status affects the frequency of EBV detection in the saliva. We found that HIV-1 infection was significantly associated with EBV detection in oral swabs, increasing the frequency of observation 1.27-fold (CI = 1.10-1.47, p-value = 0.001, Fig 1A) and increasing the median quantity of EBV detected in oral swabs by 1.61 log<sub>10</sub> genome copies (CI = 1.29-1.93, p-value <0.001, Fig 1B). We also saw large variability in participant viral loads over time (Fig 1C). In HIV-1 uninfected participants, the median percentage of swabs positive for EBV was 86% (range 0-100%, interquartile range (IQR) 33%) while in HIV-1 co-infected participants the median percentage of swabs positive for EBV was 100% (range 27-100%, IQR 0%). Of swabs that tested positive for EBV, viral loads varied over time by a median of 3.49 orders of

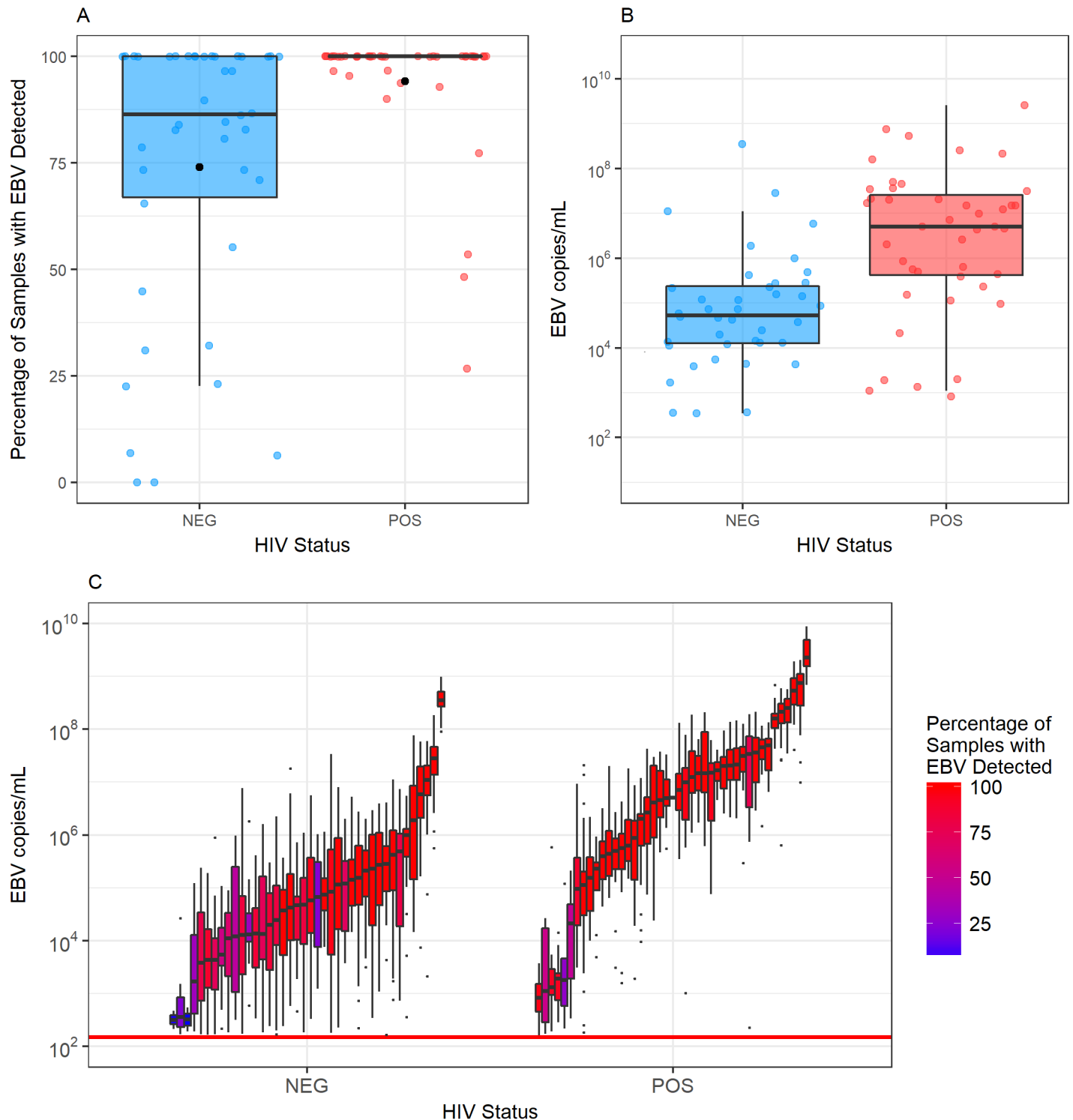
magnitude (range 0.34-5.27, IQR 1.32) within individual HIV-1 uninfected participants and a median of 2.30 orders of magnitude (range 0.95-5.93, IQR 1.15) within individual HIV-1 co-infected participants.

## The effect of CD4+ T count and HIV-1 plasma RNA levels on oral EBV shedding in HIV-1 co-infected individuals

We next analysed how CD4+ T cell counts and HIV-1 plasma RNA levels impacted the frequency of EBV detection in oral swabs of HIV-1 infected participants (Table 1). CD4+ T cell count significantly affected the frequency of EBV detection in oral swabs. Each additional 100-cell increase in CD4+ T cells/mm<sup>3</sup> was associated with a 5% reduction in the odds of EBV detection. This is consistent with cell-mediated immunity conferring partial but incomplete control of EBV replication. We found no significant association between HIV-1 viral load in plasma and the frequency of EBV detection in saliva.

**Table 1. Effects of plasma HIV-1 load and CD4+ T cell count on the frequency of EBV detection.** The effects of each 100-cell increase in CD4+ T cells/mm<sup>3</sup> and each log<sub>10</sub> increase in HIV-1 RNA on the incidence rate ratio (IRR) of EBV detection, their confidence intervals (CI) and their p-values are shown.

Trait	IRR	95% CI	p-value
HIV-1 RNA	1.04	0.98-1.12	0.201
CD4+ T cell	0.95	0.92-0.99	0.008



**Fig 1. Impact of HIV-1 infection on EBV detection** A. Percentages of saliva samples that tested positive for EBV for each participant stratified by HIV-1 status. Black dots indicate the percentage of samples that tested positive for EBV when pooling participant samples. B. Median EBV viral loads/ml in oral swabs testing positive for EBV for each participant stratified by HIV-1 status. C. Distributions of participants' oral swab viral loads. Each box and whisker represents the viral loads of EBV-positive oral swabs for an individual participant. The percent of oral swabs that tested positive for EBV for each participant is indicated by the colour of the box. The red horizontal line represents the threshold of EBV detection (150 EBV copies/ml).

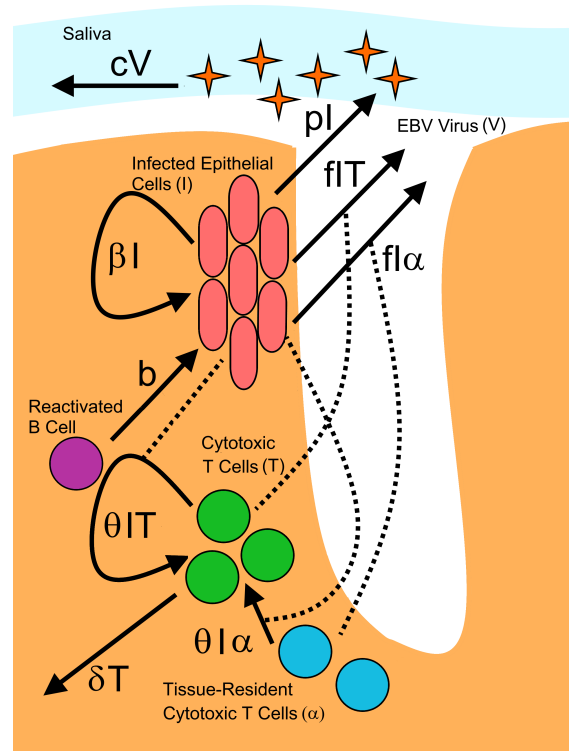
## Mathematical Model of EBV Shedding in the Tonsils

To obtain mechanistic insights into chronic oral EBV shedding and to better understand the drivers of lytic replication and transmission, we constructed a mathematical model that captures the relevant anatomic, virologic and immunologic features of oral EBV infection. In chronically infected individuals, EBV is shed in all areas of Waldeyer's ring including the palatine, lingual, tubal tonsils and adenoids [9]. Most of the tonsillar area is composed of stratified squamous epithelium or ciliated pseudostratified columnar epithelium, arranged into a series of crypts or folds, allowing for a large surface area [32]. The epithelium is often only one cell thick where EBV can transcytose to reach the underlying lymphoid tissue where B cells and germinal centres are found [33]. The palatine tonsils have an estimated surface area of 295 cm<sup>2</sup>, arranged into approximately 20 crypts [32], while the lingual tonsil area is composed of 35-100 crypts, and the adenoids are composed of a series of folds in lymphoid tissue [34]. By estimating that each palatine tonsil is approximately  $\frac{1}{12}$  of the entire surface area of Waldeyer's ring, a series of 240 crypts, each serving as sites where EBV infection may occur, can represent the entire region. We assumed that the dynamics of each crypt are independent from each other and explicitly modelled the dynamics of infected epithelial cells, the immune response, and viral load within each crypt.

The dynamics within each crypt are shown in Fig 2. All parameters were assumed to remain constant over the course of study, with some varying between participants to account for differences in EBV shedding and HIV-1 infection status. We assumed that latently infected B cells reactivate and infect the epithelium at a constant rate  $b$ . Infected epithelial cells,  $I$ , infect other epithelial cells through cell-to-cell contact at a constant per-capita rate  $\beta$ . As the total number of cells within a single crypt is large in comparison to the expected number that can become infected with EBV, we assumed that target cell number is not limiting. Epithelial infection causes the recruitment and proliferation of EBV-specific cytotoxic T cells,  $T$ , at a per-capita rate  $\theta I$ . Cytotoxic T cells kill infected epithelial cells following the law of mass action at a rate  $fIT$ . We assumed that cytotoxic T cells die or leave the tonsils at a per-capita rate  $\delta$ . Independent of infection, we assumed a constant number of EBV-specific cytotoxic T cells,  $\alpha$ , are tissue-resident. Like  $T$ , these cells can kill infected epithelial cells and stimulate the proliferation of new EBV-specific cytotoxic T cells; however, while population  $T$  leaves the system over time, these tissue-resident T cells remain within the tissue and do not recirculate. This means there are always immune cells present to respond to new infection and tissue is never entirely unprotected, allowing for faster control of infection. EBV virions,  $V$ , are produced by infected epithelial cells, enter saliva for at a per-capita rate  $p$  and are cleared at a per-capita rate  $c$ . In this model, we assumed the main contributors to virus in the saliva are infected epithelial cells and thus we did not directly model virions produced by infected B cells [10]. With the propagation of EBV infection shown to be 800-fold more efficient through cell-to-cell contact rather than through free virus, we also chose to assume all new epithelial cell infection is caused by cell-to-cell contact [10, 35].

The concentration of EBV detected in the saliva of participants was highly variable, and frequently undetectable. Therefore, we chose to implement our model in a stochastic framework in order to capture these traits (Methods). Our model assumptions were used to build a chemical master equation system [36] that describes all system reactions within a single crypt, as follows:

$$\begin{aligned}
 I &\rightarrow I + 1 && \text{with rate } b + \beta I && (1) \\
 I &\rightarrow I - 1 && \text{with rate } fI(T + \alpha) && (2) \\
 T &\rightarrow T + 1 && \text{with rate } \theta I(T + \alpha) && (3) \\
 T &\rightarrow T - 1 && \text{with rate } \delta T && (4) \\
 V &\rightarrow V + 1 && \text{with rate } pI && (5) \\
 V &\rightarrow V - 1 && \text{with rate } cV. && (6)
 \end{aligned}$$



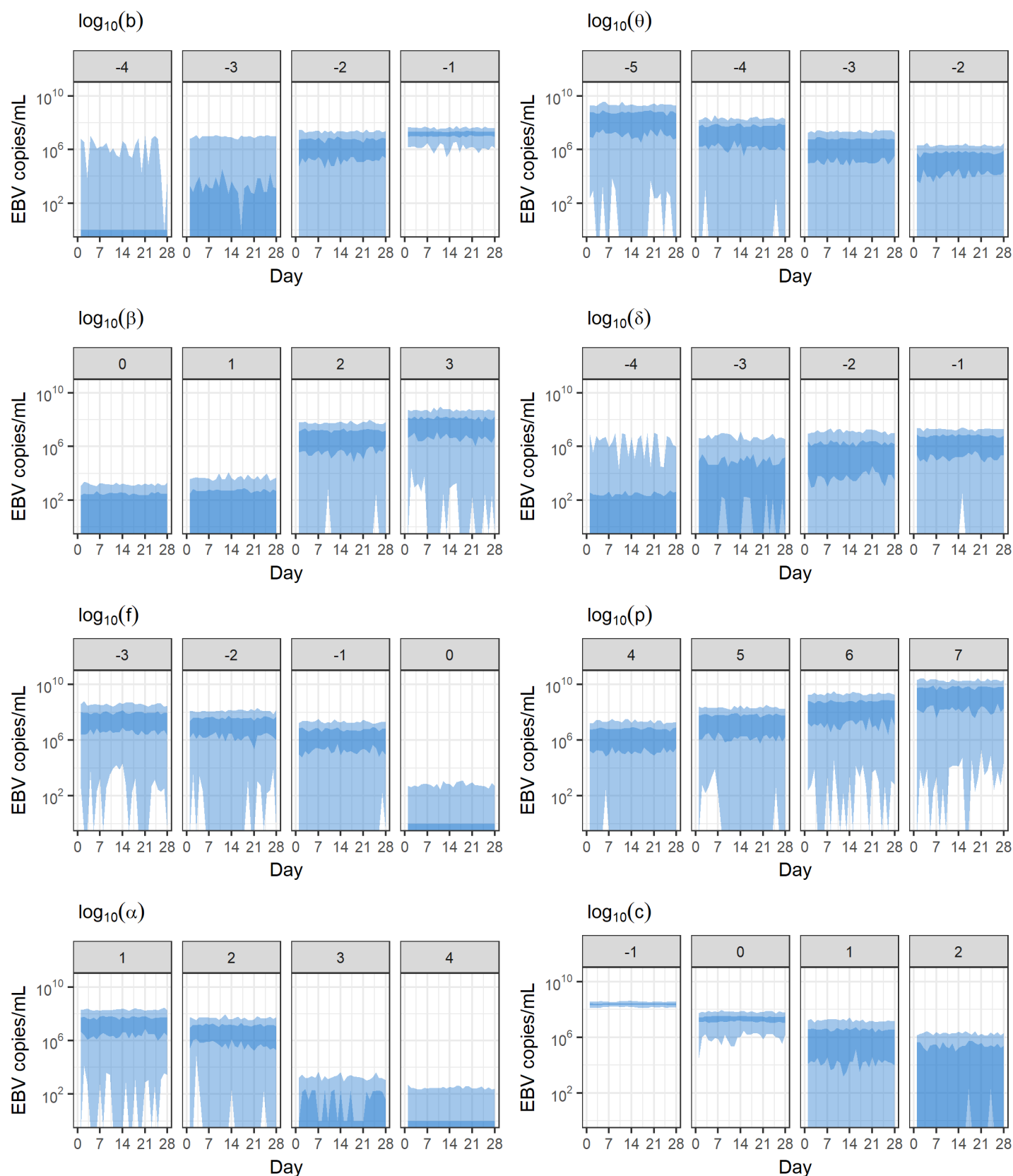
**Fig 2. Description of single crypt dynamics.** Waldeyer's ring is represented as a series of 240 individual crypts in which infection dynamics occur. Within each crypt, the population dynamics of infected epithelial cells (I), cytotoxic T cells (T) and EBV (V) are described. The salivary viral load is represented by the total viral load aggregated across all crypts.

## Basic model analysis

To understand how the different parameters of our mathematical model affect simulated EBV shedding in saliva, we first conducted a literature review to find previously estimated or measured values of our parameters (Table 2). We then performed a simple univariate analysis to evaluate how changes in these parameters influence the characteristics of model simulations. The starting set of parameters was chosen based on current estimates in the literature; we then performed multiple model simulations, varying each parameter in turn over 4 orders of magnitude to observe how sensitive the model is to these changes (Fig 3). Since we observed large variability in the viral loads of different cohort participants (Fig 1C), our aim was to select a fixed set of parameters

that would allow for the generation of a wide range of viral loads. Increases in parameter  $b$  or  $\delta$  increased the median viral load of our model simulations; however, the maximum viral load remained stable. In contrast,  $\beta$  appeared to control the maximum viral load reached in simulations, with higher  $\beta$  values causing higher viral loads and larger variance in the viral load. Parameters  $f$ ,  $\alpha$ , and  $\theta$ , all governing the strength of the cellular immune response, appeared to have similar effects on the simulation, all causing comparable decreases in viral loads as the magnitude of the parameter increased. Intuitively, increases in parameter  $p$ , governing the viral burst size, increased the viral loads in simulations. Low viral loads and periods of viral extinction (as seen in many HIV-1 uninfected participants) were uncommon with higher values of  $p$ . Thus, these higher values of  $p$  are improbable. Lastly, increases in parameter  $c$  reduced the viral loads and increased the variance. From this analysis, we chose our fixed parameter values to be  $\beta = 50$ ,  $f = 0.1$ ,  $\alpha = 200$ ,  $\delta = 0.1$ ,  $p = 10^4$ , and  $c = 6$ , which generally agreed with published estimates. We allowed  $b$  and  $\theta$  to vary during fitting as these two parameters are the most likely to be affected by HIV-1 infection status.





**Fig 3. Evaluation of individual model parameters.** Each parameter of the model was examined to determine how it influences the simulation of oral viral shedding over time. Parameters were varied one at a time over 4 orders of magnitude while keeping all others constant, and 100 simulations were run for each parameter set. Viral quantity over time is shown; 25-75% quartiles are shown in dark blue while 0-100% quartiles are shown in light blue. When not varying,  $b = 0.01$ ,  $\beta = 50$ ,  $f = 0.1$ ,  $\alpha = 200$ ,  $\theta = 0.001$ ,  $\delta = 0.1$ ,  $p = 10^4$ , and  $c = 6$ .

**Table 2. Parameter values used in the model.** Parameters that remained fixed throughout cohort data fitting were chosen based on published values and univariate analysis. \* The discrepancy between the published and model values for parameter  $\beta$  is due to previous models separately accounting for infection due to cell-free virus. \*\* The discrepancy between the published and model values for parameter  $\delta$  is due to previous models not accounting for tissue-resident cytotoxic T cells as a separate population.

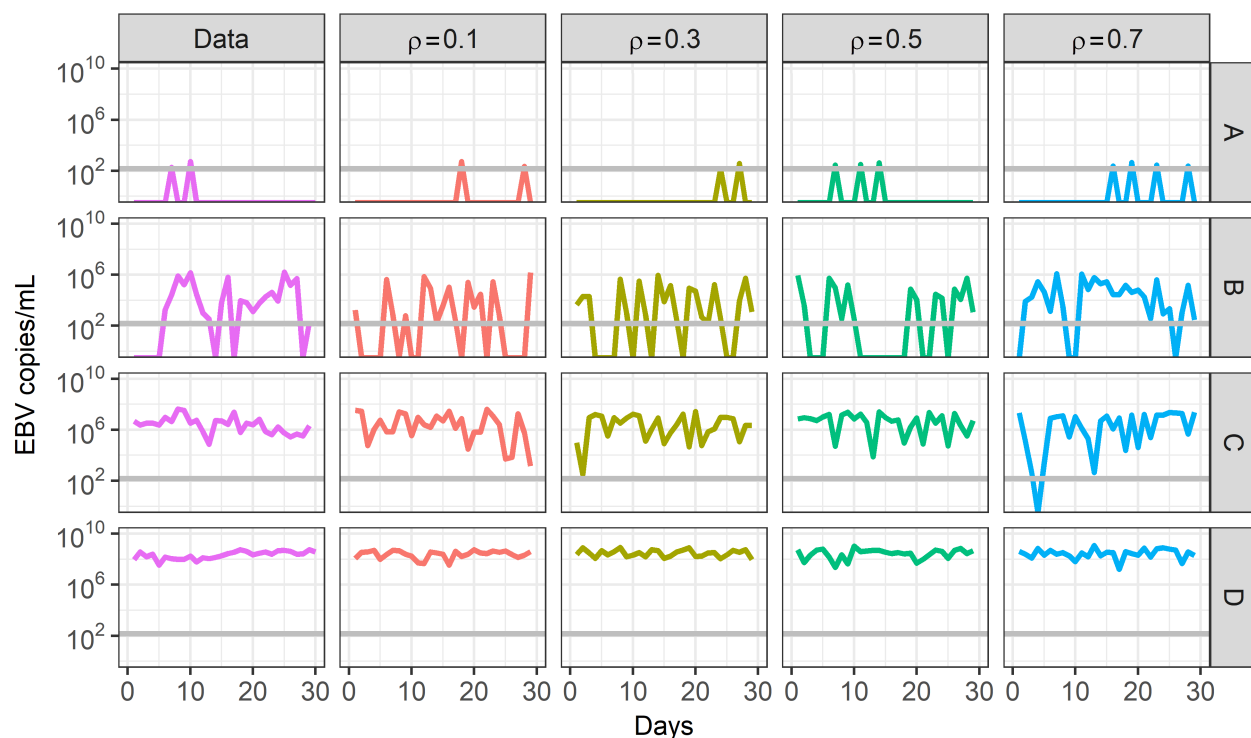
Parameter	Units	Description	Published Values	Model Values
$b$	cell day <sup>-1</sup>	rate of B cell reactivation causing new lytic epithelial infection	-	fitted, see text
$\beta$	day <sup>-1</sup>	per-capita rate of cell-to-cell infection	1.6 [37]	50*
$f$	day <sup>-1</sup> cell <sup>-1</sup>	per-capita death rate of infected epithelial cells due to the effect of cytotoxic T cells	$1.8 \times 10^{-9} - 0.2$ [27, 38, 39]	0.1
$\alpha$	cell	number of tissue-resident cytotoxic T cells	-	200
$\theta$	day <sup>-1</sup> cell <sup>-1</sup>	per-capita proliferation rate of cytotoxic T cells dependent on the number of infected epithelial cells	-	fitted, see text
$\delta$	day <sup>-1</sup>	per-capita death rate of cytotoxic T cells	$9.5 \times 10^{-3}$ [27, 40]	0.1**
$p$	virion day <sup>-1</sup> ml <sup>-1</sup> cell <sup>-1</sup>	production rate of EBV virions by infected epithelial cells	$10^2 - 10^5$ [10, 27]	$10^4$
$c$	day <sup>-1</sup>	per-capita clearance rate of EBV virions	2 [27, 41]	6

## Mathematical model fits clinical data well and simulates oral shedding data with high fidelity

We fit our mathematical model to each participant's data using Approximate Bayesian Computation (ABC). As our model is stochastic, this fitting process involved finding parameter values that produced simulations with similar summary statistics to the data rather than simulations that directly matched the curve of the data. The traits of each participant were captured using five summary statistics: the percentage of EBV-positive swabs, the median, maximum, and variance of detectable viral loads, and the number of peaks in viral load, a peak being defined as when the directly preceding and following time points have lower viral loads. The goodness of fit was assessed by the statistic  $\rho_{i,j}$ , which is defined as

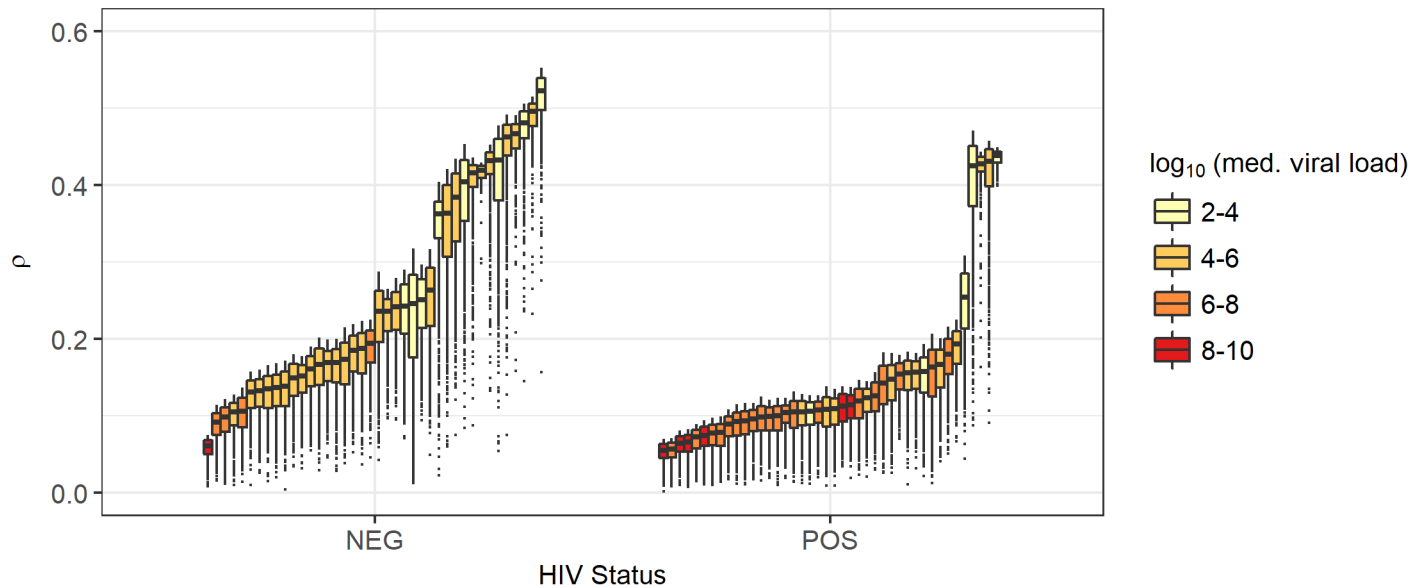
$$\rho_{i,j} = \frac{1}{5} \sum_{k=1}^5 \left| \frac{D_{i,k} - \hat{D}_{j,k}}{D_{i,k}} \right| \quad (7)$$

for participant  $i$  and parameter set  $j$ . Here  $D_{i,k}$  is the  $k^{th}$  summary statistic for the data of participant  $i$  and  $\hat{D}_{j,k}$  is the  $k^{th}$  summary statistic for simulations using parameter set  $j$ . Lower  $\rho$  values indicate a better fit between the model simulation and the data. Full details are given in the Methods. Examples of 4 participants' shedding data and model simulations with  $\rho$  values varying between 0.1 and 0.7 are shown in Fig 4. At these low  $\rho$  values, all simulations capture the summary statistics of the participants quite well, but a general improvement in fit quality is apparent as  $\rho$  approaches 0.1.



**Fig 4. Comparison of EBV shedding patterns in Ugandan cohort participants and model simulations.** The EBV shedding patterns of four representative participants are shown. Participants A and B are uninfected with HIV-1 while participants C and D are co-infected with HIV-1. Model simulations that fit these data with varying success ( $\rho = 0.1, 0.3, 0.5, 0.7$ ) are shown. Lower  $\rho$  values indicate a better fit to the summary statistics of the data. The horizontal grey line indicates the threshold of detection (150 EBV copies/ml) when measuring EBV loads in participant saliva samples.

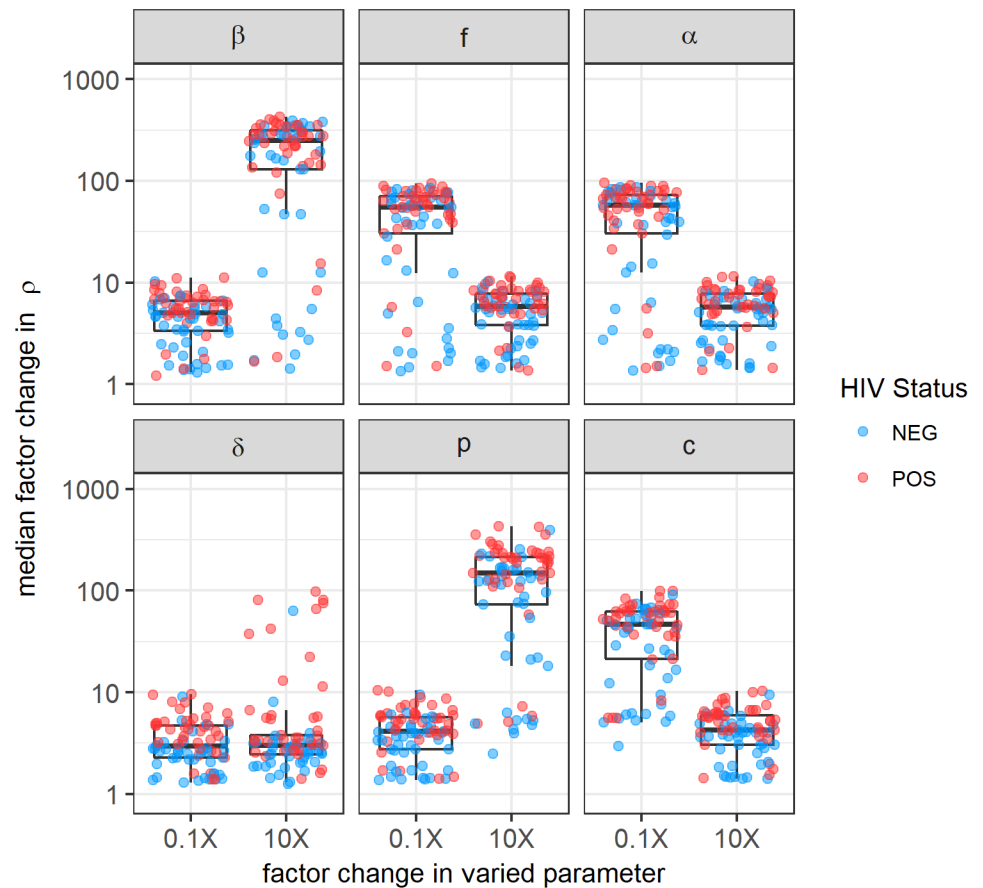
Among all 85 participants' data, we were able to fit parameters to 82. Of the 3 participants whose data could not be fit, 2 participants had no EBV detected in any of their saliva swabs, and 1 participant had only 1 swab taken. Fig 5 shows the distribution of  $\rho$  values (Equation 7) calculated for 1000 optimal parameter sets for each participant that was fit by the model. The  $\rho$  values accepted during fitting ranged between 0.002 and 0.558, indicating that accepted model fits matched the data at least as well as the examples shown in Fig 4. In general, our model fits were slightly better for participants with a high median viral load. Using generalized estimating equations (GEE), and assuming a Gaussian distribution for  $\rho$  values, we found that each  $\log_{10}$  decrease in a participant's median viral load increased  $\rho$  by 0.051 (95% CI = 0.039-0.063, p-value < 0.001).



**Fig 5. Goodness of fit of mathematical model to participant data.** Datasets from 82 study participants were fit by our mathematical model. Each participant's dataset is represented by 1000 parameter sets that best fit model simulations to the data. Goodness of fit for each parameter set is measured by a  $\rho$  value as defined in Equation 7. Boxes indicate the interquartile range and whiskers indicate the 95% range. Lower values indicate a better fit between the data and the model. The colour of the box indicates the median EBV viral load detected in the saliva of that participant.

### Sensitivity analysis of the fitted model

After fitting our model to data (Methods), we performed a sensitivity analysis to confirm that we chose suitable values for the parameters that remained fixed throughout the fitting process. Parameters  $b$  and  $\theta$  were fixed at their fit values, while those that had remained fixed during fitting were varied over two orders of magnitude. A simulation was run for each new set of parameters and the goodness of fit was compared to the original. Results are shown in Fig 6. In all cases and for all participants, the median  $\rho$  value generated from parameter sets was higher and, therefore, a worse fit than our fixed parameter choices, providing confidence that the values chosen for our fixed parameters provided a consistently good fit.



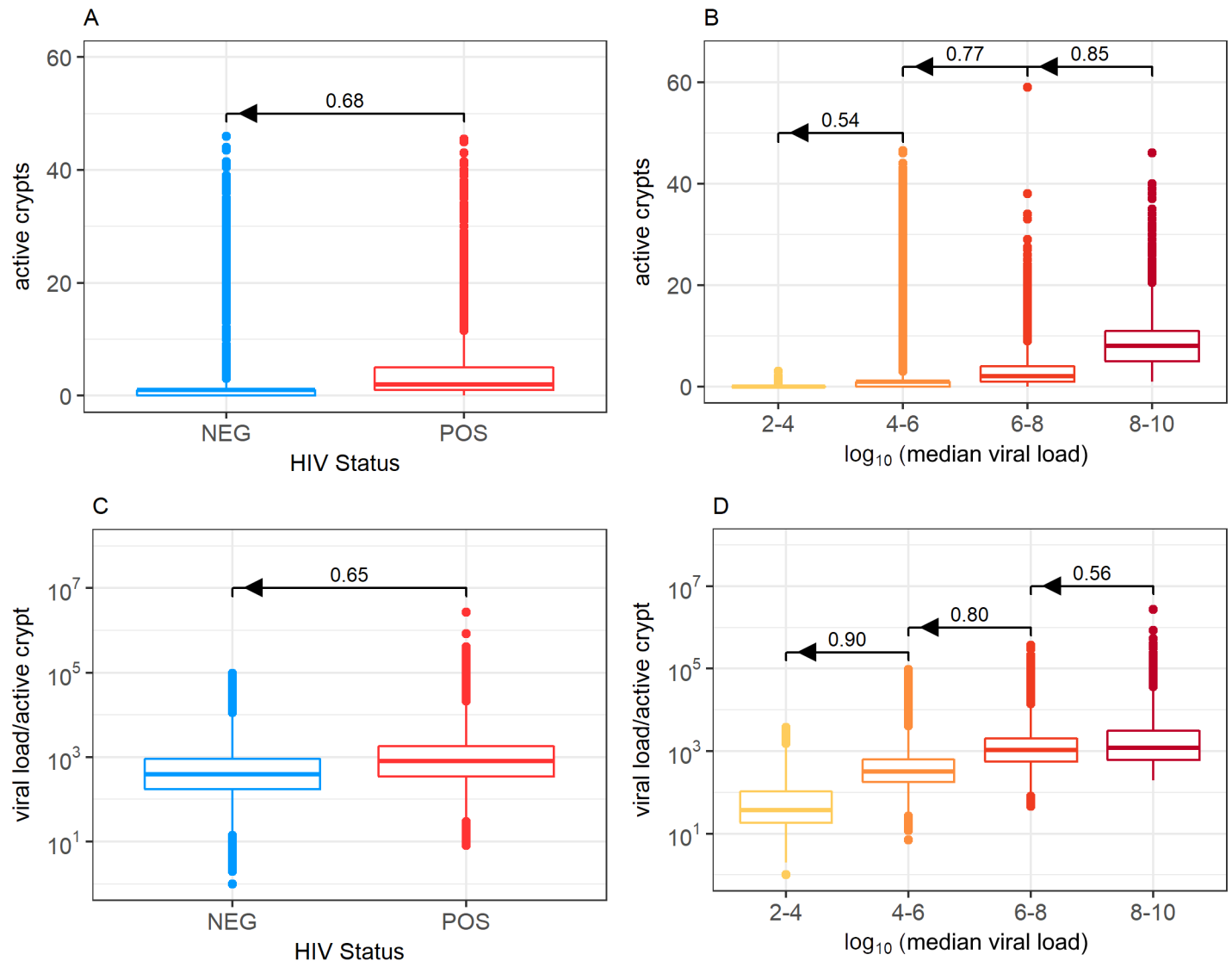
**Fig 6. Sensitivity analysis of fit parameters.** For the 1000 parameter sets selected by the ABC fitting algorithm for each participant, we varied parameters  $\beta$ ,  $f$ ,  $\alpha$ ,  $\delta$ ,  $p$ , and  $c$  (the parameters that remained fixed during fitting) over two orders of magnitude to observe whether changes in these values could have improved the model fit. Each parameter was set to equal 0.1 and 10 times the value it was set to during fitting and one simulation of each new parameter set was performed. The median factor-change in the resulting  $\rho$  value for each participant is shown on the y-axis. In all cases, the new parameter values led to worse fits (factor change in  $\rho$  exceeded one).

## High EBV viral loads are caused by large numbers of actively infected crypts

When examining the model-predicted viral dynamics in simulated individuals, we noticed that high viral loads in the saliva were caused by multiple crypts actively producing virus at the same time (Fig 7). Using importance sampling (Methods) on the results of our ABC fitting algorithm, we achieved a representation of how the number of actively infected crypts and the amount of virus produced by these crypts differ among individuals stratified by HIV-1 infection status.

The distributions for the median number of crypts HIV-1 co-infected and HIV-1 uninfected individuals have over time are shown in Fig 7A. Over time, an HIV-1 co-infected person will have a higher median number of actively infected crypts than an HIV-1 uninfected person with probability 0.68 (Fig 7A). While these distributions vary greatly (IQR of 1 and 4 for HIV-1 uninfected and HIV-1 co-infected individuals respectively), over time HIV-1 uninfected persons are expected to have a median of 1

crypt within their tonsils actively producing virus, while HIV-1 co-infected individuals are expected to have 2. These results match well with previous estimates that indicate individuals have  $\leq 3$  independent plaques of oral epithelial infection at any given time [10]. Infected crypts within an HIV-1 infected individual are also shown to produce more virus. Over time, an HIV-1 co-infected individual is expected to have a higher median viral load within their actively infected crypts than an HIV-1 uninfected individual with probability 0.65 (Fig 7C). Our simulations predict a median of 798 EBV DNA copies per active crypt in HIV-1 co-infected individuals and 389 EBV DNA copies per active crypt in HIV-1 uninfected individuals. These distributions again vary greatly, leading to overlap (IQR of 753 and 1472 for HIV-1 uninfected and HIV-1 co-infected individuals respectively). The overlap between the distributions of traits for HIV-1 co-infected and HIV-1 uninfected individuals implies that HIV-1 status alone does not determine a distinct phenotype. Thus, we chose to also examine the behaviour of traits when stratifying our simulations according to participants' median viral loads in the saliva. Higher median EBV viral loads detected in saliva correlate well with higher numbers of actively infected crypts (Fig 7B) and higher viral loads per actively infected crypt (Fig 7D). Together, these results indicate that high EBV loads in the saliva are caused by more frequent and extensive infection in tonsillar crypts.



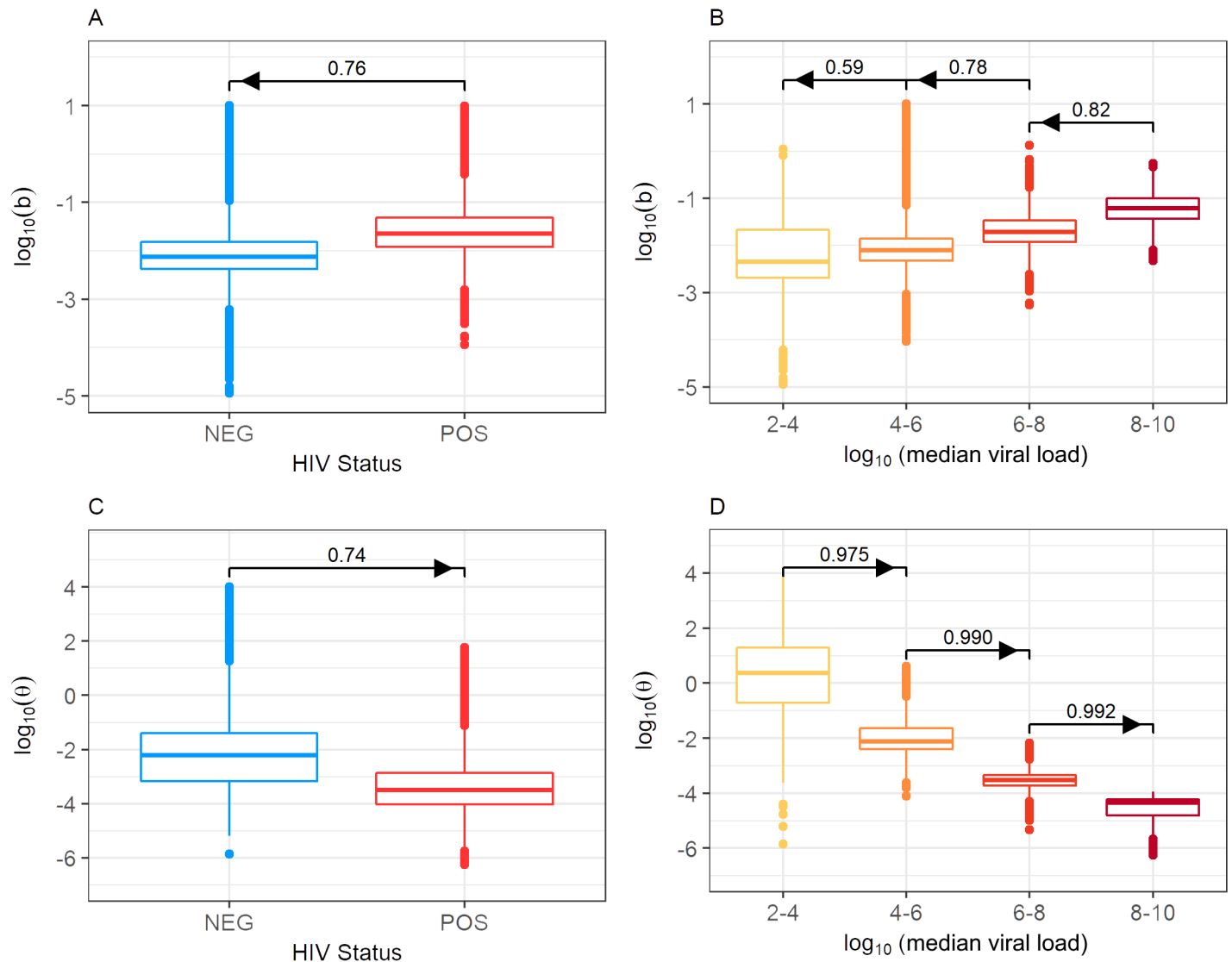
**Fig 7. Predicted numbers of active crypts and viral load per active crypt.** Distributions of the median number of crypts actively producing EBV (A and B) and the median EBV viral load produced by an active crypt at any given time (C and D) are shown stratified by participant HIV-1 status and EBV median viral load in the saliva. Increases in median salivary EBV viral load are caused by a higher number of crypts having active (B), and more extensive (D) infection. This trend translates to HIV-1 co-infected participants having more infected crypts infected, and each infected crypt producing more virus. Bars above boxplots indicate the probability that a randomly selected individual of one group has a higher parameter value (be it number of active crypts or viral load/active crypt) than a randomly selected individual in a second group. Arrows show the direction of comparison. We see that HIV-1 uninfected individuals usually have more actively infected crypts and more virus per active crypt than HIV-1 co-infected individuals. Similarly, we see that individuals with higher median viral loads in their saliva usually have more actively infected crypts and more virus per active crypt than individuals with lower median viral loads.

## Greater oral EBV shedding in HIV-1 co-infection is due to increased B cell reactivation and weaker cellular immune response

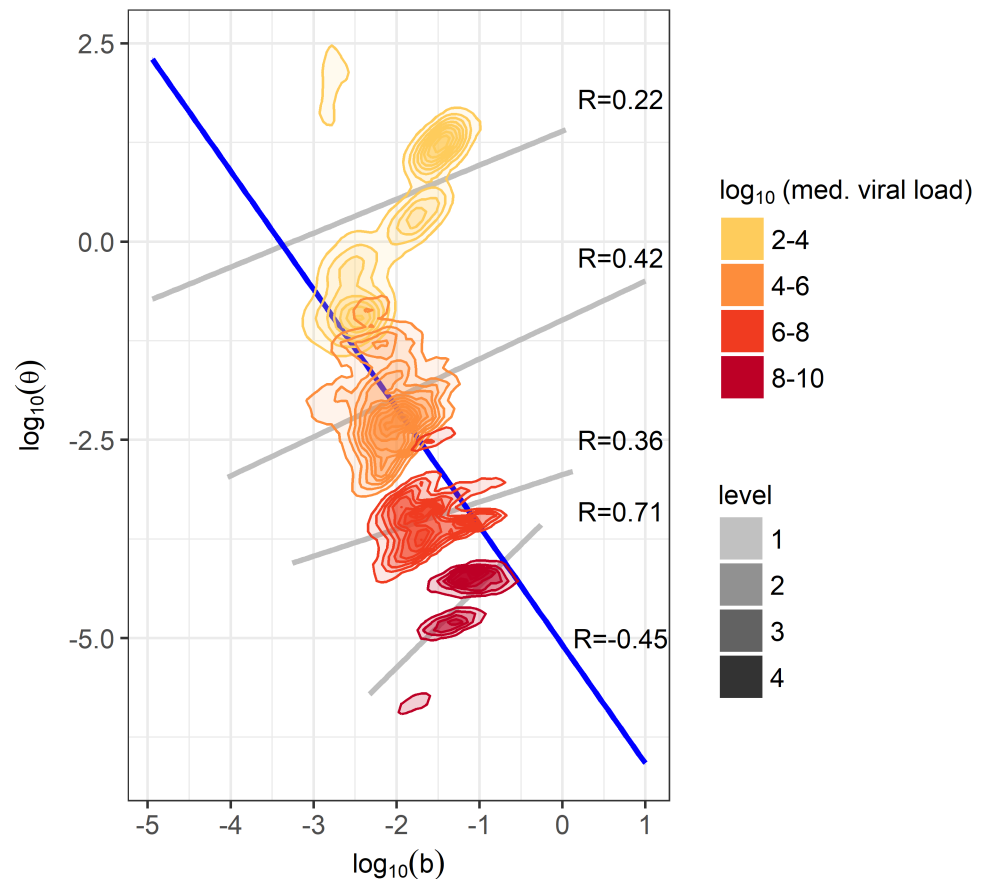
We next looked at parameter-specific differences between individuals of different infection status groups and determined distributions for parameters  $b$  (rate of B cell reactivation causing new lytic epithelial infection) and  $\theta$  (rate of EBV-specific cytotoxic T cell proliferation and recruitment). Parameter distributions stratified by HIV-1 infection status and median EBV viral load in saliva are shown in Fig 8. While these distributions display substantial variability and overlap, changes in both  $b$  and  $\theta$  account for the the generally higher EBV viral loads observed in the saliva of HIV-1 co-infected individuals. HIV-1 co-infected individuals appear to have higher rates of B cell reactivation. The median value of  $b$  was found to be 2.9 times higher in HIV-1 co-infected individuals, equalling 0.007/day in HIV-1 uninfected individuals (IQR for  $b$  of 0.01) and 0.02/day in HIV-1 co-infected individuals (IQR for  $b$  of 0.04). Together, these distributions indicate that an HIV-1 co-infected individual will have a higher  $b$  value than an HIV-1 uninfected individual with probability 0.76. HIV-1 co-infected individuals also appear to have lower rates of EBV-specific cytotoxic T cell immune response. The median value of  $\theta$  was 19.7 times lower in HIV-1 co-infected individuals, equalling  $3.1 \times 10^{-4}$  in HIV-1 co-infected individuals (IQR for  $\theta$  of 0.04) and  $6.1 \times 10^{-3}$  in HIV-1 uninfected individuals (IQR for  $\theta$  of 0.001). Together these distributions indicate that an HIV-1 co-infected individual will have a lower  $\theta$  value than an HIV-1 uninfected individual with probability 0.74. We also note that, in general, increases in median EBV viral load in saliva correlate with increases in the rate at which B cell reactivation occurs and causes infection ( $b$ ). Similarly, there is an inverse correlation between the strength of the EBV-specific cytotoxic T cell immune response ( $\theta$ ) and the median viral load (Fig 8B and 8D).

We also sought to quantify the correlation between  $b$  and  $\theta$  (Fig 9). When looking at within-group correlation of participants with similar median viral loads,  $b$  and  $\theta$  have moderate positive correlation ( $R = 0.22, 0.42, 0.36$  and  $0.71$  for increasing viral load groups), indicating that  $b$  and  $\theta$  can counter-balance to cause similar viral loads. However, when observing all data,  $b$  and  $\theta$  are moderately negatively correlated ( $R = 0.45$ ), indicating the highest viral loads are found in individuals whose parameters feature high  $b$  and low  $\theta$  values.





**Fig 8. Distribution of parameters  $b$  and  $\theta$ , stratified by HIV-1 status and median EBV viral load in saliva.** Fitting our mathematical model to participant data revealed that parameter  $b$  is usually greater in HIV-1 co-infected participants (A) and increases with median EBV viral load (B), and parameter  $\theta$  is usually lower in HIV-1 co-infected participants (C) and decreases with median saliva EBV viral load (D). Bars above boxplots indicate the probability that a randomly selected individual of one group has a higher parameter value (be it  $b$  or  $\theta$ ) than a randomly selected individual in a second group. Arrows show the direction of comparison.



**Fig 9. Correlation between parameters  $b$  and  $\theta$ .** Obtained densities of parameters  $b$  and  $\theta$  are plotted, stratified by median oral EBV viral load group. As  $b$  and  $\theta$  have opposite effects on viral load, positive correlations are seen within each group (grey lines). However, across all participants,  $b$  and  $\theta$  are negatively correlated (blue line).

We next examined whether rates of B cell reactivation and cellular immune response could be predicted by the CD4+ T cell count and HIV-1 plasma viral load in HIV-1 co-infected participants. A generalized linear model (Methods) revealed that each  $\log_{10}$  increase in HIV-1 RNA copies/ml significantly decreased the value of  $\theta$  but did not significantly affect  $b$ . However, each 100-cell/ $\text{mm}^3$  increase in CD4+ T cell count significantly changed the value of  $\theta$  and  $b$ , increasing  $\theta$  and decreasing  $b$  (Table 3).

**Table 3. Effect of CD4+ T cell count and HIV-1 load on parameters  $b$  and  $\theta$ .** In HIV-1 co-infected participants, median values of parameters  $b$  and  $\theta$  are influenced by the CD4+ T cell count and HIV-1 plasma viral load. The fold-change (FC) in the fit  $b$  and  $\theta$  values, for every  $\log_{10}$  increase in HIV-1 RNA copies/mL and every 100 CD4+ T cell/ $\text{mm}^3$  increase is shown.

	Trait	FC	95% CI	p-value
$\theta$	HIV-1 RNA	0.16	0.06-0.42	<0.001
	CD4+ T cell	1.84	1.05-3.22	0.040
$b$	HIV-1 RNA	1.21	0.89-1.63	0.226
	CD4+ T cell	0.80	0.69-0.92	0.004

## Application of our model to a distinct data set from a North American cohort

To test its generalizability, we applied our model to a previously described set of data from a cohort of 26 participants in Seattle, Washington, who underwent daily oral EBV sampling and testing using the same methods as the Ugandan cohort described above [22]. A total of 1323 swabs were collected during the 8-week period of the study, with a median of 55 swabs per participant (3-61 swabs). 16 (62%) of these participants were HIV-1 co-infected and if on HAART were required to remain on a stable regimen throughout the study. None were receiving other antiviral drugs at the time of enrollment. While the objective of the Seattle study was to analyse the effects of valganciclovir on daily EBV oral shedding, we restricted our analysis to the data from the eight-week period when participants received placebo.

Participants of the Seattle cohort showed significantly lower oral EBV viral loads than those of the Uganda cohort. Given a positive swab, HIV-1 uninfected, EBV shedding participants had a mean of 1.4  $\log_{10}$ -lower EBV viral load, while HIV-1 co-infected, EBV shedding participants had a mean of 2.0  $\log_{10}$ -lower viral load with p-values<0.001 for both cases. Nonetheless, both sets of data serve as good representations of EBV shedding in the saliva and we expect the host-pathogen interactions occurring within the tonsils to be the same. Our mathematical model fit the data again with high fidelity (Fig S3 in Supporting Information) and produced similar results in terms of the number of infected crypts (Fig S4A and S4B in Supporting Information), virus produced by crypts (Fig S4C and S4D in Supporting Information), and the values for parameters  $b$  and  $\theta$  when stratified by HIV-1 status and median EBV load (Fig S5 in Supporting Information). Full details are given in the Supporting Information. Overall, we find that our modelling approach and findings from the Uganda cohort are strongly supported by analysis of the Seattle cohort data.

## Discussion

By modelling the replication patterns of EBV in saliva in HIV-1 co-infected and uninfected individuals, we were able to evaluate the potential mechanisms that explain why persons with HIV-1 have worse EBV infections and are more susceptible to EBV-related malignancies. Specifically, our model indicates that increased oral EBV shedding with HIV-1 co-infection is due to greater reactivation of EBV-infected B cells as well as impaired EBV-specific cytotoxic T cell immune control.

It has been previously reported that persons infected with HIV-1 have higher oral EBV shedding than HIV-1 uninfected individuals [3, 4, 23–25, 42, 43]. However, few data sets allow detailed representation of the dynamics of oral EBV shedding in HIV-1

co-infected and uninfected participants over time. We also found that HIV-1 infection was associated with a significantly increased frequency and quantity of oral EBV shedding. There was also a statistically significant association between higher CD4+ T cell counts and a lower frequency of EBV shedding in HIV-1 co-infected participants.

B cell activation and plasma cell differentiation have been shown to induce EBV reactivation and lytic replication [12]. Furthermore, it is clear that HIV-1 infection is associated with increased levels of B cell activation [18, 44]. As such, we hypothesized that, in addition to impaired cellular immune control of EBV infection, an increase in EBV reactivation from latently-infected B cells in HIV-1 co-infected individuals contributes to higher levels of oral EBV replication. This prompted us to develop a mathematical model to describe the dynamics of virus, infected cells, and the cellular immune response within the tonsillar epithelium. We then fit this model to participant data, seeking to determine the drivers of these differences.

While previous mathematical models have examined the within-host dynamics of EBV infection [10, 26–29], none have examined the differences between HIV-1 infected and HIV-1 uninfected individuals [38, 45–47]. Strengths of our approach include the incorporation of granular quantitative EBV shedding measurements from two independent cohorts, each made up of HIV-1 co-infected and uninfected individuals, and the availability of CD4+ T cell counts and HIV-1 plasma viral load data from co-infected participants. Limitations include the lack of data on EBV-specific T cell responses and direct measures of B cell activation in these individuals.

Our mathematical model was based on representing the tonsillar epithelium as a series of crypts, each serving as a potential site of epithelial infection and viral shedding. In this way, a single tonsillar crypt behaves similarly to how individual herpes simplex virus (HSV)-2 lesions have been modelled in the past, with reactivation of EBV-infected B cells being analogous to the release of HSV-2 from infected neurons, sparking new epithelial lesions [38, 45–47]. These previous models all captured the stochastic patterns of HSV-2 shedding well, which appear similar to the patterns of oral EBV shedding. Our model assumes that all tonsillar crypts are independent from one another. In reality, virus from one infected crypt may spill over and seed infection in a neighbouring crypt, rather than EBV entering an uninfected crypt purely via the reactivation of B cells as we assume in our model. By not accounting for this, our predicted B cell reactivation rates are likely higher than their true biological values. However, assuming the amount of viral spill-over into neighbouring crypts is proportional to the viral load, our qualitative comparison remains valid. Furthermore, modelling the tonsillar crypts as multiple, segregated sites of infection was essential for simulating the high variability in viral load seen over time in cohort data. In simulations without this spatial separation, viral loads and levels of immune surveillance within an individual often equilibrated over time and no longer matched the stochastic nature of participant data. This result indicates that while virus and immune cells may travel between crypts, this effect is likely minimal. While a fully-spatial model including the mobility of virus and immune cells throughout the tonsils would be ideal, it could only be parameterized speculatively, and would be computationally intensive. Our intermediate strategy of a crypt-level simulation is simple enough to be computationally feasible but complex enough to retain the inherent stochasticity and spatial diversity of EBV infection dynamics within the tonsils.

Using our mathematical model, we found that HIV-1 co-infected individuals are expected to have more actively infected crypts producing more EBV at any given time than HIV-1 uninfected individuals. Also, by specifically including the dynamics of cellular immune control and rates of B cell reactivation leading to new epithelial infection, we saw that both factors contribute to higher EBV viral loads in HIV-1 co-infected individuals. These results are consistent with previous observations that

cellular immune control of EBV infection in HIV-1 co-infected individuals is impaired [48]. Similarly, HIV-1 infection is known to dysregulate B cell function and increase B cell activation, which can result directly from HIV-1 itself as well as more frequent and severe co-infections compared to HIV-1-uninfected individuals [17, 18, 49, 50]. Indeed, our model demonstrated that higher CD4+ T cell counts were associated with significantly higher predicted rates of EBV-specific cytotoxic T cell activity and significantly lower predicted rates of EBV-infected B cell reactivation. Furthermore, increases in HIV-1 plasma RNA levels were associated with significantly lower predicted rates of EBV-specific cytotoxic T cell activity.

Importantly, our results have implications for strategies to prevent EBV infection and disease. EBV-specific cellular immunity is recognized as critical for controlling EBV replication and preventing EBV-associated malignancies. [48, 51–55]. Independent of restoring EBV-specific cellular immune responses, strategies to reduce B cell reactivation in EBV-infected persons might limit viral replication, transmission, and related malignancies [56, 57].

# Methods

## Cohorts and samples

Men and women aged 18 to 65 were enrolled in the Uganda cohort as previously described and were followed for four weeks [31]. Eligible HIV-1 seropositive persons had a CD4+ T cell count greater than 200 cells/mm<sup>3</sup> and were not taking antiretroviral therapy, in accordance with the WHO guidelines at that time [58]. Men aged 24 to 66 were enrolled in the Seattle cohort as previously described [21, 22]. As the Seattle cohort shedding data was obtained from a randomized placebo-controlled cross-over trial of valnaciclovir, only data collected while participants were receiving placebo was used for this study. Both participants and pill administrators were unaware of group assignments. Participants of both cohorts did not take any drugs with anti-herpesvirus activity during the study period. Self-collected daily oropharyngeal swabs for both cohorts were collected by swabbing the oral mucosa and pharynx with a Dacron swab and were then placed in a vial containing 1 ml of 1X digestion buffer, stored at room temperature, and returned at weekly (Ugandan cohort) or bi-weekly (Seattle cohort) clinical visits. In the Uganda cohort, focused physical exams and collection of genital and plasma samples were performed at weekly clinic visits. This data is described in the Supporting Information. All participants provided either written or verbal informed consent. Study procedures for both cohorts were approved by the University of Washington Human Subjects Review Board. Additional approval for the Ugandan cohort study was given by the Makerere University Research and Ethics Committee and the Uganda National Council for Science and Technology.

## Laboratory testing

Commercially available immunoassays were used to ascertain HIV-1 and EBV serostatus (Inverness Medical Innovations, Inc and Wampole® for the Ugandan cohort and Abbott Laboratories for the Seattle Cohort [22]). For the Ugandan cohort, CD4+ T cell counts, and plasma HIV-1 RNA levels were determined in HIV-1-infected persons at the Makerere University-John Hopkins University laboratory using standard cell sorting techniques and the Amplicor HIV-1 monitor test (Roche, version 1.5), respectively. For both cohorts, DNA was extracted from mucosal swabs and plasma [59], and real-time quantitative polymerase chain reaction (qPCR) was performed using specific primers to detect EBV [60], with positive and negative controls as previously described [19, 59]. Mucosal samples with greater than 150 copies/ml and plasma samples with greater than 50 copies/ml herpesvirus DNA/ml were considered positive [61].

## Statistical analyses of data

The frequency of mucosal shedding and viremia was defined as the proportion of samples testing positive for EBV. The frequency of mucosal shedding was first compared in HIV-1 co-infected and uninfected persons. To do this, we used generalized estimating equations (GEE) and assumed frequencies followed a Poisson distribution. In HIV-1 seropositive persons, frequencies of mucosal shedding and viremia were also modelled with GEE allowing for continuous adjustment for each 100 cell/mm<sup>3</sup> increase in CD4+ T cell count and each log<sub>10</sub> increase in HIV-1 RNA. Again, we assumed the frequency of shedding follows a Poisson distribution. Finally, GEE were used to compare the quantities of virus shed in mucosal samples in HIV-1 co-infected and HIV-1 uninfected persons assuming a Gaussian distribution. In all tests, two-sided p-values ≤0.05 were considered statistically significant.

## Mathematical model simulations

Based on the reactions of the model, we applied the tau leaping algorithm to stochastically simulate the dynamics of each crypt [62]. With this algorithm, a small, constant-sized time step is taken, and the number of occurrences of each reaction is stochastically chosen following a Poisson or Multinomial distribution depending on the independence of the reaction. One long simulation is performed, which is then divided into 240 sections to represent the dynamics of each of the 240 crypts. Specifically, the simulation is run out to a time of

$$W_{\text{init}} + 240(L_i + W_{\text{crypt}}) \quad (8)$$

and crypt  $c$ 's dynamics are taken from the time interval

$$t \in [W_{\text{init}} + (c - 1)(L_i + W_{\text{crypt}}), W_{\text{init}} + cL_i] \quad (9)$$

where  $W_{\text{init}}$  represents the time necessary to remove the effects of the initial conditions on the simulation,  $L_i$  represents the duration of which participant  $i$  had oral swabs taken, and  $W_{\text{crypt}}$  represents the time necessary to make the dynamics of one crypt quasi-independent of the next. With immune cell decay ( $\delta T$ ) acting as the slowest rate in the model ( $\delta = 0.1/(\text{day-cell})$ ), we let both  $W_{\text{crypt}}$  and  $W_{\text{init}}$  equal 120 days, so that if infection in one crypt occurred, only an expected  $1/10^6$  of the responding immune cells would carry on to the next crypt's dynamics.

Viral loads from each crypt are added together to get the model-predicted amount of virus seen in the saliva over time. As the qPCR threshold of detection was 150 copies/ml for the data used, whenever the total simulated viral load in the saliva dropped below 150 copies/ml, we set the output to zero.

## Model fitting using Approximate Bayesian Computation

We next fit parameters to daily quantitative oral EBV shedding data from our cohort participants. We used Approximate Bayesian Computation (ABC), where summary statistics of the data and model simulations are compared to determine which parameters allow the model to best fit the data. We used the R package EasyABC to execute sequential ABC, following Lenormand's algorithm [63]. Here, uniform priors for each parameter are set, and an initial  $n$  number of simulations are run, each with a different set of parameters randomly chosen from the priors. The algorithm calculates summary statistics, chosen by the user, for each parameter set ( $\hat{D}$ ) and compares how well they match with the summary statistics of the data ( $D$ ) by calculating a distance measure,  $\rho(D, \hat{D})$ . The best-matching  $\phi$  percent of these simulations are kept, with the parameters of the chosen simulations used to build new priors. This process repeats until the distance between the summary statistics of the data and simulations is minimized. We executed this algorithm for the data of each individual participant. We chose to capture the trends of the data using 5 summary statistics: the frequency of positive swabs, the median, maximum, and variance of detectable viral loads, and the number of peaks in viral loads, a peak being defined as when the directly preceding and following time points have lower viral loads. The associated  $\rho$  value for participant  $i$  and parameter set  $j$  ( $\rho_{i,j}$ ) is defined as in Equation 7 of the main text:

$$\rho_{i,j} = \frac{1}{5} \sum_{k=1}^5 \left| \frac{D_{i,k} - \hat{D}_{j,k}}{D_{i,k}} \right| \quad (10)$$

where  $D_{i,k}$  is the  $k^{\text{th}}$  summary statistic for the data of participant  $i$  and  $\hat{D}_{j,k}$  is the  $k^{\text{th}}$  summary statistic for parameter set  $j$ . Using the Lenormand algorithm, 1000 parameter



sets that minimize  $\rho_{i,j}$  were selected for each participant. During this fitting, we fixed all parameters except  $b$  and  $\theta$  as these are the two parameters that are likely most affected by HIV-1 infection. For each chosen parameter set, we also calculated the median number of actively infected crypts and the median amount of virus produced by an actively infected crypt, so we could later compare how these traits differ between participant groups.

Since we lack information on immune cell presence in tonsillar crypts and can only fit the model to data on viral load, we had to censor simulations where the cytotoxic T cell level became unrealistically high. Whenever a parameter set led to a simulation where cytotoxic T cell count within a tonsillar crypt was greater than  $10^5$  cells, we prevented it from being selected by the ABC algorithm, ensuring only biologically relevant simulations were being considered.

## Determining the posterior distribution of parameters and differences between participant groups

We combined the results of our ABC fitting algorithm to compare how the posterior distributions of parameters  $b$  and  $\theta$ , the number of actively infected crypts, and the amount of virus produced vary between different participant groups.

As some parameter values selected by the ABC fitting algorithm fit the data better than others (i.e. have lower  $\rho$  values), we approximated the posterior distributions of our parameter sets by performing importance sampling on the raw posterior distributions [64,65]. To do this, we weighted each output parameter set by the reciprocal of its  $\rho$  value. By weighting inversely to  $\rho$ , we assume our model is a correct representation of viral dynamics in the tonsils and put greater importance on those parameter sets that fit the data well.

To determine the posterior distributions of parameters in HIV-1 co-infected and uninfected groups ( $X_A$  and  $X_B$  respectively), the probability of each parameter set ( $x(i, j)$ ) serving in each posterior is set to

$$P(X_A = x_{i,j}) = P(i \in A) \frac{1}{\rho_{i,j}} \frac{1}{q_A} \quad (11)$$

$$P(X_B = x_{i,j}) = P(i \in B) \frac{1}{\rho_{i,j}} \frac{1}{q_B} \quad (12)$$

where  $A$  and  $B$  are the sets of indices of participants who are HIV-1 uninfected and co-infected, respectively, and we define the normalization factors

$$q_A = \sum_{\forall i \in A} \sum_{\forall j} \frac{1}{\rho_{i,j}} \quad \text{and} \quad q_B = \sum_{\forall i \in B} \sum_{\forall j} \frac{1}{\rho_{i,j}}. \quad (13)$$

Note that

$$\sum_{\forall i} \sum_{\forall j} P(X_A = x_{i,j}) = 1 \quad \text{and} \quad \sum_{\forall i} \sum_{\forall j} P(X_B = x_{i,j}) = 1. \quad (14)$$

We took  $10^5$  draws from each distribution and plotted the resulting data to obtain graphical representations of the posterior parameter distributions for parameters  $b$  and  $\theta$ , the number of actively infected crypts, and the virus produced, for HIV-1 co-infected and uninfected participants. The above process was repeated where instead of stratifying by HIV-1 status, participants were stratified by median EBV load in order to produce similar plots.

We also performed importance sampling on the raw posterior distributions for individual participants. Using these, we were able to calculate the mean parameter



values for  $b$  and  $\theta$  for each participant. Means of parameters  $b$  and  $\theta$  in HIV-1 co-infected participants were then modelled as functions of the participants' CD4+ T cell count and HIV-1 RNA load. This was done using GLM allowing for continuous adjustment for each 100 cell/mm<sup>3</sup> increase in CD4+ T cell count and each log<sub>10</sub> increase in HIV-1 RNA. Parameters  $b$  and  $\theta$  were assumed to follow a Gaussian distribution. In these tests, two-sided p-values  $\leq 0.05$  were considered statistically significant.

## Sensitivity analysis of model parameters

We performed two sensitivity analyses to evaluate how changes in parameter values affect the results of the model. First, to initially determine acceptable values for parameters, we performed a univariate analysis, starting with an initial set of parameters, varying one parameter at a time, and running 100 simulations for each parameter set to gain a representation of how EBV viral dynamics behave. From this analysis we selected parameter values for  $\beta$ ,  $f$ ,  $\alpha$ ,  $\delta$ ,  $p$ , and  $c$  which would remain fixed throughout the ABC data fitting process while leaving the parameters of most interest,  $b$  and  $\theta$ , free.

After completing ABC, we performed another univariate analysis to observe whether our choices for fixed parameter values were correct. By letting  $b$  and  $\theta$  equal values selected by the ABC algorithm and then individually varying the parameters that were fixed, we checked whether different values of our fixed parameters would have improved the model's fit.

# References

1. Cohen JI, Fauci AS, Varmus H, Nabel GJ. Epstein-Barr virus: an important vaccine target for cancer prevention. *Science Translational Medicine*. 2011;3(107):107fs7.
2. Shiels MS, Engels EA. Evolving epidemiology of HIV-associated malignancies. *Current Opinion in HIV and AIDS*. 2017;12(1):6–11.
3. Ferbas J, Rahman MA, Kingsley LA, Armstrong JA, Ho M, Zhou SY, et al. Frequent oropharyngeal shedding of Epstein-Barr virus in homosexual men during early HIV infection. *AIDS*. 1992;6(11):1273–1278.
4. Diaz-Mitoma F, Ruiz A, Flowerdew G, Houston S, Romanowski B, Kovithavongs T, et al. High levels of Epstein-Barr virus in the oropharynx: a predictor of disease progression in human immunodeficiency virus infection. *Journal of Medical Virology*. 1990;31(2):69–75.
5. Matrajt L, Gantt S, Mayer BT, Krantz EM, Orem J, Wald A, et al. Virus and host-specific differences in oral human herpesvirus shedding kinetics among Ugandan women and children. *Scientific Reports*. 2017;7(1):13105–017–12994–0.
6. Ghebrekidan H, Ruden U, Cox S, Wahren B, Grandien M. Prevalence of herpes simplex virus types 1 and 2, cytomegalovirus, and varicella-zoster virus infections in Eritrea. *Journal of Clinical Virology*. 1999;12(1):53–64.
7. Biggar RJ, Henle W, Fleisher G, Bocker J, Lennette ET, Henle G. Primary Epstein-Barr virus infections in African infants. I. Decline of maternal antibodies and time of infection. *International Journal of Cancer*. 1978;22(3):239–243.
8. Zheng QY, Huynh KT, van Zuylen WJ, Craig ME, Rawlinson WD. Cytomegalovirus infection in day care centres: A systematic review and meta-analysis of prevalence of infection in children. *Reviews in Medical Virology*. 2019;29(1):e2011.
9. Laichalk LL, Hochberg D, Babcock GJ, Freeman RB, Thorley-Lawson DA. The dispersal of mucosal memory B cells: evidence from persistent EBV infection. *Immunity*. 2002;16(5):745–754.
10. Hadinoto V, Shapiro M, Sun CC, Thorley-Lawson DA. The dynamics of EBV shedding implicate a central role for epithelial cells in amplifying viral output. *PLoS Pathogens*. 2009;5(7):e1000496.
11. Thorley-Lawson DA. Epstein-Barr virus: exploiting the immune system. *Nature Reviews Immunology*. 2001;1(1):75–82.
12. Thorley-Lawson DA. In: Münz C, editor. *EBV Persistence—Introducing the Virus*. Cham: Springer International Publishing; 2015. p. 151–209.
13. Tovey MG, Lenoir G, Begon-Lours J. Activation of latent Epstein-Barr virus by antibody to human IgM. *Nature*. 1978;276(5685):270–272.
14. Takada K. Cross-linking of cell surface immunoglobulins induces Epstein-Barr virus in Burkitt lymphoma lines. *International Journal of Cancer*. 1984;33(1):27–32.

15. Laichalk LL, Thorley-Lawson DA. Terminal differentiation into plasma cells initiates the replicative cycle of Epstein-Barr virus in vivo. *Journal of Virology*. 2005;79(2):1296–1307. 550  
551  
552
16. Gaidano G, Carbone A, Dalla-Favera R. Pathogenesis of AIDS-related lymphomas. *American Journal of Pathology*. 1998;152(3):113–153. 553  
554
17. Righetti E, Ballon G, Ometto L, Cattelan AM, Menin C, Zanchetta M, et al. Dynamics of Epstein-Barr virus in HIV-1-infected subjects on highly active antiretroviral therapy. *AIDS*. 2002;16(1):63–73. 555  
556  
557
18. Petrara MR, Cattelan AM, Zanchetta M, Sasset L, Freguja R, Gianesin K, et al. Epstein-Barr virus load and immune activation in human immunodeficiency virus type 1-infected patients. *Journal of Clinical Virology*. 2012;53(3):195–200. 558  
559  
560
19. Pauk J, Huang ML, Brodie SJ, Wald A, Koelle DM, Schacker T, et al. Mucosal shedding of human herpesvirus 8 in men. *New England Journal of Medicine*. 2000;343(19):1369–1377. 561  
562  
563
20. Mark KE, Wald A, Magaret AS, Selke S, Olin L, Huang ML, et al. Rapidly cleared episodes of herpes simplex virus reactivation in immunocompetent adults. *The Journal of Infectious Diseases*. 2008;198(8):1141–1149. 564  
565  
566
21. Casper C, Krantz EM, Corey L, Kuntz SR, Wang J, Selke S, et al. Valganciclovir for suppression of human herpesvirus-8 replication: a randomized, double-blind, placebo-controlled, crossover trial. *The Journal of Infectious Diseases*. 2008;198(1):23–30. 567  
568  
569  
570
22. Yager JE, Magaret AS, Kuntz SR, Selke S, Huang ML, Corey L, et al. Valganciclovir for the Suppression of Epstein-Barr Virus Replication. *The Journal of Infectious Diseases*. 2017;216(2):198–202. 571  
572  
573
23. Miller CS, Berger JR, Mootoor Y, Avdiushko SA, Zhu H, Kryscio RJ. High prevalence of multiple human herpesviruses in saliva from human immunodeficiency virus-infected persons in the era of highly active antiretroviral therapy. *Journal of Clinical Microbiology*. 2006;44(7):2409–2415. 574  
575  
576  
577
24. Griffin E, Krantz E, Selke S, Huang ML, Wald A. Oral mucosal reactivation rates of herpesviruses among HIV-1 seropositive persons. *Journal of Medical Virology*. 2008;80(7):1153–1159. 578  
579  
580
25. Yan Y, Ren Y, Chen R, Hu J, Ji Y, Yang J, et al. Evaluation of Epstein-Barr Virus Salivary Shedding in HIV/AIDS Patients and HAART Use: A Retrospective Cohort Study. *Virologica Sinica*. 2018;33(3):227–233. 581  
582  
583
26. Thorley-Lawson DA, Hawkins JB, Tracy SI, Shapiro M. The pathogenesis of Epstein-Barr virus persistent infection. *Current Opinion in Virology*. 2013;3(3):227–232. 584  
585  
586
27. Huynh GT, Rong L. Modeling the dynamics of virus shedding into the saliva of Epstein-Barr virus positive individuals. *Journal of Theoretical Biology*. 2012;310:105–114. 587  
588  
589
28. Huynh GT, Adler FR. Alternating host cell tropism shapes the persistence, evolution and coexistence of Epstein-Barr virus infections in human. *Bulletin of Mathematical Biology*. 2011;73(8):1754–1773. 590  
591  
592

29. Hawkins JB, Delgado-Eckert E, Thorley-Lawson DA, Shapiro M. The cycle of EBV infection explains persistence, the sizes of the infected cell populations and which come under CTL regulation. *PLoS Pathogens*. 2013;9(10):e1003685. 593  
594  
595
30. Duca K, Shapiro M, Delgado-Eckert E, Hadinoto V, Jarrah A, Laubenbacher R, et al. A virtual look at Epstein-Barr virus infection: biological interpretations. *PLoS Pathogens*. 2007;10(3):1388–1400. 596  
597  
598
31. Johnston C, Orem J, Okuku F, Kalinaki M, Saracino M, Katongole-Mbidde E, et al. Impact of HIV infection and Kaposi sarcoma on human herpesvirus-8 mucosal replication and dissemination in Uganda. *PloS One*. 2009;4(1):e4222. 599  
600  
601
32. Perry M, Whyte A. Immunology of the tonsils. *Immunology Today*. 1998;19(9):414–421. 602  
603
33. Tugizov SM, Herrera R, Palefsky JM. Epstein-Barr Virus Transcytosis through Polarized Oral Epithelial Cells. *Journal of Virology*. 2013;87(14):8179–8194. 604  
605  
doi:10.1128/JVI.00443-13. 606
34. Lewis DE, Harriman GR, Blutt SE. Organization of the immune system. In: Rich RR, Fleisher TA, Shearer WT, Schroeder HW, Frew AJ, Weyand CM, editors. *Clinical Immunology*. 3rd ed. Edinburgh: Mosby; 2008. p. 17 – 38. 607  
608  
609
35. Imai S, Nishikawa J, Takada K. Cell-to-cell contact as an efficient mode of Epstein-Barr virus infection of diverse human epithelial cells. *Journal of Virology*. 1998;72(5):4371–4378. 610  
611  
612
36. Gillespie DT. A rigorous derivation of the chemical master equation. *Physica A*. 1992;188:404–425. 613  
614
37. Iwami S, Takeuchi JS, Nakaoka S, Mammano F, Clavel F, Inaba H, et al. Cell-to-cell infection by HIV contributes over half of virus infection. *eLife*. 2015;4:10.7554/eLife.08150. 615  
616  
617
38. Schiffer JT, Abu-Raddad L, Mark KE, Zhu J, Selke S, Magaret A, et al. Frequent release of low amounts of herpes simplex virus from neurons: results of a mathematical model. *Science Translational Medicine*. 2009;1(7):7ra16. 618  
619  
620
39. Schiffer JT, Abu-Raddad L, Mark KE, Zhu J, Selke S, Koelle DM, et al. Mucosal host immune response predicts the severity and duration of herpes simplex virus-2 genital tract shedding episodes. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;107(44):18973–18978. 621  
622  
623  
624
40. Hadinoto V, Shapiro M, Greenough TC, Sullivan JL, Luzuriaga K, Thorley-Lawson DA. On the dynamics of acute EBV infection and the pathogenesis of infectious mononucleosis. *Blood*. 2008;111(3):1420–1427. 625  
626  
627  
doi:10.1182/blood-2007-06-093278. 628
41. Shapiro M, Duca K, Lee K, Delgado-Eckert E, Hawlins J, Jarrah A, et al. A virtual look at Epstein-Barr virus infection: Simulation mechanism. *Journal of Theoretical Biology*. 2008;252(4):633–648. 629  
630  
631
42. Lucht E, Brytting M, Bjerregaard L, Julander I, Linde A. Shedding of cytomegalovirus and herpesviruses 6, 7, and 8 in saliva of human immunodeficiency virus type 1-infected patients and healthy controls. *Clinical Infectious Diseases*. 1998;27(1):137–141. 632  
633  
634  
635

43. Fidouh-Houhou N, Duval X, Bissuel F, Bourbonneux V, Flandre P, Ecobichon JL, et al. Salivary cytomegalovirus (CMV) shedding, glycoprotein B genotype distribution, and CMV disease in human immunodeficiency virus-seropositive patients. *Clinical Infectious Diseases*. 2001;33(8):1406–1411. 636
44. Marks MA, Rabkin CS, Engels EA, Busch E, Kopp W, Rager H, et al. Markers of microbial translocation and risk of AIDS-related lymphoma. *AIDS*. 2013;27(3):469–474. 637
45. Schiffer JT, Swan D, Al Sallaq R, Magaret A, Johnston C, Mark KE, et al. Rapid localized spread and immunologic containment define Herpes simplex virus-2 reactivation in the human genital tract. *eLife*. 2013;2:e00288. 638
46. Byrne CM, Gantt S, Coombs D. Effects of spatiotemporal HSV-2 lesion dynamics and antiviral treatment on the risk of HIV-1 acquisition. *PLoS Computational Biology*. 2018;14(4):e1006129. 639
47. Schiffer JT, Swan DA, Roychoudhury P, Lund JM, Prlic M, Zhu J, et al. A Fixed Spatial Structure of CD8(+) T Cells in Tissue during Chronic HSV-2 Infection. *Journal of Immunology*. 2018;201(5):1522–1535. 640
48. van Baarle D, Hovenkamp E, Callan MF, Wolthers KC, Kostense S, Tan LC, et al. Dysfunctional Epstein-Barr virus (EBV)-specific CD8(+) T lymphocytes and increased EBV load in HIV-1 infected individuals progressing to AIDS-related non-Hodgkin lymphoma. *Blood*. 2001;98(1):146–155. 641
49. Moir S, Malaspina A, Ogwaro KM, Donoghue ET, Hallahan CW, Ehler LA, et al. HIV-1 induces phenotypic and functional perturbations of B cells in chronically infected individuals. *Proceedings of the National Academy of Sciences of the United States of America*. 2001;98(18):10362–10367. 642
50. Amadori A, Chieco-Bianchi L. B-cell activation and HIV-1 infection: deeds and misdeeds. *Immunology Today*. 1990;11(10):374–379. 643
51. Kersten MJ, Klein MR, Holwerda AM, Miedema F, van Oers MH. Epstein-Barr virus-specific cytotoxic T cell responses in HIV-1 infection: different kinetics in patients progressing to opportunistic infection or non-Hodgkin's lymphoma. *Journal of Clinical Investigation*. 1997;99(7):1525–1533. 644
52. Smets F, Latinne D, Bazin H, Reding R, Otte JB, Buts JP, et al. Ratio between Epstein-Barr viral load and anti-Epstein-Barr virus specific T-cell response as a predictive marker of posttransplant lymphoproliferative disease. *Transplantation*. 2002;73(10):1603–1610. 645
53. Heslop HE, Ng CY, Li C, Smith CA, Loftin SK, Krance RA, et al. Long-term restoration of immunity against Epstein-Barr virus infection by adoptive transfer of gene-modified virus-specific T lymphocytes. *Nature Medicine*. 1996;2(5):551–555. 646
54. Rooney CM, Smith CA, Ng CY, Loftin S, Li C, Krance RA, et al. Use of gene-modified virus-specific T lymphocytes to control Epstein-Barr-virus-related lymphoproliferation. *Lancet*. 1995;345(8941):9–13. 647
55. Rooney CM, Smith CA, Ng CY, Loftin SK, Sixbey JW, Gan Y, et al. Infusion of cytotoxic T cells for the prevention and treatment of Epstein-Barr virus-induced lymphoma in allogeneic transplant recipients. *Blood*. 1998;92(5):1549–1555. 648

56. Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nature medicine*. 2015;21(8):922–926. 680
57. Du W, Goldstein R, Jiang Y, Aly O, Cerchietti L, Melnick AM, et al. A Virtual B Cell Lymphoma Model to Predict Effective Combination Therapy. *Blood*. 2014;124(21):928–928. 681
58. Scaling up antiretroviral therapy in resource-limited settings: guidelines for a public health approach. Executive summary. April 2002. IAPAC Monthly. 2002;8(6):168–175. 682
59. Casper C, Krantz E, Selke S, Kuntz SR, Wang J, Huang ML, et al. Frequent and asymptomatic oropharyngeal shedding of human herpesvirus 8 among immunocompetent men. *The Journal of Infectious Diseases*. 2007;195(1):30–36. 683
60. Kimura H, Morita M, Yabuta Y, Kuzushima K, Kato K, Kojima S, et al. Quantitative analysis of Epstein-Barr virus load by using a real-time PCR assay. *Journal of Clinical Microbiology*. 1999;37(1):132–136. 684
61. Magaret AS, Wald A, Huang ML, Selke S, Corey L. Optimizing PCR positivity criterion for detection of herpes simplex virus DNA on skin and mucosa. *Journal of Clinical Microbiology*. 2007;45(5):1618–1620. 685
62. Gillespie DT. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*. 2001;115(4):1716–1733. doi:10.1063/1.1378322. 686
63. Lenormand M, Jabot F, Deffuant G. Adaptive approximate Bayesian computation for complex models. *Computational Statistics*. 2012;28(6):2777–2796. 687
64. Prangle D. Lazy ABC. *Statistics and Computing*. 2016;26(1):171–185. 688
65. Tran MN, Kohn R. Exact ABC using Importance Sampling. *ArXiv E-prints*. 2015;ArXiv:1509.08076. 689