Title

Intra- and Inter-cellular Modeling of Dynamic Interaction between Zika Virus and Its Naturally Occurring Defective Viral Genomes

Journal

Journal of Virology, 95(22)

ISSN

0022-538X

Authors

Sharov, Vadim
Rezelj, Veronica V
Galatenko, Vladimir V
et al.

Publication Date

2021-10-27

DOI

10.1128/jvi.00977-21

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

Intra- and Inter-cellular Modeling of Dynamic Interaction between Zika Virus and Its Naturally Occurring Defective Viral Genomes

Vadim Sharov,[a,g] Veronica V. Rezelj,[b] Vladimir V. Galatenko,[a*] Avi Titievsky,[a] Julia Panov,[a,f] Konstantin Chumakov,[d,e] Raul Andino,[c] Marco Vignuzzi,[b] Leonid Brodsky[a]

[a]Tauber Bioinformatics Research Center, University of Haifa, Haifa, Israel
[b]Institut Pasteur, Viral Populations and Pathogenesis Unit, Centre National de la Recherche Scientifique UMR 3569, Paris, France
[c]Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, California, USA
[d]Office of Vaccines Research and Review, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland, USA
[e]Global Virus Network, Baltimore, Maryland, USA
[f]Sagol Department of Neurobiology, University of Haifa, Haifa, Israel
[g]Laboratory of Genomic Research and Biotechnology, Federal Research Center Krasnoyarsk Science Center of the Siberian Branch of the Russian Academy of Sciences, Krasnoyarsk, Russian Federation

Vadim Sharov (computational pipeline) and Veronica V. Rezelj (experimental data and biological interpretations) contributed equally to this work. Author order was based on the contribution to the manuscript. Marco Vignuzzi and Leonid Brodsky are senior authors of this study.

ABSTRACT Here, we examine *in silico* the infection dynamics and interactions of two Zika virus (ZIKV) genomes: one is the full-length ZIKV genome (wild type [WT]), and the other is one of the naturally occurring defective viral genomes (DVGs), which can replicate in the presence of the WT genome, appears under high-MOI (multiplicity of infection) passaging conditions, and carries a deletion encompassing part of the structural and NS1 protein-coding region. Ordinary differential equations (ODEs) were used to simulate the infection of cells by virus particles and the intracellular replication of the WT and DVG genomes that produce these particles. For each virus passage in Vero and C6/36 cell cultures, the rates of the simulated processes were fitted to two types of observations: virus titer data and the assembled haplotypes of the replicate passage samples. We studied the consistency of the model with the experimental data across all passages of infection in each cell type separately as well as the sensitivity of the model's parameters. We also determined which simulated processes of virus evolution are the most important for the adaptation of the WT and DVG interplay in these two disparate cell culture environments. Our results demonstrate that in the majority of passages, the rates of DVG production are higher inC6/36 cells than in Vero cells, which might result in tolerance and therefore drive the persistence of the mosquito vector in the context of ZIKV infection. Additionally, the model simulations showed a slower accumulation of infected cells under higher activation of the DVG-associated processes, which indicates a potential role of DVGs in virus attenuation.

IMPORTANCE One of the ideas for lessening Zika pathogenicity is the addition of its natural or engineered defective virus genomes (DVGs) (have no pathogenicity) to the infection pool: a DVG is redirecting the wild-type (WT)-associated virus development resources toward its own maturation. The mathematical model presented here, attuned to the data from interplays between WT Zika viruses and their natural DVGs in mammalian and mosquito cells, provides evidence that the loss of uninfected cells is attenuated by the DVG development processes. This model enabled us to estimate the rates of virus development processes in the WT/DVG interplay, determine the key processes, and show that the key processes are faster in mosquito cells than in mammalian ones. In general, the presented model and its detailed study suggest in what important virus development processes

Citation Sharov V, Rezelj VV, Galatenko VV, Titievsky A, Panov J, Chumakov K, Andino R, Vignuzzi M, Brodsky L. 2021. Intra- and inter-cellular modeling of dynamic interaction between Zika virus and its naturally occurring defective viral genomes. J Virol 95:e00977-21. https://doi.org/10.1128/JVI.00977-21.

Editor Rebecca Ellis Dutch, University of Kentucky College of Medicine

Address correspondence to Marco Vignuzzi, marco.vignuzzi@pasteur.fr, or Leonid Brodsky, lbrodsky@research.haifa.ac.il.

* Present address: Vladimir V. Galatenko, Evotec International GmbH, Göttingen, Germany.

Received 15 June 2021
Accepted 26 August 2021
Accepted manuscript posted online 1 September 2021
Published 27 October 2021

the therapeutically efficient DVG might compete with the WT; this may help in assembling engineered DVGs for ZIKV and other flaviviruses.

KEYWORDS defective virus genomes, mathematical modeling, Zika virus

Zika virus (ZIKV) is an arthropod-borne virus transmitted to humans through the bite of infected mosquitoes. Although most infections are asymptomatic, the virus can cause symptoms ranging from flu-like illness with mild fever, body aches, and rashes to severe birth defects in pregnant women and acute-onset paralysis (1). ZIKV is an enveloped virus with a positive-sense, single-stranded RNA genome. Its genome is capped and contains 5′ and 3′ untranslated regions (UTRs) flanking a single open reading frame (ORF), which is translated into a single polyprotein, cleaved by host and viral proteases at the endoplasmic reticulum (ER) membrane into three structural proteins (C, PrM, and E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4, NS4B, and NS5). Our group has recently shown that noncanonical, deleterious forms of the ZIKV genome are naturally generated during virus replication *in vitro* and can be propagated with the virus population (2). These molecules, called defective viral genomes (DVGs), are generated through genetic recombination and lack some function(s) required for their autonomous replication and propagation, which in some cases can be provided by wild-type (WT) virus. How this close interplay between ZIKV DVGs and WT virus can influence intracellular dynamics has not been studied.

Our previous analysis of high-throughput genome sequencing data showed that a significant number of DVGs carrying a similar deletion appear naturally under high virus load conditions, suggesting that these DVGs are an active component of the ZIKV population that may influence infection kinetics (2). Mathematical modeling of the dynamics of viral infections is an efficient tool to dissect the mechanisms of infection processes. Insights from the mathematical model can then be used to develop and optimize antiviral strategies (3, 4). A conventional and widely used method for modeling the dynamics of viral infection is based on ordinary differential equations (ODEs), whose parameters are fitted to the experimental observations. Typically, the empirical target data consist of a virus titer that is observed over the timeline of infection. In this study, this classic approach was expanded in two aspects. First, the proposed ODE5 model takes into account an interplay between wild-type ZIKV and one of its naturally occurring DVGs on both the intracellular and between-cell levels. Second, taking advantage of next-generation sequencing (NGS) of virus genomes, we phylodynamically reconstructed a time course "demography" of the virus population in each passage. More specifically, the across-passage trajectory of the effective population size was fitted to sequences of haplotype genomes representing major ZIKV quasispecies in the corresponding passage (5). These two types of parameter fitting to data were combined to find sets of parameters that provide good fitting of the time course virus demography to the observed virus titer and matching phylogeny of the haplotype genomes assembled from the NGS data.

The suggested ODE5 model takes into account the processes of infection "between cells" when ZIKV and DVG particles bind to the cell membrane and release the viral genomes as well as within-cell processes of viral genome replication and packaging. This type of model provides the intracellular dynamics of viral genome replication together with the dynamics of the interaction between infected and healthy cells in a cell culture passage (6, 7). The goals of this *in silico* study were to understand

1. how WT-DVG ZIKV coinfection affects the replication and production of each infectious agent as well as the rate of transition of cells from an uninfected to a WT- and/or DVG-infected state;
2. what parameters of the model are critical, and what are their biologically meaningful ranges that achieve the observed coexistence of the two virus genomes;
3. what fitted parameters of the model are significantly different between WT and DVG infection processes in C6/36 (mosquito) compared to Vero (mammalian) cell cultures;

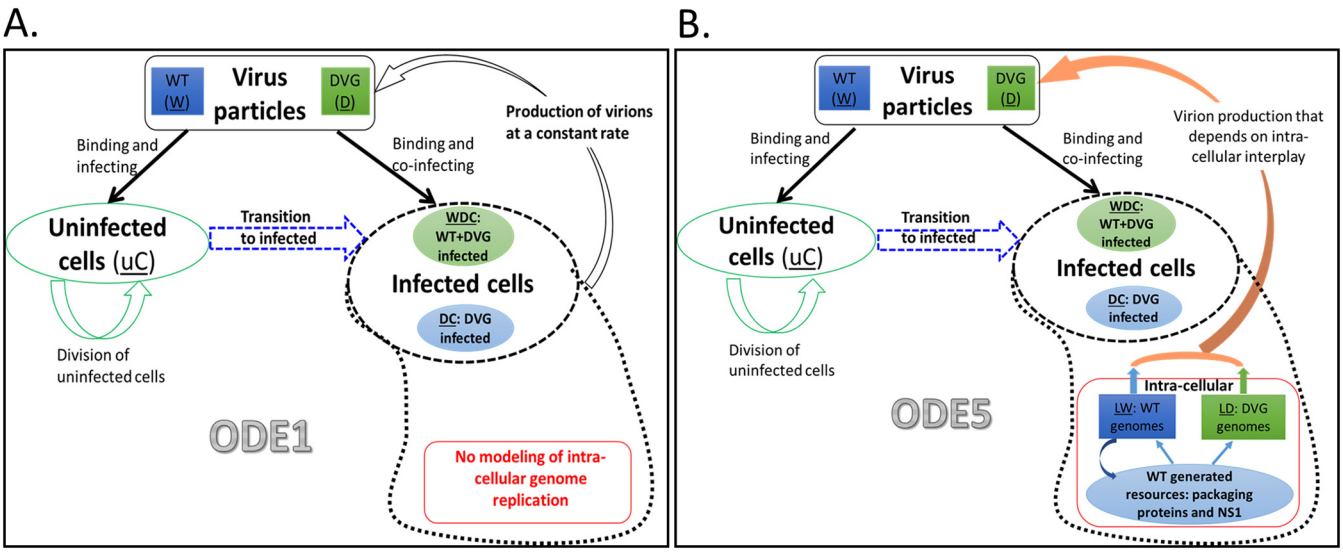

FIG 1 Schematic representation of the ODE models. (A) Flow chart of the ODE1 model for intercellular dynamics between cells and WT and defective viral genome (DVG) virus particles. The ODE1 model includes the following variables: uninfected cells (*uC*), cells infected with virus particles, and two types of virus particles. The two types of virus particles include virus particles with wild-type genomes (*W*) and particles with defective genomes (*D*). Cells infected by virus particles can be of two types, *WDC* or *DC*, where *WDC* stands for the number of cells infected by both wild-type and defective viruses and *DC* stands for the number of cells infected by DVGs only. We did not introduce WT-only-infected cells because a cell cannot be infected by only WT virus since defective genomes are generated as by-products of WT replication. (B) Flow chart of the ODE5 model. The ODE5 model introduces the intracellular processes, in which WT+DVG- and DVG-infected cells generate virus particles. The new variables introduced into this ODE5 model are as follows: the intracellular virus genome quantity (LW_{WD} [the total number of WT genomes in WT- or WT+DVG-infected cells] and LD_D and LD_{WD} [the total numbers of DVG-type genomes in *DC* and *WDC* cells, respectively]) and the quantity of packaging proteins and cotranslated NS1 protein (P_{WD} [the total amount of proteins in cells infected by both types of viruses]).

4. what ODE parameters ensure efficient inhibition of ZIKV infectivity; and
5. how parameter perturbations influence the time course patterns of the ODE solutions.

In the modeling presented below, instead of genetically engineering defective virus genomes, as was done in other models, we considered the naturally occurring DVGs of ZIKV that interact with the WT viruses. In addition, we took into consideration a series of 11 cell culture passages in a number of cell replicates during $11 \times 60 = 660$ h of infection in total. The solutions provided by the proposed ODE model with the parameters in biologically meaningful ranges showed a significant difference between C6/36 and Vero cells in the rates of processes associated with the production of DVG particles and their infectivity. These rates are higher in C6/36 than in Vero cells in the majority of passages/replicates of the experiments. We also found that even without taking into account the immune response of the cells, according to ODE5 simulations, the activation of the DVG-associated processes slows down the loss of uninfected cells.

RESULTS

We apply two modeling approaches to simulate the dynamics of wild-type ZIKV and its naturally occurring defective viral genome in arthropod and vertebrate cell cultures. The computational models were fitted to experimental data in every serial passage as well as for each cell type. One model (ODE1) assumes a constant rate of virus particle production by any cell. Another model (ODE5) takes into account intracellular processes of virus genome replication and particle formation. These intracellular processes vary across time depending on the abundance of each type of genome inside the cell as well as the cumulative storage of the supporting proteins thereby produced. A schematic representation of the simulated virus development processes in the two models is depicted in Fig. 1.

ODE1 model and its comparison with ODE5. A detailed description of the ODE1 model can be found in Supplemental Material #1. This model failed to accurately predict the experimental results under biologically meaningful values of its parameters.

Thus, from here on, we focus only on the ODE5 model and provide a detailed comparison of the results of the two models and their estimated parameters in Supplemental Material #4.

ODE5 model introduces the intracellular dynamics of virus production without cell resistance. The introduction of major intracellular virus-related processes into the model provided a deeper understanding of the interplay between WT and DVG-type viruses and improved the modeling of virion production in comparison to the previous ODE1 model, where a uniform rate of virus production was assumed (see Supplemental Material #4). In the ODE5 model that includes intracellular events, we considered a specific type of naturally occurring DVG that, as we have previously shown (2), is encapsidated and propagated with the virus population in both the Vero and C6/36 cell environments. This DVG lacks genes encoding the full set of packaging proteins and also lacks the cotranslated NS1 gene, which is crucial for the replication of the virus genome (2). Although these DVGs cannot self-replicate, complementation by WT virus enables their packaging and propagation within a virus population by providing functions that they have lost. Certain DVGs interfere with virus replication by competing for viral and/or host resources. The model does not assume any cell resistance to ZIKV infection because the Vero and C6/36 cell cultures that were used in the experimental studies do not have intact primary antiviral defense mechanisms in their respective environments, namely, the interferon or RNA interference (RNAi) response, respectively (8–10).

Equations for the ODE5 model can be found in Supplemental Material #2.

The model considers the following intracellular events: translation, replication, and packaging of virus genomes. All of these events are described in simplified terms without extensively going into detail. Intracellular processes in cells infected only with DVG virus are simpler due to the absence of the WT/DVG intracellular virus interplay. Figure 1B illustrates the flow of cells, virus particles, intracellular virus genomes, and proteins modeled by this ODE5 model.

Here, the uninfected cells (*uC*) proliferate with the maximum rate of b1, with decreasing cell multiplication in dependence to the total number of accumulated cells. This decrease of cell multiplication is regulated by the *C*_m constant and is conceptually similar to a logistic-like limit on the total number of cells (11). The uninfected cells (*uC*) in the model die at rate d1, and the infected cells of both types (*WDC* and *DC*) die at rate g1. The cells become infected with some likelihood after contact with the WT or DVG virions with total binding attempt rates k1 and k2, where 1/k1 and 1/k2 represent the probabilities of the successful completion of an infection for one binding attempt. A score of <1 for a contact [see ratios like $W/(W + uC)$] regulates the likelihood of an interaction between the virus particle and a cell: when the number of particles is substantially higher than the number of cells, the score is close to 1. Finally, a1 and a2 are the rates of production of WT and DVG virus particles by a cell infected by both virus types, respectively.

Cell division as well as translation and replication processes are modeled in ODE5 using a logistic-like limitation approach. Virus genomes and proteins in cells infected by both types of viruses (WT+DVG) are assumed to be uniformly distributed among cells. Thus, an individual cell infected by WT viruses is assumed to contain $\frac{LW_{WD}}{WDC}$ viruses and $\frac{P_{WD}}{WDC}$ of the packaging and NS1 proteins. The rate of free WT virus production of one cell is $a1 \cdot \frac{LW_{WD}}{WDC} \cdot \frac{P_{WD}}{WDC}$, and the term representing the total production of free virions by these cells is $a1 \cdot \frac{LW_{WD} \cdot P_{WD}}{WDC}$. It is similar for the production of the DVG-type particles by a WT+DVG-infected cell: $a2 \cdot \frac{LD_{WD} \cdot P_{WD}}{WDC}$.

When a WT virus particle infects a cell that is already infected with a DVG particle, these cells are reassigned from the pool associated with the DVG-infected cells to the pool associated with the cells simultaneously infected by viruses of both types. The transition pool associated with cells infected by DVG viruses contains no viral proteins due to the lack of replication and expression of genes from the DVGs in the absence of supporting WT genomes. The flow of cells associated with the transition of the genomic pool from the DVG-infected cells to cells infected by both types of viruses is expressed by the term $k1 \cdot \frac{W}{W+DC} \cdot LD_D$.

RNA and/or protein degradation as well as the death of infected cells reduce the

pools of intracellular genomes and proteins associated with these types of cells. The death of *WDC* cells infected by WT viruses is represented by the term $g1 \cdot WDC$, and the per-cell amount of packaging and NS1 proteins is $\frac{P_W}{WDC}$, so the term expressing the reduction of the total protein quantity in these cells due to cell death is $g1 \cdot WDC \cdot \frac{P_W}{WDC} = g1 \cdot P_W$. Taking into consideration unknown protein degradation, a rate of reduction in the protein pool is denoted by the term $g3 \cdot P_W$. A similar reduction of the pool of intracellular genomes due to cell death and RNA degradation is denoted by the terms $g2 \cdot LW_{WD}$, $g2 \cdot LD_{WD}$, and $g2 \cdot LD_D$.

Consistency of the ODE5 model with experimental data and comparison of the most fitted parameters in Vero and C6/36 cell cultures. In each passage of the Vero and C6/36 cell cultures, Markov chain Monte Carlo (MCMC) fitting of the model parameters provided a solution consistent with empirical observations of the virus titer and virus genome "phylogeny" (see "Fitting ODE solutions to the experimental data: estimation of model parameters" in Materials and Methods). The averages and standard deviations of the 10 most fitted parameters across all passages in two sets of Vero cell triplicates (VeroA and VeroB) and the C6/36 cell triplicate are specified in Table S1 in Supplemental Material #3. These parameters seem biologically adequate. For example, the model predicts that without virus infection, the estimated cell division rate (b1) results in the doubling of the cell number in approximately 20 to 30 h. This prediction is consistent with previously reported observations (12, 13).

For passage 6 (p6), taken as an example, the resulting solution curves for counts of cells, virus particles, intracellular virus genomes, and proteins are shown in Fig. S1 in Supplemental Material #3. In both cell cultures (Vero and C6/36), the predicted amounts of virus particles agreed well with the observed distribution of the virus titers across the passage replicates (see Fig. S1E and F in Supplemental Material #3). For the same solutions, the virus population "demographics" across time were also phylodynamically well fitted to the divergence of haplotype genomes assembled from NGS data in each replicate of every passage under the assumption that virus genome quasispecies came from the previous passage (see Fig. S1G and H in Supplemental Material #3).

Comparing the most fitted solutions of the ODE5 model to VeroA and VeroB with the solutions fitted to the C6/36 cell lines, we found substantial differences in four of the model parameters between the vertebrate (Vero) and invertebrate (C6/36) cell lines. First, principal-component analysis (PCA) of the most fitted parameters in two sets of Vero triplicates (VeroA and VeroB) and the triplicate of C6/36 cells yielded specificities of the r1 and r2 (rates of replication of the WT and DVG, respectively) and t1 (rate of WT translation) parameters, which were clearly distinct from other parameters based on principal component 1 (PC1) (Fig. 2A). PC2 loadings again gave separation of the r1, r2, and t1 parameters from all other parameters (see Fig. S2A in Supplemental Material 3). In addition, PC3 loading gave separation of all three above-mentioned parameters (r1, r2, and t1) and parameter a2, the rate of DVG production, as clearly separated from other parameters (see Fig. S2B in Supplemental Material #3). Second, differences in the values of these four model parameters between fitting to C6/36 cells and Vero cells were most evident in passages p4 and p10 (Fig. 2B and C). In these passages, the key parameters r1, r2, t1, and a2 tended to be higher in C6/36 cells than in Vero cells (Fig. 2D and E). This observation in two passages was supported by the average values of key parameters across all passages: the average parameters r1, r2, t1, and a2 across all passages were higher in C6/36 cells than in Vero cells (Fig. 2F). Together, these results suggested that the DVG viruses interfered more in the infection kinetics at the intracellular level in C6/36 cells than in Vero cells. DVGs could interfere with the replication of the full-length viral genomes (WT) by competing for essential replication machinery and thus causing the reduction in the production of WT virus. The parameters r1 and t1 determined the rates at which NS1 and coat proteins from WT genomes were produced and thus also defined the interfering effect of DVGs. Altogether, the higher rate of interference of DVGs with WT virus production could explain the persistence of arbovirus infection in invertebrate cells (2, 14, 15).

In studying the changes over time in ZIKV demography (i.e., changes in the sequences of haplotype genomes representing major ZIKV quasispecies of the virus population) in

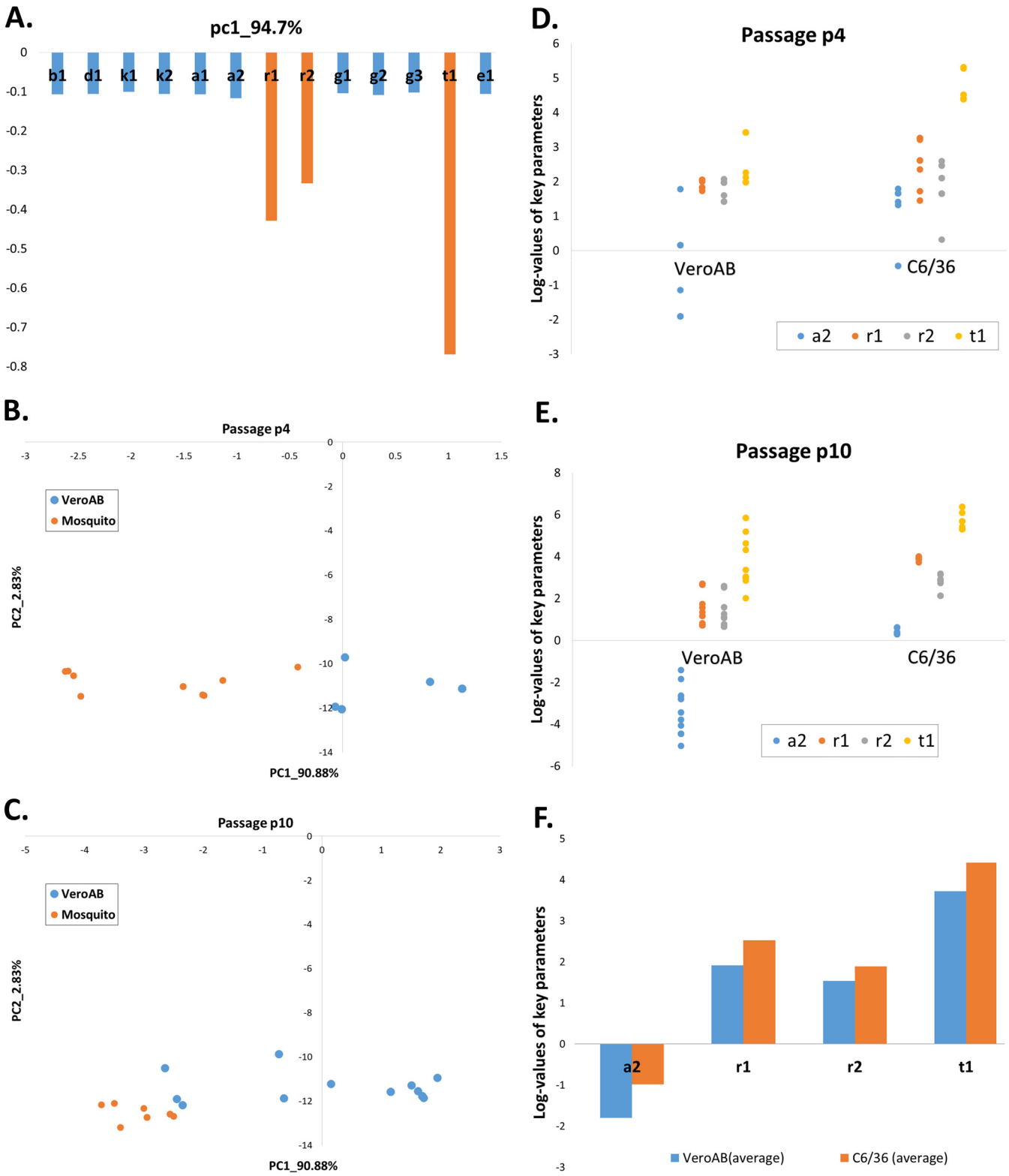

FIG 2 PCA-based detection of key parameters for model-to-data fitting in C6/36 and Vero cell lines. (A) The key parameters with the highest absolute loading on the PC1 component had the highest input in the multidimensional variability of parameter sets taken from the most fitted solutions in two Vero cell data sets (VeroA and VeroB [VeroAB]) and one C6/36 data set across all passages. The maximum loadings of key parameters are highlighted in orange. (B and C) In at least two passages (p4 and p10), the solutions (as their parameter sets) most fitted to the two Vero data sets (VeroAB) and the C6/36 virus titer data were separated in the space of key parameters based on PCs: PC1 was the most separating component. (D and E) Log values of key parameters across the most fitted solutions to the VeroAB and C6/36 data. (F) Log values of key parameters averaged across the best solutions in two cell cultures demonstrated higher values of the parameters in fitting ODE5 to observations in C6/36 cell culture than in Vero cell culture, indicating that the key virus adaptation processes were faster in mosquito cells.

each passage reconstructed by the ODE5 model, we observed slow evolution of quasispecies across passages: the number of haplotypes was slow growing across passages in both cell cultures.

An equilibrium state of uninfected cells cannot be achieved in ODE5 simulations by modulating only the DVG-associated parameters. From a practical perspective, one of the important questions that this mathematical model of ZIKV can help elucidate is whether changes to DVG-related parameters can alter infection dynamics and whether this will ensure that a significant amount of uninfected cells remain after viral infection ceases. This question is of special interest as it has been previously suggested that naturally occurring DVGs may act as interfering particles preventing the WT virus from infecting all host cells and thus providing an equilibrium virus-host interaction where the infection is prolonged but is not fatal to the host (2, 15, 16).

A study of the ODE5 model, however, shows that such an equilibrium state, when a significant level of uninfected cells is kept constant, cannot be achieved by altering only the DVG-related parameters while keeping other parameters unchanged. Almost all parameters of the model have to be adjusted to achieve the quasi-equilibrium solution fitted to a constant level of uninfected cells in the long run (see Fig. S5 in Supplemental Material #5).

ODE5 model: parameter sensitivity and interdependence. Another important question when studying mathematical models of virus infections is the question of the sensitivity of parameters. In determining how best to tackle the disease, it is necessary to know the relative importance of the different factors influencing its mechanisms of production and evolution. We performed regression-factorial analysis to explore the sensitivities of individual parameters and their pair interactions in a neighborhood of the most fitted parameter sets across passages for vertebrate (Vero) and invertebrate (C6/36) cell lines. The sensitivity is with respect to the deviation of the ODE solutions from the observations. Factorial variation of the ODE5 parameters was performed in a series of $3^5//27$ (F_{12}, F_{13}, F_{14}, F_{15}) partial designs (see "Sensitivity of model parameters and their pair interactions in the solution-to-data fitting" in Materials and Methods). The overall design allows several orthogonal estimations of all of the main effects and pair interactions and the subsequent evaluation of the sensitivity (statistical significance) of these effects.

Using this approach, we identified parameters with the most significant impact on the fitness of the model to the empirical data. ODE5 model fitness to Vero virus titer data was sensitive to two parameters: the rates of replication of the WT genome (r1) and the DVG (r2) (see Fig. S6A and B in Supplemental Material #6). Similarly to the Vero cell line, the fitness of the ODE5 model to the C6/36 cell line data (see Fig. S6C in Supplemental Material #6) was sensitive to the same two parameters: the rates of replication of the WT genome (r1) and the DVG (r2).

For pair interactions of the parameters (factors), we found that an interaction of the WT and DVG replication rates (r1:r2) as well as an interaction between WT virus genome replication and protein degradation (r1:g3) had significant impacts on the model fitness to empirical data from both cell cultures (see Fig. S6D and E in Supplemental Material #6). In C6/36 cell culture, in addition to these two interaction effects, an interaction of the rate of WT virus replication and the rate of cell infection (r1:k1) had a significant impact on ODE5 fitness to empirical data (see Fig. S6F in Supplemental Material #6).

Differences in the rates of DVG-associated processes between vertebrate (Vero) and invertebrate (C6/36) cell lines. We collected the log values of the parameters of the 10 most fitted ODE5 solutions to virus titer and haplotype observations from 11 passages in three replicates of the C6/36 cell experiments and in two randomly selected subsets of triplicates of the infected mammalian Vero cells (VeroA and VeroB). Thus, we collected 110 values (cases) of each type of parameter (b1, d1, k1, k2, a1, a2, r1, r2, g1, g2, g3, t1, and e1) in every replicate. Next, we categorized the log values of each parameter into three categories: the lower quartile, the upper quartile, and the two quartiles in the middle, which were united in one interval. For each parameter separately, in every category, we calculated a difference between the counts of the

parameter's cases in C6/36 and Vero cells belonging to this quartile category. In each category, differences in the counts of cases of a given parameter between C6/36 and Vero cells were z-normalized using the averages and standard deviations of the parameter's case counts from two studied Vero cell subsets of replicates in this quartile category. Figure S7 in Supplemental Material #7 depicts the normalized C6/36-Vero differences of counts of the parameter's cases in each category quartile across 13 parameter types. In the upper quartile (index "3" in the parameter-type notation), we observed substantial C6/36-Vero differences in the counts of cases belonging to this category for parameters of the DVG cycle-associated process:

- r1(3), the replication rate of WT genomes, which are "helpers" in producing DVGs and DVG particles
- r2(3), the replication rate of DVGs
- k2(3), the rate of cell infection by DVG particles
- a2(3), the rate of the DVG particle production by the WT+DVG-infected cells
- t1(3), the protein translation rates from WT genomes: structural and NS1 proteins are needed in the DVG production cycle but cannot be translated from DVG genomes

In the lower quartile of the parameter values, we found that C6/36 parameters for g1(1), which is the protein degradation rate, were significantly higher than the parameters for Vero cell lines.

In conclusion, the higher rates of the DVG-associated processes in C6/36 cells match our previous observation of higher abundances of DVGs with the specific deletion of structural proteins and the NS1 part of the virus genome in the C6/36 cell culture than in Vero cells (2).

Variations of time course behaviors of ODE5 solutions under parameter perturbations. (i) "Super-parameters" of the model and distinctive responses of DVG- and WT-associated groups of variables to super-parameter alterations. We studied the changes in the patterns of the time course trajectories of nine model variables in response to ±20% variation of the ODE5 model parameters with respect to the baseline data-fitted parameter values that were fitted to the Vero passage 6 data: passage 6 was taken as representing the robust infection process in the middle of the series of passages. The almost orthogonal factorial design (see "Analysis of time course patterns of ODE5 simulations under parameter perturbations" in Materials and Methods; see also Table S3 in the supplemental material) consisting of 417 combinations of the 20% up/down deviations of 13 model parameters from baseline values was used for simulating the corresponding 417 ODE5 solutions for nine model variables. These variables are as follows:

- number of uninfected cells (*uC*),
- number of cells infected with defective virus (*DC*),
- number of cells infected with either WT or WT+DVG virus (*WDC*),
- counts of either WT or defective virus particles (variables *W* and *D*),
- counts of intracellular genomes of WT and DVG types of viruses in two types of infected cells (LD_D [DVG genomes in *D*-infected cells], LD_{WD} [DVG genomes in WT+DVG-infected cells], and LW_{WD} [WT genomes in WT+DVG-infected cells]), and
- counts of translated viral proteins (P_{WD}).

By presenting nine simulated trajectories of each ODE5 solution as a distribution of the 9-mer vectors in 1,000 time points of the time course, we used the Kullback-Leibler relative entropy (Kullback-Leibler divergence [KLD]) measure as a distance between pairs of ODE5 solutions under parameter perturbations. Utilizing singular value decomposition (SVD) -based two-dimensional (2D) scaling, the 417 ODE5 solutions were presented as points in the 2D space of two SVD main components, where the distances between points reflect the Kullback-Leibler distances between 417 solutions (see "Analysis of time course patterns of ODE5 simulations under parameter perturbations" in Materials and Methods). Figure 3A depicts the positions of 417 solutions on the plane of SVD1 and SVD2 components: the most fitted to the Vero p6 virus titer solution is denoted "b1," and three solutions at the edges of the point distribution are denoted "b2," "b78," and "b186": the indexes indicate

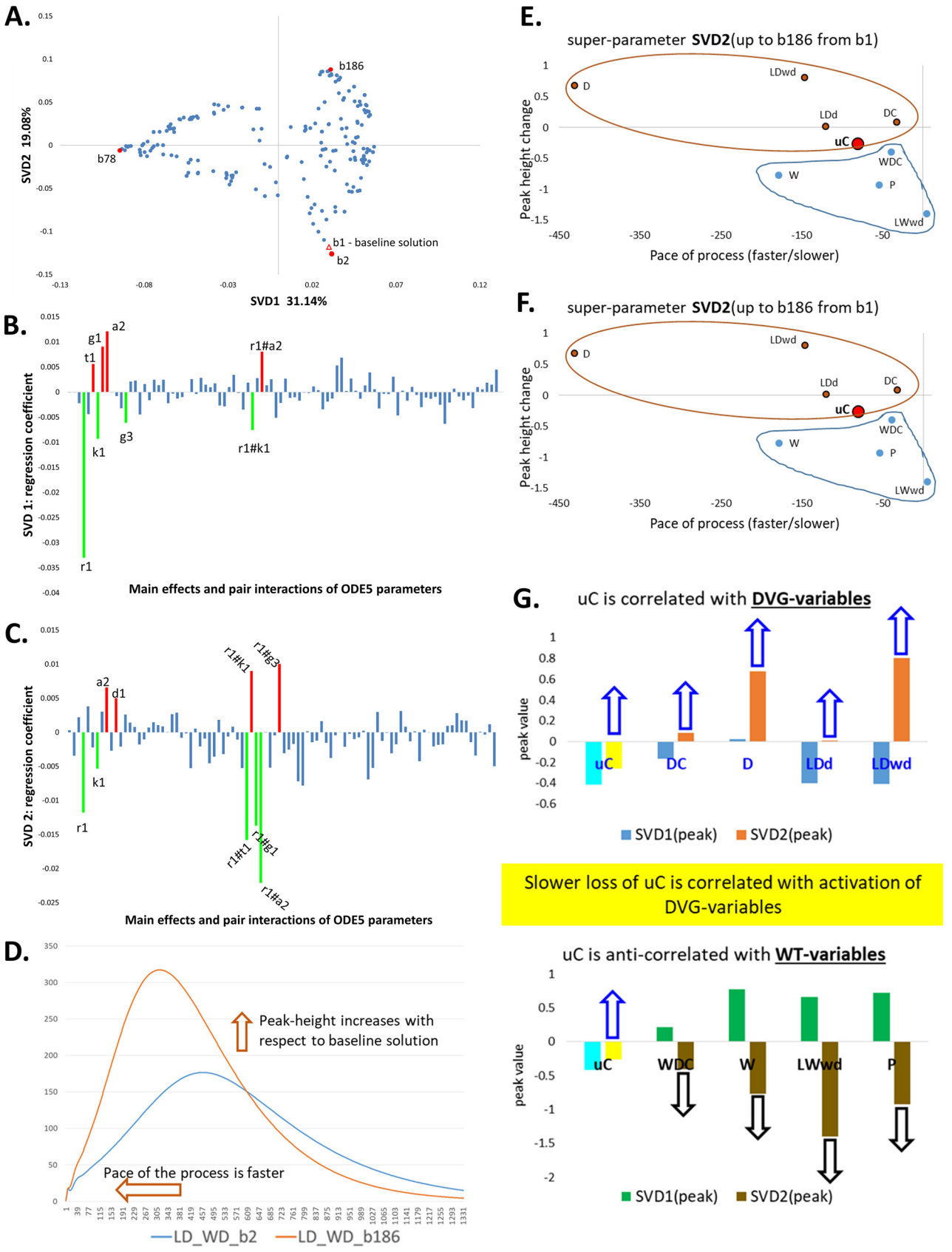

FIG 3 Analysis of ODE5 solutions under parameter perturbations. (A) Distribution of 417 time course ODE5 solutions on the SVD1-SVD2 plane. Solution "b1" is the baseline solution most fitted to the Vero passage p6 virus titer data. Solutions "b2," "b78," and "b186" are solutions at the

(Continued on next page)

TABLE 1 Changes in the pace and peak height of time course trajectories of nine solution variables

Variable	Change across SVD1 (pace)	Change across SVD1 (peak)	Change across SVD2 (pace)	Change across SVD2 (peak)
uC	−51	−0.41	−80	−0.26
DC	−77	−0.16	−32	0.08
WDC	−58	0.22	−39	−0.40
W	−229	0.77	−178	−0.77
D	0	0.02	−430	0.68
LD_D	−59	−0.41	−120	0.009
LW_{WD}	−12	0.66	5	−1.39
LD_{WD}	−19	−0.41	−146	0.8
P	−59	0.72	−54	−0.92

the number of the row in the parameter variation design (Table S3) that is associated with this solution.

The SVD1 and SVD2 components can be considered "super-parameters" of the ODE5 modeling because their values are specific to the time course behavior of the corresponding ODE5 solution vis-à-vis the behavior of the other solution. In order to link SVD1/SVD2 super-parameters with model parameters, the factorial linear regression equations were estimated separately for SVD1 and SVD2 as dependent variables of the regression against 13 model parameters and their pair interactions as independent variables of the regression. Since the covariation matrix of the design has almost equal values on its diagonal (see "Analysis of time course patterns of ODE5 simulations under parameter perturbations" in Materials and Methods), the values of the regression coefficients show inputs of the model parameters and their pair interactions into the approximation of two super-parameters: SVD1 and SVD2. The rates r1 (WT replication), k1 (infection of cells by WT particles), a2 (packaging of DVGs), t1 (translation of proteins), g3 (protein degradation), and g1 (death of WT+DVG-infected cells) together with pair interactions of r1 with other rates of this list make the highest impacts on SVD1/SVD2 values and, therefore, the positions of the ODE5 solutions on the super-parameter plane (Fig. 3B and C).

(ii) Interpretation of super-parameters via two main characteristics of time course solution behavior and distinctive responses of DVG- and WT-associated groups of variables to super-parameter alterations. The main characteristics of changes in the time course behavior of any variable of a solution are the increase/decrease of the peak height as a ratio of its change from the baseline peak height value as well as the increase/decrease of the pace of the process measured by changes in the time to reach the peak (Fig. 3D).

We calculated changes in the process pace and peak height of the time course behavior of nine solution variables while moving along SVD1 (from b1/b2 to b78) and along SVD2 (from b1/b2 to b186). These changes in the peak height and pace characteristics are presented in Table 1 and Fig. S8.

The transition along SVD1 provided a clear separation of WT-associated variables from DVG-associated variables in the pace peak coordinates (Fig. 3E). Indeed, the WT-associated variables (biological processes) are shifted up from the DVG-associated ones: peak heights of WT variables increase with respect to the b1/b2 solution along the SVD1 direction from b1/b2 to b78, and peak heights of DVG variables decrease in regard to b1/b2.

FIG 3 Legend (Continued)
edges of the point distribution on the SVD1-SVD2 plane. (B) SVD1 super-parameter as the linear combination of ODE5 parameters and their pair interactions. Denoted in red are the ODE5 parameters and their pair interactions that have the largest positive contribution to the SVD1 super-parameter. Denoted in green are the ODE5 parameters that make the largest negative contribution to the SVD1 super-parameter. (C) SVD2 super-parameter as the linear combination of ODE5 parameters and their pair interactions. Denoted in red are the ODE5 parameters and their pair interactions that have the largest positive contribution to the SVD2 super-parameter. Denoted in green are the ODE5 parameters and their pair interactions that have the largest negative contribution to the SVD2 super-parameter. (D) Main characteristics of alterations in the time course behaviors of each of nine solution variables: the pace of the process and peak height. (E) Transition of a position of the ODE5 solution along SVD1: clear separation of the WT-associated variables (denoted in blue) from the DVG-associated variables (denoted in brown) by their peak-height coordinates. (F) Transition of a position of the ODE5 solution along SVD2: clear separation of WT-associated variables (denoted in blue) from DVG-associated variables (denoted in brown) by their peak-height coordinates. (G) Changes in the rate of the loss of uninfected cells (*uC*) when moving from the transition along SVD1 to the transition along SVD2 (from the second column to the fourth column in Table 1) are correlated with changes in DVG variables and are inversely correlated with changes in WT variables.

The transition along SVD2 has an opposite effect (Fig. 3F). Namely, the WT-associated variables are shifted down from the DVG-associated ones: peak heights of WT variables decrease with respect to b1/b2 when moving along SVD2 from b1/b2 to b186, and peak heights of DVG variables increase with regard to b1/b2.

Thus, the dynamic pattern-based understanding of the super-parameters is as follows: with respect to the baseline solution, the SVD1 change to the left increases the peak heights of WT variables (correspondingly decreasing the peak heights of DVG variables), and the change of SVD2 influences the peak heights of the WT and DVG variables in the inverse direction.

A variable of the highest biological interest is "*uC*," the number of uninfected cells. The question is whether its time course behavior is linked either to DVG variables or to WT variables or whether it is independent of both. Table 1 shows that under a change of movement from along SVD1 to along SVD2 (from column 2 to column 4 in Table S3), a change in the half-peak value of *uC* is correlated with changes in the peak heights of *D*-variables. This correlation is depicted in Fig. 3G.

Therefore, one can conclude that the slower loss of the number of uninfected cells (*uC*) is correlated with the activation of DVG variables.

DISCUSSION

During the last several decades, mathematical modeling of the dynamics of viral infections has developed in several directions. In particular, multiscale models that include intracellular dynamics have been suggested (17, 18), as have models that simultaneously consider two types of virus infections (19). The most recent literature exploring the interplay between the WT virus and DVGs in cell cultures (20–22). model relatively short time spans of virus infection and consider only genetically engineered, and not naturally occurring, DVGs. In the ODE5 model presented above, we consider and monitor the naturally occurring DVGs of ZIKV that interact with the WT viruses. In addition, we take into consideration a series of 11 cell culture passages and many biological replicates during $11 \times 60 = 660$ h of infection in total.

Zika virus is an arthropod-borne virus and, as such, survives in nature in a transmission cycle between the invertebrate vector (mosquito) and a vertebrate host (e.g., humans). In our study, the same type of model was fitted to the development of two naturally interplaying variants of ZIKV in invertebrate and vertebrate cells (Vero and C6/36 cell lines). The comparison of the simulated rates of biological processes fitted to the empirical data from the two cell cultures allowed us to clarify which strategy of virus adaptation is typical in each cell type and to reveal how the interplay between the WT genome and naturally occurring DVGs influences these adaptation strategies.

The presence of defective viral genomes has been previously linked to persistent infection. Indeed, an arbovirus may not "want" to kill its vector but rather will establish persistent infection so that a mosquito can spread disease throughout its life span. DVGs have been shown to serve as the templates for the generation of viral DNA in insects (14), which in turn is pivotal for the establishment of persistent infection (13).

Genetic variability allows viruses to adapt to new environments. DVGs are part of the virus population as a whole and thus may influence virus adaptation. In parallel to fitting the ODE models to the observed virus titer data in each passage, we added the richness of the high-throughput NGS virus genome data sequenced in each passage as a target for the model fitting. This double-target fitting of the model in each passage allowed us to reconstruct the virus and cell time course "demography" from both perspectives: the number of virus particles and ongoing genomic mutation processes.

The predominant DVG type that we found in our previous work to accumulate following virus passaging involved DVGs with a deletion in the NS1 protein but conserving the open reading frame intact following the deletion site (2). NS1 is an essential component of the virus life cycle, forming a part of the viral replicase. The accumulation of a DVG lacking a protein with an essential role in replication can be explained by the ability of this protein to function in *trans*. Previous work with other flaviviruses such as Kunjin virus or yellow fever virus showed that NS1 is one of the few nonstructural genes that can be *trans*-complemented

for replication (23–25). In our previous work, we have confirmed that while this type of DVG is unable to replicate *per se*, *trans*-complementation by wild-type virus of the NS1 protein rescues DVG replication (2). Furthermore, we have also shown that the DVG can be packaged and form virions that are able to infect new cells. Thus, unlike a plethora of other DVGs that may form but cannot replicate or be packaged, by lacking NS1 while conserving the open reading frame of the rest of the replicase proteins, DVGs lacking NS1 are able to replicate in the presence of complementing wild-type virus and thus accumulate throughout passaging.

We stated that the intercellular dynamics presented in an ODE1 model is similar to those in previously described ODE-based coinfection models (19) and, in particular, models describing the interplay between standard (wild-type) and defective viruses at the intercellular level (21, 26). However, there are certain differences in comparison with previous studies by Frensing et al. (26) and Tapia et al. (21) associated with both cells and viruses. Taking into account the properties of Zika virus and the modeled time span, we did not include degradation of free virus particles or loss of infectivity (and, consequently, noninfectious virions) in our model. Furthermore, as models assume that cells infected by wild-type virus produce defective virus genomes, making such cells coinfected by wild-type and defective viruses, we did not consider a separate pool of cells infected by wild-type virus alone but combined these cells and coinfected cells into one pool. Due to differences in experimental setups, our model also does not include terms associated with the continuous feeding of cells and the dilution of the reactor content and replaces exponential cell proliferation by a model where proliferation becomes slower with an increase in the total number of accumulated cells. The latter allows splitting the cell proliferation rate and the cell death rate for uninfected cells into two separate parameters. Importantly, we also have a different type of term representing infection of cells by virus particles: the added denominator, in particular, prevents the growth of the cell infection rate to infinity when the number of virus particles is constantly small and the number of cells is increased.

If ODE1 is used as the whole model in simulations, then, similar to the studies by Frensing et al. (26) and Tapia et al. (21), it assumes a constant rate of virus particle production by infected cells. This assumption is generally incorrect. The virus particle production rate depends on the number of intracellular virus genome copies and the amounts of virus-related proteins, such as coat proteins, which vary over time in an infected cell. Shirogane et al. presented an ODE-based model describing the intracellular competition of wild-type and defective viruses (20), which we essentially used to define the ODE5 model by combining the ODE1 intercellular level with an additional, intracellular, modeling level. Similar to the model of Shirogane et al. (20), we introduced amounts of wild-type and defective intracellular genomes as model variables. To define model equations, we combined intra- and intercellular levels, and, in contrast to Shirogane et al. (20), who modeled infection in one "typical" cell, we considered total amounts of virus genomes for entire pools of cells, averaging these amounts over a number of cells in the pool. A similar approach was used for virus-related proteins, packaging and cotranslated NS1 proteins. The amounts of virus-related proteins played the same role in our combined intra- and inter-cell ODE5 model as the amounts of free capsids in the model of Shirogane et al. Instead of introducing "a generic set of resources" required for virus replication, we described virus replication by terms that decreased the replication rate (in comparison with the standard exponential replication model) with an accumulation of virus genomes by infected cells. The structure of these terms is similar to the structure of the term that describes the proliferation of uninfected cells in our two-level model.

The introduction of the intracellular level may be considered an analogue of time delays in inter-cell DDE models (27, 28). However, the intracellular level is more biologically relevant, giving deeper insight into the infection processes than with a delay-based modification of a standard one-level model. In particular, the introduction of the intracellular level leads to a nonconstant rate of virus particle production by infected cells, which is affected by the interplay between the WT virus and its DVGs.

We did not introduce cell resistance to infection to our ODE5 model as was done by

Shirogane et al. (20) because the Vero and C6/36 cell cultures that were used in the experimental study have impaired interferon and RNA interference responses, respectively. Obviously, cell resistance provides one more mechanism for the WT/DVG virus interplay. Indeed, irrespective of resistance, in the cells containing both WT virus and naturally occurring DVGs, competition for packaging (Pr-M-E) and NS1 (produced by WT virus but lacking in the DVG) proteins could result in a reduction of WT virion production. Importantly, competition would affect WT virion production only if the levels of Pr-M-E and/or NS1 produced by WT were below a threshold required for the efficient packaging and replication of both genomes. In the case of no resistance, DVGs in cells infected only by DVGs do not participate in the competition because (i) there is no WT genome to provide proteins lacking in the DVG and (ii) due to the lack of NS1, the DVG is unable to replicate and would thus degrade over time. It is possible that if resistance were incorporated into the model, it would make DVG infection of a cell a part of the competition since it potentially moves the uninfected nonresistant cells into the resistant cell pool, complicating their further infection by a WT virus. The impairment of a specific interferon pathway in Vero cells does not mean that they are not capable of any innate immunity response. There may be other mechanisms that may still be unknown that respond to virus infection and result in its restriction. The same applies to small interfering RNA (siRNA) mechanisms in insect cells.

The solutions provided by the proposed ODE model with the parameters in biologically meaningful ranges showed a difference between C6/36 and Vero cells in the rates of processes associated with the production of DVG particles and their infectivity. These rates are higher in C6/36 than in Vero cells in the majority of passages/replicates of the studied experiments. It is possible that the higher production of naturally occurring DVGs in the invertebrate cell environment regulates the virus-host equilibrium, contributing to tolerance and persistence in the mosquito vector. Mosquitoes have adapted to carry long-lasting viral infections without seemingly compromising survival, allowing them to transmit the virus throughout their life span. One may hypothesize that this DVG-based host-saving mechanism of virus infections is not adapted yet to vertebrate cells.

In this paper, we introduced two types of ODE-based multiscale mathematical models for the dynamics of viral infections. Both models simulate the interplay of two ZIKV forms: the wild type and one of the naturally occurring satellite defective virus genomes that lacks genes encoding the full set of packaging proteins and the cotranslated NS1 gene, which is crucial for the replication of the virus genome (DVG). One model (ODE1) simulates the WT/DVG infection dynamics on the intracellular level, where each infected cell produces WT and/or DVG virus particles at constant rates specific for each virus type (see "ODE1 model and its comparison with ODE5" in Results). The other model (ODE5), in addition to the intracellular level, introduces the highly interdependent intracellular replication dynamics of the two viruses (WT and DVG) (see "ODE5 model introduces the intracellular dynamics of virus production without cell resistance" in Results).

We demonstrate that the model that includes the production of virus particles controlled by intracellular processes is more consistent than the model with a constant rate of particle production: consistent with both the observations of virus titers across passages as well as NGS data (see "Consistency of the ODE5 model with experimental data and comparison of the most fitted parameters in Vero and C6/36 cell cultures" in Results). In addition, the ODE5 model parameters of the most fitted simulations remain in biologically meaningful ranges, unlike the ODE1 parameters (see Supplemental Material #4). In application to other types of viruses, this result indicates the importance of the intracellular replication dynamics for simulating virus development data.

Utilizing the sensitivity analysis, we determined several parameters as being crucial for fitting the ODE5 simulations to virus observations (see "ODE5 model: parameter sensitivity and interdependence" in Results). These crucial parameters are intracellular replication rates of the WT genome and the DVG (r1 and r2), which are dependent on each other in the fitted solutions. We also found that the rates of cell infection by WT particles (a1) and DVG ones (a2) coupled with the intracellular protein production rate from WT genomes (t1) are other

key parameters of the development of the WT+DVG system. Perhaps the most important result of the sensitivity analysis is the close interdependence of all five key parameters: their interaction effects substantially control virus development dynamics.

From the sensitivity study and analysis of parameter fitting to the viral titer and haplotype phylogeny observations across passages, we selected the most influential ("key") parameters of our ODE5 model, which are the intracellular rates of replication of the WT genome and the DVG (r1 and r2), the protein production rate (t1), and a rate of DVG particle production (a2). Next, we associated the factorially designed variations of all parameters with the positions of dynamic solutions on the SVD plane of mutual conditional entropy distances (Kullback-Leibler) between solutions. Regression dependencies of the solution's SVD positions from the model parameters showed that substantially the same key parameters and their pair interactions define the position of a solution on the SVD plane. Since we additionally linked the position of a solution on the SVD plane with dynamic characteristics of the solution curve (peak height and pace of the modeled process), we obtained associations of the key parameter modifications with visually assessed dynamic characteristics of the corresponding solution curves.

With respect to different cell cultures, we found that the rates of the DVG-associated processes are higher in the majority of passages in C6/36 cell culture experiments than in Vero cell lines. In the interplay between the helper (WT) and the satellite (DVG) viruses, which includes competition for resources within cells infected by both virus types, this modeling conclusion about the higher competitiveness of DVGs (satellite genomes) in mosquito C6/36 cells might explain the tolerance of the mosquito organism to Zika virus infection. This suggestion is in agreement with previous experimental studies showing that Sindbis virus DVGs serve as a template for virus DNA forms, thus modulating antiviral immunity through the RNAi pathway (15) and consequently driving persistence (14, 29). Therefore, our model predicts that a similar process is applicable to other viruses, such as Zika virus. Indeed, our previous work (2) has shown that experimental injection of mosquitoes with DVG RNA affects virus dissemination and transmission, suggesting that DVGs play an important role in modulating WT dynamics. Additionally, analysis of the ODE5 solutions under variation of parameter values (see "Variations of time course behaviors of ODE5 solutions under parameter perturbations" in Results) suggests that even without an immune response of the cells, the activation of the DVG-associated processes slows down the loss of uninfected cells. One may conclude that if they are present in high enough numbers, the naturally occurring ZIKV defective genomes could attenuate the pathogenicity of the virus.

We believe that the utilization of the model can provide insights into the dependence of the infection dynamics on the characteristics of the viruses (specifically, the presence of variants that interplay with the WT virus), which are important for the design of live prophylactic vaccines.

MATERIALS AND METHODS

Experimental data. The experimental data used here were related to virus titers throughout passages and NGS reads and were described in our previous study (2), which describes virus passages in both Vero and C6/36 cells.

The ZIKV strain used in this study was the prototype African MR-766 strain, derived from a previously described infectious clone (30). Wild-type ZIKV was recovered by transfecting HEK-293T cells (TransIT-LT1 transfection reagent; Mirus Bio) according to the manufacturer's instructions. At 4 days posttransfection (p.t.), the cell culture supernatant was clarified by centrifugation. Virus stocks were then prepared in Vero cells by infecting the cells with a multiplicity of infection (MOI) of 0.01 PFU/cell. At 5 days postinfection (p.i.), the cell culture supernatant was clarified by centrifugation. Virus titers were determined by plaque assays in Vero and C6/36 cells.

For the virus passages, ZIKV stocks were used to inoculate Vero or C6/36 cells at a low (0.01 PFU/cell) or high (20 PFU/cell) MOI. At 3 or 5 days p.i. (for Vero or C6/36 cells, respectively), the cell culture supernatant was collected, clarified by centrifugation, and used for the next passage. After passage 1 (p1), blind passaging was performed in 12-well plates, using 5 μl or 300 μl of the virus per well from the previous passage for low- and high-MOI conditions, respectively.

RNA from the clarified cell culture supernatants was extracted using Tri reagent (Sigma). Purified viral RNA was quantified using a Qubit fluorimeter (Invitrogen) and used for sequencing by next-generation RNA sequencing. Libraries were prepared using the NEBNext Ultra II RNA library preparation kit for Illumina (New

England BioLabs). The prepared libraries were loaded into a NextSeq500/550 mid-output kit v2.5 (Illumina) and sequenced in a NextSeq500 system for 151 cycles and 8 nucleotides of index.

Phylogenetic data: assembly of quasispecies haplotypes from raw NGS data. Raw sequencing reads were aligned to the ZIKV reference genome (GenBank accession no. KY989511.1) using the Bowtie2 algorithm (31). Next, assembly of the quasispecies haplotypes was performed by the ShoRAH procedure (32). The main idea of the procedure is to cluster the reads inside each position of the sliding window across multiple alignments of reads over the virus genome. The clusters are linked along the genome and give partial multiple alignments associated with the virus quasispecies. Haplotype sequences of the quasispecies used in the phylodynamics procedure were assembled from the most abundant partial multiple alignments.

Numerical solution of ODEs. Since the studied ODEs were stiff, the ODIENT boost package (33) implementation of the Rosenbrock method (34) was used for the numerical solution. Additionally, a number of standard numerical modifications were introduced into the ODEs. Namely, the regularization term ε was included in all denominators that could otherwise be arbitrarily close to zero, and cutoffs that prevent values from entering impossible zones (e.g., that prevent quantities from dropping below zero) were added.

Fitting ODE solutions to the experimental data: estimation of model parameters. Parameter estimation from fitting the ODE solution to the virus titer was performed by the MCMC method using the deBInfer R package (35). Also, we used phylodynamic PhyDyn software (5) for parameter estimation by MCMC fitting the ODE-simulated time course of virus demography to the divergence of the haplotype sequences. Each of the two parameter estimation procedures was passage specific and targeted to the corresponding data type, where the data were collected from triple replicates of the passage. Two sets of triplicates were randomly selected from 12 total Vero cell replicates in order to be compatible with 3 replicates in the C6/36 cell culture. From each passage in two cell cultures (C6/36 and Vero cells), 10 sets of parameters providing the best fit of the model to the virus titer and haplotype phylogenetics data were taken to the next steps of the analysis. The recovery of parameters was performed by the iterative use of both methods, one after another in each iteration, with the most fitted parameter set as a start in the alternative method or the next iteration, until a reasonable convergence of the parameter sets was reached.

Whenever logarithm is mentioned throughout this article, the natural logarithm is meant.

Sensitivity of model parameters and their pair interactions in the solution-to-data fitting. Factorial variations of the parameters were performed in a multidimensional neighborhood of the parameter set providing the most fitted solution to the observations. The factorial design with factors (parameters) on 3 levels was based on an orthogonal design for the main effects of five factors and four pair interactions of the first factor with the four other factors in 27 experiments: the design $3^5//27$ (F_{12}, F_{13}, F_{14}, F_{15}) adopted from the method of Brodsky (36). For the 13 parameters of the ODE model, the design was repeated several times. Namely, for each parameter as the first in the 5 sets of parameters, a subset of length 4 out of the other 12 parameters was combined with the current first one in the $p_1 + (p_2, p_3, p_4, p_5)$ $3^5//27$ design. Each design estimated the p_1p_2, p_1p_3, p_1p_4, p_1p_5 interaction effects and the main effects of the selected p_1, p_2, p_3, p_4, p_5 parameters on the deviation of the ODE solution from the data. A series of the $p_1 + (p_2, p_3, p_4, p_5)$ $3^5//27$ designs was constructed in order to obtain an estimation of each p_kp_m interaction effect ($k,m = 1, \ldots, 13$, $k \neq m$) and each $p_1, p_2, \ldots, p_{13}$ main effect. The constructed series of designs allowed estimating the main and pair interaction effects in several replicates for all 13 parameters of the most extended ODE5 model for all combinations of parameters on three levels, which is measured in the thousands. Indeed, the alternative full factorial experiment for 13 parameters, each on 3 levels, contained $3^{13} = 1{,}594{,}323$ parameter combinations in total.

Under the multiple almost-orthogonal design described above, the same effect was estimated several times, each time in the $3^5//27$ design for a specific partial factorial model. A distribution of all the estimated effects of the same parameter/pair interaction allowed assessing the statistical significance of the impact of the parameter or pair interaction on the fitting of the ODE solution to the empirical data.

Interdependence of model parameters in providing successful fitting to data. The interdependence of the model parameters in reaching a successful data approximation was evaluated as a pair interaction effect in estimating the sensitivity of parameters by factorial regression (36). The significant pair interaction effect specifies the interdependence of the corresponding two parameters in affecting an ODE solution fitted to the data.

Analysis of time course patterns of ODE5 simulations under parameter perturbations. The time course behavior of nine ODE5 variables can be presented as a distribution of 9-mer points, where each point consists of simulated values of the ODE5 variables at a given time point. With recording a simulation every 3.6 min during 60 h, the ODE5 simulation run was presented as a distribution of a thousand 9-mer data points.

A variation of ODE5 time course solutions under 20% alterations of 13 equation parameters was measured by Kullback-Leibler divergences (KLDs) between pairs of the 1,000-point distributions. Each distribution corresponded to an ODE5 solution simulated under a particular combination of the $\pm$20% changes in the equation parameters. Altogether, a design of 417 parameter combinations on two levels was used for generating ODE5 simulations presented as distributions of a thousand 9-mer vectors. In the factorial regression model with 13 factors, each on two levels, the factorial design allowed almost-orthogonal estimations for the impacts of the parameters and their pair interactions on the variations of any dependent variable (36). The KLD values between every two distributions of the thousand 9-mer vectors were calculated using the software described previously (37, 38).

Using SVD-based two-dimensional scaling (39), the symmetrized KLD values between ODE5 solutions were converted into distances between 417 points on the plane of two main SVD components. In this way, the SVD scaling of KLD values gave a position for each of the 417 ODE5-simulated solutions on two principal SVD components.

In the next step, we used the position of 417 ODE5 solutions on each SVD component separately as an independent variable in the factorial regression analysis with the 13 ODE5 parameters as factors. Thus, for each SVD component separately, the factorial regression analysis provided estimations of the contributions of parameters and their pair interactions to the predicted SVD component value. Therefore, two factorial regression approximations for two components gave a position of each ODE5 solution on the SVD plane.

We considered two SVD components as two "super-parameters" of the ODE5 model because they describe a major portion of the variability of the ODE5 time course patterns. These super-parameters are expressed as linear combinations of 13 original parameters of differential equations and $(13 \cdot 12)/2 = 78$ pair interactions of these parameters by the estimated regression equations.

z-score normalization of observations and parameter estimations. For normalizing measures with different variability ranges, we used the z-score of a quantity defined as the deviation from the expected quantity divided by the evaluated standard deviation of the variability of the corresponding measure.

Data availability. The NGS data set is available in the NCBI Sequence Read Archive under accession number PRJNA703982.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 1.9 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Stepan Nersisyan, Alexei Galatenko, and Vladimir Staroverov (Tauber Bioinformatics Research Center, University of Haifa, Israel, and Moscow State University, Moscow, Russia) for participation in establishing the ODE computational pipeline, valuable comments, and discussions.

This work was funded by the DARPA Intercept program managed by Seth Cohen and administered through DARPA cooperative agreement no. HR0011-17-2-0023 and through contract no. HR0011-17-2-0027 (the content of the information does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred). This work also received funding from the Laszlo N. Tauber Family Foundation, the Fondation de Recherche Médicale (FRM EQU201903007777), and Agence Nationale de Recherche Laboratoire d'Excellence grant ANR-10-LABX-62-IBEID.

REFERENCES

1. Bogoch II, Brady OJ, Kraemer MUG, German M, Creatore MI, Kulkarni MA, Brownstein JS, Mekaru SR, Hay SI, Groot E, Watts A, Khan K. 2016. Anticipating the international spread of Zika virus from Brazil. Lancet 387:335–336. https://doi.org/10.1016/S0140-6736(16)00080-5.
2. Rezelj VV, Carrau L, Merwaiss F, Levi LI, Erazo D, Tran QD, Henrion-Lacritick A, Gausson V, Suzuki Y, Shengjuler D, Meyer B, Vallet T, Weger-Lucarelli J, Bernhauerová V, Titievsky A, Sharov V, Pietropaoli S, Diaz-Salinas MA, Legros V, Pardigon N, Barba-Spaeth G, Brodsky L, Saleh M-C, Vignuzzi M. 2021. Defective viral genomes as therapeutic interfering particles against flavivirus infection in mammalian and mosquito hosts. Nat Commun 12:2290. https://doi.org/10.1038/s41467-021-22341-7.
3. Ribeiro RM, Lo A, Perelson AS. 2002. Dynamics of hepatitis B virus infection. Microbes Infect 4:829–835. https://doi.org/10.1016/S1286-4579(02)01603-9.
4. Boianelli A, Nguyen VK, Ebensen T, Schulze K, Wilk E, Sharma N, Stegemann-Koniszewski S, Bruder D, Toapanta FR, Guzmán CA, Meyer-Hermann M, Hernandez-Vargas EA. 2015. Modeling influenza virus infection: a roadmap for influenza research. Viruses 7:5274–5304. https://doi.org/10.3390/v7102875.
5. Volz EM, Siveroni I. 2018. Bayesian phylodynamic inference with complex models. PLoS Comput Biol 14:e1006546. https://doi.org/10.1371/journal.pcbi.1006546.
6. Reddy B, Yin J. 1999. Quantitative intracellular kinetics of HIV type 1. AIDS Res Hum Retroviruses 15:273–283. https://doi.org/10.1089/088922299311457.
7. Heldt FS, Frensing T, Reichl U. 2012. Modeling the intracellular dynamics of influenza virus replication to understand the control of viral RNA synthesis. J Virol 86:7806–7817. https://doi.org/10.1128/JVI.00080-12.
8. Desmyter J, Melnick JL, Rawls WE. 1968. Defectiveness of interferon production and of rubella virus interference in a line of African green monkey kidney cells (Vero). J Virol 2:955–961. https://doi.org/10.1128/JVI.2.10.955-961.1968.
9. Mosca JD, Pitha PM. 1986. Transcriptional and posttranscriptional regulation of exogenous human beta interferon gene in simian cells defective in interferon synthesis. Mol Cell Biol 6:2279–2283. https://doi.org/10.1128/MCB.6.6.2279.
10. Brackney DE, Scott JC, Sagawa F, Woodward JE, Miller NA, Schilkey FD, Mudge J, Wilusz J, Olson KE, Blair CD, Ebel GD. 2010. C6/36 Aedes albopictus cells have a dysfunctional antiviral RNA interference response. PLoS Negl Trop Dis 4:e856. https://doi.org/10.1371/journal.pntd.0000856.
11. De Boer RJ, Perelson AS. Target cell limited and immune control models of HIV infection: a Comparison. J Theor Biol 1998;190:201–214. https://doi.org/10.1006/jtbi.1997.0548.
12. HeLa Cell Line. 27 June 2020. HeLa cell culture information and resources. http://www.hela-transfection.com/cell-culture-information/.
13. Das AB, Loying P, Bose B. 2012. Human recombinant Cripto-1 increases doubling time and reduces proliferation of HeLa cells independent of pro-proliferation pathways. Cancer Lett 318:189–198. https://doi.org/10.1016/j.canlet.2011.12.013.
14. Goic B, Stapleford KA, Frangeul L, Doucet AJ, Gausson V, Blanc H, Schemmel-Jofre N, Cristofari G, Lambrechts L, Vignuzzi M, Saleh M-C. 2016. Virus-derived DNA drives mosquito vector tolerance to arboviral infection. Nat Commun 7:12410. https://doi.org/10.1038/ncomms12410.
15. Poirier EZ, Goic B, Tomé-Poderti L, Frangeul L, Boussier J, Gausson V, Blanc H, Vallet T, Loyd H, Levi LI, Lanciano S, Baron C, Merkling SH, Lambrechts L, Mirouze M, Carpenter S, Vignuzzi M, Saleh M-C. 2018. Dicer-2-dependent generation of viral DNA from defective genomes of RNA viruses modulates antiviral immunity in insects. Cell Host Microbe 23:353–365.e8. https://doi.org/10.1016/j.chom.2018.02.001.
16. Vignuzzi M, López CB. 2019. Defective viral genomes are key drivers of the virus-host interaction. Nat Microbiol 4:1075–1087. https://doi.org/10.1038/s41564-019-0465-y.
17. Rong L, Guedj J, Dahari H, Coffield DJ, Levi M, Smith P, Perelson AS. 2013. Analysis of hepatitis C virus decline during treatment with the protease

inhibitor danoprevir using a multiscale model. PLoS Comput Biol 9: e1002959. https://doi.org/10.1371/journal.pcbi.1002959.
18. Manna K, Chakrabarty SP. 2015. Chronic hepatitis B infection and HBV DNA-containing capsids: modeling and analysis. Commun Nonlinear Sci Numer Simul 22:383–395. https://doi.org/10.1016/j.cnsns.2014.08.036.
19. de Sousa BC, Cunha C. 2010. Development of mathematical models for the analysis of hepatitis delta virus viral dynamics. PLoS One 5:e12512. https://doi.org/10.1371/journal.pone.0012512.
20. Shirogane Y, Rousseau E, Voznica J, Rouzine IM, Bianco S, Andino R. 2019. Experimental and mathematical insights on the competition between poliovirus and a defective interfering genome. bioRxiv 519751. https://doi.org/10.1101/519751.
21. Tapia F, Laske T, Wasik MA, Rammhold M, Genzel Y, Reichl U. 2019. Production of defective interfering particles of influenza A virus in parallel continuous cultures at two residence times—insights from qPCR measurements and viral dynamics modeling. Front Bioeng Biotechnol 7:275. https://doi.org/10.3389/fbioe.2019.00275.
22. Koelle K, Farrell AP, Brooke CB, Ke R. 2019. Within-host infectious disease models accommodating cellular coinfection, with an application to influenza. Virus Evol 5:vez018. https://doi.org/10.1093/ve/vez018.
23. Lindenbach BD, Rice CM. 1997. *trans*-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. J Virol 71:9608–9617. https://doi.org/10.1128/jvi.71.12.9608-9617.1997.
24. Lindenbach BD, Rice CM. 1999. Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. J Virol 73:4611–4621. https://doi.org/10.1128/jvi.73.6.4611-4621.1999.
25. Khromykh AA, Sedlak PL, Westaway EG. 2000. *cis*- and *trans*-acting elements in flavivirus RNA replication. J Virol 74:3253–3263. https://doi.org/10.1128/jvi.74.7.3253-3263.2000.
26. Frensing T, Heldt FS, Pflugmacher A, Behrendt I, Jordan I, Flockerzi D, Genzel Y, Reichl U. 2013. Continuous influenza virus production in cell culture shows a periodic accumulation of defective interfering particles. PLoS One 8:e72288. https://doi.org/10.1371/journal.pone.0072288.
27. Banks HT, Bortz DM, Holte SE. 2003. Incorporation of variability into the modeling of viral delays in HIV infection dynamics. Math Biosci 183:63–91. https://doi.org/10.1016/S0025-5564(02)00218-3.
28. Alshorman A, Wang X, Joseph Meyer M, Rong L. 2017. Analysis of HIV models with two time delays. J Biol Dyn 11:40–64. https://doi.org/10.1080/17513758.2016.1148202.
29. Tassetto M, Kunitomi M, Andino R. 2017. Circulating immune cells mediate a systemic RNAi-based adaptive antiviral response in Drosophila. Cell 169:314–325.e13. https://doi.org/10.1016/j.cell.2017.03.033.
30. Schwarz MC, Sourisseau M, Espino MM, Gray ES, Chambers MT, Tortorella D, Evans MJ. 2016. Rescue of the 1947 Zika virus prototype strain with a cytomegalovirus promoter-driven cDNA clone. mSphere 1:e00246-16. https://doi.org/10.1128/mSphere.00246-16.
31. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
32. Zagordi O, Bhattacharya A, Eriksson N, Beerenwinkel N. 2011. ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. BMC Bioinformatics 12:119. https://doi.org/10.1186/1471-2105-12-119.
33. Ahnert K, Mulansky M. 2011. Odeint—solving ordinary differential equations in C++. AIP Conf Proc 1389:1586–1589. https://doi.org/10.1063/1.3637934.
34. Rosenbrock HH. 1963. Some general implicit processes for the numerical solution of differential equations. Comput J 5:329–330. https://doi.org/10.1093/comjnl/5.4.329.
35. Boersch-Supan PH, Ryan SJ, Johnson LR. 2017. deBInfer: Bayesian inference for dynamical models of biological systems in R. Methods Ecol Evol 8:511–518. https://doi.org/10.1111/2041-210X.12679.
36. Brodsky SZ. 2019. Mathematical foundations of the factorial design of experiments. Model Assist Stat Appl 14:121–173. https://doi.org/10.3233/MAS-190460.
37. Pérez-Cruz F. 2008. Kullback-Leibler divergence estimation of continuous distributions. Proc IEEE Int Symp Infect Theory 2008:1666–1670. https://doi.org/10.1109/ISIT.2008.4595271.
38. Dehban A. 18 May 2021. KL divergence for multivariate samples. GitHub. https://gist.github.com/atabakd/ed0f7581f8510c8587bc2f41a094b518.
39. Becavin C, Tchitchek N, Mintsa-Eya C, Lesne A, Benecke A. 2011. Improving the efficiency of multidimensional scaling in the analysis of high-dimensional data using singular value decomposition. Bioinformatics 27: 1413–1421. https://doi.org/10.1093/bioinformatics/btr143.