UC Irvine
UC Irvine Previously Published Works

Title

Effect of synaptic cell-to-cell transmission and recombination on the evolution of double mutants in HIV

Permalink

https://escholarship.org/uc/item/6v3376f3

Journal

Journal of The Royal Society Interface, 17(164)

ISSN

1742-5689

Authors

Kreger, Jesse
Komarova, Natalia L
Wodarz, Dominik

Publication Date

2020-03-01

DOI

10.1098/rsif.2019.0832

Peer reviewed

Effect of synaptic cell-to-cell transmission and recombination on the evolution of double mutants in HIV

Jesse Kreger[1], Natalia L. Komarova[1] and Dominik Wodarz[1,2]

[1]Department of Mathematics, and [2]Department of Population Health and Disease Prevention Program in Public Health, Susan and Henry Samueli College of Health Sciences, University of California, Irvine, CA 92697, USA

JK, 0000-0001-6438-171X; NLK, 0000-0003-4876-0343; DW, 0000-0002-8017-3707

Recombination in HIV infection can impact virus evolution *in vivo* in complex ways, as has been shown both experimentally and mathematically. The effect of free virus versus synaptic, cell-to-cell transmission on the evolution of double mutants, however, has not been investigated. Here, we do so by using a stochastic agent-based model. Consistent with data, we assume spatial constraints for synaptic but not for free-virus transmission. Two important effects of the viral spread mode are observed: (i) for disadvantageous mutants, synaptic transmission protects against detrimental effects of recombination on double mutant persistence. Under free virus transmission, recombination increases double mutant levels for negative epistasis, but reduces them for positive epistasis. This reduction for positive epistasis is much diminished under predominantly synaptic transmission, and recombination can, in fact, lead to increased mutant levels. (ii) The mode of virus spread also directly influences the evolutionary fate of double mutants. For disadvantageous mutants, double mutant production is the predominant driving force, and hence synaptic transmission leads to highest double mutant levels due to increased transmission efficiency. For advantageous mutants, double mutant spread is the most important force, and hence free virus transmission leads to fastest invasion due to better mixing. For neutral mutants, both production and spread of double mutants are important, and hence an optimal mixture of free virus and synaptic transmission maximizes double mutant fractions. Therefore, both free virus and synaptic transmission can enhance or delay double mutant evolution. Implications for drug resistance in HIV are discussed.

1. Introduction

Virus evolution *in vivo* is a central characteristic of human immunodeficiency virus (HIV-1) infection [1–3]. Viral evolutionary processes have been shown to drive disease progression through a variety of mechanisms, including evolution of immune escape or evolution towards virus strains with faster replication kinetics, increased cytopathicity and broader cell tropism [1]. A relatively high mutation rate of HIV-1 [4], together with a high turnover of the virus during the chronic phase of the infection [5–7], certainly contributes to the generation and emergence of mutants that drive this disease. These mutational processes are also implicated in the evolution of drug resistance during anti-viral therapy.

In addition to mutations, another mechanism that contributes to virus evolution is recombination [8–10]. HIV is a diploid virus containing two copies of genomic RNA. If cells are infected simultaneously by different virus strains [8], two different viral genomes can be packaged into the same virus particle. When this virus infects a new target cell, recombination between these two genomes can occur during reverse transcription, when the viral DNA is generated. Recombination has the potential to bring two separate point mutations together in a single virus genome that previously were present in different genomes. Recombination has been shown experimentally to play an important role in HIV-1 infection [10,11] in situations where the accumulation of two or more

mutations is required to achieve a given phenotypic effect. Examples are the generation of virus mutants that are simultaneously resistant against two or more drugs, or mutants that have escaped two or more immune cell clones.

The process of recombination requires the infection of cells with two or more viruses that are genetically different [8]. In HIV, the multiple infection of cells has been shown to be promoted by direct cell-to-cell transmission of the virus, through the formation of virological synapses [12–15]. Many viruses are transferred simultaneously from the source cell to the target cell, several of which can successfully integrate into the new host cell, making this an efficient mode of infection. Furthermore, experiments have shown that if cells are infected with two distinct virus strains, synaptic transmission promotes the repeated co-transmission of these different strains from one cell to the next [16,17], which can promote the occurrence of recombination. This was demonstrated both *in vitro* [16] and *in vivo* [17] using HIV-1 infection of humanized mice. *In vivo*, data also suggest that the process of synaptic transmission is spatially restricted, meaning that transmission likely occurs to neighbouring target cells [17].

The effect of viral recombination on the *in vivo* evolution of HIV has been investigated with mathematical models, revealing a wealth of results, in particular in the context of drug-resistant viruses. In [18], recombination was found to be detrimental to the doubly resistant virus. In [19], the role of recombination was reported to depend on the relative fitness characteristics of single and double mutants, but for most plausible scenarios it was established that recombination slowed down the evolution of resistance. In the models of [20,21], it was determined that recombination was beneficial for double mutants. In [22], it was clarified that the results strongly depend on the model formulation. In particular, a distinction was made between (i) population genetic (constant population) and population dynamic models, and (ii) stochastic and deterministic models. The model employed in [22] combines a population dynamic description with stochasticity, and finds that recombination decelerates the emergence of drug resistance.

In the present paper, we focus on the evolutionary dynamics of double mutant evolution in HIV infection, and how this is influenced by the mode of virus spread (synaptic versus free virus transmission) and the occurrence of recombination. Just as in [22], we use a stochastic, population dynamic model. In contrast to the above paper, however, we do not use a combined model where 'pre-treatment' and 'treatment' regimes are both included, but instead focus in more general terms on disadvantageous, advantageous and neutral mutants. We consider fitness landscapes that range from maximal positive to maximal negative epistasis, expressed by a parameter that ranges from zero to one. Times to double mutant invasion and the fraction of double mutants at defined time points are recorded in the presence and absence of recombination, and for a variety of different virus transmission strategies that range from 100% synaptic to 100% free virus transmission.

2. Modelling virus evolution

2.1. Stochastic modelling of spatially restricted synaptic virus transmission

Virus dynamics can be modelled by using ordinary differential equations (ODEs) [23,24]. Extensions of those models that include both free virus and synaptic transmission modes, as well as multiple infection have been since investigated, see [25–31].

In vivo data from humanized mice indicate that synaptic transmission results in spatial clusters of infected cells [17]. *In vitro* data suggest that synaptic transmission can result in productive infection of targets cells and result in high multiplicity of infection [32]. In order to explicitly include spatial dynamics of cell-to-cell transmission, we turn to a stochastic agent-based model. This includes both free virus and cell-to-cell transmission, and is adaptable to make either transmission process spatial or non-spatial. We consider a $\mathcal{N} \times \mathcal{N}$ two-dimensional grid, where each grid point can be empty, contain an uninfected cell, or contain an infected cell. Infected cells can contain any natural number of virus copies. For each time step, we randomly make $\mathcal{N}^2$ updates to the grid according to the following rules:

— empty grid points can become uninfected cells with probability λ;
— uninfected cells can die with probability d; and
— infected cells can die with probability a, infect another cell by free virus transmission with probability β, or infect another cell by cell-to-cell transmission (with S copies of the virus) with probability γ. During the infection processes, a target spot is chosen randomly either from the entire grid or the local neighbourhood. If that target spot contains a susceptible cell (uninfected or already infected), the infection event proceeds, otherwise it is aborted.

We assume that synaptic transmission can only occur to one of the eight nearest neighbours, while free virus transmission can occur to any cell on the grid. Basic simulations of this model can be seen in electronic supplementary material, figures S3 and S4.

It is straightforward to extend the agent-based model to include two virus strains that compete for the same target cell population (also see corresponding ODEs (5-6) of electronic supplementary material, §1). In this setting, a cell can be infected by i copies of virus strain A and j copies of virus strain B. If a cell containing both virus strains is chosen for an infection event, the probability to transmit a given virus strain is proportional to the fraction of the strain among all viruses in the cell if the two strains are neutral with respect to each other. If the two strains have different replication rates, the fitness difference is implemented during the infection event, which can correspond to different rates of reverse transcription. That is, an infecting strain is again chosen randomly with a probability that is proportional to its fraction in the cell. A disadvantageous/advantageous mutant would then have a lower/higher probability to infect the chosen target cell upon infection. In this way, a strain's fitness is independent of whether or not it is contained within a coinfected cell. This method of modelling fitness is one choice among many, others are explored further in [25]. Additional assumptions could, in principle, be included here, such as complementation or inhibition among viruses within the same cell. Due to the complexity of the dynamics considered, however, our aim was to first study those in a simpler setting where such higher level interactions do not occur. Those more complex scenarios can be explored in future work, based on the detailed insights that are generated in the current paper.

In the neutral case, drift is observed with the eventual fixation of one of the virus strains. If the two virus strains have

different fitness, the strain with the larger basic reproductive ratio [23] wins. Both of these cases can be seen in electronic supplementary material, figure S3.

2.2. Mutations and recombinations

We consider a virus population that can mutate at two different sites, denoted by a and b. Simulations are started with unmutated wild-type cells, ab. Single-mutant viruses (Ab or aB) can be generated during infection by point mutations, which occur with a probability μ per site. Each single-mutant can, in turn, mutate further to give rise to a double mutant AB. Note that a wild-type virus can directly mutate into a double mutant with a probability μ^2 if both sites mutate during the same reverse transcription event. The model also takes into account back-mutations, which again occur with a probability μ during an infection event. All the possible mutation events are illustrated in electronic supplementary material, figure S1.

Apart from mutations, however, a double mutant can also be generated through the recombination of different single-mutant viruses. This is implemented as follows. When viruses from a given source cell are chosen to infect a target cell, two virus genomes are randomly chosen with a probability that is proportional to the fraction of their abundance in the cell. The first virus genome that is chosen is the template from which reverse transcription is initiated. If no recombination occurs, reverse transcription is assumed to proceed on this genome only. Recombination is assumed to occur with a probability ρ. In this case, the reverse-transcribed virus is assumed to be a recombinant, the identity of which depends on the two infecting genomes. Electronic supplementary material, figure S2 list all recombination events that can occur.

There are two recombination processes in particular that are important: (i) $Ab + aB \rightarrow AB$ with probability $\rho/2$ (or ab with probability $\rho/2$), and (ii) $ab + AB \rightarrow Ab$ with probability $\rho/2$ or aB with probability $\rho/2$. These processes capture two roles of recombination that have been previously discussed in the literature [19]. Recombination between two single mutants can promote the generation of the double mutant, but recombination can also break up a double mutant upon recombination with the wild-type virus.

2.3. Simulations of the model and parameter values

We initialize the infection by randomly and uniformly spreading an equilibrium number of infected cells across the grid. These cells are singly infected with the wild-type. We used a mutation rate of 3×10^{-5} [4] and a recombination rate $\rho = 0$, 0.1, 0.2 and 0.5. Most of the parameters of this system are unknown. The average lifespan of productively infected cells is around 2 days[1] [7], and the basic reproductive ratio (R_0) of HIV (and SIV) has been estimated to be around 8 [33,34]. In our simulations, the fitness of viruses varies, depending on whether the virus is wild-type, a one-hit mutant, or a two-hit mutant. Therefore, we chose parameters (provided in figure legends) such that the base-line $R_0 \sim 5$, which is in the correct order of magnitude. For advantageous mutants (single and double), the value of R_0 increases, depending on the assumptions about the fitness landscape. For disadvantageous mutants, the value of R_0 decreases, depending on the nature of the fitness landscape. R_0 calculations are given for the free virus transmission scenario

(mass action) in electronic supplementary material, §1. For the spatial scenario (synaptic transmission), the expression for R_0 is currently not worked out, but for the same parameters is lower than for mass action. We assume that the uninfected cell death rate is half of the infected cell death rate, so that the average lifespan of uninfected cells is around 4 days. The lifespan of susceptible T cells *in vivo* has been shown to be heterogeneous, with lifespans ranging from a few days to several weeks, depending on the subpopulation under consideration [35].

In general, the multiplicity of infection in such models depends on virus load. For non-spatial, free virus transmission, the number of cells infected with i viruses correlates with the ith power of the singly infected cell population. In spatially structured models including synaptic transmission, the number of multiply infected cells correlates linearly with the number of singly infected cells. For the parameter values considered here, the average equilibrium multiplicities were as follows. In regimes where the basic reproductive ratio of the virus is around 8, the average multiplicity of infection in cells lies between 4 and 14, depending on how prevalent synaptic transmission is assumed to be (free virus transmission only leads to an average MOI of 4, while synaptic transmission only results in an average MOI of 14). Widely varying estimates for average infection multiplicities have been published [9,36–38], and there is some uncertainty about that. While some of these papers suggest that the above-quoted MOI range is too high, reference [38] showed that in a minority subset of T cells, up to 175 viruses were transmitted, likely due to synaptic infection processes. This might imply a relatively large number of integration events in such cells, even if many of the transmitted viruses fail. To investigate scenarios in which the average infection multiplicity is on the lower end (between 1 and 3, depending on viral transmission mode), we modified the model to track time since infection and assumed that the probability of superinfection declines over time due to receptor down-modulation [29]. This is described in electronic supplementary material, §4.

To investigate the relative contribution of free virus transmission (β) and cell-to-cell transmission (γ) we ran the model for different combinations of $\beta + \gamma = c$, where c is a fixed constant, ranging from purely synaptic to purely free virus transmission. The average outcome of the simulations was determined, including the average generation rate of double mutants, the average fraction of double mutants at a specific time point, and the time until the double mutant population grew to 90%.

3. Generation and spread dynamics of the double mutant

We will present all results for a range of transmission mode combinations, ranging from 100% synaptic transmission to 100% free virus transmission. In this section, however, we will mainly discuss under what fitness landscapes and assumptions recombination generally promotes the presence of double mutant populations, and when it works against them. The subsequent section will then discuss in more detail how these basic patterns are modulated by synaptic versus free virus transmission.
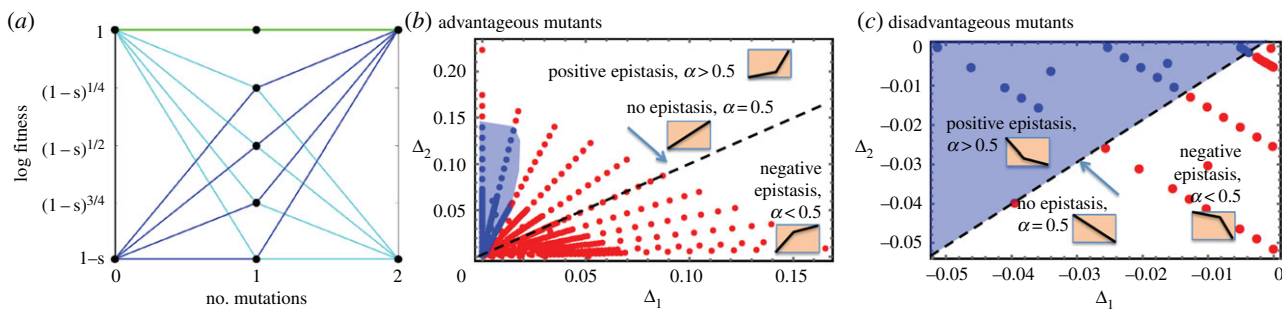
Figure 1. Summary of different scenarios. (a) Examples of fitness landscapes used in the simulations neutral (green), disadvantageous (cyan) and advantageous (dark blue). Without epistasis the single-mutant fitness is given by $f = (1-s)^{1/2}$. For negative and positive epistasis examples in the figure, it is given by $(1-s)^{1/4}$ and $(1-s)^{3/4}$, respectively. For the extreme form of positive epistasis, single-mutant fitness is the same as that of wild-types, $1-s$. (b,c) Role of recombination for different fitness landscapes. The horizontal axis is Δ_1 and the vertical axis is Δ_2, which are the relative log fitness values of single and double mutants, respectively. Each of the dots corresponds to a particular fitness landscape. (b) Advantageous mutants: red dots correspond to runs in which recombination accelerated double-hit mutant invasion to 90%, while blue dots indicate that recombination slowed down invasion. The boundary between the regions represents where there are no significant results one way or another. (c) Disadvantageous mutants: red dots indicate that recombination increased the double mutant fraction at $T = 10^5$, while blue dots mean that recombination reduced the double mutant fraction. Blue shading marks the regions where recombination suppresses double-hit mutants. The dashed black line corresponds to the cases of no epistasis ($\alpha = 0.5$) and separates the regions with positive epistasis ($\alpha > 0.5$) and negative epistasis ($\alpha < 0.5$). For both (b,c), we fixed the probability of free-virus transmission at 40% ($\beta = 0.04$); the rest of the parameters are as in figure 2. The determination of whether recombination suppressed or enhanced double-hit mutants was made by a statistical comparison of the averages over many runs, using the t-test.

3.1. Fitness landscapes and epistasis

Our investigation will span a variety of fitness landscapes, including neutral, advantageous and disadvantageous mutants. Let us assume that a mutation in site A or B results in an identical change in the fitness of the virus. Then, possible fitness landscapes can be separated into three groups for both advantageous and disadvantageous mutants [19]: negative epistasis, no epistasis and positive epistasis, see figure 1a for examples of these.

Note that each of the landscapes with advantageous mutants can be written as a triple of numbers,

$$((1-s), (1-s)^\alpha, 1), \tag{3.1}$$

which represent fitness values of the wild-types, one hit mutants and double mutants respectively. Here, $s > 0$ measures the amount of advantage, and α represents epistasis. We have $\alpha > 1/2$ for positive epistasis, $\alpha < 1/2$ for negative epistasis, and $\alpha = 1/2$ for no epistasis landscapes. Define the relative (log) fitness of the one-hit mutants compared to that of wild-types, $\Delta_1 = \ln(1-s)^\alpha - \ln(1-s) = (1-\alpha)|\ln(1-s)|$, and the relative (log) fitness of the two-hit mutants compared to that of one-hit mutants, $\Delta_2 = \ln(1) - \ln(1-s)^\alpha = \alpha|\ln(1-s)|$. Note that the sum of the two coordinates, $\Delta_1 + \Delta_2 = |\ln(1-s)|$ represents the relative log fitness of the two-hit mutants compared to the wild-types.

Similarly, each of the landscapes with disadvantageous mutants presented in figure 1a can be written as a triple of numbers,

$$(1, (1-s)^\alpha, (1-s)), \tag{3.2}$$

where $s > 0$ measures the amount of disadvantage. The relative (log) fitness of the one-hit mutants compared to that of wild-types, $\Delta_1 = \ln(1-s)^\alpha - \ln(1) = \alpha\ln(1-s)$. The relative (log) fitness of the two-hit mutants compared to that of one-hit mutants, $\Delta_2 = \ln(1-s) - \ln(1-s)^\alpha = (1-\alpha)\ln(1-s)$. Again, the sum of the two coordinates, $\Delta_1 + \Delta_2 = \ln(1-s)$, represents the relative log fitness of the two-hit mutants compared to the wild-types.

3.2. Advantageous mutants

A reasonable measure of double mutant success is the time it takes for the double-hit mutant to reach 90% of all infected cells. The following factors trade-off to determine whether recombination boosts or suppresses double mutant spread: (i) recombination between single mutants increases the rate at which double mutants are generated; (ii) recombination between double mutants and wild-type can break apart double mutants; (iii) the strength of selection of the double mutant defines how long the previous two factors are at play.

The net effect of recombination depends on the degree of the selective advantage, parameter s. For stronger advantages (larger values of s), recombination reduces the time to double mutant invasion (figure 2b). For lower selective advantages (lower s), however, recombination increases the time to double mutant invasion (figure 2a). The stronger the selective advantage, the quicker the double mutants spread at the expense of the wild-type, and the less likely it is that detrimental recombination events with the wild-type virus occur. This is illustrated with specific realizations of the stochastic dynamics in electronic supplementary material, §3 and figure S6.

The selective advantage threshold below which recombination slows double mutant invasion depends on the nature of the fitness landscape, in particular the value of α. This is summarized in figure 1b. The horizontal axis is Δ_1 (fitness difference between single-mutant and wild-type) and the vertical axis is Δ_2 (fitness difference between double and single mutants). Each point in this coordinate system corresponds to a unique fitness landscape. The red colour means that recombination events promote double mutant invasion, and blue means that they suppress this process. This picture has been composed by assuming 40% free virus transmission and 60% synaptic transmission. Arrays of points radially fanning out of the origin correspond to landscapes with the same level of epistasis (the same value of α) but different selection strength (the closer to the origin, the lower s). We observe that for any level of epistasis, for sufficiently high fitness advantage, recombinations are advantageous for double-hit mutants. As we decrease fitness s, however, there comes a point where recombinations no longer enhance
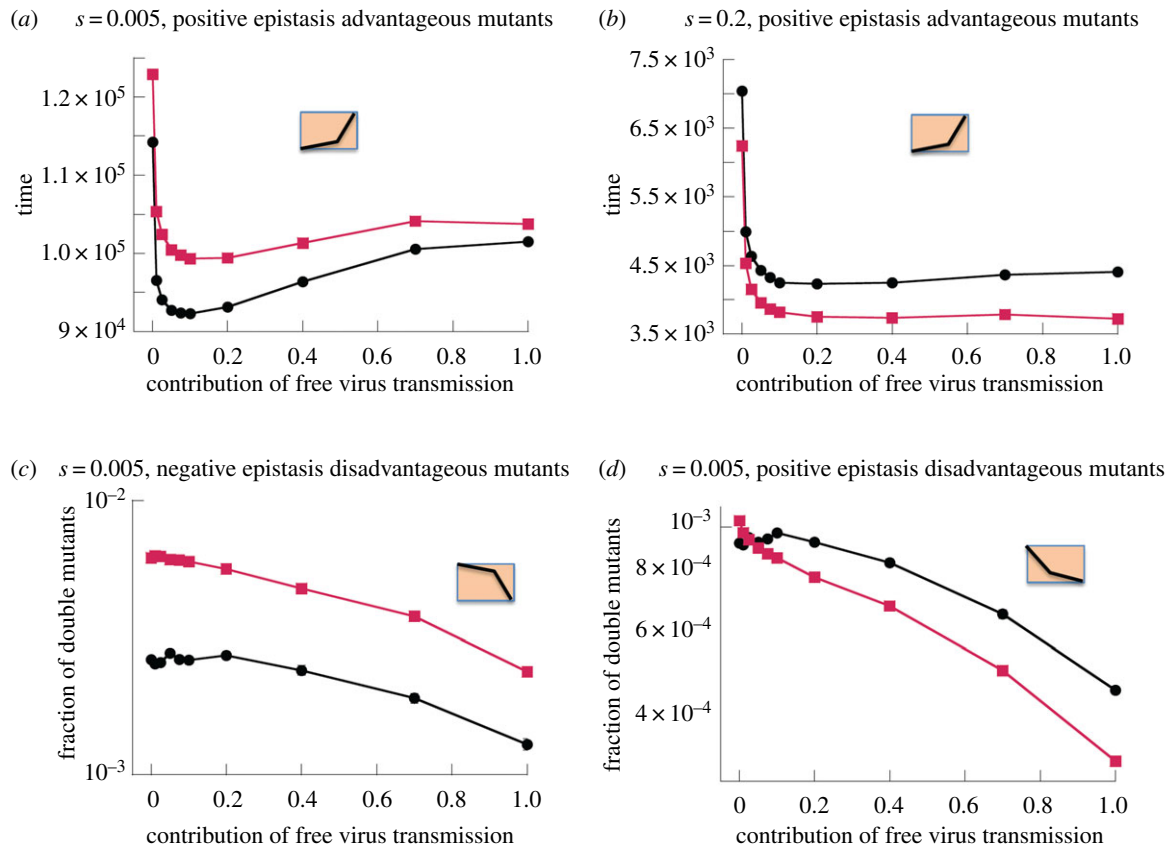
Figure 2. The role of recombination in (*a,b*) advantageous and (*c,d*) disadvantageous mutant dynamics; fitness landscapes are shown schematically in the insets. Red: with recombinations, and black: without recombination. (*a,b*) The time until the advantageous mutant reaches 90%, as a function of the fraction of free virus transmission. The means and standard errors are shown. For equivalent plots with standard deviations, see electronic supplementary material, §5.3. (*a*) $s = 0.005$, $\alpha = 0.75$, (*b*) $s = 0.2$, $\alpha = 0.75$. (*c,d*) The fraction of disadvantageous mutants at time $T = 10^5$, as a function of the fraction of free virus transmission. The means and standard errors are shown. (*c*) $s = 0.005$, $\alpha = 0.25$ (*d*) $s = 0.005$, $\alpha = 0.75$. The parameters are: $\beta + \gamma = 0.1$, $S = 3$, $\lambda = 1$, $d = 0.01$, $a = 0.02$, $\mathcal{N} = 100$, $\mu = 3 \times 10^{-5}$. All averages are based on at least 10^4 simulations.

double mutants but instead suppress them. In other words, any radial line will enter the blue region if it is sufficiently close to the origin. Recombinations can suppress the double mutant population significantly even for relatively large fitness advantages (large value of s) if α is relatively large and converges to one, i.e. for large positive epistasis (points close to the vertical axis in figure 1*b*). For lower values of α (weaker positive epistasis, no epistasis and negative epistasis), however, the transition happens for progressively smaller values of s. For large negative epistasis, the transition happens for very small values of s (for example, calculations show that for $\alpha = 0.25$, the blue region starts at about $s \approx 10^{-5}$, which is too small to see clearly in the figure and irrelevant for practical purposes, because such mutants are effectively neutral and take on average very long times to rise). The intuitive explanation for these observations is that lower values of α result in a more pronounced fitness advantage of single-hit mutants compared to wild-type virus. This, in turn, results in a faster exclusion of the wild-type virus population, and thus reduces the chances that recombinations break the double mutants. Hence, the parameter regime in which recombinations have a net negative effect on the double mutant population becomes more restrictive.

3.3. Disadvantageous mutants

A selective disadvantage leads to competitive exclusion in the absence of mutational processes. In the presence of mutational processes, disadvantageous mutants on average persist at an equilibrium level determined by the balance between mutation and selection. Hence, we determined the average fraction of double mutants at a time when this equilibrium has been reached, for different combinations of synaptic and free virus transmission (electronic supplementary material, §3.2).

Recombination increases the double mutant population at the selection–mutation balance for negative epistasis (figure 2*c*), but tends to reduce it for positive epistasis if a sufficient amount of free virus transmission is assumed to occur (figure 2*d*). Similar results have been reported in the context of HIV drug resistance evolution [22]. If most virus transmission, however, occurs through the synaptic route, figure 2*d* suggests that the opposite becomes true: now, recombination can increase the mutant levels for positive epistasis as well. This will be explored in more detail below.

These trends are further illustrated in figure 1*c*, assuming that a mixture of free virus and synaptic transmission occurs: as the parameter α is increased, the effect of recombination on the equilibrium level of double mutants changes from beneficial to detrimental. For the particular mixture of synaptic and free virus transmission chosen in this figure, recombination increases double mutant levels for negative epistasis (red region), and suppresses double mutant levels for positive epistasis (blue region). An increase in the parameter α results in a lower fitness of single mutants relative to the wild-type virus. This results in a higher prevalence of the wild-type virus, and thus in higher chances for the wild-type to recombine with and break apart the
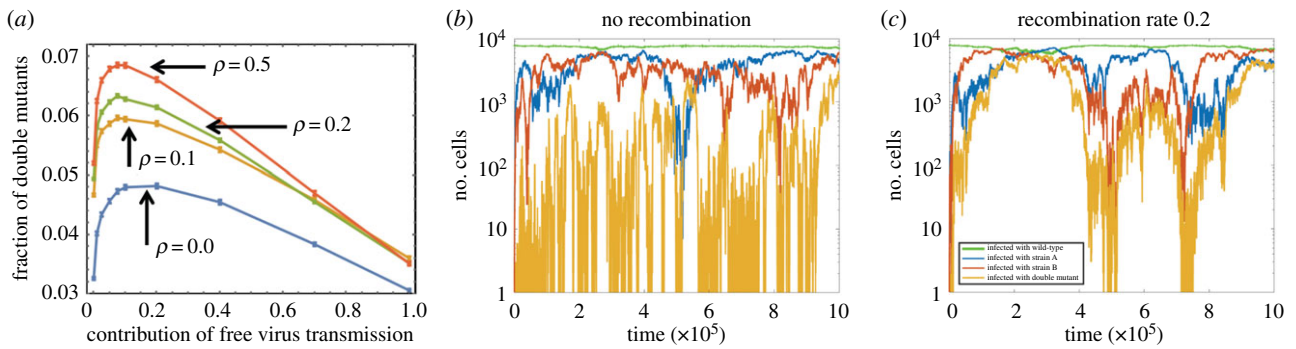
Figure 3. Neutral mutants. (a) The fraction of mutants after $T = 10^5$ steps. The horizontal axis is the fraction of free virus transmission. Four values of ρ are presented from $\rho = 0$ (no recombinations) to $\rho = 0.5$ (maximal recombinations). The means of at least 40 000 runs at each location and standard errors are shown. (b,c) The dynamics of cell populations, typical time series: (b) no recombinations, (c) with recombinations; we used $\beta = 0.04$ and $\gamma = 0.06$ (40% free virus transmission and 60% synaptic transmission). The other parameters are: $S = 3$, $\lambda = 1$, $d = 0.01$, $a = 0.02$, $\mathcal{N} = 100$, $\mu = 3 \times 10^{-5}$, $\rho = 0.2$.

double mutant; see electronic supplementary material, §3.3 for an ODE approximation of these results.

3.4. Neutral mutants

It follows from the above analysis that recombination delays the drift of neutral mutants towards dominance. Consider very weakly advantageous mutants in figure 1b. We can see that the origin is contained in the blue region, that is, as the selective advantage $s \to 0$, we expect recombinations to delay the rise of double mutants.

For neutral mutants, however, the rise to dominance will take a very long time. Interestingly, different results are obtained if we look at the fraction of double mutants at an early time point T relative to when the double mutant strain has reached its average equilibrium. Figure 3a shows that recombination increases the fraction of double mutants at time T. This can be understood by considering the early versus long-term dynamics of neutral mutants. In the long-run, the populations will converge to a state where all four virus strains fluctuate around comparable fractions. This steady state is the same whether recombination occurs or not. The speed with which the double mutant rises towards this steady state, however, is influenced by the occurrence of recombination (figure 3b,c). Initially, the populations of single mutants are generated by mutations and rise by drift. In the absence of recombinations, double mutants are created and destroyed by mutations and also experience drift (figure 3b). In the presence of recombinations, however, double-hit mutants initially enjoy positive selection due to relatively frequent recombination events between complimentary single hit mutants (which greatly outweigh the 'breaking' recombination events of the double mutants with the wild-type, due to the low levels of the former population). This can be seen in figure 3c. Once the levels of double-hit mutants increase, however, the 'making' and 'breaking' recombination events begin to balance each other and the dynamics return to neutral.

4. Mode of viral transmission and the effect of recombination on double mutant populations

The last section examined under what fitness landscapes recombination promotes or hinders the existence of double mutants.

For advantageous and neutral mutants, these results remain robustly independent of the mode of virus transmission (electronic supplementary material, §5.1). For disadvantageous mutants, however, we noted that results can change if most virus transmission is assumed to be synaptic. Figure 2 showed that while for smaller values of α (negative epistasis, figure 2c), recombination leads to an increase in double mutant levels, for large values of α (positive epistasis, figure 2d), the opposite occurred and recombination reduced the double mutant levels. At the same time, however, figure 2d indicated that if most virus transmission occurs through the synaptic pathway, recombination remains helpful for the double mutant population even for positive epistasis. This is explored in more detail in figure 4, which plots the equilibrium level of a disadvantageous mutant as a function of the parameter α for both extreme transmission modes: 100% free virus and 100% synaptic. If only free virus transmission occurs (figure 4a), recombination increases the double mutant fraction for $\alpha < 0.5$ (negative epistasis), while it decreases it for $\alpha > 0.5$ (positive epistasis). By contrast, if only synaptic transmission occurs (panel (b)), recombination always increases the number of double mutants, regardless of the value of α, although the double mutant levels in the presence and absence of recombination become practically indistinguishable for large values of α (strong positive epistasis). Electronic supplementary material, figure S11 contains further simulations showing the robustness of these patterns for different levels of mutant disadvantage. Similar patterns hold for lower infection multiplicities (electronic supplementary material, §4). While for lower multiplicities, the equilibrium fraction of double mutants can still be slightly reduced by recombination for purely synaptic transmission, this reduction is much less than in the presence of only free virus transmission, thus confirming the protective effect of synaptic transmission even in the low multiplicity scenario. Therefore, if positive epistasis is present, as is suggested for drug resistance mutations in HIV [39], a prevalence of synaptic transmission can protect against the negative effects of recombination on the level at which drug-resistant mutations pre-exist before the start of treatment.

The intuitive explanation for the detrimental effect of recombination on the double mutant population at larger values of α was given in the previous section: for larger values of α, the fitness of single mutants relative to wild-type viruses becomes lower. This leads both to a slower rate of double mutant production, and to a higher prevalence of wild-type viruses that
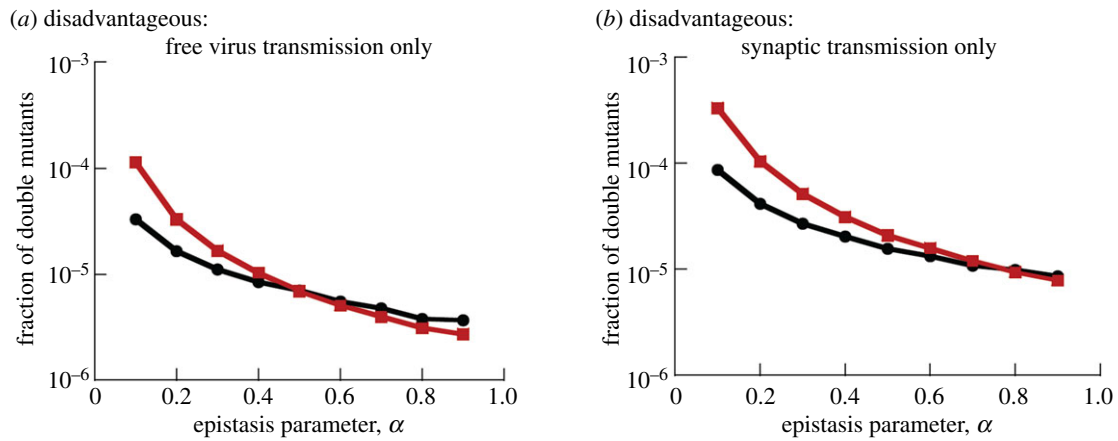
Figure 4. The role of recombinations under different transmission modes, for disadvantageous mutants. Shown is the temporal average of the fraction of double mutants at selection–mutation balance, as a function of that parameter α, defining the nature and extent of epistasis. Red denotes simulations with recombination and black without recombinations. (a) Free virus transmission only, (b) synaptic transmission only. $s = 0.05$, and other parameters are as in figure 2. Standard errors are too small to see. For equivalent plots with standard deviations, see electronic supplementary materials, §5.3.

can recombine with the double mutant and break it. If most of virus transmission occurs through virological synapses, however, the spatially restricted virus spread that is assumed to occur with this transmission mode results in the generation of single and double mutant 'clusters' or 'islands'. Single-mutant islands protect them from being outcompeted, resulting in larger numbers and thus a higher rate of double mutant generation. Double mutant islands isolate them from contact with wild-type virus, which prevents those detrimental recombination events from occurring. These dynamics are similar to the effect of 'mutant islands' discussed in [40,41].

5. Mode of viral transmission and the rate of double mutant emergence

All simulations were performed for varying combinations of synaptic and free virus transmission, yet we have so far not discussed the effect of this itself on the emergence of the double mutant population. A number of factors trade-off to determine what combination of synaptic and free virus transmission is optimal for the double mutant population. On the one hand, synaptic transmission results in the simultaneous transfer of multiple viruses from the source cell to the target cell, which increases the rate at which mutants are generated, and increases the rate of co-transmission of genetically different viruses, which in turn promotes the occurrence of recombination. On the other hand, if synaptic transmission is spatially restricted, as indicated by data [17], the rate at which the number of infected cells increases is slower under this mode of transmission, and it is less likely that genetically different strains come together in the same cell. For the different mutant types, the net effect is as follows:

Disadvantageous mutants. For disadvantageous mutants, more synaptic transmission tends to increase the equilibrium levels of double mutants at the selection–mutation balance (figure 2c,d; electronic supplementary material, figure S9a,b for the model with limited multiplicity). The main driving force responsible for the abundance of double mutants is production. This is maximized by synaptic transmission, because under this mechanism, there are more possibilities for mutations. Spread to higher levels is not an important force for disadvantageous mutants.

Advantageous mutants. In the case of advantageous mutants, the rate of double mutant invasion tends to be increased by free virus transmission, and purely synaptic transmission results in the slowest rate of invasion (figure 2a,b). The reason is that in this scenario, the spread of the double mutant from low to high numbers is the driving process, and this is slower for synaptic transmission, which is assumed to be spatially restricted [17]. While increasing the contribution of free virus transmission generally speeds up mutant invasion, this trend can weaken or reverse for larger fractions of free virus transmission, which can result in a shallow optimum, see figure 2a,b and electronic supplementary material, figure S9c–e for the model with limited multiplicity. The reason is that in the absence of significant synaptic transmission, fewer overall infection, and hence reverse transcription, events occur, which delays mutant production.

Neutral mutants. In the neutral case, and also for very weakly advantageous and disadvantageous mutants, a mixture of both free virus and cell-to-cell transmission maximizes the fraction of cells infected with the double mutant (figure 3a; electronic supplementary material, figure S9f for the model with limited multiplicity). This result is similar to what was observed in electronic supplementary material, §2, where the generation time of double-hit mutants was studied. Here, we observe that this holds even in the absence of recombination, and is more pronounced. The reason is that while more synaptic transmission results in the simultaneous transfer of multiple viruses, and hence in more chances to mutate, it also slows down the increase of the infected cell population due to the assumed spatial restriction. In this scenario, both production and spread play important roles.

6. Discussion and conclusion

We aimed to comprehensively analyse the effect of recombination on double mutant evolution in the context of HIV, depending on the details of fitness landscapes and the assumptions about the mode of viral spread (relative importance of synaptic versus free virus transmission). This is different from previous approaches, which focused more specifically on the evolution of drug resistance in HIV in
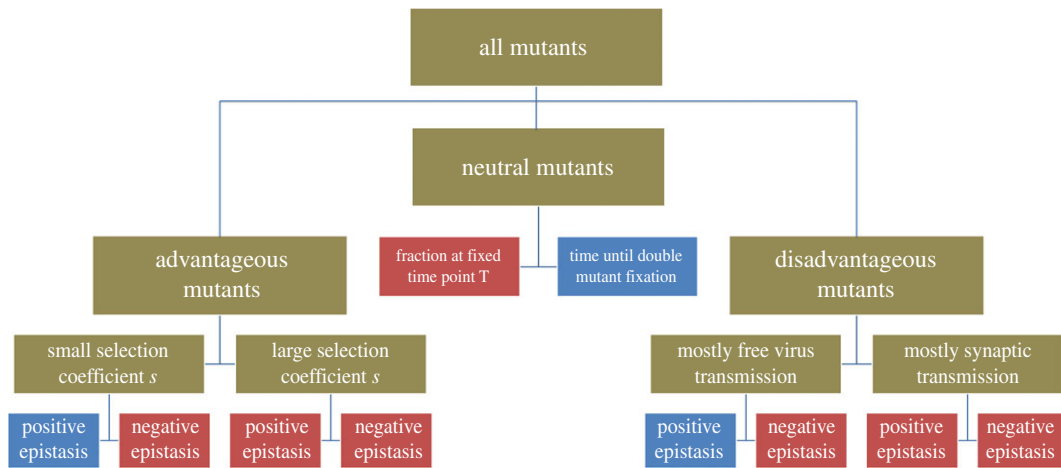
Figure 5. An overview of the main results. Here, red boxes correspond to situations in which recombination accelerated double mutant evolution, whereas blue boxes indicate that recombination slowed down double mutant evolution.

the context of only free virus transmission, and concentrated on specific fitness landscapes characterized by positive or negative epistasis. Our approach characterized the fitness landscape by the parameter α, which could be continuously varied from 0 to 1, thus capturing all fitness landscapes ranging from negative to positive epistasis for advantageous and disadvantageous mutants. The constraint in our fitness landscapes was that the two different single-hit mutants were assumed to have identical fitness. Another constraint of our analysis was that certain parameters of the system were kept constant throughout this analysis, such as the production rate and death rate of target cells or the death rate of infected cells, due to the complexity of the scenarios considered. We did perform selective simulations assuming different values for these parameters (see §5.4; electronic supplementary material, table S1) and did not find a qualitative change in the results, but an exhaustive analysis of the entire parameter space was not feasible. A tree showing an overview of the main results can be seen in figure 5.

The opposing effect of recombination to make and break double mutants played out as follows in the model analysed here: for advantageous mutants, recombination largely accelerates double mutant invasion except for cases of very strong positive epistasis with an intermediate fitness advantage of the mutants, or in cases where the fitness advantage becomes relatively low. The mode of viral spread does not modulate these patterns. If the mutants are disadvantageous, however, the mode of virus spread can significantly influence the effect of recombination on the equilibrium level of double mutants at selection–mutation balance. If the contribution of free virus transmission to virus spread lies above a threshold, recombination increases the double mutant population for negative epistasis, but decreases it for positive epistasis. If the dominant mode of virus spread is synaptic transmission, however, the negative effect of recombination for positive epistasis is greatly reduced, indicating a protective effect on the persistence of disadvantageous double mutants. In fact, for higher multiplicities, recombination increases double mutant levels even for positive epistasis if synaptic transmission is the dominant mode of virus spread. Finally, for neutral mutants, we observed that recombination always delays the rise of double mutants to dominance, but at the same time increases double mutant fractions measured at relatively early time points in the dynamics. Interestingly, neutral

double mutant dynamics in the presence of recombination are characterized by an 'advantageous' initial growth phase before converging to neutral drift, which explains the positive effect of recombination on early double mutant fractions.

These findings have implications for the pre-existence of multi-drug-resistant HIV mutants before the start of therapy. In the absence of treatment, resistant mutants typically carry a fitness cost. Moreover, evidence for positive epistasis has been observed in HIV resistance evolution [39]. Therefore, a relatively high rate of synaptic transmission could significantly increase the chances that multi-drug resistant virus mutants are present at selection–mutation balance before treatment is initiated.

The way in which the mode of virus spread was observed to influence the rate of double mutant emergence was driven by two opposing effects: synaptic virus transmission increases double mutant production in our model, but slows down double mutant spread due to the experimentally supported assumption that synaptic transmission is associated with spatially clustered dynamics [17]. For disadvantageous mutants, production is the main driving force, and hence purely synaptic transmission results in the highest mutant levels. For advantageous mutants, double mutant spread is a crucial factor, and hence, free virus transmission tends to speed up mutant invasion in our model. For neutral mutants, both production and spread are similarly important, and hence, there is an optimal combination of free virus and synaptic transmission that maximizes double mutant fractions.

This again reinforces the notion that synaptic transmission promotes the pre-existence of drug resistance mutants before therapy, since such mutants tend to be disadvantageous. More generally, our results suggest that synaptic transmission increases the persistence of disadvantageous mutants at selection–mutation balance, which could later become advantageous due to changes in selection pressures. At the same time, however, synaptic transmission is predicted to slow down the invasion of mutants that have escaped two (or more) immune response specificities (CD8 T cell or B cell responses), since escape mutations are advantageous. The model suggests that the invasion of such mutants is promoted by free virus transmission. Escape from two immune cell clones with different specificities is likely characterized by positive epistasis. If a virus population is controlled by two immune cell clones with different

specificities, escape from one of them probably leads to an incremental fitness increase, while escape from both clones can result in a significant loss of control. In this scenario, the model further suggests that recombination promotes the rise of double immune escape mutants if escape leads to a sufficiently large fitness advantage, which is likely to be true. Finally, the presence of neutral genetic diversity is predicted to be maximized by an optimal balance between synaptic and free virus transmission. This discussion emphasizes that the mode of virus spread does not have a universally positive or negative effect on mutant persistence or emergence, and that the effect depends on the exact nature of the mutants under consideration.

There is some controversy in the literature about the average multiplicity of infected cells *in vivo*. While some papers reported significant levels of multiple infection, especially in tissue compartments [9,38], other publications found an infection multiplicity close to one, both in the blood and tissues [36,37]. Reasons for the discrepancy could be the methodology that was used to measure multiplicity, and also the T cell subsets that were taken into account during this analysis. In the light of data that document an important contribution of recombination to the *in vivo* evolution of HIV [8–11], it is likely that a sufficient amount of multiple infection occurs. An important role of multiple infection is further suggested by studies that document a very efficient infection process during synaptic cell-to-cell transmission, resulting in the simultaneous transfer of multiple viruses from the source cell to the target cell [12,13,15]. Furthermore, the frequent co-transmission of different virus strains was observed both *in vitro* and *in vivo* [16,17]. An important point in our analysis was that a model with reduced infection multiplicity due to a declining ability to super-infect over time resulted in similar insights.

Another crucial assumption of our model was that synaptic cell-to-cell transmission was characterized by spatially restricted virus spread. While imaging studies have shown an ability of immune cells to move about within tissues [42], our work on humanized mice demonstrated that virus spread in the presence of synaptic, cell-to-cell transmission, was characterized by the spatial clustering of infected cells [17], which supports the assumption we made. If it were assumed that synaptic transmission follows mass action law, then several results would change, since synaptic transmission would no longer give rise to slower virus spread than free virus transmission. While there is evidence that HIV-infected cells are motile, we assume that they are static within the context of the agent-based model and that mixing effects occur through non-spatial transmission [43]. A certain amount of cell migration could be incorporated into the model, and this would be an interesting future extension of the current work. While we have investigated the effect of spatially restricted virus transmission in the context of a 2D model (which might best represent *in vitro* conditions), the geometry of cell arrangements in the lymphatic tissues is more complex and in fact not documented in detail. A three-dimensional version of this model could be

a next step when increasing complexity, but a biologically more realistic computational description will require more detailed data to be collected that characterize the exact spatial arrangement of T cells and their migration patterns.

The relative contribution of synaptic and free virus transmission to virus spread *in vivo* is still not well understood. *In vitro,* experiments have estimated that the two transmission modes contribute approximately equally to virus spread [44], but conditions *in vivo* are likely significantly different, and this could have a large impact on these dynamics. Our results indicate that both free virus and synaptic transmission have important and different effects on double mutant populations, depending on the nature of the mutants. Free virus transmission promotes the invasion of advantageous double mutants, while synaptic transmission promotes the existence of disadvantageous double mutants at selection–mutation balance. Furthermore, we observed synaptic transmission to protect against negative effects of recombination for disadvantageous double mutants characterized by positive epistasis. These selective forces likely shape the balance between synaptic and free virus transmission towards which HIV has evolved.

An important next step will be to address some of the insights obtained from our modelling with data. Because our models examine the role of synaptic and free virus transmission for double mutant evolution, and investigate the impact of recombination on these dynamics, this requires a system that can be easily manipulated in these respects. We have previously analysed HIV dynamics *in vitro* under static conditions, where both free virus and synaptic transmission occurs, as well as under gentle shaking conditions, where synaptic transmission is largely disrupted and most virus spread occurs via the release of free viruses [44]. Viruses that are labelled with two different fluorescent reporter genes, and upon recombination give rise to a third and distinct fluorescent colour can be grown under these conditions [8]. This will allow us to experimentally parse the effect of different transmission modes on double mutant evolution and on the rate of recombination, and to relate experimental results to modelling predictions that are presented here. This is subject to ongoing work.

Endnote

[1] Note that the death rate parameter used in our simulations corresponds to the time units of hours.

References

1. Kimata JT, Kuller L, Anderson DB, Dailey P, Overbaugh J. 1999 Emerging cytopathic and antigenic simian immunodeficiency virus variants influence AIDS progression. *Nat. Med.* **5**, 535–541. (doi:10.1038/ 8414)

2. Hirsch VM. 1999 Evolution of the fittest ends in tragedy. *Nat. Med.* **5**, 488–489. (doi:10.1038/8369)

3. Mikhail M, Wang B, Lemey P, Beckthold B, Vandamme A-M, Gill MJ, Saksena NK. 2005 Role of viral evolutionary rate in HIV-1 disease progression in a linked cohort. *Retrovirology* **2**, 41. (doi:10.1186/1742-4690-2-41)

4. Mansky LM, Temin HM. 1995 Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69**, 5087–5094. (doi:10.1128/JVI.69.8.5087-5094.1995)

5. Wei X *et al.* 1995 Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**, 117–122. (doi:10.1038/373117a0)

6. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. 1995 Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123–126. (doi:10.1038/373123a0)

7. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. 1996 HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582–1586. (doi:10.1126/science.271.5255.1582)

8. Levy DN, Aldrovandi GM, Kutsch O, Shaw GM. 2004 Dynamics of HIV-1 recombination in its natural target cells. *Proc. Natl Acad. Sci. USA* **101**, 4204–4209. (doi:10.1073/pnas.0306764101)

9. Jung A, Maier R, Vartanian J-P, Bocharov G, Jung V, Fischer U, Meese E, Wain-Hobson S, Meyerhans A. 2002 Recombination: multiply infected spleen cells in HIV patients. *Nature* **418**, 144. (doi:10.1038/418144a)

10. Moutouh L, Corbeil J, Richman DD. 1996 Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc. Natl Acad. Sci. USA* **93**, 6106–6111. (doi:10.1073/pnas.93.12.6106)

11. Mostowy R, Kouyos RD, Fouchet D, Bonhoeffer S. 2011 The role of recombination for the coevolutionary dynamics of HIV and the immune response. *PLoS ONE* **6**, e16052. (doi:10.1371/journal.pone.0016052)

12. Hübner W *et al.* 2009 Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. *Science* **323**, 1743–1747. (doi:10.1126/science.1167525)

13. Chen P, Hübner W, Spinelli MA, Chen BK. 2007 Predominant mode of human immunodeficiency virus transfer between T cells is mediated by sustained Env-dependent neutralization-resistant virological synapses. *J. Virol.* **81**, 12 582–12 595. (doi:10.1128/JVI.00381-07)

14. Agosto LM, Uchil PD, Mothes W. 2015 HIV cell-to-cell transmission: effects on pathogenesis and antiretroviral therapy. *Trends Microbiol.* **23**, 289–295. (doi:10.1016/j.tim.2015.02.003)

15. Sattentau Q. 2008 Avoiding the void: cell-to-cell spread of human viruses. *Nat. Rev. Microbiol.* **6**, 815–826. (doi:10.1038/nrmicro1972)

16. Del Portillo A, Tripodi J, Najfeld V, Wodarz D, Levy DN, Chen BK. 2011 Multiploid inheritance of HIV-1 during cell-to-cell infection. *J. Virol.* **85**, 7169–7176. (doi:10.1128/JVI.00231-11)

17. Law KM, Komarova NL, Yewdall AW, Lee RK, Herrera OL, Wodarz D, Chen BK. 2016 *In vivo* HIV-1 cell-to-cell transmission promotes multicopy micro-compartmentalized infection. *Cell Rep.* **15**, 2771–2783. (doi:10.1016/j.celrep.2016.05.059)

18. Fraser C. 2005 HIV recombination: what is the impact on antiretroviral therapy? *J. R. Soc. Interface* **2**, 489–503. (doi:10.1098/rsif.2005.0064)

19. Bretscher MT, Althaus CL, Müller V, Bonhoeffer S. 2004 Recombination in HIV and the evolution of drug resistance: for better or for worse? *Bioessays* **26**, 180–188. (doi:10.1002/bies.10386)

20. Althaus CL, Bonhoeffer S. 2005 Stochastic interplay between mutation and recombination during the acquisition of drug resistance mutations in human immunodeficiency virus type 1. *J. Virol.* **79**, 13 572–13 578. (doi:10.1128/JVI.79.21.13572-13578.2005)

21. Carvajal-Rodriguez A, Crandall KA, Posada D. 2007 Recombination favors the evolution of drug resistance in HIV-1 during antiretroviral therapy. *Infect. Genet. Evol.* **7**, 476–483. (doi:10.1016/j.meegid.2007.02.001)

22. Kouyos RD, Fouchet D, Bonhoeffer S. 2009 Recombination and drug resistance in HIV: population dynamics and stochasticity. *Epidemics* **1**, 58–69. (doi:10.1016/j.epidem.2008.11.001)

23. Nowak M, May RM. 2000 *Virus dynamics: mathematical principles of immunology and virology*. Oxford, UK: Oxford University Press.

24. Perelson AS. 2002 Modelling viral and immune system dynamics. *Nat. Rev. Immunol.* **2**, 28–36. (doi:10.1038/nri700)

25. Komarova NL, Wodarz D. 2013 Virus dynamics in the presence of synaptic transmission. *Math. Biosci.* **242**, 161–171. (doi:10.1016/j.mbs.2013.01.003)

26. Komarova NL, Levy DN, Wodarz D. 2012 Effect of synaptic transmission on viral fitness in HIV infection. *PLoS ONE* **7**, e48361. (doi:10.1371/journal.pone.0048361)

27. Komarova NL, Levy DN, Wodarz D. 2013 Synaptic transmission and the susceptibility of HIV infection to anti-viral drugs. *Sci. Rep.* **3**, 2103. (doi:10.1038/srep02103)

28. Asatryan A, Wodarz D, Komarova NL. 2015 New virus dynamics in the presence of multiple infection. *J. Theor. Biol.* **377**, 98–109. (doi:10.1016/j.jtbi.2015.04.014)

29. Dixit NM, Perelson AS. 2005 HIV dynamics with multiple infections of target cells. *Proc. Natl Acad. Sci. USA* **102**, 8198–8203. (doi:10.1073/pnas.0407498102)

30. Dixit NM, Perelson AS. 2004 Multiplicity of human immunodeficiency virus infections in lymphoid tissue. *J. Virol.* **78**, 8942–8945. (doi:10.1128/JVI.78.16.8942-8945.2004)

31. Dimitrov DS, Willey RL, Sato H, Chang L-J, Blumenthal R, Martin MA. 1993 Quantitation of human immunodeficiency virus type 1 infection kinetics. *J. Virol.* **67**, 2182–2190. (doi:10.1128/JVI.67.4.2182-2190.1993)

32. Russell RA, Martin N, Mitar I, Jones E, Sattentau QJ. 2013 Multiple proviral integration events after virological synapse-mediated HIV-1 spread. *Virology* **443**, 143–149. (doi:10.1016/j.virol.2013.05.005)

33. Ribeiro RM, Qin L, Chavez LL, Li D, Self SG, Perelson AS. 2010 Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. *J. Virol.* **84**, 6096–6102. (doi:10.1128/JVI.00127-10)

34. Nowak MA *et al.* 1997 Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. *J. Virol.* **71**, 7518–7525. (doi:10.1128/JVI.71.10.7518-7525.1997)

35. Gossel G, Hogan T, Cownden D, Seddon B, Yates AJ. 2017 Memory CD4 T cell subsets are kinetically heterogeneous and replenished from naive T cells at high levels. *Elife* **6**, e23013. (doi:10.7554/eLife.23013)

36. Josefsson L *et al.* 2011 Majority of CD4$^+$ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. *Proc. Natl Acad. Sci. USA* **108**, 11 199–11 204. (doi:10.1073/pnas.1107729108)

37. Josefsson L *et al.* 2013 Single cell analysis of lymph node tissue from HIV-1 infected patients reveals that the majority of CD4$^+$ T-cells contain one HIV-1 DNA molecule. *PLoS Pathog.* **9**, e1003432. (doi:10.1371/journal.ppat.1003432)

38. Sigal A, Kim JT, Balazs AB, Dekel E, Mayo A, Milo R, Baltimore D. 2011 Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. *Nature* **477**, 95–98. (doi:10.1038/nature10347)

39. Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM, Petropoulos CJ. 2004 Evidence for positive epistasis in HIV-1. *Science* **306**, 1547–1550. (doi:10.1126/science.1101786)

40. Komarova NL. 2007 Loss- and gain-of-function mutations in cancer: mass-action, spatial and hierarchical models. *J. Stat. Phys.* **128**, 413–446. (doi:10.1007/s10955-006-9238-0)

41. Komarova NL, Shahriyari L, Wodarz D. 2014 Complex role of space in the crossing of fitness valleys by asexual populations. *J. R. Soc. Interface* **11**, 20140014. (doi:10.1098/rsif.2014.0014)

42. Germain RN, Robey EA, Cahalan MD. 2012 A decade of imaging cellular motility and interaction dynamics in the immune system. *Science* **336**, 1676–1681. (doi:10.1126/science.1221063)

43. Murooka TT, Deruaz M, Marangoni F, Vrbanac VD, Seung E, von Andrian UH, Tager AM, Luster AD, Mempel TR. 2012 HIV-infected T cells are migratory vehicles for viral dissemination. *Nature* **490**, 283–287. (doi:10.1038/nature11398)

44. Komarova NL, Anghelina D, Voznesensky I, Trinité B, Levy DN, Wodarz D. 2013 Relative contribution of free-virus and synaptic transmission to the spread of HIV-1 through target cell populations. *Biol. Lett.* **9**, 20121049. (doi:10.1098/rsbl.2012.1049)