Effects of Epidemiological Structure on the Transient Evolution of HIV Virulence

Sang Woo Park[1] Benjamin M. Bolker[1,2,3,*]

1 Department of Mathematics & Statistics, McMaster University, Hamilton, Ontario, Canada

2 Department of Biology, McMaster University, Hamilton, Ontario, Canada

3 Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada

* bolker@mcmaster.ca

Abstract

The evolutionary dynamics of parasite virulence over the course of an emerging epidemic have important implications both for our basic understanding of epidemiological dynamics and, potentially, for the outcomes of public health interventions. Changes in the fitness landscape will generally select for higher virulence during the early phase of an epidemic, but quantitative outcomes may depend sensitively on biological details and the structure of mathematical models used to capture them. Fraser and co-workers have proposed a series of models for eco-evolutionary dynamics of HIV that are relatively detailed in their portrayal of the tradeoffs between transmission and virulence (mediated by set-point viral load, SPVL) and their heritability between hosts. However, these models use implicit representations of the transmission process that drastically simplify the partnership dynamics that previous research has found to be critical in driving epidemics of sexually transmitted diseases. We explore models that combine HIV virulence tradeoffs with a range of epidemiological structures, modeling partnership formation and dissolution and allowing for individuals to transmit disease outside of partnerships. We assess summary statistics such as the minimum time of progression to

AIDS and the time at which this extreme value (peak virulence) occurs across all models for a range of partnership dynamic parameters applicable to the HIV epidemic in sub-Saharan Africa. In order to account for the different interpretations of parameters across model structures, we scale all parameter sets to constrain the simulated epidemic growth rate to be identical, matching a realistic baseline value. Although virulence trajectories are broadly similar across model structures — an initial increase in virulence/decrease in progression time, followed by a minimum and an increase to an equilibrium progression time slightly higher than the minimum — the timing and magnitude of the minimum vary considerably. Models of intermediate complexity predicted much lower peak virulence (15 years to progress to AIDS) compared to both more realistic models and simple random-mixing models with no partnership structure at all (both approx. 7.25 years to progress to AIDS). In this range of models, the simplest random-mixing structure is actually the best approximation to the most realistic model; this surprising outcome occurs because the dominance of extra-pair contact in the realistic model tends to mask the effects of partnership structure.

Author Summary

Pathogens such as HIV can evolve rapidly in response to changes in their environments; such changes include both increases in disease prevalence over the course of the epidemic and treatment interventions. While researchers have successfully used computational models to explore these evolutionary dynamics, these models often neglect details such as the formation and dissolution of sexual partnerships; other research has shown that these processes can strongly affect epidemic outcomes. We built and compared models that used different methods to model both partnership dynamics and sexual contact outside of partnerships. Models of intermediate complexity predicted much lower peak virulence (15 years to progress to AIDS) compared to both more realistic models and simple random-mixing models with no partnership structure at all (both approx. 7.25 years to progress to AIDS); extra-pair contact tended to wash out the effects of epidemiological structure. The large differences in evolutionary dynamics among different epidemiological models suggests that researchers trying to predict the evolution of pathogens should proceed with caution.

Introduction

The evolution of pathogen virulence has both theoretical and, potentially, practical importance. In general, evolutionary theory suggests that disease strains that can reproduce more — where reproduction is defined the amount of *between-host* transmission, or the number of new hosts infected — will increase in prevalence. Pathogens can increase their net reproduction rate either by increasing their birth rate, the rate (per infected host) at which they infect new hosts, or by decreasing their death rate, the rate at which hosts recover or die from disease. The trade-off theory [1, 2] — postulates that the birth and death rate are both linked to the rate at which the pathogen reproduces within the host, and that the pathogen evolution will thus strike a balance between the pathogen's rate of transmission to new hosts and its rate of killing its host (or of provoking the host's immune system to eliminate it). Some biologists have criticized the tradeoff theory [3, 4], but others have successfully applied it to a variety of host-pathogen systems [5–8]. Fraser *et al.* have applied these ideas in a particularly interesting way by showing that HIV appears to satisfy the prerequisites of the tradeoff theory: in studies of discordant couples (i.e. long-term sexual partnerships with one infected and one uninfected partner), HIV virulence as measured by the rate of progression to AIDS was both heritable and covaried with the set-point viral load (i.e., the characteristic virus load measured in blood during the intermediate stage of infection), which in turn predicted the probability of transmission [9, 10]. Subsequent studies [11, 12] used these data to parameterize mechanistic models of HIV virulence evolution, suggesting that HIV invading a novel population would initially evolve increased virulence, peaking after approximately 100-200 years and then declining slightly to a long-stable virulence level.

The work of Shirreff *et al.* [11], and particularly the predicted transient peak in HIV virulence midway through the epidemic, highlights the importance of interactions between epidemiological and evolutionary factors [13, 14]. However, despite these studies' attention to detail at the individual or physiological level, the epidemiological structures used in these models are relatively simple.

As we discuss in detail below, existing models of HIV eco-evolutionary dynamics either use implicit models that incorporate the average effects of within-couple sexual

contact — without representing the explicit dynamics of pair formation and dissolution or accounting for extra-partnership contact — or use an agent-based formulation with parameters that effectively lead to random mixing among infected and uninfected individuals. Here we explore the effects of incorporating *explicit* epidemiological structure in eco-evolutionary models.

We add complexity to the epidemiological model following the general approach of Champredon *et al.* [15], which is in turn based on work of Dietz and Hadeler [16]; individuals join and leave partnerships at a specified rate, and can have sexual contact both within and outside of established partnerships. In order to explore how virulence evolution depends on epidemiological structure, we consider a series of models with increasing levels of complexity. In order to avoid dependence of the results on a particular set of parameters — as we explain below, finding matching sets of parameters across models with widely differing epidemiological structures is challenging — we evaluate our models across a wide range of parameters, again following Champredon *et al.* [15] in using a Latin hypercube design. For each model run, we compute a set of metrics (minimum progression time/peak virulence, timing of maximum virulence, equilibrium virulence) that summarize the evolutionary trajectory of a simulated HIV epidemic.

As our primary goal is to explore how different epidemiological structures (i.e. partnership dynamics and contact structures) affect our conclusions about the evolution of virulence, our models use a simplified description of within-host dynamics and heritability derived from Shirreff *et al.*'s multi-strain evolutionary model [11]. Like Shirreff *et al.*, we use a simple susceptible-infected-susceptible demographic formulation; rather than modeling birth and death (or more specifically, recruitment into the sexually active population and death), we assume that whenever an individual dies from infection, another enters the susceptible compartment.

Materials and Methods

Infection dynamics

Like Shirreff *et al.* [11], we focus on the evolution of mean $\log_{10}$ set-point viral load, SPVL (which we denote as α), rather than the rate of progression to AIDS itself (we refer to SPVL as "virulence" hereafter). In contrast to Shirreff *et al.*, we use a single-stage disease model instead of accounting explicitly for progression through the three main stages of HIV infection (primary, asymptomatic, and disease), and we use a simple exponentially distributed infectious period instead of a more realistic Weibull-distributed infectious period. We account for varying transmission rates and durations of each disease stage by summing the durations of three stages (again based on Shirreff *et al.*'s model) and taking the duration-weighted average of transmission rates of three stages. Thus the within-couple transmission rate, β, for our models is given by:

$$\beta(\alpha) = \frac{D_P\beta_P + D_A(\alpha)\beta_A(\alpha) + D_D\beta_D}{D_P + D_A(\alpha) + D_D}, \tag{1}$$

where the duration of infection (D_P and D_D) and rate of transmission (β_P and β_D) of the Primary and Disease stages of infection are independent of the host's SPVL. Following Shirreff *et al.*, the duration of infection (D_A) and rate of transmission (β_A) for the Asymptomatic stage are Hill functions of the SPVL:

$$\begin{aligned} D_A(\alpha) &= \frac{D_{\max} D_{50}^{D_k}}{V_\alpha^{D_k} + D_{50}^{D_k}}, \\ \beta_A(\alpha) &= \frac{\beta_{\max} V_\alpha^{\beta_k}}{V_\alpha^{\beta_k} + \beta_{50}^{\beta_k}}, \end{aligned} \tag{2}$$

where $V_\alpha = 10^\alpha$. The uncoupled and extra-couple transmission rates are scaled by multiplying the within-couple transmission rate β by the contact ratios c_u/c_w and c_e/c_w (see Appendix). Simplifying the model of HIV pathogenesis from three stages to a single stage could affect our conclusions about the evolution of virulence (e.g. Kretzschmar and Dietz [17] show that pair formation dynamics and multiple stages of infectivity have interactive effects on $\mathcal{R}_0$). However, our simplified model produces results that are qualitatively similar to those of Shirreff *et al.*'s [11] model; when our

model is calibrated to have a similar initial epidemic growth rate r, the peak $\log_{10}$ SPVL occurs at the same time ($\approx$ 200 years) but slightly higher (4.7 vs. 4.3); the relative height of the peak is **compute difference and fill this in** (Fig 1).

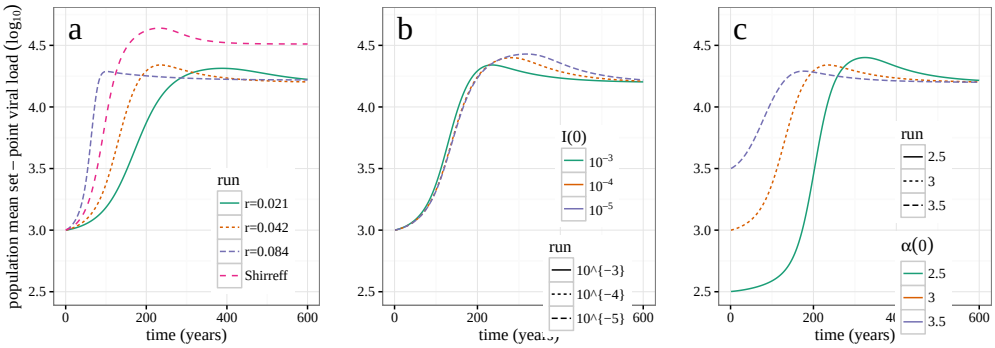

Fig 1. Baseline dynamics. Time series of mean population $\log_{10}$ SPVL. (a) Contrast with three-stage Shirreff model and effects of varying exponential growth rate, r. (b) Effects of varying initial infectious density $I(0)$. (c) Effects of varying initial mean virulence $\alpha(0)$. The orange curve in panel (a) shows that our simplified model can produce similar trajectories to those Shirreff *et al.*'s model [11]. In order to compare the models, we match the initial exponential growth rate (see main text "Simulation runs"). Furthermore, we experiment with wide ranges of initial values (panels b and c) to show that evolutionary dynamics are not overly sensitive to starting parameter values, which we hold constant in our simulations.

Mutation

Like Shirreff *et al.* [11] we incorporate a between-host mutation process in the SPVL. We simplify Shirreff *et al.*'s evolutionary model by using a one-to-one genotype-phenotype mapping rather than allowing for variation in phenotypes of a single genotype. The mutational process in our model is directly taken from Shirreff *et al.*. Over the course of infection, mutation occurs within the host. However, it is assumed that SPVL of the strain transmitted by an infected individual is determined by the SPVL at the time of infection and is not further affected by within-host mutation. Instead, the mutational effect takes place in a single step at the time of transmission. First, the distribution of $\log_{10}$ SPVL is discretized into a vector:

$$\alpha_i = \alpha_{\min} + (\alpha_{\max} - \alpha_{\min})\frac{i-1}{n-1} \qquad i = 1, 2, 3, \ldots n. \tag{3}$$

We have experimented with varying degrees of discretization in the strain distribution (i.e., values of n); in our model runs comparing results with Shirreff *et al.* [11] (Fig 1) we

use $n = 51$ (i.e. a bin width of 0.1 $\log_{10}$ SPVL for α), but reducing n to 21 (bin width = 0.25 $\log_{10}$ SPVL) makes little difference; we use this coarser grid for all other simulations reported.

We construct an n by n mutational matrix, M — which is multiplied with the transmission term — so that M_{ij} is the probability that a newly infected individual will have $\log_{10}$ SPVL of α_j given that the infector has $\log_{10}$ SPVL of α_i. Finally, the probabilities are normalized so that each row sums to 1:

$$M_{ij} = \frac{\Phi(\alpha_j + d/2; i) - \Phi(\alpha_j - d/2; i)}{\Phi(\alpha_{\max} + d/2; i) - \Phi(\alpha_{\min} - d/2; i)}, \tag{4}$$

where $\Phi(x; i)$ is the Gaussian cumulative distribution function with mean α_i and variance of σ_M^2, and $d = (\alpha_{\max} - \alpha_{\min})/(n-1)$. Transmission rate and disease induced mortality rates are discretized into a vector as well:

$$\begin{aligned} \beta_i &= \beta(\alpha_i), \\ \lambda_i &= \frac{1}{D_P + D_A(\alpha_i) + D_D}. \end{aligned} \tag{5}$$

Contact structure and partnership dynamics

We developed seven multi-strain evolutionary models covering a gamut between Champredon *et al.*'s relatively realistic [15] and Shirreff *et al.*'s relatively simple [11] epidemiological structures, each of which is based on different assumptions regarding contact structure and partnership dynamics. Specifically, we focus on the effects of the assumptions of (1) instantaneous vs. non-instantaneous partnership formation, (2) zero vs. positive extra-partnership sexual contact and transmission and (3) homogeneous vs. heterogeneous levels of sexual activity on the evolution of mean $\log_{10}$ SPVL.

Our first four models consider explicit partnership dynamics and are based on Champredon *et al.*'s model [15]. The first two ("pair-formation" or "pairform" for short) assume non-instantaneous partnership formation (i.e. individuals spend some time uncoupled, outside of partnerships) and consist of five states that are classified by infection status and partnership status. S is the number of single (uncoupled) susceptible individuals, and I is the number of single infected individuals. SS is the number of concordant negative (susceptible-susceptible) couples, SI is the number of

serodiscordant (susceptible-infected) couples, and II is the number of concordant positive (infected-infected) couples. The first ("pairform+epc") includes extra-partnership contact (with both uncoupled individuals and individuals in other partnerships) whereas the second ("pairform") only considers within-couple transmission.

The next two models, which are intended to bridge the gap between models with fully explicit pair-formation dynamics and simpler, implicit models, assume instantaneous partnership formation ("instswitch"). The compartmental structure thus omits the single states S and I, comprising only the three partnered states: SS, SI, and II. Like the first two models, this pair of models differs in their inclusion of extra-pair contact: the third model ("instswitch+epc") includes extra-partnership contact (now only with individuals in other partnerships, since uncoupled individuals do not exist in this model) while the fourth ("instswitch") only considers within-couple transmission.

The next two models represent extreme simplifications of sexual partnership dynamics. One ("implicit") is an implicit serial monogamy model based on the epidemiological model used by Shirreff *et al.* [11]. It is actually a random mixing model that consist of only two states, S and I, and does not consider explicit partnership dynamics. However, to reflect the effect of partnership structure, it uses an adjusted transmission rate that is derived from an approximation of the basic reproduction number of a serial monogamy model with instantaneous pair formation [18]. The second model of this pair ("random") is a simple random-mixing model.

Lastly, we add a model of heterogeneity in sexual activity to the pairform+epc model ("hetero"). Individuals are divided into different risk groups based on the sexual activity level; we scale all aspects of sexual activity, assuming that sexual activity level in both within- and extra-couple contacts is directly proportional to number of non-cohabiting (extra-couple and uncoupled) partners per year [19] (see Appendix for model details). We assume random activity-weighted mixing between risk groups **cite old Anderson-May HIV paper?**. While this model lacks some important elements, such as age-structured mixing patterns, needed for realistic models of HIV transmission in sub-Saharan Africa, we see it as a first step toward assessing the effects of epidemiological complexity. As even these models are pushing the limits of compartmental-based models, adding further complexity will probably require a shift to

an agent-based model framework, as well as considerable effort in model calibration [12, 20].

The base model (i.e. pairform+epc) for the first four models and the heterogeneous model is an extension of Champredon *et al.*'s model [15]. Individuals in single compartment acquire a partner at a rate ρ, and partnerships dissolve at a rate c. Infected individuals in a discordant partnership infect their susceptible partner at a rate β (within-couple transmission rate) and susceptible individuals outside the partnership at a rate c_e (extra-couple transmission rate). Likewise, a single infected individual can infect any susceptible individuals at a rate c_u through uncoupled mixing. Extra-couple and uncoupled transmission are modeled in the same way as Champredon *et al.*'s model. All the details have been adapted to a multi-strain scenario. The second through fourth models (pairform, instswitch+epc, instswitch) are derived from the base model by simplifying epidemiological processes (partnership formation and uncoupled/extra-couple contact). Model details are explained in the appendix.

Latin hypercube sampling

Despite considerable effort [15, 18], the parameters determining the rate and structure of sexual partnership change and contact are still very uncertain; this led Champredon *et al.* [15] to adopt a Latin hypercube sampling (LHS) strategy [21] that evaluates model outcomes over a range of parameter values. In order to make sure that our comparisons among models apply across the entire space of reasonable parameter values, and in order to evaluate the differential sensitivity of different model structures to parameter values, we follow a similar protocol and perform LHS over a parameter set including both the early- and late-stage transmission and duration parameters (β_P, D_P, β_D, D_D) and contact/partnership parameters (ρ, c, c_u/c_w, and c_e/c_w). For the heterogeneity model, mean and squared coefficient of variation (CV) for number of non-cohabiting partners are sampled as well. We do not allow for uncertainties in parameters that are directly related to the evolutionary process ($\beta_{\max}$, β_{50}, β_k, $D_{\max}$, D_{50}, D_k, σ_M), using Shirreff *et al.*'s point estimates throughout [11].

Latin hypercube sampling is done as in Champredon *et al.* [15]. For each parameter,

z, its range is divided into $N = 1000$ equal intervals on a log scale:

$$z_i = \exp\left(\log(z_{\min}) + [\log(z_{\max}) - \log(z_{\min})]\frac{i-1}{N-1}\right) \qquad i = 1, 2, 3, \ldots, N. \qquad (6)$$

Random permutations of these vectors form columns in a sample parameter matrix; each row contains a different parameter set that is used for one simulation run.

Table 1 gives the ranges of the model parameters used for LHS. Parameter ranges regarding contact and partnership dynamics (ρ, c, and c_e/c_w) are taken from Champredon *et al.* [15], whereas those regarding infection (β_P, D_P, β_D, and D_D) are taken from Hollingsworth *et al.* [18]. The remaining parameters are taken from Shirreff *et al.* [11].

The one completely new parameter in our model, the ratio of uncoupled to within-couple transmission c_u/c_w, is needed to more flexibly contrast uncoupled and extra-couple transmission dynamics within multi-strain models (Appendix S1); we need to pick a reasonable range for it. Champredon *et al.* [15] assume that the effective within-couple contact rate and effective uncoupled contact rate have the same range of 0.05 - 0.25. Given Champredon *et al.*'s parameter range, the possible maximum and minimum values of c_u/c_w are 5 and 1/5. Therefore, we use 1/5-5 as the range for the parameter c_u/c_w. Although this adds more uncertainty to the parameter c_u — Champredon *et al.*'s range implies a 5-fold difference whereas ours gives a 25-fold difference — we consider the wider range appropriate, as little is not much known about the uncoupled transmission rate.

add something about choice of heterogeneity

Simulation runs

One of the most difficult parts of model comparison is finding parameter sets that are commensurate with many different model structures. For the most part, our models are too complex to easily derive analytical correspondences among them. Given a numerical criterion, such as r (initial exponential growth rate) or $\mathcal{R}_0$ (intrinsic reproductive number), we can adjust one or more parameters by brute force to ensure that all of the models match according to that criterion. While $\mathcal{R}_0$ is often considered the most fundamental property of an epidemic, and might thus seem to be a natural matching

Table 1. Parameter ranges/values. Values of c and ρ are doubled from those given by Champredon *et al.*; we keep track of individuals in the model, while they keep track of couples. Starred (*) parameters (used in Fig 1), and descriptions of Hill function coefficients, are taken from [11].

Notation	Description	Range/Value	Source
ρ	Partnership formation rate	1/10-2/5 per year	[15]
c	Partnership dissolution rate	1/15-1/5 (1.25*) per year	[15]
c_u/c_w	Relative contact rate for uncoupled transmission	1/5-5	Assumption
c_e/c_w	Relative contact rate extra-couple	0.01-1	[15]
β_P	Rate of transmission during primary infection	1.31-5.09 (2.76*) per year	[18]
β_D	Rate of transmission during high transmission disease stage	0.413-1.28 (0.76*) per year	[18]
D_P	Duration of primary infection	1.23/12-6/12 (0.25*) years	[18]
D_D	Duration of high transmission disease stage	4.81/12-14/12 (0.75*) years	[18]
$\beta_{\max}$	Maximum rate of transmission during asymptomatic stage	0.317 per year	[11]
β_{50}	SPVL at which infectiousness is half maximum	13938 copies per ml	[11]
β_k	Hill coefficient: steepness of increase in infectiousness as a function of SPVL	1.02	[11]
$D_{\max}$	Duration of primary infection	25.4 years	[11]
D_{50}	SPVL at which duration of asymptomatic infection is half maximum	3058 copies per ml	[11]
D_k	Hill coefficient: steepness of decrease in duration as a function of SPVL	0.41	[11]
σ_M	Mutation standard deviation of $\log_{10}$ SPVL	0.12	[11]
$\alpha_{\min}$	Minimum $\log_{10}$ SPVL	2	[11]
$\alpha_{\max}$	Maximum $\log_{10}$ SPVL	7	[11]
n	Number of strains	21 (51*)	Assumption
μ	Mean number of non-cohabiting sexual partners	0.103 - 1.206	[19]
κ	Squared coefficient of variation of number of non-cohabiting sexual partners	0.01 - 100	Assumption

criterion, here we focus on matching the initial growth rate r for several reasons. First, our primary interest is in the transient evolutionary dynamics of virulence, which are more strongly affected by r than $\mathcal{R}_0$. Second, r is more directly observable in real epidemics; r can be estimated by fitting an exponential curve to the initial incidence or prevalence curves [22], while $\mathcal{R}_0$ typically requires either (1) knowledge of *all* epidemic parameters or (2) calculations based on r and knowledge of the serial interval or

generation interval of the disease [23]. Thus, we scale parameters so that every run has the same initial exponential growth rate of incidence.

In order to allow for all models to have equal initial exponential growth rate, r, we need to pick a parameter, s, such that $\lim_{s\to 0} r(s) = 0$ and $\lim_{s\to\infty} r(s) = \infty$. As adjusting either partnership change rate (i.e. partnership formation and dissolution rate) or transmission rate fails this requirement for some of our models, we scaled partnership change rate and dissolution rate by the same factor of γ: $\beta_{\text{adj}} = \gamma\beta_{\text{base}}$, $c_{\text{adj}} = \gamma c_{\text{base}}$, $\rho_{\text{adj}} = \gamma\rho_{\text{base}}$. Since transmission rate is adjusted by the scale of γ, uncoupled and extra-couple transmission rates are adjusted as well. For the instantaneous-switching and implicit models, none of which track single individuals, only the transmission rate and partnership dissolution rate (in this case equivalent to the partnership change rate) are adjusted.

We run each model for each of 1000 parameter sets chosen by Latin hypercube sampling, with fixed starting conditions of mean $\log_{10}$ SPVL of 3 and epidemic size of 10^{-4}. After each run, initial exponential growth rate is calculated. Then, parameters are scaled so that the initial exponential growth rate is scaled to 0.04, which is approximately equal to that implied by Shirreff *et al.*'s original models.

The calibration runs for each parameter set are slightly simplified. We run each model for only 500 years (full simulations are run for 4000 years). We use a 4/5 order Runge-Kutta method (`ode45` from the `deSolve` package [24]) for all simulations.

Although each disease strain's core characteristic is its SPVL, the SPVL has one-to-one correspondences (based on eq. 2) with both the expected time to progression to AIDS and with the rate (probability per unit time) of HIV transmission. Because the time to progression (measured in years) is easier to interpret than SPVL (measured in $\log_{10}$ SPVL units), we summarize the virulence trajectories for each model run in terms of time to progression rather than SPVL. Because the time to progression is inversely related to SPVL (increasing SPVL decreases the time to progression), the time to progression is technically measuring inverse virulence rather than virulence (we did not think that reporting virulence as the rate of progression to AIDS, in units of years^{-1}, would help interpretability). For each model we derive the following summary statistics: minimum expected time to progression; time at which this minimum occurs (corresponding to peak virulence — this is also the time at which the maximum rate of

progression, maximum SPVL, and maximum transmission rate occur); equilibrium time to progression; and the ratio of progression time at its minimum to the equilibrium value. Equilibrium progression time is calculated after 4000 years of simulated time. Although most simulations reach equilibrium much earlier, we set our time horizon at a much later date as some simulation runs have slow rate of evolution depending on the parameter set and model assumptions.

Knowing the minimum progression time, timing of the minimum progression time/peak virulence, and equilibrium progression time provide sufficient detail to identify the overall shape of the virulence trajectory. In particular, knowing the timing of the peak virulence (how many years into the epidemic the virulence peaks) can help epidemiologists guess whether the virulence of an emerging pathogen is likely (1) to have peaked early, possibly even before the pathogen is detected spreading in the population, and decline over the remaining course of the epidemic; (2) to increase, peak, and decline over the foreseeable future; or (3) to increase very slowly, peaking only in the far future. To the extent that our simplistic model for HIV reflects reality, we would take the peak time of 150-300 years (Figure 1c) to mean that, in the absence of treatment, the epidemic would probably still be increasing in virulence.

Results

Our simplifications of Shirreff *et al.*'s model [11] reproduce its qualitative behaviour — in particular, its predictions of virulence dynamics — reasonably well. As r decreases from 0.084 to 0.42 (the latter value matching the initial rate of increase in prevalence in Shirreff *et al.*'s full model) the initial trajectory of increasing virulence brackets the rate from the original model (Fig 1a). However, our model produces lower peak virulence (≈ 4.3 vs. ≈ 4.6 $\log_{10}$ SPVL) and equilibrium virulence (≈ 4.25 vs. ≈ 4.5 $\log_{10}$ SPVL) than Shirreff's, even for matching initial incidence trajectories (i.e., $r = 0.042$ year^{-1}).

Changing the initial infectious density ($I(0)$), while it produces the expected changes in the initial epidemic trajectory (Supplementary material), has little effect on the virulence trajectory, making the virulence peaks slightly later and larger as $I(0)$ decreases. Decreasing $I(0)$ allows a longer epidemic phase before the transition to endemic dynamics (Fig 1b). Decreasing the initial virulence similarly but more strongly

leads to progressively later, larger peaks in virulence (Fig 1c).

Across the entire range of parameters covered by the LHS analysis, all of the classes of models we considered produce qualitatively similar virulence trajectories (Fig 2). Although the speed of virulence evolution varies, leading to wide variation in the minimum progression time (means ranging from approximately **fill in**), virulence peaks in all models between 200 and 300 years.

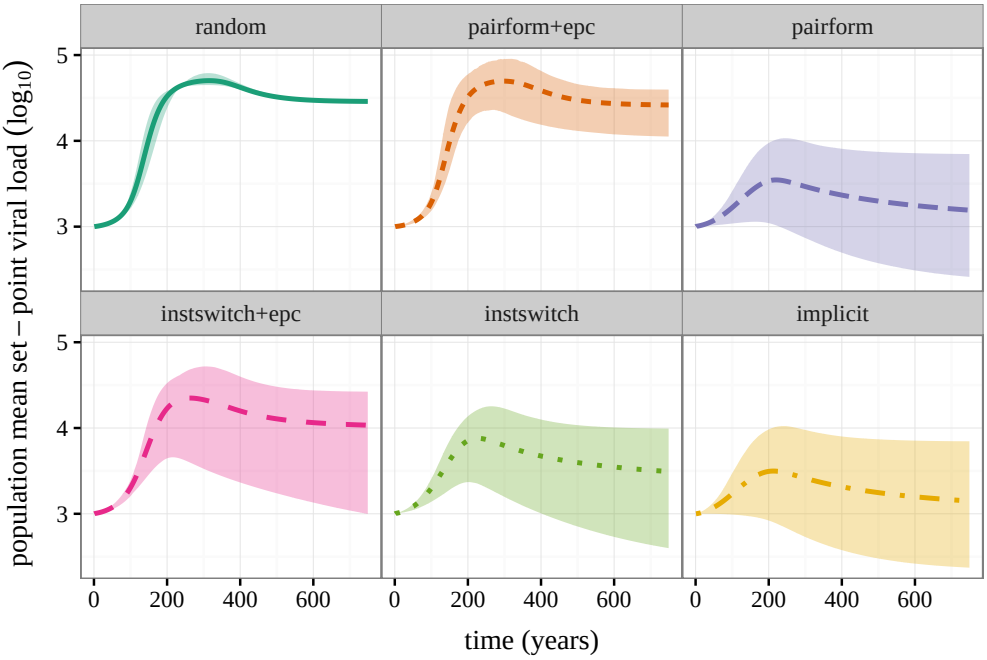

Fig 2. Envelopes of virulence trajectories (expected time of progression to AIDS) under all models. All models were run until $t = 4000$ years; truncated series are shown here.

Our chosen summary statistics (peak time, minimum progression time, equilibrium progression time, and relative progression time) all vary considerably across models 3. We first consider the models of intermediate realism: implicit, instantaneous-switching with and without extra-pair contact, and pair formation without extra-pair contact. Some parameter sets for these models lead to low equilibrium virulence ($\approx$ 18 years to progression); these same sets lead to correspondingly low peak virulence (16 years to progression) and early peak times (before 200 years). At the opposite extreme, parameter sets that produce high equilibrium virulence (8 years to progression) also produce late peaks ($>$ 200 years) and high peak virulence (4 years to progression). The pair-formation without extra-pair contact and implicit models occasionally have

parameter sets that select for such low virulence across the board that they never exceed their initial virulence, leading to a tail of peak times near zero.

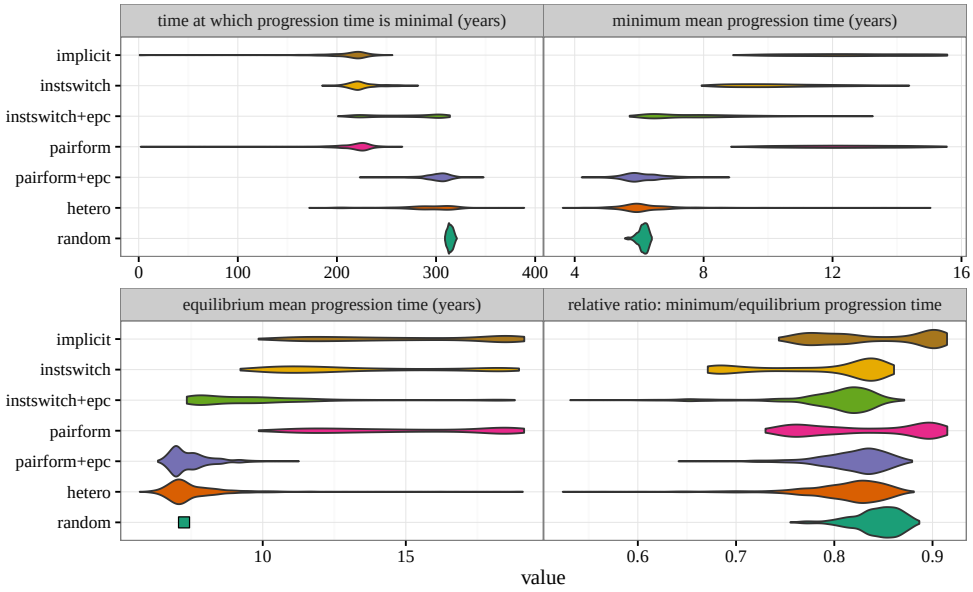

Fig 3. Univariate distributions of summary statistics. The distribution of equilibrium virulence for the random mixing model is very narrow, and has been replaced by a point in order to preserve the vertical axis scaling.

The most striking aspect of the univariate comparisons in Fig 3, (and the bivariate comparisons in Fig 4) is the similarity between the results of the least (random mixing) and the most complex (pair formation with extra-pair contact, pairform+epc with heterogeneity) models. The random-mixing model has the lowest variability, because it is unaffected by uncertainty in pair formation and extra-pair contact parameters, but otherwise the virulence dynamics of these three extreme models are remarkably similar. This phenomenon is driven by the strong effects of extra-pair contact in the model with explicit pair formation and extra-pair contact ("pairform+epc" in Figs 2-5). When individuals spend time uncoupled between partnerships, and when these single individuals can transmit disease to coupled individuals, the resulting unstructured mixing overwhelms the effect of structured mixing within couples, leading to mixing that is effectively close to random. Once unstructured mixing is strong, adding realistic heterogeneity of mixing to the model has little effect other than increasing the variability in the outcomes.

These differences are practically as well as scientifically important. The

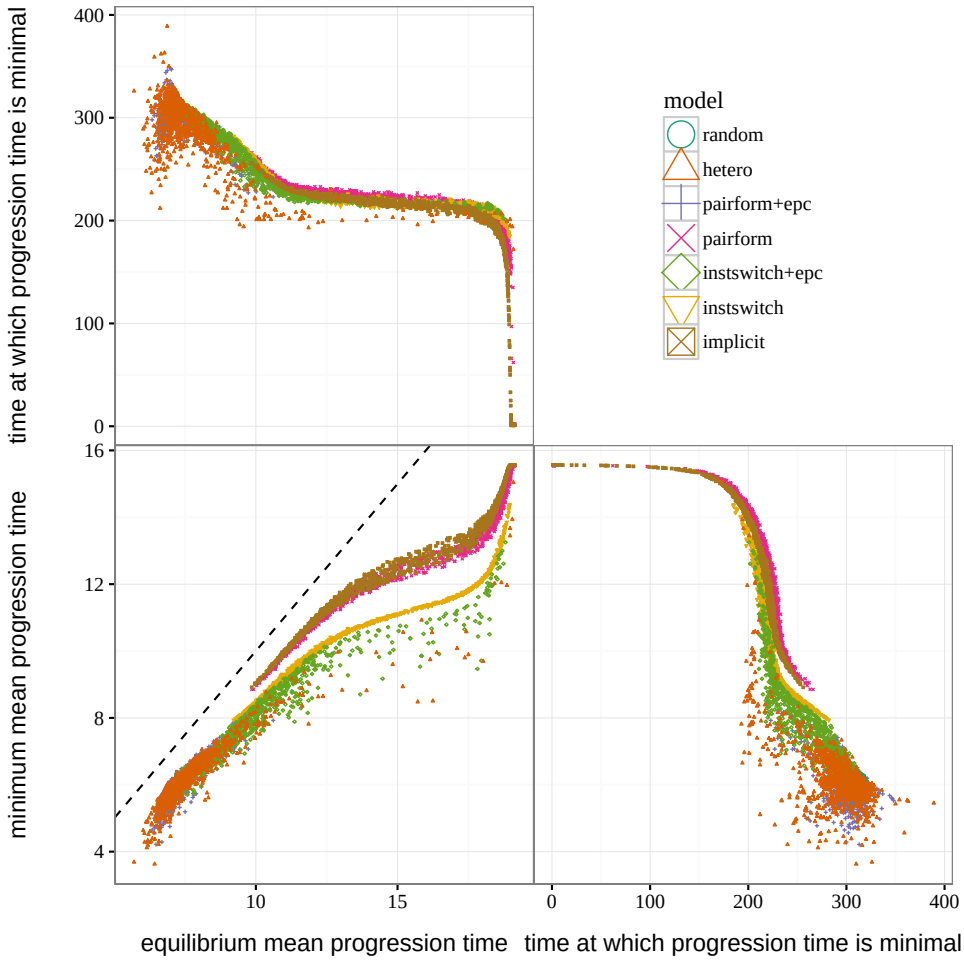

Fig 4. Pairs plot: bivariate relationships among summary statistics for each model structure. Dashed line in equilibrium vs. peak virulence plot shows 1:1 line.

random-mixing and pairform+epc models predict minimum times to progression (at the virulence peak) of 5.7 (95% CI 6.1-6.3) and 6.0 (5.0-7.7) years respectively, while the implicit model gives progression times about twice as long: 12.5 (9.5-15.6) years. The corresponding differences in within-couple transmission probability are even more extreme, about a fourfold difference: 0.249 (0.24-0.26) and 0.252 (0.19-0.28) per year for the random and pairform+epc models vs. 0.059 (0.02-0.13) per year for the implicit model (Figs ???? **fill in!** show univariate distributions on the epidemiological scales of progression time and transmission probability, for all summary measures and all models).

The bivariate relationships (Fig 4) help distinguish the results of different models

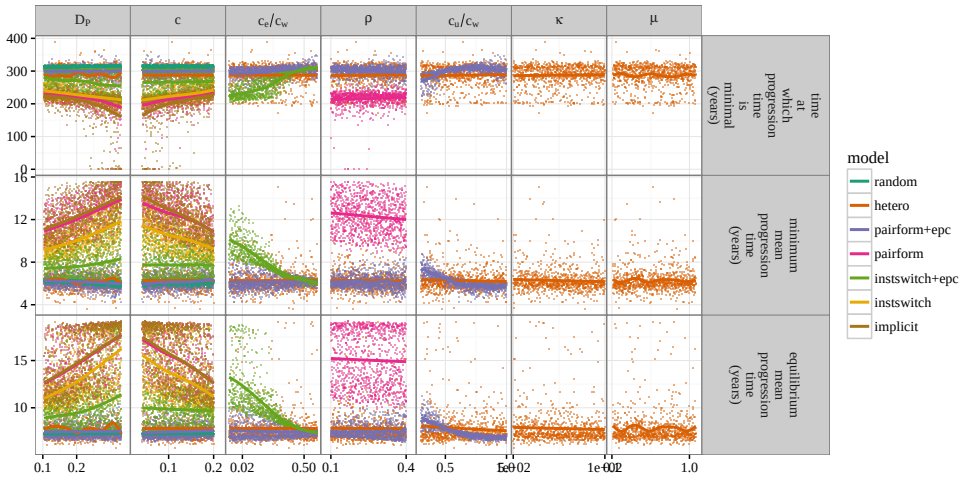

Fig 5. Sensitivity plot. For each parameter in the Latin hypercube sample and each summary statistic, shows the distribution (points) and trend (smooth line) of the summary statistic as a function of the *unscaled* parameter value, i.e. prior to adjusting the parameters to achieve the standard initial epidemic growth rate.

with similar univariate dynamical summaries. While the relationship between equilibrium virulence and peak time is similar for all model structures (top left panel), the other relationships are more separate. In particular, the implicit and pair-formation (without extra-pair contact) are very similar to each other, but distinct from the other models. We still do not have a convincing explanation for this distinction; we would have expected the implicit model to be most similar to the the instantaneous-switching model without extra-pair contact, which most closely matches its derivation. However, we note that the implicit model derivation is based on defining the force of infection to match a scaled version of $\mathcal{R}_0$, and as such would be expected to match the equilibrium behaviour but not necessarily the epidemic-phase behaviour of a model with explicit partnership dynamics.

Finally, the sensitivity plots (Fig 5) show the effects of each parameter on the summary statistics. In almost every case the effects of the parameters are monotonic; note that the plot shows the effects of the *unscaled* parameters, i.e. before they have been calibrated to achieve a standard initial epidemic growth rate. Increases in the transmission rates (β_P, β_D) and durations (D_P, D_D) in the primary and disease stages generally decrease the equilibrium virulence, peak virulence, and peak time, although the random and pair-formation+epc models have high, relatively constant values with respect to these parameters.

The partnership dissolution rate (c), which essentially acts as a contact rate in the model, increases virulence and peak time in almost all cases, although the pair-formation+epc model is again relatively insensitive. The ratio of extra-pair to within-pair contact (c_e/c_w) affects virulence in the instantaneous-switching+epc model, but not the pair-formation+epc model (probably because the uncoupled individuals present in the pair-formation+epc model make extra-pair contact by coupled individuals less important). Surprisingly, neither of the pair-formation models is particularly sensitive to the rate of partnership formation (ρ); the rate of uncoupled contact increases virulence and peak time in the pair-formation+epc model, which is the only model to which it applies.

Discussion

All models must simplify the world. Many constraints — such as data availability, computation time, or code complexity — drive the need for parsimony, with different constraints applying in different contexts. The critical question that modelers must ask is whether the simplified model gives adequate answers, or whether the simplifications lead to qualitative or quantitative errors. This question is especially important for modelers who are hoping that their conclusions will guide management decisions.

In the particular example of HIV virulence eco-evolutionary dynamics and epidemiological, we reach the slightly ironic conclusion that the effort put into building a more realistic model essentially cancels out, leaving us in the same position as if we had ignored the problem of epidemiological complexity entirely and used a naive random-mixing contact model. In Herbeck *et al.*'s [12] network model of partnerships, the partnership duration is set to 1 day — very unrealistic in epidemic terms, but perhaps actually more true to real-world HIV epidemiological dynamics than a model with realistic partnership durations that neglects extra-pair contact [25]. Making the model even more realistic might make the random-mixing model less appropriate. For example, our model forms partnerships randomly, and assumes that extra-pair contact is randomly mixing across the population; one could instead model extra-pair contact as arising from multiple concurrent partnerships (some, such as contact with sex workers, of very short duration) and/or more structured partnership formation (by age, ethnicity,

or behaviour group). The effects of other realistic complications such as explicit modeling of two sexes (both in contact structure and differential transmission probabilities), temporal and spatial variation in epidemic processes, or presence of genetic variation in hosts are harder to predict.

Parameterization is one of the biggest challenges of epidemiological modeling. In addition to following Champredon *et al.* [15] by doing Latin hypercube sampling across a wide range of epidemiological parameters, we calibrated each set of parameters to the same initial epidemic growth rate, chosen to match the results of previous models [11]. Previous models in this area have drawn their parameters from cohort studies from the 1990s [18, 26] rather than doing any explicit calibration to epidemic curves, but they give reasonable order-of-magnitude growth rates (≈ 0.04 year^{-1}) for the early stages of the HIV epidemic (although considerably lower than estimates of $\approx 0.07 - 0.1$ year^{-1} based on population genetic reconstructions [27]). However, our reason for calibrating was not to match any specific observed epidemic, but rather to make sure that we were making meaningful comparisons across a range of models with radically different epidemiological structures, and hence involving different interpretations of the same quantitative parameters. For example, in models with instantaneous switching the partnership dissolution rate c is identical to the partnership formation rate; in models with explicit partnership formation, the partnership formation rate is also c at equilibrium, but might vary over the course of an epidemic. It is not obvious whether models with equal parameters but different structures should be directly compared; calibration solves this problem.

More generally, any model that wants to be taken seriously for management and forecasting purposes should be calibrated to *all* available data, using informative priors to incorporate both realistic distributions of uncertainty for all parameters from independent measurements [28] and calibration from population-level observations of epidemic trajectories. Such a procedure would also be an improvement on the common — although not universal — practice, which we have followed here, of assessing uncertainty over uniform ranges rather than using distributions that allow more continuous variation in support over the range of a parameter.

Researchers have documented that HIV virulence and set-point viral load are changing, on time scales comparable to those portrayed here (e.g., compare Fig 2 to

Herbeck *et al.*'s estimated rate of change of 1.3 $\log_{10}$ SPVL per century [95% CI -0.1 to 3] [29]), and have begun to build relatively realistic models that attempt to describe how interventions such as mass antiretroviral therapy (ART) can be expected to change the trajectory of virulence evolution [25, 30, 31]. While these efforts are well-intentioned, we caution that epidemiological and other structural details that are currently omitted from these models could significantly change their conclusions.

Acknowledgements

We would like to thank Christophe Fraser and David Champredon for access to simulation code; this work was funded by NSERC Discovery Grant 386590-2010.

Supporting Information

Appendix S1: model details Since we use multi-strain models in which the distribution of $\log_{10}$ SPVL has been discretized into a vector, we use a matrix notation to describe our models. The five states described in the *Methods* section are replaced with the following notations: S, I_i, SS, SI_i, II_{ij}, where the subscripts denote the strain with which an individual is infected. For example, I_i is number of infected individuals with $\log_{10}$ SPVL of α_i, and II_{ij} is the number of concordant, HIV-positive couples in which the two partners have $\log_{10}$ SPVL of α_i and α_j (independent of order; II_{ij} is synonymous with II_{ji}). Below, we use the Kronecker delta (i.e. $\delta_{ij} = 1$ if $i = j$ and 1 otherwise) in a slightly non-standard fashion as an exponent, e.g. $2^{\delta_{ij}}$, to set a value to 2 when $i = j$ and 1 otherwise.

Models 1 ("pairform+epc") and 2 ("pairform")

Partnership dynamics

Single individuals acquire partners at per-person rate ρ. Partnership formation rate for S, I_i are ρS and ρI_i, respectively. We follow Champredon *et al.*'s results [15] and assume that single individuals are distributed into coupled states with pair-formation (PF) rates as follows:

$$\begin{aligned}
\mathrm{PF}(SS) &= \frac{\rho S \cdot S}{2(S + \sum_k I_k)}, \\
\mathrm{PF}(SI_i) &= \frac{\rho S \cdot I_i}{S + \sum_k I_k}, \\
\mathrm{PF}(II_{ij}) &= \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot \frac{\rho I_i \cdot I_j}{S + \sum_k I_k}.
\end{aligned} \tag{7}$$

Partnerships dissolve at per-partnership rate c: the dissolution rates for SS, SI_i, and II_{ij} pairs are cSS, cSI_i, and cII_{ij} respectively. Unlike single strain model, where both individuals leaving the II partnership would enter I, we have to account for strains which the individuals in concordant partnership are infected with (i.e. both partners in II_{ii} enter I_i whereas one partner in II_{ij} enters the I_i compartment while the other enters I_j). Thus, coupled individuals are distributed into single states through partnership dissolution (DS):

$$\begin{aligned}
\mathrm{DS}(S) &= 2cSS + \sum_k cSI_k, \\
\mathrm{DS}(I_i) &= cSI_i + \sum_k 2^{\delta_{ik}} cII_{ik}.
\end{aligned} \tag{8}$$

Combining the partnership formation and dissolution processes yields the following equation:

$$\begin{aligned}
S' &= -\rho S + 2cSS + \sum_k cSI_k \\
I_i' &= -\rho I_i + cSI_i + \sum_k 2^{\delta_{ik}} cII_{ik} \\
SS' &= \frac{\rho S \cdot S}{2(S + \sum_k I_k)} - cSS \\
SI_i' &= \frac{\rho S \cdot I_i}{S + \sum_k I_k} - cSI_i \\
II_{ij}' &= \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot \frac{\rho I_i \cdot I_j}{S + \sum_k I_k} - cII_{ij}
\end{aligned} \tag{9}$$

Pair-formation models: infection dynamics

Within-couple transmission (WT) occurs in both models. An infected partner in SI partnership transmits virus to a susceptible partner at per-partnership rate β: $\mathrm{WT}(SI_i) = -\beta_i SI_i$. Since we assume that mutation occurs, II_{ij}, where $i \neq j$, can be formed from both SI_i and SI_j partnership: $\mathrm{WT}(II_{ij}) = M_{ij}\beta_i SI_i + M_{ji}\beta_j SI_j$. On the

other hand, II_{ii} can only be formed from an SI_i partnership: $\mathrm{WT}(II_{ii}) = M_{ii}\beta_i SI_i$. Using the Kronecker delta notation, we obtain the following set of equations that describes within-couple transmission dynamics:

$$\begin{aligned} \mathrm{WT}(SI_i) &= -\beta_i SI_i, \\ \mathrm{WT}(II_{ij}) &= \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot (M_{ij}\beta_i SI_i + M_{ji}.\beta_j SI_j) \end{aligned} \tag{10}$$

Champredon *et al.* [15] define the proportion of infectious extra-couple and uncoupled contact through the following term:

$$P = \frac{c_u I + c_e(SI + 2II)}{c_u(S + I) + 2c_e(SS + SI + II)}. \tag{11}$$

The effective uncoupled, c_u, and extra couple, c_e, contact rates are the product of two terms: uncoupled/extra-couple contact rate $\times$ rate of transmission per contact. Therefore, the transmission rate per contact term in c_u and c_e is canceled out in the equation above. Using this property, we modify the equation above as follows:

$$P = \frac{r_u I + r_e(SI + 2II)}{r_u(S + I) + 2r_e(SS + SI + II)}, \tag{12}$$

where $r_u = c_u/c_w$ and $r_e = c_e/c_w$ are the relative uncoupled/extra-couple contact rates. This simplification is useful in a multi-strain model since we cannot multiply a vector with a single value (e.g. $c_u S$ in denominator) if we use Champredon *et al.*'s equation in its original form. Extending the above equation to multi-strain model so that P_i represents the proportion of the extra-couple and uncoupled contact of an infected individual with strain i, we obtain:

$$P_i = \frac{r_u I_i + r_e(SI_i + \sum_k (II_{ik} + \delta_{ik} II_{ik}))}{r_u(S + \sum_k I_k) + r_e(2SS + \sum_k 2SI_k + \sum_l \sum_k 2^{\delta}_{lk} II_{lk})}. \tag{13}$$

Using the equation above, we can model extra-pair transmission (ET). For convenience, uncoupled and extra-couple transmission rates, c_u and c_e, will be replaced with $U_i = r_u\beta_i$ and $E_i = r_e\beta_i$ hereafter.

Single susceptible individuals become infected through uncoupled contact at per-person rate $\sum_k P_k U_k$ and enter single infected state. Through mutation, newly infected individuals are distributed into single infected compartments with different

strains: $\mathrm{ET}(I_i) = \sum_k M_{ki} P_k U_k S$. Either partner in an SS partnership becomes infected at per-person rate $\sum_k P_k E_k$, and partnership state changes to an SI partnership at the total rate of $\sum_i 2 P_i E_i SS$. The formation of SI_i partnerships is similar to the process through which single susceptible individuals are distributed into single infected compartments: $\mathrm{ET}(SI_i) = \sum_k 2 M_{ki} P_k E_k SS$. Lastly, the susceptible partner in an SI partnership becomes infected from extra-couple contacts at a per-person rate of $\sum_k P_k E_k$, and partnership state changes to an II partnership. As in the previous cases, SI_i partnerships are lost at a rate of $\sum_k P_k E_k SI_i$. The mutation process is similar to that of within-couple transmission. The only difference is that the $\log_{10}$ SPVL of a newly infected partner is not determined by its social partner but from an extra-couple partner (i.e. the term P_i):
$\mathrm{ET}(II_{ij}) = (\frac{1}{2})^{\delta_{ij}} (\sum_k (M_{kj} P_k E_k SI_i + M_{ki} P_k E_k SI_j))$. Combining these equations we get the following set of equations that describe all the transmission dynamics:

$$
\begin{aligned}
S' &= -\sum_k P_k U_k S, \\
I_i' &= \sum_k M_{ki} P_k U_k S, \\
SS' &= -\sum_i 2 P_i E_i SS, \\
SI_i' &= \sum_k 2 M_{ki} P_k E_k SS - \beta_i SI_i - \sum_k P_k E_k SI_i, \\
II_{ij}' &= \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot (M_{ij} \beta_i SI_i + M_{ji}, \beta_j SI_j) + \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot (\sum_k (M_{kj} P_k E_k SI_i \\
&\quad + M_{ki} P_k E_k SI_j)).
\end{aligned}
\tag{14}
$$

Pair formation models: Disease induced mortality

The per-person disease induced mortality (DM) rate, λ, is given by taking the reciprocal of the total duration of the infection: $\lambda_i = 1/(D_A + D_P(\alpha_i) + D_D)$. Since we assume an SIS formulation, where infected individuals that die from infection are immediately replaced by an individual in the single susceptible compartment, we obtain the following equation for single infected individuals:

$$\begin{aligned} \mathrm{DM}(S) &= \sum_k \lambda_k I_k, \\ \mathrm{DM}(I_i) &= -\lambda_i I_i. \end{aligned} \tag{15}$$

If an infected individual in a partnership dies, the partnership dissolves. Thus, an SI_i partnership dissolves at per-partnership rate λ_i, and the susceptible partner enters the single susceptible compartment at rate $\lambda_i SI_i$ (due to the SIS formulation, the infected partner that dies also gives rise, at an equal rate, to an individual entering the single susceptible compartment):

$$\begin{aligned} \mathrm{DM}(S) &= \sum_k 2\lambda_k SI_k, \\ \mathrm{DM}(SI_i) &= -\lambda_i SI_i. \end{aligned} \tag{16}$$

Similarly, since II_{ij} partnerships are composed of two infected partners, they dissolve at a per-partnership rate $(\lambda_i + \lambda_j)$. However, two cases, when $i \neq j$ and $i = j$, must be considered separately. When an II_{ij} partnership dissolves due to disease-induced mortality, where $i \neq j$, the death of the partner with strain i causes its partner to enter I_j compartment at rate $\lambda_j II_{ij}$, and vice versa. When an $II_i i$ partnership dissolves, the death of either partner causes the other partner to enter the I_i compartment at rate $\lambda_i II_{ii}$, which sums up to $2\lambda_i II_{ii}$. Combining these dynamics yields:

$$\begin{aligned} \mathrm{DM}(S) &= \sum_l \sum_k 2^{\delta_{lk}} \lambda_k II_{lk}, \\ \mathrm{DM}(I_i) &= \sum_k 2^{\delta_{ik}} \lambda_k II_{ik}, \\ \mathrm{DM}(II_{ij}) &= -(\lambda_i + \lambda_j) II_{ij}. \end{aligned} \tag{17}$$

Finally, combining all these equations give us the full model, which is Model 1. We can simply drop the uncoupled and extra-couple transmission terms to obtain model 2:

$$
\begin{aligned}
S' =& -\rho S + 2cSS + \sum_k cSI_k - \sum_k P_k U_k S + \sum_k \lambda_k I_k \\
&+ \sum_k 2\lambda_k SI_k + \sum_l \sum_k 2^{\delta_{lk}} \lambda_k II_{lk} \\
I_i' =& -\rho I_i + cSI_i + \sum_k 2^{\delta_{ik}} cII_{ik} + \sum_k M_{ki} P_k U_k S - \lambda_i I_i \\
&+ \sum_k 2^{\delta_{ik}} \lambda_k II_{ik} \\
SS' =& \frac{\rho S \cdot S}{2(S + \sum_k I_k)} - cSS - \sum_i 2P_i E_i SS \\
SI_i' =& \frac{\rho S \cdot I_i}{S + \sum_k I_k} - cSI_i - \beta_i SI_i + \sum_k 2M_{ki} P_k E_k SS - \sum_k P_k E_k SI_i \\
&- \lambda_i SI_i \\
II_{ij}' =& \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot \frac{\rho I_i \cdot I_j}{(S + \sum_k I_k)} - cII_{ij} + \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot (M_{ij}\beta_i SI_i + M_{ji}\beta_j SI_j) \\
&+ \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot (\sum_k (M_{kj} P_k E_k SI_i + M_{ki} P_k E_k SI_j)) - (\lambda_i + \lambda_j) II_{ij}
\end{aligned}
\tag{18}
$$

Models 3 ("instswitch") and 4 ("instswitch")

Partnership dynamics

Since model 3 and 4 assume instantaneous partnership formation, there are only three states: SS, SI_i, and II_{ij}. Partnership dissolution rates are equal to those of model 1 and 2: $\mathrm{DS}(SS) = -cSS$, $\mathrm{DS}(SI_i) = -cSI_i$, and $\mathrm{DS}(II_{ij}) = -cII_{ij}$. Once individuals leave a partnership, they are instantaneously distributed into coupled states. In order to make the equations simpler, we introduce the following two terms: X and Y_i, where X denotes the number of susceptible individuals that leave the partnership at a given time, and Y_i the number of infected individuals with $\log_{10}$ SPVL of α_i who leave partnership at a given time. These temporarily single individuals then form couples through the same partnership formation rule described in the previous section:

$$
\begin{aligned}
X &= 2cSS + \sum_k cSI_k \\
Y_i &= cSI_i + \sum_k 2^{\delta_{ik}} cII_{ik} \\
SS' &= -cSS + \frac{X^2}{2(X + \sum_k Y_k)} \\
SI'_i &= -cSI_i + \frac{XY_i}{X + \sum_k Y_k} \\
II'_{ij} &= -cII_{ij} + \left(\frac{1}{2}\right)^{\delta_{ij}} \frac{Y_i Y_j}{X + \sum_k Y_k}.
\end{aligned}
\tag{19}
$$

Instantaneous-switching models: Infection dynamics

Model 3 and 4 share the within-couple transmission term with model 1 and 2. Since there is no single (uncoupled) state, only extra-couple transmission exists:

$$
P_i = \frac{r_e(SI_i + \sum_k (II_{ik} + \delta_{ik} II_{ik}))}{r_e(2SS + \sum_k 2SI_k + \sum_l \sum_k (2^{\delta}_{kl} II_{lk}))}. \tag{20}
$$

Movement from SS state to SI state and SI to SS is modeled through the same equation that is used in model 1 and 2.

Instantaneous-switching models: Disease induced mortality

Disease induced mortality is modeled similar to model 1 and 2. However, as single state does not exist in model 3 and 4, individuals that has left their partnerships due to death of their partners enter temporary compartments and form partners instantly:

$$
\begin{aligned}
X &= \sum_k 2\lambda_k SI_k + \sum_l \sum_k 2^{\delta_{lk}} \lambda_k II_{lk}, \\
Y_i &= \sum_k 2^{\delta_{ik}} \lambda_k II_{ik}, \\
SS' &= = \frac{X^2}{2(X + \sum_k Y_k)}, \\
SI'_i &= -\lambda_i SI_i + \frac{XY_i}{X + \sum_k Y_k}, \\
II'_{ij} &= -(\lambda_i + \lambda_j) II_{ij} + \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot \frac{Y_i Y_j}{X + \sum_k Y_k}.
\end{aligned}
\tag{21}
$$

Combining all these dynamics, we have model 3. If we remove extra-couple transmission, we have model 4.

$$
\begin{aligned}
X =& 2cSS + \sum_k cSI_k + \sum_k 2\lambda_k SI_k + \sum_l \sum_k 2^{\delta_{lk}} \lambda_k II_{lk}, \\
Y_i =& cSI_i + \sum_k 2^{\delta_{ik}} cII_{ik} + \sum_k 2^{\delta_{ik}} \lambda_k II_{ik}, \\
SS' =& -cSS + \frac{X^2}{2(X + \sum_k Y_k)} - \sum_i 2P_i E_i SS, \\
SI_i' =& -cSI_i + \frac{XY_i}{X + \sum_k Y_k} - \beta_i SI_i + \sum_k 2M_{ki} P_k E_k SS \\
& - \sum_k P_k E_k SI_i - \lambda_i SI_i, \\
II_{ij}' & - cII_{ij} + \left(\frac{1}{2}\right)^{\delta_{ij}} \frac{Y_i Y_j}{X + \sum_k Y_k} + \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot (M_{ij}\beta_i SI_i + M_{ji}\beta_j SI_j) \\
& + \left(\frac{1}{2}\right)^{\delta_{ij}} \cdot (\sum_k (M_{kj} P_k E_k SI_i + M_{ki} P_k E_k SI_j)) - (\lambda_i + \lambda_j) II_{ij}.
\end{aligned}
\tag{22}
$$

Implicit model

Following [11], Model 5 is an implicit instantaneous partnership formation model that uses an adjusted transmission rate, β^*, that is derived from [18]'s approximate basic reproduction number:

$$
\beta_i^* = \frac{c\beta_i}{c + \beta_i + \lambda_i}. \tag{23}
$$

Thus, we get the following model:

$$
\begin{aligned}
S' &= \sum_k \lambda_k I_k - \sum_k \beta_k^* SI_k. \\
I_i' &= \sum_k M_{ki} \beta_k^* SI_k - \lambda_i I_i.
\end{aligned}
\tag{24}
$$

Random-mixing model

Model 6 is a random mixing model. It is modeled in a same way as model 5 without the adjusted transmission rate:

$$
\begin{aligned}
S' &= \sum_k \lambda_k I_k - \sum_k \beta_k SI_k, \\
I_i' &= \sum_k M_{ki} \beta_k SI_k - \lambda_i I_i.
\end{aligned}
\tag{25}
$$

Heteregenous model

We extend "pairform+epc model" by allowing for heterogeneity in sexual behaviour. Since pairform+epc model captures four distinct sexual behaviours – pair formation, pair dissolution, extra-couple mixing, and uncoupled mixing – we assume that all four parameters that model the mentioned behaviours are scaled by the same factor based on the risk group. In other words, an individual in a higher risk is more likely to form a stable partnership, leave a stable partnership, and engage in a extra-couple/uncoupled mixing. We denote this scaling parameter as φ_i where i is the risk group. For simplicity, we assume that the transmission rate per partnership is not affected by sexual behaviour.

Partnership dynamics

Individuals in a risk group i leave single state at per-person rate $\varphi_i\rho$. Let $XY_{ij,kl}$ be a coupled state where X and Y are the infection status (susceptible or infected) of each partner, k and l are the risk groups X and Y belong to respectively, and i and j are the strains of an infected partner. If a partner is susceptible, strain index is replaced by $\cdot$. For example, $SI_{\cdot j,kl}$ is the number of partners where the susceptible partner is in risk group k and infected partner is in risk group l and has $\log_{10}$ SPVL of α_j. For simplicity, we assume that people mix randomly. Then, we can write the partnership formation process as follows:

$$\mathrm{PF}(XY_{ij,kl}) = \left(\frac{1}{2}\right)^{\delta_{ij}\delta_{kl}} \frac{\varphi_k\rho X_{i,k}\varphi_l\rho Y_{j,l}}{\sum_m \varphi_m\rho(S_{\cdot,m} + \sum_n I_{n,m})} \tag{26}$$

For dissolution process, individual in risk group i expects to leave partnership at a rate $\varphi_i c$. If partnership is formed between two individuals from a different risk group, rate at which they leave the partnership differs. We resolve this conflict by assuming that a partnership dissolution rate of a couple is equal to the average of that of two partners. Therefore, $XY_{ij,kl}$ dissolve at per-partnership rate $\frac{\varphi_k+\varphi_l}{2}c$, and both $X_{i,k}$ and $Y_{j,l}$ partners return to single state at the same rate.

Heterogeneous models: Infection dynamics

Since we assume that the rate of transmission per partnership stays constant across different risk groups, within-couple infection process is similar to other models:

$$\begin{aligned} WT(SI_{\cdot,j,kl} &= -\beta_j SI_{\cdot,j,kl} \\ WT(II_{ij,kl}) &= \left(\frac{1}{2}\right)^{\delta_{ij}\delta_{kl}} \cdot (M_{ji}\beta_j SI_{\cdot,j,kl} + M_{ij}\beta_i SI_{\cdot,i,lk}) \end{aligned} \tag{27}$$

Note that $II_{ij,kl}$ can be formed from two types of partnerships: 1) Infected partner with $\log_{10}$ SPVL of α_j and risk group of l infects a susceptible partner in risk group k, yielding $\log_{10}$ SPVL of α_i through mutation. 2) Infected partner with $\log_{10}$ SPVL of α_i and risk group of k infected a susceptible partner in risk group l, yielding $\log_{10}$ SPVL of α_j through mutation. On the other hand, if $i = j$ and $k = l$, $II_{ii,kk}$ can only be formed from $SI_{\cdot,i,kk}$ partnership, which is resolved by $\left(\frac{1}{2}\right)^{\delta_{ij}\delta_{kl}}$.

Heterogeneous extra-couple and uncoupled contact process is similar to partnership formation process. Relative uncoupled/extra-couple contact rates are scaled by the factor of φ_i, where i is the risk group. First, we define Q_i as the total rate of uncoupled/extra couple contact by individuals in risk group k:

$$\begin{aligned} Q_i =& \varphi_i r_u (S_{\cdot,i} + \sum_j I_{j,i}) + \varphi_i r_e \Bigg(\sum_k 2^{\delta_{ik}} SS_{\cdot,ik} + \\ & \sum_l \sum_j (SI_{\cdot,j,il} + SI_{\cdot,j,li}) + \sum_j \sum_l \sum_k 2^{\delta_{kl}\delta_{ij}} II_{kl,ij} \Bigg) \end{aligned} \tag{28}$$

We now define $P_{k,i}$ as the proportion of the extra-couple and uncoupled contact that arises from an infected individual from risk group i with $\log_{10}$ SPVL of α_k:

$$P_{k,i} = \frac{\varphi_i r_u I_{k,i} + \varphi_i r_e (SI_{k,i} + \sum_j \sum_l 2^{\delta_{kl}\delta_{ij}} II_{kl,ij})}{\sum_j Q_j} \tag{29}$$

Since the relative uncoupled/extra couple contact ratio are scaled by the factor of φ_i, uncoupled and extra-couple transmission rates are scaled by the same factor as well: $U_{k,i} = \varphi_i r_u \beta_k$ and $E_{k,i} = \varphi_i r_e \beta_k$. Once again, we assume random mixing between individuals. Then, a susceptible individual in risk group i becomes infected through extra-couple and uncoupled contact at a per capita rate of $\sum_j \sum_k P_{k,j} X_{k,i}$. Once infected, individuals are distributed into each Strain categories through mutation.

Heterogeneous model: Disease induced mortality

Disease induced mortality is not affected by the sexual behaviour of an individual.

Initial distribution of infected individuals

We follow Champredon *et al.*'s result to calculate the initial distribution of infected individuals. For model 1 and 2, we have disease equilibrium state of $S^* = \frac{c}{c+\rho}$ and $SS^* = \frac{1-S^*}{2}$. We let $\epsilon = 10^{-4}$, which is the total number of infected individuals in the beginning of simulation and D be the vector such that D_i represent the proportion of individuals with $\log_{10}$ SPVL of i. Y_i is taken from normal distribution with mean 3 and is normalized so that $\sum_i D_i = 1$. Then, we have the following initial distribution of each states:

$$\begin{aligned} S(0) &= (1-\epsilon)S^*, \\ SS(0) &= (1-\epsilon)^2 SS^*, \\ SI_i(0) &= 2\epsilon(1-\epsilon)SS^* D_i, \\ I_i(0) &= \epsilon S^* D_i, \\ II_{ij}(0) &= \left(\frac{1}{2}\right)^{\delta_{ij}} 2\epsilon^2 SS^* D_i D_j. \end{aligned} \tag{30}$$

Since model 3 and 4 do not have single state, $SS^* = 1$ at the disease free equilibrium and the initial distribution becomes as follows:

$$\begin{aligned} SS(0) &= (1-\epsilon)^2 SS^*, \\ SI_i(0) &= 2\epsilon(1-\epsilon)SS^* D_i, \\ II_{ij}(0) &= \left(\frac{1}{2}\right)^{\delta_{ij}} 2\epsilon^2 SS^* D_i D_j. \end{aligned} \tag{31}$$

As model 5 is an implicit model, which does not consider different stages of partnership, we have the following initial distribution:

$$\begin{aligned} S(0) &= 1-\epsilon, \\ I_i(0) &= \epsilon D_i. \end{aligned} \tag{32}$$

Model 6 has the same distribution of initial infected individuals as model 5.

Lastly, for heterogeneity model, we assume that the risk distribution of the

population follows gamma distribution and calculate the shape and scale parameters given the mean and squared coefficient of variation. Using the shape and scale parameterws, we define gamma quantile function $Q(p)$ and $p_j = p_{\min} + (p_{\max} - p_{\min})\frac{j-1}{n_r+1}$, where n_r is number of risk groups and $j = 1, 2, 3, \ldots, n_r + 1$. Since $Q(1) = \infty$, we set $p_{\max} = 0.99$ and $p_{\min} = 0.01$. Then, we define $\varphi_i = \frac{Q(p_j)+Q(p_{j+1})}{2}$. We define R_i as the proportion of individuals in risk group i in a disease free equilibrium and assume R_i is equal for all i, i.e. $R_i = \frac{1}{n_r}$. In order to start the simulation in a quasi-equilibrium state, we first run the model with the following initial state:

$$\begin{aligned} S_{\cdot,i}(0) &= (1-\epsilon)R_i, \\ I_{k,i}(0) &= \epsilon D_k R_i, \\ SS_{\cdot,ij}(0) &= SI_{\cdot k,ij}(0) = II_{kl,ij}(0) = 0. \end{aligned} \tag{33}$$

For this particular simulation, we disregard infection process as well as disease induced mortality in order to preserve the strain distribution of infected individuals. Furthermore, since the scaling parameter, γ, does not affect the risk group distribution in the absence of disease transmission, we increase the scaling parameter to 5 ($\gamma = 5$) to speed up the simulation and run the model for 50 years. After the model has reached its quasi-equilibrium state, we take this distribution of susceptible and infected individuals as the initial state of the actual simulation.

Appendix S2: dynamics of transmission and virulence This section presents

References

1. Anderson RM, May RM. Coevolution of hosts and parasites. Parasitology. 1982;85(02):411–426.

2. Ebert D. The evolution and expression of parasite virulence. In: Stearns SC, editor. Evolution in Health & Disease. New York: Oxford University Press, Oxford, UK; 1999. p. 161–172.

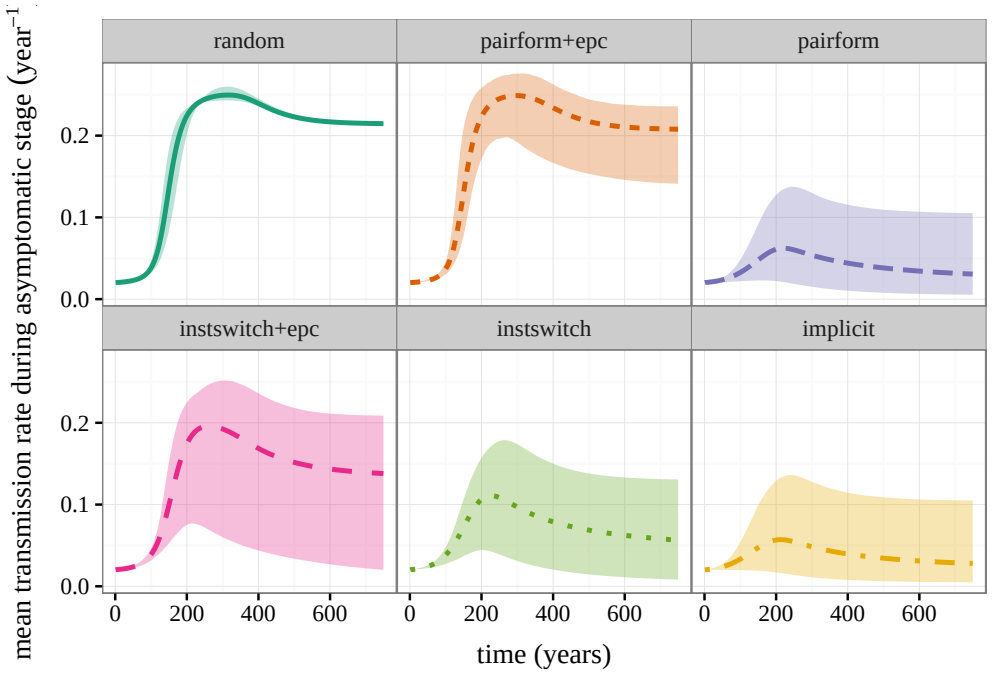

Fig 6. Envelopes of transmission trajectories under all models. This figure matches Fig 2, but displays the envelope of population-mean transmission probabilities rather than $\log_{10}$ SPVL over time for each model.

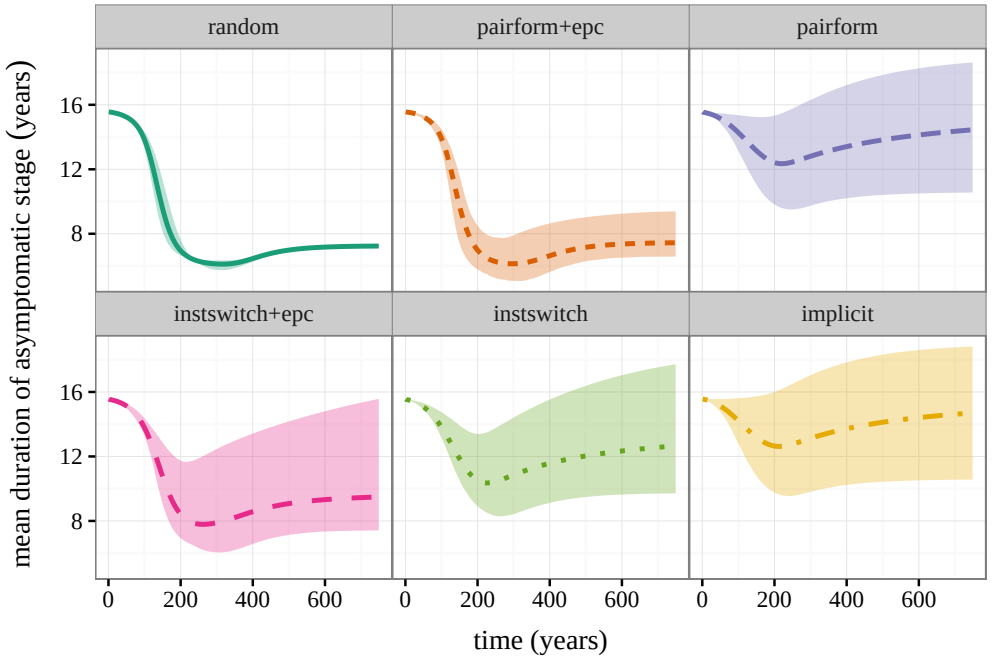

Fig 7. Envelopes of progression trajectories under all models. This figure matches Fig 2, but displays the envelope of population-mean expected time of progression to AIDS (i.e., length of intermediate HIV phase) rather than $\log_{10}$ SPVL over time for each model.

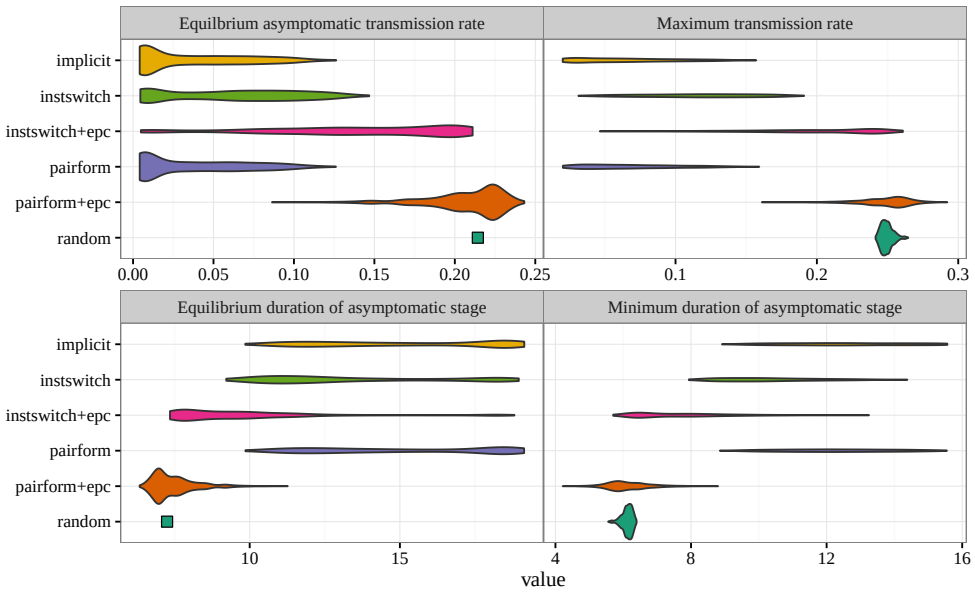

Fig 8. Univariate distributions of transmission probabilities and progression. This figure matches Fig 3, but displays the univariate distributions for the transmission probability and progression time at the virulence peak and at the epidemiological equilibrium, rather than the distributions of $\log_{10}$ SPVL.

3. Ebert D, Bull JJ. Challenging the trade-off model for the evolution of virulence: is virulence management feasible? Trends Microbiol. 2003;11(1):15–20.

4. Alizon S, Michalakis Y. Adaptive virulence evolution: the good old fitness-based approach. Trends in Ecology & Evolution. 2015;30(5):248–254. doi:10.1016/j.tree.2015.02.009.

5. Dwyer G, Levin SA, Buttel L. A simulation model of the population dynamics and evolution of myxomatosis. Ecol Monog. 1990;60:423–447.

6. Mackinnon MJ, Read AF. Genetic relationships between parasite virulence and transmission in the rodent malaria *Plasmodium chabaudi*. Evolution. 1999; p. 689–703.

7. Jensen KH, Little T, Skorping A, Ebert D. Empirical support for optimal virulence in a castrating parasite. PLoS Biol. 2006;4(7):e197.

8. De Roode JC, Yates AJ, Altizer S. Virulence-transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. Proceedings of the National Academy of Sciences. 2008;105(21):7489–7494.

9. Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP. Variation in HIV-1 set-point viral load: Epidemiological analysis and an evolutionary hypothesis. PNAS. 2007;104:17441–17446.

10. Fraser C, Lythgoe K, Leventhal GE, Shirreff G, Hollingsworth TD, Alizon S, et al. Virulence and Pathogenesis of HIV-1 Infection: An Evolutionary Perspective. Science. 2014;343(6177):1243727. doi:10.1126/science.1243727.

11. Shirreff G, Pellis L, Laeyendecker O, Fraser C. Transmission Selects for HIV-1 Strains of Intermediate Virulence: A Modelling Approach. PLoS Computational Biology. 2011;7(10):e1002185. doi:10.1371/journal.pcbi.1002185.

12. Herbeck JT, Mittler JE, Gottlieb GS, Mullins JI. An HIV Epidemic Model Based on Viral Load Dynamics: Value in Assessing Empirical Trends in HIV Virulence and Community Viral Load. PLoS Comput Biol. 2014;10(6):e1003673.

13. Day T, Proulx SR. A General Theory for the Evolutionary Dynamics of Virulence. The American Naturalist. 2004;163(4):E40–E63. doi:10.1086/382548.

14. Alizon S. The Price equation framework to study disease within-host evolution. Journal of Evolutionary Biology. 2009;22(5):1123–1132. doi:10.1111/j.1420-9101.2009.01726.x.

15. Champredon D, Bellan S, Dushoff J. HIV Sexual Transmission Is Predominantly Driven by Single Individuals Rather than Discordant Couples: A Model-Based Approach. PLoS ONE. 2013;8(12):e82906. doi:10.1371/journal.pone.0082906.

16. Dietz K, Hadeler KP. Epidemiological models for sexually transmitted diseases. Journal of Mathematical Biology. 1988;26(1):1–25.

17. Kretzschmar M, Dietz K. The effect of pair formation and variable infectivity on the spread of an infection without recovery. Mathematical Biosciences. 1998;148(1):83–113.

18. Hollingsworth TD, Anderson RM, Fraser C. HIV-1 Transmission, by Stage of Infection. Journal of Infectious Diseases. 2008;198(5):687–693. doi:10.1086/590501.

19. Omori R, Chemaitelly H, Abu-Raddad LJ. Dynamics of non-cohabiting sex partnering in sub-Saharan Africa: a modelling study with implications for HIV transmission. Sexually transmitted infections. 2015; p. sextrans–2014.

20. Delva W, Leventhal GE, Helleringer S. Connecting the dots: network data and models in HIV epidemiology. AIDS. 2016;30(13):2009–2020. doi:10.1097/QAD.0000000000001184.

21. Blower SM, Hartel D, Dowlatabadi H, Anderson RM, May RM. Drugs, Sex and HIV: A Mathematical Model for New York City. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 1991;331(1260):171–187. doi:10.1098/rstb.1991.0006.

22. Ma J, Dushoff J, Bolker BM, Earn DJD. Estimating Initial Epidemic Growth Rates. Bulletin of Mathematical Biology. 2014;76(1):245–260. doi:10.1007/s11538-013-9918-2.

23. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proceedings of the Royal Society of London B: Biological Sciences. 2007;274(1609):599–604. doi:10.1098/rspb.2006.3754.

24. Soetaert K, Petzoldt T, Setzer RW. Solving Differential Equations in R: Package deSolve. Journal of Statistical Software. 2010;33(9):1–25. doi:10.18637/jss.v033.i09.

25. Herbeck J, Mittler J, Gottlieb G, Goodreau S, Murphy J, Cori A, et al. Evolution of HIV virulence in response to widespread scale up of antiretroviral therapy: a modeling study. bioRxiv. 2016; p. 039560.

26. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, et al. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. Journal of Infectious Diseases. 2005;191(9):1403–1409.

27. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. The early spread and epidemic ignition of HIV-1 in human populations. Science (New York, NY). 2014;346(6205):56–61. doi:10.1126/science.1256739.

28. Elderd BD, Dukic VM, Dwyer G. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. Proceedings of the National Academy of Sciences. 2006;103(42):15693 –15697. doi:10.1073/pnas.0600816103.

29. Herbeck JT, Müller V, Maust BS, Ledergerber B, Torti C, Di Giambenedetto S, et al. Is the virulence of HIV changing? A meta-analysis of trends in prognostic markers of HIV disease progression and transmission. AIDS (London, England). 2012;26(2):193–205. doi:10.1097/QAD.0b013e32834db418.

30. Payne R, Muenchhoff M, Mann J, Roberts HE, Matthews P, Adland E, et al. Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. Proceedings of the National Academy of Sciences. 2014;111(50):E5393–E5400. doi:10.1073/pnas.1413339111.

31. Roberts HE, Goulder PJ, McLean AR. The impact of antiretroviral therapy on population-level virulence evolution of HIV-1. Journal of The Royal Society Interface. 2015;12(113):20150888.