

RESEARCH ARTICLE

HIV Competition Dynamics over Sexual Networks: First Comer Advantage Conserves Founder Effects

Bence Ferdinandy¹, Enys Mones¹, Tamás Vicsek^{1,2}, Viktor Müller^{3,4*}

1 Department of Biological Physics, Eötvös Loránd University, Budapest, Hungary, **2** MTA-ELTE Statistical and Biological Physics Research Group, Eötvös Loránd University and the Hungarian Academy of Sciences, Budapest, Hungary, **3** MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Eötvös Loránd University and the Hungarian Academy of Sciences, Budapest, Hungary, **4** Parmenides Center for the Conceptual Foundations of Science, Pullach/Munich, Germany

* mueller.viktor@gmail.com



OPEN ACCESS

Citation: Ferdinandy B, Mones E, Vicsek T, Müller V (2015) HIV Competition Dynamics over Sexual Networks: First Comer Advantage Conserves Founder Effects. *PLoS Comput Biol* 11(2): e1004093. doi:10.1371/journal.pcbi.1004093

Editor: Christian L. Althaus, University of Bern, SWITZERLAND

Received: August 19, 2014

Accepted: December 16, 2014

Published: February 5, 2015

Copyright: © 2015 Ferdinandy et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: VM is a Fellow of the Parmenides Center for the Conceptual Foundations of Science (<http://www.parmenides-foundation.org/>). This work was partly supported by the EU ERC COLLMOT project (grant No. 227878, erc.europa.eu/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Outside Africa, the global phylogeography of HIV is characterized by compartmentalized local epidemics that are typically dominated by a single subtype, which indicates strong founder effects. We hypothesized that the competition of viral strains at the epidemic level may involve an advantage of the resident strain that was the first to colonize a population. Such an effect would slow down the invasion of new strains, and thus also the diversification of the epidemic. We developed a stochastic modelling framework to simulate HIV epidemics over dynamic contact networks. We simulated epidemics in which the second strain was introduced into a population where the first strain had established a steady-state epidemic, and assessed whether, and on what time scale, the second strain was able to spread in the population. Simulations were parameterized based on empirical data; we tested scenarios with varying levels of overall prevalence. The spread of the second strain occurred on a much slower time scale compared with the initial expansion of the first strain. With strains of equal transmission efficiency, the second strain was unable to invade on a time scale relevant for the history of the HIV pandemic. To become dominant over a time scale of decades, the second strain needed considerable (>25%) advantage in transmission efficiency over the resident strain. The inhibition effect was weaker if the second strain was introduced while the first strain was still in its growth phase. We also tested how possible mechanisms of interference (inhibition of superinfection, depletion of highly connected hubs in the network, one-time acute peak of infectiousness) contribute to the inhibition effect. Our simulations confirmed a strong first comer advantage in the competition dynamics of HIV at the population level, which may explain the global phylogeography of the virus and may influence the future evolution of the pandemic.

Author Summary

The African epicentre of the HIV pandemic is home to a vast array of divergent viruses; however, local epidemics in other parts of the world are typically dominated by a single variant (subtype) of the virus, with different subtypes found in the different regions. This pattern indicates that local epidemics outside Africa have been started by the introduction of single “founder” viruses in the susceptible populations. However, how these patterns persisted over several decades in the face of international migration requires further explanation. By analyzing simulated epidemics, we demonstrated that an epidemic established by the first successful founder strain can inhibit the introduction and slow down the subsequent spread of further virus strains by several mechanisms of interference. Our results have implications for the global evolution of the HIV pandemic: the fast expansion of subtypes benefited from a “first comer advantage,” and founder viruses may have been selected by random sampling, rather than due to superior transmissibility/fitness; the fast expansion of these (possibly) suboptimal virus strains may have considerably delayed the spread of more transmissible HIV variants; however, the future evolution of the pandemic is likely to be characterized by a slow expansion of viral strains with increased transmission potential.

Introduction

The global pandemic of Human Immunodeficiency Virus (HIV) infections is being driven mainly by the group M lineage of HIV-1, which crossed the species barrier from chimpanzees to humans about 100 years ago [1,2]. By the time it started to spread beyond its epicentre in Central Africa, the virus had already accumulated considerable sequence diversity [2], and distinct divergent clades initiated a series of rapid expansions that gave rise to the subtypes of HIV-1 group M [3,4]. The global molecular diversity of the pandemic still bears the clear footprint of the strong founder effects that characterized this initial expansion. While diversity is very high near the epicentre of the epidemic in Central Africa, the epidemics of other regions are typically characterized by the dominance of at most a few subtypes or circulating recombinant forms (CRFs) [5]. The countries where more than one subtype is prevalent tend to be characterized by parallel, compartmentalized epidemics with distinct subtypes infecting different risk or ethnic groups [6–9], and transmission chains rarely cross national borders [10]. While the global spatial distribution of HIV subtypes is not completely static, the diversification of the epidemic and shifts between subtypes occur very slowly in most regions [5].

Understanding the factors that set the time scale of HIV competition dynamics at the population level has great practical relevance. Subtypes differ in both transmissibility [11–13] and the rate of disease progression [14,15], and further variation in these traits is bound to exist within the subtypes and in the vast diversity of unique recombinant forms (URFs) and unclassified basal lineages in Central Africa [16–18]. Virus variants that have higher transmission potential are likely to be spreading at the expense of less efficient strains, and epidemics may expand as the original variants are gradually replaced by “fitter” viral lineages. The risk and pace of these processes needs to be better characterized.

We developed a simple model of sexually transmitted HIV epidemics that allowed us to monitor the competition dynamics of distinct virus strains with varying rates of transmission. In sexually transmitted epidemics, HIV is transmitted over the network of sexual contacts, which tends to include a very limited subset of all possible contacts, i.e. the host population is very far from “free mixing”. To be able to detect effects arising from network structure, we

implemented an individual based simulation that allowed us to track a dynamic network of contacts between the individuals, and to model HIV transmission along the links of the network. The model was parameterized based on data from generalized heterosexual epidemics in Africa.

Results

Our model tracked a dynamic network of contacts between three types of nodes (individuals): male, female, and female sex worker (FSW). Stochastic processes were implemented for the formation and dissolution of links (heterosexual relationships); sexual acts over the links; transmission of HIV in serodiscordant acts; and death due to AIDS. Natural (non-AIDS related) turnover was implemented by tracking the age of individuals, and replacing individuals at the end of sexual activity; the network of contacts was parameterized based on contemporary data including variability in the promiscuity of individuals. The transmissibility of HIV depended on disease stage.

We simulated a simple scenario of competition between two strains of the virus. To assess the maximum potential for a “first comer advantage”, we started the simulations with one of the strains (the founder, or “resident” strain) and let the epidemics attain steady-state prevalence before introducing the second (“invader”) virus strain. The transmission rate of the invader strain was equal to or greater than that of the resident strain, and its chance and pace of growth was assessed in relation to its transmission advantage over the resident strain.

We hypothesized that the effect on the spread of the invader strain may depend on the prevalence of the resident strain, and have therefore considered two scenarios, where the steady-state prevalence of the resident strain was around 0.03 and 0.2, respectively. The two scenarios were set by changing the baseline rate of transmissibility (see [Materials and Methods/Parameters](#)); all other parameters were kept constant. [Fig. 1](#) shows the time course of multiple simulations for two selected cases where the invader virus had equal ([Fig. 1A,C](#)) or 25 percent greater ([Fig. 1B,D](#)) transmission rate compared with the resident strain in the high ([Fig. 1A-B](#)) or low ([Fig. 1C-D](#)) prevalence scenarios ([S1 Fig.](#) depicts cases with further values of the transmission rate advantage). The resident strain attains steady-state prevalence in about 84 and 74 years in the low- and high-prevalence scenario, respectively. With equal transmission rate, the invader strain shows no appreciable growth in a hundred years in the high-prevalence scenario ([Fig. 1A](#)), and grows, but remains in strong minority over the same time span in the majority of the simulations with the low-prevalence scenario ([Fig. 1C](#)). A 25 percent advantage in the transmission rate allowed the invader virus to outgrow the resident strain in both scenarios ([Fig. 1B,D](#)), but it still took a median of 60 and 104 years until the prevalence of the invader strain reached that of the resident strain in the low- and high-prevalence setting; due to its higher transmissibility, the invader strain was then able to attain higher steady-state prevalence compared with the initial steady state of the resident strain. Compared with the initial expansion of the resident strain, the expansion of the invader was much slower in all cases. In addition, 66.4 and 68.3 percent of the simulations with equal transmissibility of the invader resulted in the extinction of the invader virus in the low- and high-prevalence scenarios, respectively; extinction occurred in 2.4 and 1.2 percent of the cases when the invader had 25 percent transmission advantage. In contrast, with our settings the initial introduction of the resident virus was nearly always (in 998/1000 and 1000/1000 independent simulation runs of the low and high-prevalence settings, respectively) able to establish an epidemic that grew to steady state.

Mechanisms of interference

Our simulation framework allowed us to investigate three potential mechanisms of interference between the resident and the invader strains. First, infection with one HIV strain may afford

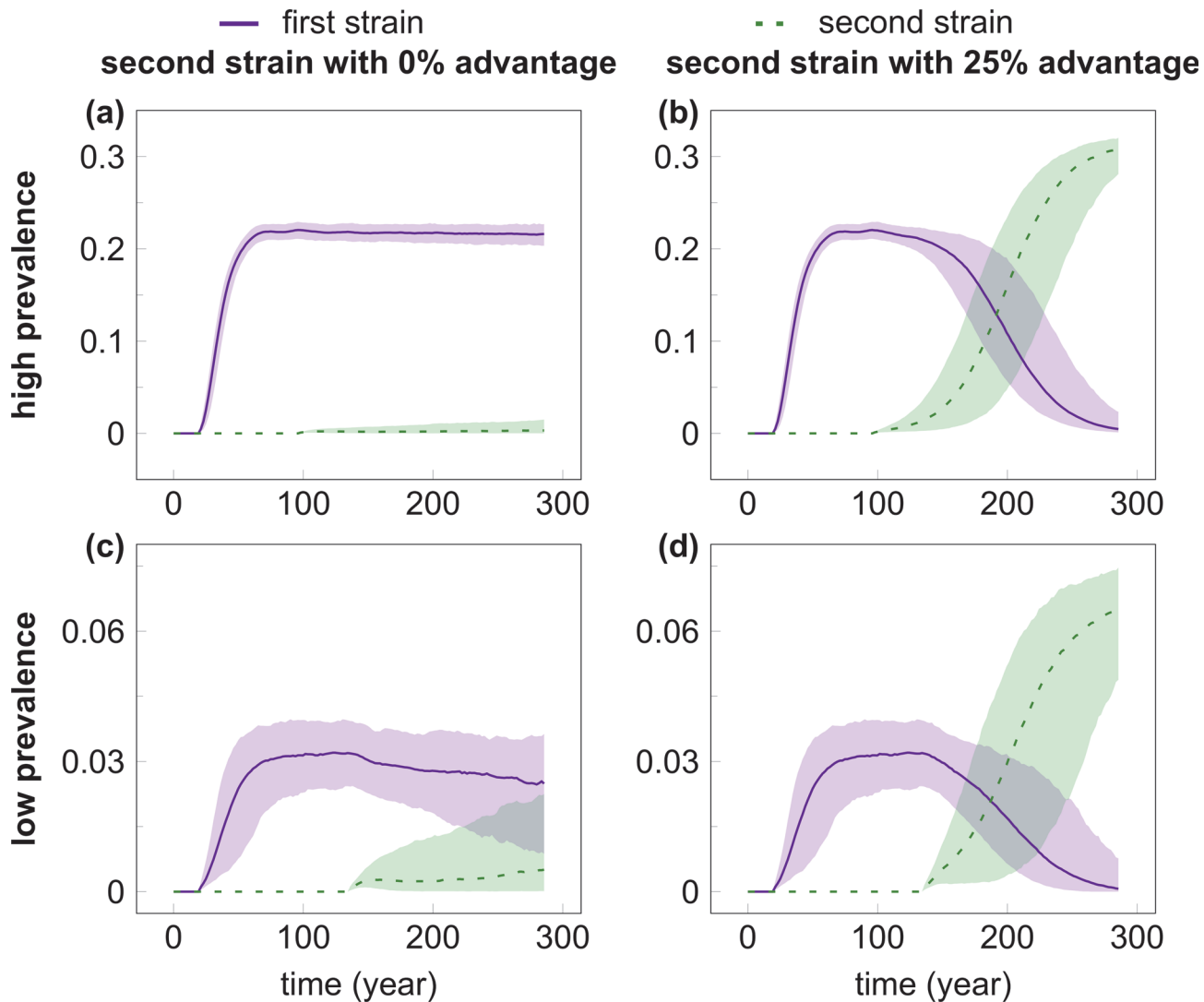


Fig 1. The time course of infection prevalence in multiple simulations of the competition of HIV strains. The invader virus had equal (A, C) or 25% greater (B, D) transmission rate compared with the resident strain in the high (A, B) or low (C, D) prevalence scenarios. The resident strain (solid purple line) was introduced in the population at Week 1000 (to allow the network to attain steady state); the invader strain (dashed green line) was introduced in the population when the first had already reached steady-state prevalence (at Week 5000 and 7000 for the high- and low-prevalence setting, respectively). Even with a 25% advantage in the transmission rate, it took the invader strain a median of 60 and 104 years to reach the prevalence of the resident strain in the low- and high-prevalence scenario, respectively. The lines show median prevalence from simulations where the invader strain did not go extinct (out of 1000 simulation runs); shading indicates the areas between the 5% and 95% quantiles. Simulation parameters were set as in [Table 1](#).

doi:10.1371/journal.pcbi.1004093.g001

some protection against superinfection with another strain: both the depletion of target cells and the induction of anti-HIV immune responses are likely to create less favourable conditions for infection compared with an uninfected individual. Because the strength of such an effect is still subject to debate (see [Discussion](#)), we used the following conservative approach in the default settings for our simulations. In a sexual act between two individuals who were both infected but with different strains of the virus, both strains had a chance to be transmitted in a two-step procedure. The first step tested successful initial transmission, which had a probability based on the transmission rate of the given strain, equivalent to the first infection of an uninfected individual. Then in the second step the superinfecting virus replaced the original strain with a probability based on the relative transmission rates of the two strains: in 50% of the

cases if both strains had equal transmission rate, and with smaller or greater probability if the superinfecting strain had lower or higher transmission rate, respectively, compared with the original strain (see [Materials and Methods](#) for details). In this scenario, the first strain had neither advantage nor disadvantage at the within-host level, and the “inhibition effect” arose only from the assumption that the infection of each individual is dominated by a single virus strain, implying replacement rather than coexistence upon superinfection (which is a reasonable simplification for the modelling of population level spread; see [Discussion](#)). This algorithm reduced the average probability of superinfection to 50% of that of initial infection, which is consistent with a recent prospective cohort study that estimated about two-fold lower hazard of superinfection compared with initial infection [19]. However, we also tested a “dual infection” scenario, in which superinfection was completely unhindered, and both strains were able to co-exist within one individual after superinfection occurred. In this scenario, in sexual acts between a dual infected and an uninfected individual both virus strains had an independent probability of being transmitted.

Second, we hypothesized that the peak of infectivity that characterizes acute infection [20,21] may not occur again upon superinfection due to depleted target cell levels and the presence of anti-HIV immune responses. If this is indeed the case then the first virus strain to colonize a population may take advantage of a rapid early wave of expansion fuelled by a high relative frequency of efficient acute stage transmissions in a largely susceptible population (the relative contribution of acute stage transmissions in our simulations is shown in [S2 Fig.](#)). In contrast, any subsequent “invader” strain is limited to the lower rates of chronic transmission that characterize mature epidemics ([22] and [S2 Fig.](#)), and even successfully superinfected individuals represent a diminished resource if acute peak infectivity cannot be repeated. We implemented this possible mechanism by keeping track of disease stage independently of the identity of the infecting strain. If superinfection occurred after the end of acute infection, the individual was assumed to remain in chronic stage and the onward transmission of the superinfecting strain occurred according to its baseline (chronic) transmission rate. When superinfection occurred during the acute stage of the initial infection, then the superinfecting strain received the benefit of enhanced acute-stage transmission for the remaining time of the acute stage, timed from the initial infection of the individual. However, some evidence indicates that superinfection can generate a new temporary peak of viremia at least in some of the cases [23]. We have therefore also tested a scenario where superinfection started a new window of enhanced acute-stage infectivity.

Third, we hypothesized that, in the absence of broadly available antiretroviral treatment (ART), the first HIV epidemics may also have an impact by selectively infecting and killing highly promiscuous individuals who form the “hubs” of the network. Such individuals have been shown to be particularly important for the spread of sexually transmitted diseases [24], and they are likely to be infected preferentially due to their larger number of contacts. Indeed, in our simulations the probability of infection was strongly related to the promiscuity (preferred contact degree) of the individuals ([Fig. 2A](#)). Using collated data from 100 simulation runs, logistic regression against log transformed contact degree (controlling also for age and gender) estimated an effect size of 2.48 (95% CI: 2.46–2.50; $p < 10^{-10}$), implying that the odds of being infected increased by a factor of $\exp(2.48)$, i.e., about 12-fold for every order of magnitude increase in the preferred contact degree (see [S3 Fig.](#) for the model fit); the effect was robust also in regressions on individual simulation runs (effect size range in 100 simulations: 2.26–2.66; $p < 10^{-10}$ for all simulation runs). As a result, an established epidemic of the resident virus strain depleted highly connected nodes of the network preferentially: the power-law exponent of the contact degree distributions (fitted to the actual number of partners in the last year) increased significantly compared with the pre-epidemic steady state ($p < 10^{-10}$, Wilcoxon rank

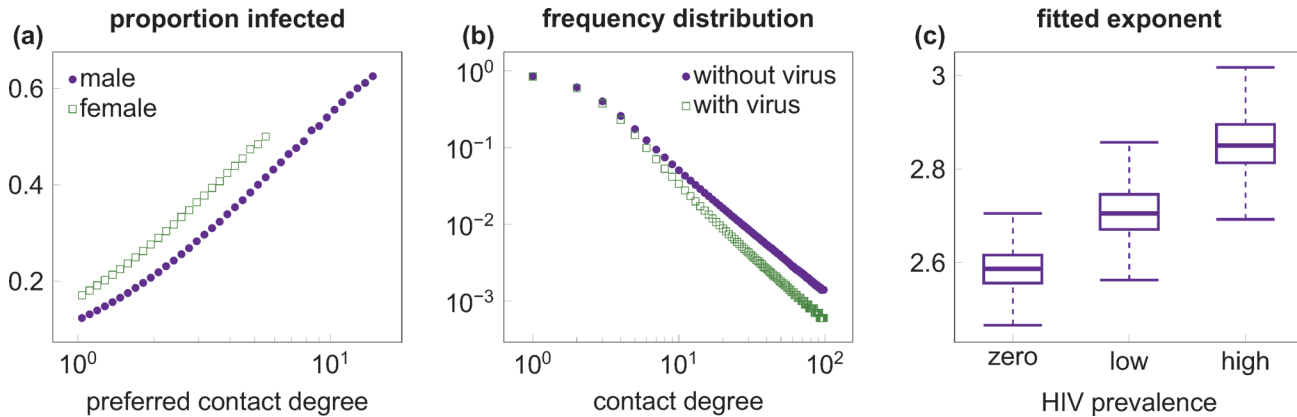


Fig 2. Preferential infection and killing of highly promiscuous individuals. (A) The ratio of infection among men (purple dots) and (non-FSW) women (green squares) increased with preferred contact degree (number of partners per year; plotted on logarithmic scale). The plot was created from 1000 independent simulation runs of single-strain epidemics of high prevalence, using logarithmic binning, right-censored at the top 1% of the male/female population (where rare classes result in strong stochastic variation). (B) The frequency distribution of the annual number of sexual contacts (realized contact degree) of males in uninfected populations (purple dots) and in populations with high-prevalence epidemics (green squares), based on median data from 1000 simulation runs. Highly promiscuous individuals were selectively depleted in the presence of the virus. (C) Boxplot of the exponents of power-law distributions fitted to male individuals in batches of 1000 independent runs with no virus, low and high prevalence epidemics, respectively. Boxes depict interquartile range, median is indicated by horizontal lines within the boxes, and whiskers extend to the farthest values that are not more than 1.5 times the box width away from the box. Medians (and IQR) of the exponents were 2.59 (2.56–2.62), 2.70 (2.67–2.75) and 2.85 (2.81–2.90) in the absence of the virus and with low or high prevalence epidemics, respectively; all pairwise comparisons between the three scenarios were statistically significant ($p < 10^{-10}$; Wilcoxon rank sum test). Selective depletion among females is shown in [S4 Fig](#). Simulation parameters were set as in [Table 1](#).

doi:10.1371/journal.pcbi.1004093.g002

sum test; [Fig. 2B-C](#) and [S4 Fig](#)), which may also have put any invader strain at a disadvantage. To assess the strength of this effect, we also implemented a scenario in which all individuals who died of AIDS were replaced by an uninfected individual with the same promiscuity (preferred contact degree) as that of the deceased individual, which preserved the degree distribution of the contact network irrespective of the epidemics.

Our strategy was thus to construct a default simulation scenario using settings that we deemed most plausible (partially inhibited superinfection, with strain replacement when superinfection is successful; one-time acute peak of infectiousness; and emergent preferential depletion of highly connected individuals), then test the effect of switching off one mechanism at a time in a series of test scenarios: i) “dual infection” with possible co-existence of the two strains in the same individual and no inhibition of superinfection; ii) “multiple acute” with repeated episodes of enhanced acute-stage infectiousness upon each successful superinfection; and iii) “fixed degrees” in which the degree distribution of the contact network was preserved. This strategy allowed us to assess the relative impact of each mechanism on the population level competition dynamics, and served also as a sensitivity analysis for relaxing the assumptions of inhibited superinfection and one-time acute peak infectiousness.

Inhibition of superinfection dominates first comer advantage

We tested eight scenarios (default and three test cases, each in both low and high prevalence settings) with levels of relative transmission advantage for the invader strain ranging between 0–50 percent. The invader strain was introduced in the population when the resident virus had attained steady-state prevalence; all combinations of scenario and transmission advantage were tested in 1000 simulation runs. We extracted several statistics to quantify the probability and rate of the expansion of the invader virus ([Fig. 3](#)).

When the transmission advantage of the invader strain was small, most simulations of the default scenario resulted in the extinction of the invader variant in both the high ([Fig. 3A](#)) and

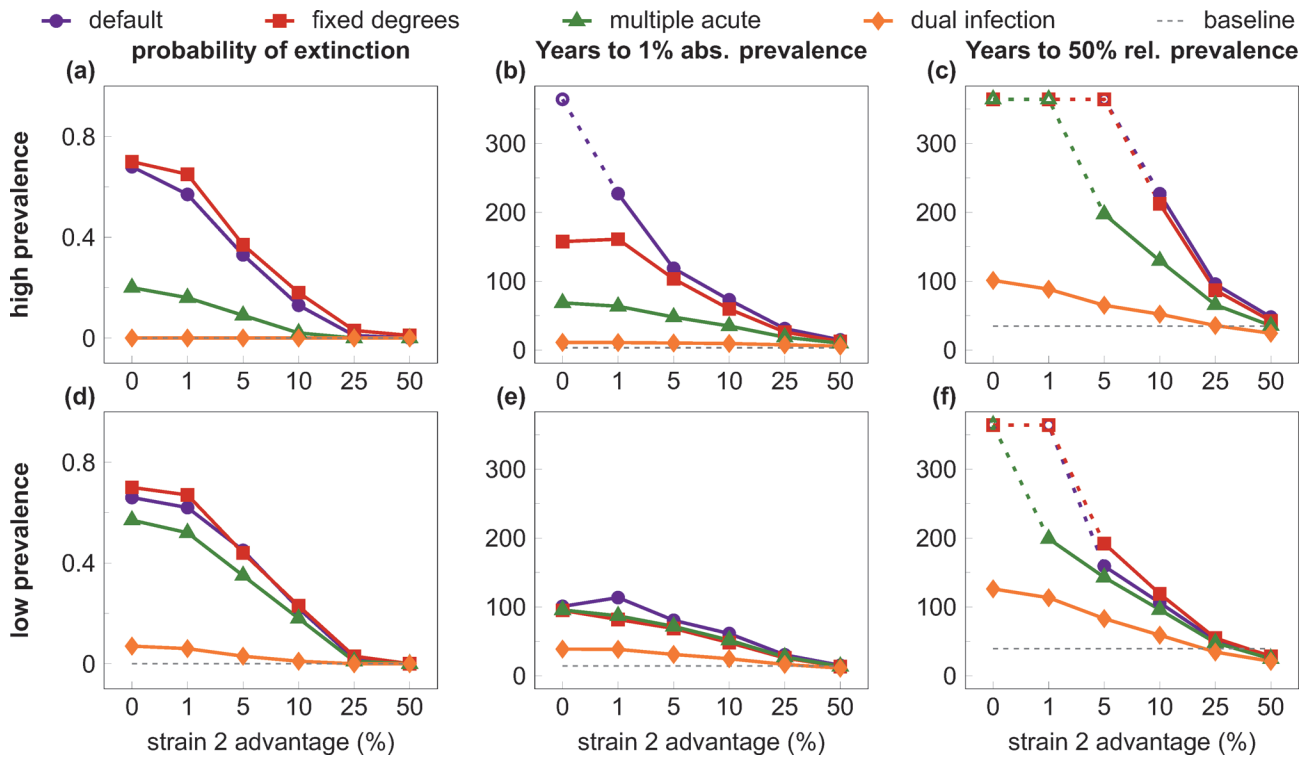


Fig 3. Quantifiers of the “first comer advantage”. All quantifiers are plotted against the relative transmission rate advantage of the second (invader) strain, with alternative scenarios to test interference mechanisms. Rows show results from the high (top row) and low prevalence (bottom row) settings; columns depict three different quantifiers; scenarios are coded by symbols and colour. In the default scenario (purple lines and dots) the invader strain faced a high risk of extinction (A, D) and experienced very slow growth to 1% absolute prevalence (B, E) and to 50% relative prevalence (C, F) at low values of transmission rate advantage, compared with the initial growth of the resident virus (dashed gray lines). The effect was largely abrogated with unhindered superinfection and co-existence (dual infection scenario; orange lines and diamonds), and, in the high-prevalence setting, partially mitigated by allowing for repeated “acute stage” peak infectivity after superinfection (multiple acute scenario; green lines and triangles); fixing the degree distribution of the contact network (fixed degrees scenario; red lines and squares) had little effect compared with the default scenario. Increasing the relative transmission rate advantage of the invader strain also decreased the inhibition effects: values comparable to the single-strain baseline were observed around 25%-50% transmission advantage. Data in B-C and E-F depict medians from 1000 simulation runs (excluding those where the invader virus went extinct). Parameters are listed in Table 1; scenarios are described in detail in the main text. The maximum length of simulations was 19,000 weeks (~365 years); empty symbols indicate where the invader strain did not reach the threshold prevalence by the end of the simulation in the majority of the cases.

doi:10.1371/journal.pcbi.1004093.g003

the low (Fig. 3D) prevalence settings. In contrast, the first (resident) strain was able to establish a stable epidemic in nearly all (>99%; dashed gray line) simulation runs when introduced into a fully susceptible population, which indicates a strong first comer advantage at the early stages of the spread of new strains. Preserving the degree distribution of the contact network (“fixed degrees”) had negligible effect compared with the default scenario; allowing multiple peaks of acute-stage infectiousness substantially reduced the probability of extinction in the high, but not in the low-prevalence setting. Finally, allowing for unhindered superinfection and coexistence (“dual infection”) reduced the probability of extinction to near zero even with no transmission advantage, illustrating that the inhibition of superinfection was the major factor in the heightened extinction risk of the invader strain. Greater relative advantage in the transmission rate reduced the risk of extinction in all scenarios, approaching zero extinction risk at around 25 percent advantage.

We defined two more quantifiers based on the time it took the invader strain to grow to selected threshold levels (in both cases we derived the statistics from the simulation runs where the invader strain did not go extinct). The time to one percent absolute prevalence (infecting one percent of the total population) was selected as a low threshold that would allow for the

detection of a new strain in a population (Fig. 3B,E). As a baseline comparison, we plotted also the median time until the resident strain attained one percent prevalence during its initial expansion (median of 14.4 and 3.3 years for the low and the high prevalence case; dashed gray lines). At small values of the transmission advantage, growing even to one percent prevalence can take a century or more in the default scenario (e.g. a median of 114 and 228 years with a transmission advantage of one percent, in the low and high-prevalence setting, respectively). The inhibition effect was stronger in the high-prevalence setting, and was gradually lost when the transmission advantage of the invader strain was increased to about 50 percent. The dominant mechanism of inhibition was again the inhibition of superinfection: allowing for dual infection abrogated most of the effect even at low values of the transmission advantage. The other two mechanisms of interference had negligible effect in the low-prevalence scenario (Fig. 3E), but had some partial effect in the high-prevalence scenario (Fig. 3B); multiple peaks of acute infectiousness had a stronger impact than fixed contact degrees also in this test case.

Finally, we also collected statistics on the time until the turning point when the invader strain accounted for 50 percent of the infections in the population (Fig. 3C,E). This time was extremely long (>300 years) when the invader strain had low transmission advantage in the default scenario, and a transmission advantage of 50 percent was needed to bring it down to a median of 27 and 48 years in the low and high-prevalence setting, respectively (in comparison, the resident strain reached 50% of its steady-state prevalence in a median of 35 and 40 years in the low and high-prevalence cases; dashed gray lines). Allowing for dual infection again had the strongest impact at lower transmission advantage, followed by allowing for multiple peaks of acute-stage infectiousness.

To understand why the impact of repeated acute-stage infectivity depended on the initial prevalence of the resident strain, we calculated the contribution of superinfection events and acute-stage transmissions to the spread of the invader strain in the various scenarios (Fig. 4). As expected, the contribution of superinfection was very low (<5%) in the low-prevalence setting, where most individuals were uninfected at the introduction of the invader strain; in contrast, many more transmissions (~20% initially) involved superinfection of carriers of the resident virus in the high-prevalence setting (Fig. 4A). Because multiple acute peaks of infectiousness take effect only when superinfection occurs, their impact on the frequency of acute transmissions was much stronger in the high-prevalence setting (Fig. 4B-C), and the increased frequency of efficient acute transmissions explains the reduced risk of extinction and faster growth of the invader strain when multiple acute peaks of infectivity were allowed in the high-prevalence scenario. In the high-prevalence setting (Fig. 4A), the relative contribution of superinfection decreases faster in the multiple acute scenario compared with the default scenario: the reason for this difference is that multiple acute peaks of infectiousness can substantially accelerate the outgrowth of the invader strain in the high-prevalence scenario (S5 Fig.), and the decline of the resident strain results in a decreasing probability that the invader (super)infects an individual who carries the resident strain.

Short head start or fast population turnover reduce first comer advantage

We next investigated what happens if the invader strain enters the population when the first strain is still in its growth phase and has not yet reached steady-state prevalence. We ran simulations where the invader was introduced when the resident strain had attained 5%, 20% or 50% of its plateau prevalence level and compared the outcome to the previous default setting (Fig. 5). As expected the first comer advantage was weaker when the second strain was introduced early in the growth phase of the first strain. However, the probability of extinction of the invader strain increased substantially already when the resident strain was at only 5% of its plateau level initially in the low-prevalence setting (Fig. 5D), or at 20% of plateau level in the high-

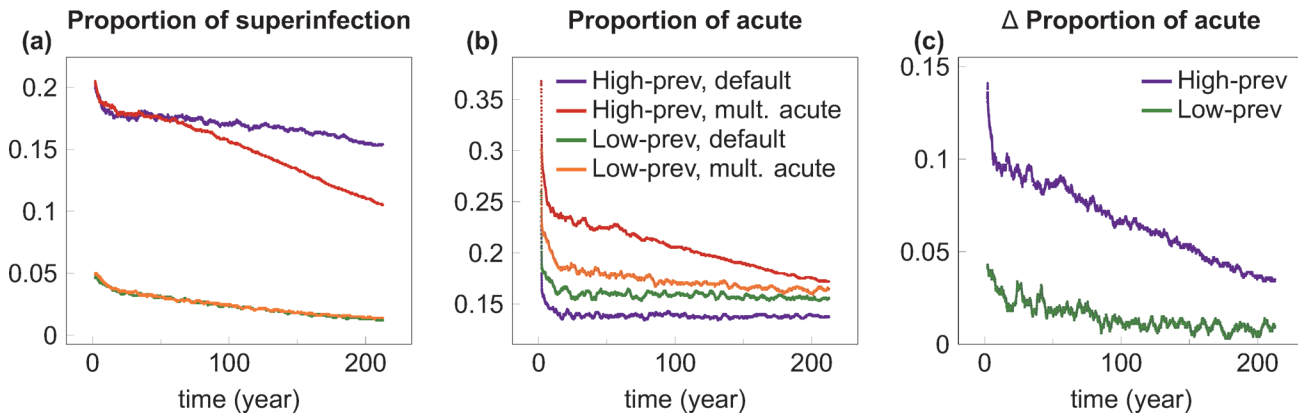


Fig 4. The contribution of superinfection events and acute-stage transmissions to the spread of the invader strain. (A) depicts the time course of the proportion of transmissions of the invader strain that involved superinfection of carriers of the resident virus. Coloured lines show smoothed proportion data for low and high prevalence epidemics using the default scenario, and the “multiple acute” scenario that allowed for repeated peaks of acute-stage infectiousness upon superinfection. In both scenarios, the contribution of superinfection was very low in the low-prevalence setting (green and orange lines), where most individuals were uninfected at the introduction of the invader strain; in contrast, many more transmissions involved superinfection in the high-prevalence setting (purple and red lines). (B) depicts the time course of the proportion of transmissions of the invader strain that originated from acute-stage transmitters in the four cases (colour coding is the same in A and B). (C) shows the difference in the proportion of acute-stage transmissions between the default and the multiple acute scenario for both prevalence settings (i.e. the distance between the red and purple, and between the green and yellow lines of Panel B). Allowing for multiple acute peaks of infectiousness greatly increased the proportion of acute-stage transmissions in the high-prevalence setting (purple line), but to a much lesser extent in the low-prevalence setting (green line). In all cases, time courses are plotted from the introduction of the invader strain into steady-state epidemics of the resident strain. Proportion data were calculated by combining transmission events recorded in 1000 simulation runs, then smoothed by averaging with a sliding window of length 100 weeks. Parameters were set as in Table 1; the transmission advantage of the invader strain was 5% in all cases.

doi:10.1371/journal.pcbi.1004093.g004

prevalence setting (Fig. 5A). The time to 50% relative prevalence of the invader strain was strongly affected when the resident strain was initially at 5% of its plateau level in the high-prevalence setting (Fig. 5C), and at 20% of plateau level in the low-prevalence setting (Fig. 5E). We thus conclude that (depending on the prevalence setting) some aspects of the first comer advantage are established relatively early in the initial expansion of the first successful strain.

We also tested the effect of faster population turnover using a residence time of 20 years (as opposed to the default of 35 years) for uninfected individuals in the population. This scenario may apply to regions that experience intense population movements and/or high rates of non-AIDS mortality. Faster population turnover had little effect on the initial risk of extinction for the invader strain, but could substantially accelerate the rate of its growth in the simulation runs where it did not go extinct (S6 Fig.). The probability of extinction is influenced by the instantaneous availability of susceptible individuals, which is not affected by the rate of turnover (at a fixed population size); however, subsequent growth depends on the continuous supply of new susceptibles, which increases with the rate of population turnover.

Case study: The expansion of HIV-1 subtype A in Uganda

While the mechanisms of interference can slow down the invasion of new strains, the global pandemic is not static and major shifts between HIV lineages have been occurring in selected regions. The best-characterized example is the expansion of HIV-1 subtype A at the expense of subtype D in Eastern Africa [5,25,26], and we used the detailed data from Uganda [26] to derive a crude estimate for the transmission advantage required for the observed expansion. Between 1994 and 2002, the estimated prevalence of subtype D decreased from 11.9% to 8.1%, and the prevalence of subtype A increased from 2.8% to 3.0% in Uganda; the overall prevalence of HIV declined from 17% to 13% over the same period [26]. The overall decline probably reflects changes in risk behaviour and/or health interventions; with stable (17%) prevalence, the

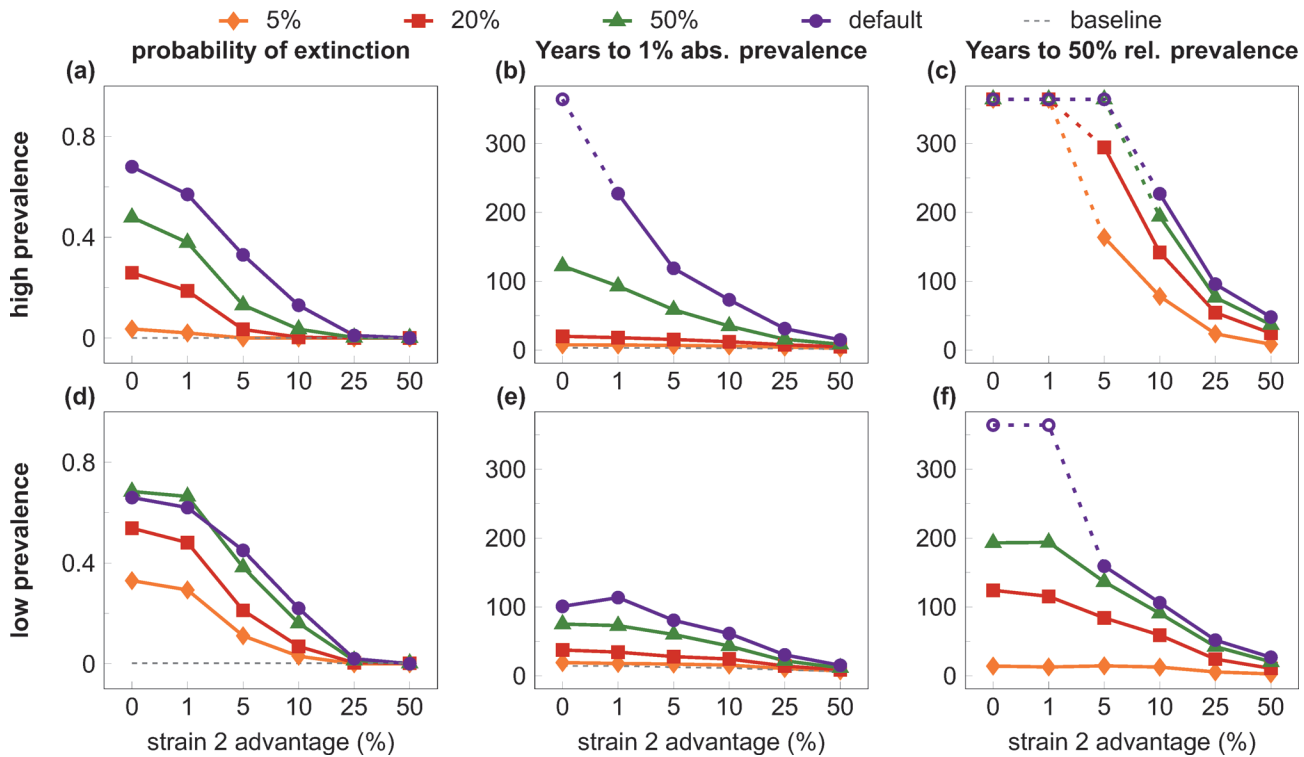


Fig 5. Quantifiers of the “first comer advantage” when the invader virus enters in the growth phase of the resident strain. Plotted are cases (coded by symbols and colour) where the invader was introduced when the resident strain had attained 5%, 20% or 50% of its plateau prevalence; in the default case the second virus was introduced at Week 7000/5000 in the low/high prevalence setting (as in Fig. 3) when the resident strain had already reached a stable plateau in its prevalence. All quantifiers are plotted against the relative transmission rate advantage of the second (invader) strain. Rows show results from the high (top row) and low prevalence (bottom row) settings; columns depict three different quantifiers. First comer advantage is weaker when the invader enters at earlier stages of the growth of the initial strain. Dashed gray lines in A-B and D-E represent the growth of the resident virus without competition; with early introduction of the invader strain, 50% relative prevalence in C and F is attained well below plateau prevalence and therefore cannot be compared to the 50% point of single-virus epidemics as a baseline. Data in B-C and E-F depict medians from 1000 simulation runs (excluding those where the invader virus went extinct). Parameters are listed in Table 1; competition dynamics followed the default scenario in all cases. The maximum length of simulations was 19,000 weeks (~365 years); empty symbols indicate where the invader strain did not reach the threshold prevalence by the end of the simulation in the majority of the cases.

doi:10.1371/journal.pcbi.1004093.g005

relative expansion of subtype A would roughly correspond to growing to 3.9% (~3*17/13) absolute prevalence, over a background prevalence comparable to that of our high-prevalence setting. Analyzing data from our high-prevalence default scenario, we found the closest match with the data when the transmission advantage of the invader strain was set to 25% (Fig. 1B), in which case the increase from 2.8% to 3.9% prevalence took 7.8 years on average (vs. 8 years in the empirical dataset). The rate of the relative expansion of subtype A observed in Uganda would thus require about 25% advantage over the resident subtype D strain, in a setting of stable overall prevalence in our simulations. Decreasing overall prevalence in the empirical data indicates a slowing turnover of infections, which requires a greater transmission advantage for the same tempo of strain replacement. This is roughly consistent with the independent empirical estimation that the overall (unadjusted) transmission rate of subtype A was 47% higher than that of subtype D in the same cohort [12].

Discussion

Both the probability and the rate of epidemic growth were strongly reduced for virus strains introduced into the steady state of a resident epidemic, when the default assumptions of partially

inhibited superinfection and one-time acute peak of transmissibility were used in our simulations. To outgrow the resident strain over a few decades (the time scale of human observations), the invader virus needed 25 percent or greater advantage in its rate of transmissibility over the resident strain. Of the three potential mechanisms of interference investigated, the direct inhibition of superinfection had the strongest effect in both prevalence settings, while one-time acute peak transmissibility had substantial effect only in the high-prevalence setting. The depletion of highly connected nodes in the network had little effect in most of the cases. In principle, a fourth mechanism of interference could also arise, because superinfected individuals (having progressed with their first infection) tend to have a shorter remaining lifespan, and therefore a shorter window of opportunity to transmit the superinfecting strain, compared with individuals who are infected for the first time. However, restarting the clock of disease progression upon superinfection had very little effect compared with the default scenario in a set of simulations (not shown); therefore, this mechanism does not seem to play an important role.

The reduction in the rate of growth of the invader strain was greater when the resident virus had higher initial prevalence, while the rate of extinction was insensitive to initial prevalence. The inhibition effect was weaker, but still considerable when the second strain was introduced while the first strain was still in the early phases of its growth, or if the (non-AIDS related) turnover of the population was faster.

Our results suggest that HIV competition dynamics is indeed characterized by a strong “first comer advantage” if the first strain to colonize a local transmission group expands to near plateau prevalence before further viral strains invade. This effect slows down the diversification of the epidemics and facilitates the persistence of founder effects. As far as we are aware, this is the first attempt to generally characterize the competition dynamics of different HIV strains over sexual networks, including multiple possible mechanisms of interference. The specific case of competition between HIV-1 and HIV-2 has been modelled in a similar framework [27], while another study looked at the competition of multiple evolving virus strains at the epidemic level without considering network structure [28]. Finally, Gross et al. [29] demonstrated that the inhibition of superinfection can preserve founder effects in the competition of equally transmissible virus strains, but have not considered network structure, alternative mechanisms of interference, or differential transmission.

The impact of the distinct mechanisms of interference may be modulated by factors that were not included in our simple model. First, heterogeneity may exist in the transmission rates across individuals and over time within the same partnership: in particular, per-contact transmission risk may decrease from the first exposures to subsequent contacts within a serodiscordant partnership (independent of the effect of acute infection) [30,31]. Such an effect can arise if the individuals highly susceptible to the virus of their partners tend to be rapidly infected, and the couples that remain serodiscordant become enriched in cases of low transmissibility over time (as reviewed in [32]). Similar effects are expected also if the partners of infected individuals can develop partially protective immunity to HIV in the exposures that do not result in transmission [33–35]. Irrespective of the mechanism, if time dependent variation in transmissibility is strain specific, then the invader virus has the advantage of being engaged in “first contact” with higher probability than the resident strain, which would decrease the first comer advantage of the latter. Second, if superinfection can generate a new “acute” temporary peak in viremia (and transmissibility) at least in some of the cases [23], then this mechanism of interference may also be weaker, which could reduce the first comer advantage (particularly in high-prevalence epidemics, according to our results). On the other hand, a detailed analysis of transmission risk in serodiscordant couples in Africa [20], and a recent phylodynamic analysis of a North American epidemic [21] have both estimated about 20-fold higher transmissibility

during acute compared with chronic HIV infection. Using a 9-fold consensus estimate [36] we may thus have underestimated the interference effect if repeated “acute” peaks of transmissibility do not (or only rarely) occur after superinfection. Third, the observed partial inhibition of superinfection may not take effect until several months from the first infection [19,37], e.g., if partially protective immune responses and/or a limiting depletion of target cells take a longer time to develop [19]. This would allow unhindered superinfection in the first few months after seroconversion, which would reduce the first comer advantage, especially if the second strain arrives while the first epidemic is still in its growth phase. Fourth, there is considerable debate on the strength of the (partial) protection from superinfection. Several studies have found zero or very low rates of superinfection [38,39] (which would implicate strong protection against superinfection), while at the other extreme some studies have found rates of superinfection comparable to those of initial infection [37,40,41] (which would indicate little or no protection). The differences in the estimates may reflect genuine variation between the study populations, but also differences in study design, inclusion criteria and sensitivity of detection [23]. Importantly, deep sequencing methods allow the detection of superinfecting strains that grow only to low levels in the superinfected individuals (e.g., [40]), and may often be lost after a transient episode of superinfection [42]. Such low-level superinfection is likely to result in onward transmission of the minority variant with much lower probability compared with the baseline rate of transmission. In the context of population level spread and competition, superinfection is likely to be relevant only when the superinfecting strain grows to dominate the virus pool of the individual, which we approximated by allowing only strain replacement (and no co-existence) in the default scenario of our simulations. Relaxing this assumption and allowing for unhindered superinfection abrogated most of the first comer advantage in our results: we therefore conclude that the strong founder effects observed in the global phylogeography of HIV are more parsimoniously explained if superinfection is partially inhibited and the transmission of more than one strain from the same individual is rare.

Our generic modelling framework could not aim to account for all the (often population specific) complexities of human population dynamics and behaviour. For simplicity, population size was kept constant in our simulations, including instantaneous replacement of individuals who died of AIDS. With this implementation population turnover increased with HIV prevalence, e.g., the rate of death/replacement was about a third higher in the high-prevalence steady state compared with an uninfected population. Given that faster turnover reduces the first comer advantage, our results can be regarded as a conservative (under) estimation of the inhibition effect. Not replacing individuals who die of AIDS results in decreasing population size, which may further inhibit the expansion of invader strains by reducing the supply of susceptible individuals. In contrast, fast population growth or immigration may dilute the inhibition effect by increasing the influx of susceptible individuals. Migration may also play a role by introducing the same invading HIV strain repeatedly from a source population: this would eventually overcome the barrier of initial extinction, but would likely have little impact on the subsequent growth of the invader strain. Furthermore, an established HIV epidemic may also affect sexual behaviour: if high-risk sexual practices and/or promiscuity decrease in response to an ongoing epidemic, the spread of subsequent invader strains may be further inhibited. Finally, the complexities of the sexual network, e.g., assortative mating may further influence the strength of the inhibition effects.

Alternative or additional mechanisms may also contribute to the preservation of founder effects. If viral adaptation occurs to host traits that vary between human populations, then a locally adapted virus strain will enjoy a selection advantage against strains adapted to other host populations (as has been observed in some model systems of host-parasite interactions [43,44]). For example, the distribution of Human Leukocyte Antigen (HLA) alleles may differ

between human populations, and local transmission may fix escape mutations against the locally frequent alleles that initially had a protective effect [45], particularly in populations with lower HLA diversity [46]; although this seems to be occurring slowly and to a limited extent where HLA diversity is high [47–49]. Location or population specific differences may exist in other host traits affecting HIV acquisition or transmission (e.g., in restriction factors [50,51] or in other components of innate immunity [52]). Each locally adapted virus strain may therefore have a competitive advantage within its established host population, and a disadvantage in other populations—which would also slow down the global mixing of variants or could even result in the long term survival of several virus strains in different populations. We note, however, that long-term co-existence of several virus strains in the same epidemic (connected transmission group) is possible only if specific conditions are fulfilled, e.g., strain-specific immunity or therapy creates frequency-dependent selection that favours the rare type. Without such specific conditions, the strain with the highest transmission potential in a given host population drives all other strains extinct in the long run: this principle of competitive exclusion holds true from simple abstract mathematical models [53,54] to complex simulations, including ours.

We parameterized our model based on data from generalized heterosexual epidemics in Africa, but it could easily be adapted to other routes of transmission and to concentrated epidemics. Furthermore, the results of our simulations can be applied not only to the competition of two distinct lineages (e.g., subtypes, or distinct clades of the same subtype [55]), but also to competition between virus variants that arise by local mutations. The general take-home message of our work posits that the expansion of the HIV pandemic to all susceptible populations across the world has made the conditions far less favourable for the spread of “novel” virus strains, irrespective of their origin.

Our results have important implications for understanding the past and for predicting the future of the HIV pandemic. The observed first comer advantage can delay evolution to “optimal virulence” [28,56] that maximizes transmissibility, and can also delay the spread of drug resistance (by onward transmission [57]) in the face of increasing selection pressure from the broadening scope of ART. Widely available ART may affect resident and invader strains equally, effectively reducing the baseline rate of transmissions and transforming a high-prevalence setting towards lower prevalence. Given that most aspects of the first comer advantage were strong in both low- and high-prevalence settings in our simulations, the broadening scope of ART may not affect this phenomenon strongly.

Because the mechanisms of first comer advantage do not operate at the front wave of an epidemic expanding into a susceptible population, we suggest that much of the (non-local) adaptation of HIV may have happened along these front waves, rather than in populations where prevalence has stabilized. Furthermore, considering that the currently dominant subtypes probably all expanded riding the wave of their first comer advantage, most or all of them may in fact possess suboptimal fitness and transmissibility. If the original founder strains of the early expansions were selected (at least partly) by “chance”, rather than due to high fitness, then even subsequent evolution may have constrained the subtypes to the local suboptima of the fitness landscape that were accessible from the initial sequence. This implies that the initial founder effects and the first comer advantage may have provided some benefit by preventing the fast global spread of the most transmissible HIV variants in the growth phase of the pandemic. However, the results also caution that the next stage of the pandemic may be characterized by a shift towards more transmissible strains over the slow time scales predicted by our model, and data from several regions indicate that this process has already started. HIV-1 subtype A is spreading at the expense of subtype D in Eastern Africa [5,25,26], and HIV-1 is expanding at the expense of HIV-2 in Western Africa [27,58]. Our results suggest that these

relatively fast replacements require a large selective advantage of the expanding strain. Indeed, subtype A is associated with higher transmissibility [12] and slower disease progression [14,15] compared with subtype D, and HIV-2 has two orders of magnitude lower replicative capacity [59] and more than 3-fold lower per contact transmissibility [60] compared with HIV-1. In comparison, within individual patients the replicative fitness (a probable correlate of transmissibility) showed only about 10% variation between the fittest and the average viral genome in a study of untreated HIV-1 infected patients [61]. Our results indicate that variations of greater magnitude are needed to drive the relatively fast replacement dynamics of the few observed cases.

While differences are expected between the currently characterized subtypes, and those more efficient at transmission are slowly gaining ground at the expense of less transmissible subtypes, major innovations and potentially higher transmissibility may be more likely to arise from the complex diversity of HIV in Central Africa and from the recombinant forms. CRFs probably emerged against the backdrop of established epidemics and their growth to detectable levels may indicate considerable selection advantage. The rapid growth of several CRFs in recent years [5,62] is consistent with this concept and is therefore cause for concern for the future of the pandemic.

The interference mechanisms and first comer advantage demonstrated in this paper may also help explain why so few cross-species transmissions of SIV to humans were able to establish epidemic HIV lineages, and why no new major HIV types or groups have emerged since the middle of the 20th century [63]. It is possible that a successful reduction or elimination of the current HIV epidemic in Africa may, by eliminating the inhibiting competition effects, increase the risk for the emergence of new HIV lineages from novel cross-species transmissions.

Finally, we note that HIV may represent a rare combination of factors relevant for the observed first comer advantage: infection lasts and remains active for life; the inhibition of superinfection does not seem to be (strongly) strain specific [64] (as opposed to other infections with serotypes that elicit type specific immunity); and infected individuals remain in the contact network for many years. Taken together, these factors may imply that the first comer advantage, and its consequence of delayed global mixing, may be particularly strong for HIV and weaker for most other pathogens. For example, a persistent infection controlled by strain specific immunity would correspond approximately to our HIV scenario with no inhibition of superinfection, in which case most of the population level effect was lost in the simulations. Non-persistent infections would tip the balance further in favour of the novel strain, because individuals recovered from the initial infection would be susceptible to the novel strain while ceasing to transmit or be susceptible to the first strain.

In all, our results suggest that the interference mechanisms of competition, possibly aided by local adaptation, can slow down the adaptation of HIV at the population level, in spite of the huge evolutionary potential of the virus. These effects may explain why strong founder effects still persist several decades after the initial global expansion of the pandemic, and may hamper the ongoing adaptation of the virus to maximize its transmissibility and also slow down the spread of drug resistance.

Materials and Methods

Simulation model

We developed a stochastic, individual-based simulation model to track the spread of HIV over a dynamic network of heterosexual contacts. The network consisted of three types of nodes (individuals): males, females and female sex workers (FSW). The model tracked the age and HIV status (stage of infection and the infecting virus type) of each individual, and for males and

(non-FSW) females also a fixed quantifier of promiscuity (preferred annual contact degree), and the number of distinct sexual partners in the last year (realized annual contact degree). Individuals entered the population at age 15 and were removed at age 50. The preferred contact degree of each individual was drawn from an empirical distribution according to the type of the node and was kept constant for the lifetime of the individual. The promiscuity of males and (non-FSW) females was characterized by continuous power-law distributions of the form $P(x) \sim x^{-\gamma}$ (with different exponents for the two sexes) parameterized based on empirical data (Table 1), and censored at both a lower cut-off (one contact per year to ensure all nodes are active in the network) and an upper cut-off. FSW had a fixed maximum number of one-time contacts per week.

The simulations had a time step of one week, and each step consisted of the following procedures: a) generation of sexual acts along the links and virus transmission, b) update of HIV status, c) birth and death dynamics of individuals, d) dissolution and formation of network links. The number of sexual acts in male-female links was drawn from a Poisson distribution (discarding zeros: no links were inactive); male-FSW links always involved a single sexual act. The probability of virus transmission to uninfected individuals was determined by the baseline transmission rate of the virus strain, amplified if the transmitting individual was in the acute stage of the infection. Newly infected individuals were immediately assigned a time to death from a uniform distribution between 3–20 years (consistent with recently estimated survival curves in ART-naïve cohorts [65]), and for each infection event we recorded: the date of the event, the strain that was transmitted, the disease stage of the transmitter, and whether the

Table 1. Parameters used for the simulation of the model.

Symbol	Description	Value [reference]
N_m	Number of men in the population	10000
N_f	Number of women in the population	10000
γ_m	Exponent of male degree distribution	2.45 ^a [71]
γ_f	Exponent of female degree distribution	3.45 ^a [71]
K_c	Number of clients per FSW per year	400 ^b [72–74]
K_{min}	Lower cutoff of annual degree distribution	1
K_{max}	Upper cutoff of annual degree distribution	1000 ^c
p_b	Probability of link breakup per week	0.05 ^d [73]
λ	Poisson parameter for the number of sex acts per week	2 [75,76]
v_1	Strain 1 per-contact transmissibility in chronic stage	0.001 or 0.002 ^e [77]
v_2	Strain 2 per-contact transmissibility in chronic stage	(1–1.5)* v_1
m_A	Transmission multiplier for acute infection	9 [36]
T_{acute}	Length of acute phase (weeks)	12 [20]
T_{age}	Duration of sexual activity (years)	35 (age 15–50)
T_{HIV}	Survival with HIV infection (range in years)	3–20 [78,79]
T_{init}	Time steps (weeks) without virus	1000
T_{single}	Time steps (weeks) with only one virus	6000 or 4000 ^e

^aUsed to generate preferred annual contact degrees; for the exponents fitted to realized contact degrees, see Fig. 2 and S4 Fig.

^bMiddle value from 600 given in [72] and 150 calculated from [73,74].

^cThe maximum realized contact degree was lower in all simulations.

^dBaseline rate for links between individuals with degree 1.

^eAlternative values used to parameterize low- and high-prevalence settings.

doi:10.1371/journal.pcbi.1004093.t001

transmission involved superinfection of an individual previously infected with the other virus type. For simplicity, the size of the population was kept constant (at 10,000 individuals of both sexes): all nodes who died of AIDS or left the network at age 50 (whichever came first) were replaced with a new individual of age 15. The preferred annual contact degree of new nodes was drawn from the power-law distribution of the respective gender at entry to the population. The links between males and females were allowed to form and break up at each time step. The baseline probability of break-up was set to yield an average duration consistent with empirical estimates (Table 1), and was scaled proportional to the average contact degree of the two nodes (such that more promiscuous individuals had shorter relationships [66]). Link formation was implemented such that all non-FSW individuals would have a yearly number of sexual contacts approximately equivalent to their preferred annual contact degree. At each time step, the nodes were assigned a number of half-links generated randomly in proportion to their preferred contact degree. Because males had greater mean promiscuity than non-FSW females, the number of half-links for males exceeded those of the females. New links were formed by first randomly connecting all female half-links to male half-links, then randomly distributing the remaining male half-links to the FSW. All runs were started with an initialization phase restricted to link formation and break-up until the sexual network settled to a steady state. FSWs had fixed promiscuity and were added one by one as long as there was a surplus of male half-links. The number of FSWs at steady-state was thus not pre-determined, but emerged to match and compensate the imbalance of male and (non-FSW) female links in each scenario.

Population level competition was simulated by implementing two virus types that were allowed to differ in their rate of transmissibility. The type of the infecting virus strain(s) was tracked for each infected individual. The first virus strain was introduced in a random sample of ten percent of all FSW after the initialization of the sexual network: this method allowed a reliable establishment of the “resident” epidemic with negligible risk of extinction. The second (invader) strain was also introduced in a sample of ten percent of all FSW (sampled from uninfected FSW) when the resident strain has attained a steady state in the population.

In the default scenario, superinfection could occur only by the replacement of the original strain with the superinfecting strain. In a sexual act between two individuals infected with different virus strains, both strains had a chance to be transmitted. Superinfection occurred if two check points were passed: initial transmission occurred according to the transmission rate of the infecting strain (modified by disease stage, if appropriate); then after successful initial transmission, the probability of superinfection was determined by the relative transmission rates (“fitness”) of the two strains as follows: $P = (v_2 / (v_1 + v_2))$, where v_1 denotes the transmission rate of the virus infecting the potential recipient and v_2 denotes the transmission rate of the strain infecting the potential transmitter. The “clock” of disease stage in the recipient was unaffected by superinfection in the default scenario; the stage of disease remained to be based on the age of infection defined by the date of the original first infection of the recipient. A new time to death was drawn randomly (from the 3–20 years range); however, it was used only if the new date of death preceded the original date determined at the initial infection: superinfection could never extend the lifespan of an individual.

We implemented three scenarios to test the potential mechanisms of interference. In the “dual infection” scenario, superinfection occurred at the same rate as initial infection (i.e., according to the transmission rate of the superinfecting strain), and the two strains co-existed in the superinfected individuals. Superinfected (dually infected) individuals were then able to transmit both virus strains independently in subsequent contacts. We did not implement a scenario in which superinfection was unhindered but co-existence not allowed, because such a situation would imply an asymmetric advantage for the superinfecting strain. In the “multiple acute” scenario, the clock of disease stage was reset upon superinfection, and the superinfecting

strain started a new episode of peak acute-stage transmissibility (however, this new episode could not extend the lifespan of the individual beyond his/her original date of death, determined after the first infection). Finally, in the “fixed degrees” scenario, new individuals were added with a preferred contact degree identical to that of the individual whom they replaced after his/her death due to AIDS.

The parameters of the sexual network were based on contemporary surveys in Africa; HIV parameters were also based on available empirical data (Table 1). The high-prevalence setting was implemented by increasing (doubling) the baseline transmission rate, consistent with the recent finding that variation in prevalence among Sub-Saharan countries can largely be explained by differences in the rate of transmission in serodiscordant couples [67].

The model was implemented in the C++ programming language. The full computer code of the simulations is available in S1 File.

Statistics

Statistical tests were performed with R [68]. Power-law exponents of the realized annual contact degrees (based on the actual numbers of sexual contacts in the last year) were fitted as described in [69], estimating the lower cutoff with Kolmogorov-Smirnov statistics, using the implementation of [70].

Supporting Information

S1 File. Computer code of simulations.

(ZIP)

S1 Fig. The growth of the invader strain at different values of relative advantage in the transmission rate. The relative advantage of the invader virus was varied from zero (top row) to 50% (bottom row) in the low (left column) or high (right column) prevalence scenarios. The resident strain (solid purple line) was introduced in the population at Week 1000 (to allow the network to attain steady state); the invader strain (dashed green line) was introduced in the population when the first had already reached steady-state prevalence (at Week 5000 and 7000 for the high- and low-prevalence setting, respectively). The lines show median prevalence from simulations where the invader strain did not go extinct (out of 1000 simulation runs); shading indicates the areas between the 5% and 95% quantiles. Simulation parameters were set as in Table 1; superinfection and replacement dynamics followed the default scenario.

(PDF)

S2 Fig. The relative contribution of acute stage transmissions over the time course of single-strain epidemics. The proportion of transmissions originating from acute-stage transmitters decreases from high levels at the beginning of the epidemics to a steady-state around 0.15 and 0.13 in the low (purple dots) and high (green dots) prevalence epidemics, respectively, over a time scale of a few decades. Proportion data were calculated by combining transmission events recorded in 1000 simulation runs, then smoothed by averaging with a sliding window of length 100 weeks. Parameters were set as in Table 1; superinfection and replacement dynamics followed the default scenario.

(PDF)

S3 Fig. The probability of infection as a function of the promiscuity (preferred contact degree) of the individuals: data and model fit. Using collated data from 100 simulation runs (2 million individuals total), we performed a logistic regression on the probability of infection in individuals using log transformed preferred contact degree, age and gender as explanatory variables. Purple and red lines show smoothed actual proportions of infecteds among females

and males, respectively, calculated with a sliding window (moving along all individuals sorted according to contact degree; each point representing the frequency of infections among 1000 individuals). Predictions from the logistic regression (plotted as orange and green lines, using the same sliding window smoothing) provide an excellent fit to the data. Effect sizes (and 95% CI) for the three factors were estimated as follows: $\log_{10}(\text{degree})$: 2.48 (95% CI: 2.46–2.50), age: 0.0460 per year (95% CI: 0.0457–0.0464), female gender: 0.420 (95% CI: 0.413–0.428); all three effects were significant at $p < 10^{-10}$. Parameters were set as in [Table 1](#); superinfection and replacement dynamics followed the default scenario.

(PDF)

S4 Fig. Preferential depletion of highly promiscuous individuals among non-FSW females.

(A) The frequency distribution of the annual number of sexual contacts (realized contact degree) in females in uninfected populations (purple dots) and in populations with high-prevalence epidemics (green squares), based on median data from 1000 simulation runs. Highly promiscuous individuals were selectively depleted in the presence of the virus. (B) Boxplot of the exponents of power-law distributions fitted to female individuals in batches of 1000 independent runs with no virus, low and high prevalence epidemics, respectively. Boxes depict interquartile range, median is indicated by horizontal lines within the boxes, and whiskers extend to the farthest values that are not more than 1.5 times the box width away from the box. Medians (and IQR) of the exponents were 3.86 (3.73–3.97), 3.99 (3.83–4.09) and 4.19 (3.99–4.32) in the absence of the virus and with low or high prevalence epidemics, respectively; all pairwise comparisons between the three scenarios were statistically significant ($p < 10^{-10}$; Wilcoxon rank sum test). Simulation parameters were set as in [Table 1](#); superinfection and replacement dynamics followed the default scenario.

(PDF)

S5 Fig. The effect of multiple acute infections on the competition of HIV strains. The figure compares the outgrowth of an invader virus with 5% transmission rate advantage in the high (top row) and low (bottom row) prevalence settings with default superinfection dynamics (left: A, C) or with repeated peaks of acute-stage infectiousness upon superinfection (right: B, D). The resident strain (solid purple line) was introduced in the population at Week 1000 (to allow the network to attain steady state); the invader strain (dashed green line) was introduced in the population when the first strain had already attained steady-state prevalence (at Week 5000 and 7000 for the high- and low-prevalence setting, respectively). Multiple acute peaks accelerate the outgrowth of the invader strain and the decline of the resident considerably in the high prevalence scenario (A vs. B), but not in the low prevalence scenario (C vs. D), where superinfection is rare. The lines show median prevalence from simulations where the invader strain did not go extinct (out of 1000 simulation runs); shading indicates the areas between the 5% and 95% quantiles. Simulation parameters were set as in [Table 1](#); scenarios are described in detail in the main text.

(PDF)

S6 Fig. The effect of population turnover on the “first comer advantage”. All quantifiers are plotted against the relative transmission rate advantage of the second (invader) strain, for two levels of population turnover: 35 (default, purple dots) or 20 years (red squares) of uninfected (sexually active) lifespan, in the high (top row) and low (bottom row) prevalence settings. Faster turnover had little effect on the probability of extinction of the invader strain, but could have a pronounced effect on its rate of growth at low values of the transmission advantage. Data in B-C and E-F depict medians from 1000 simulation runs (excluding those where the invader virus went extinct). Parameters are listed in [Table 1](#); superinfection and replacement

dynamics followed the default scenario. The maximum length of simulations was 19,000 weeks (~365 years); empty symbols indicate where the invader strain did not reach the threshold prevalence by the end of the simulation in the majority of the cases.

(PDF)

Acknowledgments

We thank João Dinis de Sousa and Prof. Douglas D. Richman for helpful discussions.

Author Contributions

Analyzed the data: BF VM. Wrote the paper: BF VM. Conceived the study: VM. Contributed to the design of the study: BF EM TV VM. Developed the modelling framework: BF EM TV VM. Implemented the model: BF.

References

1. Korber B, Muldoon M, Theiler J, Gao F, Gupta R, et al. (2000) Timing the ancestor of the HIV-1 pandemic strains. *Science* 288: 1789–1796. PMID: [10846155](#)
2. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, et al. (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 455: 661–664. doi: [10.1038/nature07390](#) PMID: [18833279](#)
3. Rambaut A, Robertson DL, Pybus OG, Peeters M, Holmes EC (2001) Human immunodeficiency virus. Phylogeny and the origin of HIV-1. *Nature* 410: 1047–1048. PMID: [11323659](#)
4. Tebit DM, Arts EJ (2011) Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect Dis* 11: 45–56. doi: [10.1016/S1473-3099\(10\)70186-9](#) PMID: [21126914](#)
5. Hemelaar J, Gouws E, Ghys PD, Osmanov S, Isolation W-UNfH, et al. (2011) Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS* 25: 679–689. doi: [10.1097/QAD.0b013e328342ff93](#) PMID: [21297424](#)
6. Ou CY, Takebe Y, Weniger BG, Luo CC, Kalish ML, et al. (1993) Independent introduction of two major HIV-1 genotypes into distinct high-risk populations in Thailand. *Lancet* 341: 1171–1174. PMID: [8098076](#)
7. Lukashov VV, Cornelissen MT, Goudsmit J, Papuashvilli MN, Rytik PG, et al. (1995) Simultaneous introduction of distinct HIV-1 subtypes into different risk groups in Russia, Byelorussia and Lithuania. *AIDS* 9: 435–439. PMID: [7639968](#)
8. van Harmelen J, Wood R, Lambrick M, Rybicki EP, Williamson AL, et al. (1997) An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. *AIDS* 11: 81–87. PMID: [9110079](#)
9. Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, et al. (2013) HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. *Retrovirology* 10: 7. doi: [10.1186/1742-4690-10-7](#) PMID: [23317093](#)
10. Frenzt D, Wensing AM, Albert J, Paraskevis D, Abecasis AB, et al. (2013) Limited cross-border infections in patients newly diagnosed with HIV in Europe. *Retrovirology* 10: 36. doi: [10.1186/1742-4690-10-36](#) PMID: [23551870](#)
11. Renjifo B, Gilbert P, Chaplin B, Msamanga G, Mwakagile D, et al. (2004) Preferential in-utero transmission of HIV-1 subtype C as compared to HIV-1 subtype A or D. *AIDS* 18: 1629–1636. PMID: [15280773](#)
12. Kiwanuka N, Laeyendecker O, Quinn TC, Wawer MJ, Shepherd J, et al. (2009) HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda. *AIDS* 23: 2479–2484. doi: [10.1097/QAD.0b013e328330cc08](#) PMID: [19841572](#)
13. Walter BL, Armitage AE, Graham SC, de Oliveira T, Skinhoj P, et al. (2009) Functional characteristics of HIV-1 subtype C compatible with increased heterosexual transmissibility. *AIDS* 23: 1047–1057. doi: [10.1097/QAD.0b013e32832a1806](#) PMID: [19390419](#)
14. Vasan A, Renjifo B, Hertzmark E, Chaplin B, Msamanga G, et al. (2006) Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. *Clin Infect Dis* 42: 843–852. PMID: [16477563](#)

15. Baeten JM, Chohan B, Lavreys L, Chohan V, McClelland RS, et al. (2007) HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. *J Infect Dis* 195: 1177–1180. PMID: [17357054](#)
16. Vidal N, Peeters M, Mulanga-Kabeya C, Nzilambi N, Robertson D, et al. (2000) Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *J Virol* 74: 10498–10507. PMID: [11044094](#)
17. Kalish ML, Robbins KE, Pieniazek D, Schaefer A, Nzilambi N, et al. (2004) Recombinant viruses and early global HIV-1 epidemic. *Emerg Infect Dis* 10: 1227–1234. PMID: [15324542](#)
18. Vidal N, Mulanga C, Bazepeo SE, Mwamba JK, Tshimpaka J, et al. (2006) HIV type 1 pol gene diversity and antiretroviral drug resistance mutations in the Democratic Republic of Congo (DRC). *AIDS Res Hum Retroviruses* 22: 202–206. PMID: [16478404](#)
19. Ronen K, McCoy CO, Matsen FA, Boyd DF, Emery S, et al. (2013) HIV-1 superinfection occurs less frequently than initial infection in a cohort of high-risk Kenyan women. *PLoS Pathog* 9: e1003593. doi: [10.1371/journal.ppat.1003593](#) PMID: [24009513](#)
20. Hollingsworth TD, Anderson RM, Fraser C (2008) HIV-1 transmission, by stage of infection. *J Infect Dis* 198: 687–693. doi: [10.1086/590501](#) PMID: [18662132](#)
21. Rasmussen DA, Volz EM, Koelle K (2014) Phylodynamic inference for structured epidemiological models. *PLoS Comput Biol* 10: e1003570. doi: [10.1371/journal.pcbi.1003570](#) PMID: [24743590](#)
22. Pinkerton SD (2007) How many sexually-acquired HIV infections in the USA are due to acute-phase HIV transmission? *AIDS* 21: 1625–1629. PMID: [17630558](#)
23. Redd AD, Quinn TC, Tobian AA (2013) Frequency and implications of HIV superinfection. *Lancet Infect Dis* 13: 622–628. doi: [10.1016/S1473-3099\(13\)70066-5](#) PMID: [23726798](#)
24. Thomas JC, Tucker MJ (1996) The development and use of the concept of a sexually transmitted disease core. *J Infect Dis* 174 Suppl 2: S134–143. PMID: [8843243](#)
25. Rainwater S, DeVange S, Sagar M, Ndinya-Achola J, Mandaliya K, et al. (2005) No evidence for rapid subtype C spread within an epidemic in which multiple subtypes and intersubtype recombinants circulate. *AIDS Res Hum Retroviruses* 21: 1060–1065. PMID: [16379611](#)
26. Conroy SA, Laeyendecker O, Redd AD, Collinson-Streng A, Kong X, et al. (2010) Changes in the distribution of HIV type 1 subtypes D and A in Rakai District, Uganda between 1994 and 2002. *AIDS Res Hum Retroviruses* 26: 1087–1091. doi: [10.1089/aid.2010.0054](#) PMID: [20925575](#)
27. Schmidt WP, Van Der Loeff MS, Aaby P, Whittle H, Bakker R, et al. (2008) Behaviour change and competitive exclusion can explain the diverging HIV-1 and HIV-2 prevalence trends in Guinea-Bissau. *Epidemiol Infect* 136: 551–561. PMID: [17559692](#)
28. Shirreff G, Pellis L, Laeyendecker O, Fraser C (2011) Transmission selects for HIV-1 strains of intermediate virulence: a modelling approach. *PLoS Comput Biol* 7: e1002185. doi: [10.1371/journal.pcbi.1002185](#) PMID: [22022243](#)
29. Gross KL, Porco TC, Grant RM (2004) HIV-1 superinfection and viral diversity. *AIDS* 18: 1513–1520. PMID: [15238769](#)
30. Downs AM, De Vincenzi I (1996) Probability of heterosexual transmission of HIV: relationship to the number of unprotected sexual contacts. European Study Group in Heterosexual Transmission of HIV. *Journal of acquired immune deficiency syndromes and human retrovirology: official publication of the International Retrovirology Association* 11: 388–395.
31. Leynaert B, Downs AM, de Vincenzi I (1998) Heterosexual transmission of human immunodeficiency virus: variability of infectivity throughout the course of infection. European Study Group on Heterosexual Transmission of HIV. *Am J Epidemiol* 148: 88–96. PMID: [9663408](#)
32. Nagelkerke NJ, Arora P, Jha P, Williams B, McKinnon L, et al. (2014) The rise and fall of HIV in high-prevalence countries: a challenge for mathematical modeling. *PLoS Comput Biol* 10: e1003459. doi: [10.1371/journal.pcbi.1003459](#) PMID: [24626088](#)
33. Kaul R, Plummer FA, Kimani J, Dong T, Kiama P, et al. (2000) HIV-1-specific mucosal CD8+ lymphocyte responses in the cervix of HIV-1-resistant prostitutes in Nairobi. *J Immunol* 164: 1602–1611. PMID: [10640781](#)
34. Nicastrì E, Ercoli L, Sarmati L, d'Ettore G, Iudicone P, et al. (2001) Human immunodeficiency virus-1 specific and natural cellular immunity in HIV seronegative subjects with multiple sexual exposures to virus. *J Med Virol* 64: 232–237. PMID: [11424109](#)
35. Promadej N, Costello C, Wernett MM, Kulkarni PS, Robison VA, et al. (2003) Broad human immunodeficiency virus (HIV)-specific T cell responses to conserved HIV proteins in HIV-seronegative women highly exposed to a single HIV-infected partner. *J Infect Dis* 187: 1053–1063. PMID: [12660919](#)

36. Boily MC, Baggaley RF, Wang L, Masse B, White RG, et al. (2009) Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *Lancet Infect Dis* 9: 118–129. doi: [10.1016/S1473-3099\(09\)70021-0](https://doi.org/10.1016/S1473-3099(09)70021-0) PMID: [19179227](https://pubmed.ncbi.nlm.nih.gov/19179227/)
37. Kraft CS, Basu D, Hawkins PA, Hraber PT, Chomba E, et al. (2012) Timing and source of subtype-C HIV-1 superinfection in the newly infected partner of Zambian couples with disparate viruses. *Retrovirology* 9: 22. doi: [10.1186/1742-4690-9-22](https://doi.org/10.1186/1742-4690-9-22) PMID: [22433432](https://pubmed.ncbi.nlm.nih.gov/22433432/)
38. Gonzales MJ, Delwart E, Rhee SY, Tsui R, Zolopa AR, et al. (2003) Lack of detectable human immunodeficiency virus type 1 superinfection during 1072 person-years of observation. *J Infect Dis* 188: 397–405. PMID: [12870121](https://pubmed.ncbi.nlm.nih.gov/12870121/)
39. Tsui R, Herring BL, Barbour JD, Grant RM, Bacchetti P, et al. (2004) Human immunodeficiency virus type 1 superinfection was not detected following 215 years of injection drug user exposure. *J Virol* 78: 94–103. PMID: [14671091](https://pubmed.ncbi.nlm.nih.gov/14671091/)
40. Redd AD, Collinson-Streng A, Martens C, Ricklefs S, Mullis CE, et al. (2011) Identification of HIV superinfection in seroconcordant couples in Rakai, Uganda, by use of next-generation deep sequencing. *J Clin Microbiol* 49: 2859–2867. doi: [10.1128/JCM.00804-11](https://doi.org/10.1128/JCM.00804-11) PMID: [21697329](https://pubmed.ncbi.nlm.nih.gov/21697329/)
41. Piantadosi A, Ngayo MO, Chohan B, Overbaugh J (2008) Examination of a second region of the HIV type 1 genome reveals additional cases of superinfection. *AIDS Res Hum Retroviruses* 24: 1221. doi: [10.1089/aid.2008.0100](https://doi.org/10.1089/aid.2008.0100) PMID: [18729772](https://pubmed.ncbi.nlm.nih.gov/18729772/)
42. Templeton AR, Kramer MG, Jarvis J, Kowalski J, Gange S, et al. (2009) Multiple-infection and recombination in HIV-1 within a longitudinal cohort of women. *Retrovirology* 6: 54. doi: [10.1186/1742-4690-6-54](https://doi.org/10.1186/1742-4690-6-54) PMID: [19493346](https://pubmed.ncbi.nlm.nih.gov/19493346/)
43. Ebert D (1994) Virulence and local adaptation of a horizontally transmitted parasite. *Science* 265: 1084–1086. PMID: [17832903](https://pubmed.ncbi.nlm.nih.gov/17832903/)
44. Lively CM, Dybdahl MF (2000) Parasite adaptation to locally common host genotypes. *Nature* 405: 679–681. PMID: [10864323](https://pubmed.ncbi.nlm.nih.gov/10864323/)
45. Kawashima Y, Pfafferoth K, Frater J, Matthews P, Payne R, et al. (2009) Adaptation of HIV-1 to human leukocyte antigen class I. *Nature* 458: 641–645. doi: [10.1038/nature07746](https://doi.org/10.1038/nature07746) PMID: [19242411](https://pubmed.ncbi.nlm.nih.gov/19242411/)
46. Furutsuki T, Hosoya N, Kawana-Tachikawa A, Tomizawa M, Odawara T, et al. (2004) Frequent transmission of cytotoxic-T-lymphocyte escape mutants of human immunodeficiency virus type 1 in the highly HLA-A24-positive Japanese population. *J Virol* 78: 8437–8445. PMID: [15280452](https://pubmed.ncbi.nlm.nih.gov/15280452/)
47. Schmid BV, Kesmir C, de Boer RJ (2008) The specificity and polymorphism of the MHC class I prevents the global adaptation of HIV-1 to the monomorphic proteasome and TAP. *PLoS One* 3: e3525. doi: [10.1371/journal.pone.0003525](https://doi.org/10.1371/journal.pone.0003525) PMID: [18949050](https://pubmed.ncbi.nlm.nih.gov/18949050/)
48. John M, Heckerman D, James I, Park LP, Carlson JM, et al. (2010) Adaptive interactions between HLA and HIV-1: highly divergent selection imposed by HLA class I molecules with common supertype motifs. *J Immunol* 184: 4368–4377. doi: [10.4049/jimmunol.0903745](https://doi.org/10.4049/jimmunol.0903745) PMID: [20231689](https://pubmed.ncbi.nlm.nih.gov/20231689/)
49. Cotton LA, Kuang XT, Le AQ, Carlson JM, Chan B, et al. (2014) Genotypic and functional impact of HIV-1 adaptation to its host population during the North American epidemic. *PLoS Genet* 10: e1004295. doi: [10.1371/journal.pgen.1004295](https://doi.org/10.1371/journal.pgen.1004295) PMID: [24762668](https://pubmed.ncbi.nlm.nih.gov/24762668/)
50. Malim MH, Bieniasz PD (2012) HIV Restriction Factors and Mechanisms of Evasion. *Cold Spring Harb Perspect Med* 2: a006940. doi: [10.1101/cshperspect.a006940](https://doi.org/10.1101/cshperspect.a006940) PMID: [22553496](https://pubmed.ncbi.nlm.nih.gov/22553496/)
51. Zhao K, Ishida Y, Oleksyk TK, Winkler CA, Roca AL (2012) Evidence for selection at HIV host susceptibility genes in a West Central African human population. *BMC Evol Biol* 12: 237. doi: [10.1186/1471-2148-12-237](https://doi.org/10.1186/1471-2148-12-237) PMID: [23217182](https://pubmed.ncbi.nlm.nih.gov/23217182/)
52. Ploquin MJ, Jacquelin B, Jochems SP, Barre-Sinoussi F, Muller-Trutwin MC (2012) Innate immunity in the control of HIV/AIDS: recent advances and open questions. *AIDS* 26: 1269–1279. doi: [10.1097/QAD.0b013e328353e46b](https://doi.org/10.1097/QAD.0b013e328353e46b) PMID: [22472855](https://pubmed.ncbi.nlm.nih.gov/22472855/)
53. Bremermann HJ, Thieme HR (1989) A competitive exclusion principle for pathogen virulence. *J Math Biol* 27: 179–190. PMID: [2723551](https://pubmed.ncbi.nlm.nih.gov/2723551/)
54. Castillo-Chavez C, Huang W, Li J (1996) Competitive Exclusion in Gonorrhea Models and Other Sexually Transmitted Diseases. *SIAM Journal on Applied Mathematics* 56: 494–508.
55. Hue S, Pillay D, Clewley JP, Pybus OG (2005) Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc Natl Acad Sci USA* 102: 4425–4429. PMID: [15767575](https://pubmed.ncbi.nlm.nih.gov/15767575/)
56. Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP (2007) Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc Natl Acad Sci USA* 104: 17441–17446. PMID: [17954909](https://pubmed.ncbi.nlm.nih.gov/17954909/)
57. Bartha I, Assel M, Sloot PM, Zazzi M, Torti C, et al. (2013) Superinfection with drug-resistant HIV is rare and does not contribute substantially to therapy failure in a large European cohort. *BMC Infect Dis* 13: 537. doi: [10.1186/1471-2334-13-537](https://doi.org/10.1186/1471-2334-13-537) PMID: [24219163](https://pubmed.ncbi.nlm.nih.gov/24219163/)

58. de Silva TI, van Tienen C, Onyango C, Jabang A, Vincent T, et al. (2013) Population dynamics of HIV-2 in rural West Africa: comparison with HIV-1 and ongoing transmission at the heart of the epidemic. *AIDS* 27: 125–134. doi: [10.1097/QAD.0b013e32835ab12c](https://doi.org/10.1097/QAD.0b013e32835ab12c) PMID: [23032414](https://pubmed.ncbi.nlm.nih.gov/23032414/)
59. Arien KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, et al. (2005) The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. *J Virol* 79: 8979–8990. PMID: [15994792](https://pubmed.ncbi.nlm.nih.gov/15994792/)
60. Gilbert PB, McKeague IW, Eisen G, Mullins C, Gueye NA, et al. (2003) Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. *Stat Med* 22: 573–593. PMID: [12590415](https://pubmed.ncbi.nlm.nih.gov/12590415/)
61. Batorsky R, Kearney MF, Palmer SE, Maldarelli F, Rouzine IM, et al. (2011) Estimate of effective recombination rate and average selection coefficient for HIV in chronic infection. *Proc Natl Acad Sci USA* 108: 5661–5666. doi: [10.1073/pnas.1102036108](https://doi.org/10.1073/pnas.1102036108) PMID: [21436045](https://pubmed.ncbi.nlm.nih.gov/21436045/)
62. Zhang M, Foley B, Schultz AK, Macke JP, Bulla I, et al. (2010) The role of recombination in the emergence of a complex and dynamic HIV epidemic. *Retrovirology* 7: 25. doi: [10.1186/1742-4690-7-25](https://doi.org/10.1186/1742-4690-7-25) PMID: [20331894](https://pubmed.ncbi.nlm.nih.gov/20331894/)
63. de Sousa JD, Müller V, Lemey P, Vandamme AM (2010) High GUD incidence in the early 20 century created a particularly permissive time window for the origin and initial spread of epidemic HIV strains. *PLoS One* 5: e9936. doi: [10.1371/journal.pone.0009936](https://doi.org/10.1371/journal.pone.0009936) PMID: [20376191](https://pubmed.ncbi.nlm.nih.gov/20376191/)
64. Piantadosi A, Chohan B, Chohan V, McClelland RS, Overbaugh J (2007) Chronic HIV-1 infection frequently fails to protect against superinfection. *PLoS Pathog* 3: e177. PMID: [18020705](https://pubmed.ncbi.nlm.nih.gov/18020705/)
65. Murray CJ, Ortblad KF, Guinovart C, Lim SS, Wolock TM, et al. (2014) Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 384: 1005–1070. doi: [10.1016/S0140-6736\(14\)60844-8](https://doi.org/10.1016/S0140-6736(14)60844-8) PMID: [25059949](https://pubmed.ncbi.nlm.nih.gov/25059949/)
66. Moslonka-Lefebvre M, Bonhoeffer S, Alizon S (2012) Weighting for sex acts to understand the spread of STI on networks. *J Theor Biol* 311: 46–53. doi: [10.1016/j.jtbi.2012.06.031](https://doi.org/10.1016/j.jtbi.2012.06.031) PMID: [22766360](https://pubmed.ncbi.nlm.nih.gov/22766360/)
67. Chemaitelly H, Awad SF, Abu-Raddad LJ (2014) The risk of HIV transmission within HIV-1 sero-discordant couples appears to vary across sub-Saharan Africa. *Epidemics* 6: 1–9. doi: [10.1016/j.epidem.2013.11.001](https://doi.org/10.1016/j.epidem.2013.11.001) PMID: [24593916](https://pubmed.ncbi.nlm.nih.gov/24593916/)
68. R Core Team (2014) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. PMID: [25566584](https://pubmed.ncbi.nlm.nih.gov/25566584/)
69. Clauset A, Shalizi C, Newman M (2009) Power-Law Distributions in Empirical Data. *SIAM Review* 51: 661–703.
70. Alstott J, Bullmore E, Plenz D (2014) Powerlaw: a Python package for analysis of heavy-tailed distributions. *PLoS One* 9: e85777. doi: [10.1371/journal.pone.0085777](https://doi.org/10.1371/journal.pone.0085777) PMID: [24489671](https://pubmed.ncbi.nlm.nih.gov/24489671/)
71. Latora V, Nyamba A, Simpore J, Sylvette B, Diane S, et al. (2006) Network of sexual contacts and sexually transmitted HIV infection in Burkina Faso. *J Med Virol* 78: 724–729. PMID: [16628573](https://pubmed.ncbi.nlm.nih.gov/16628573/)
72. Kreiss JK, Koeh D, Plummer FA, Holmes KK, Lightfoote M, et al. (1986) AIDS virus infection in Nairobi prostitutes. Spread of the epidemic to East Africa. *N Engl J Med* 314: 414–418. PMID: [3484804](https://pubmed.ncbi.nlm.nih.gov/3484804/)
73. Ferry B, Carael M, Buve A, Auvert B, Laourou M, et al. (2001) Comparison of key parameters of sexual behaviour in four African urban populations with different levels of HIV infection. *AIDS* 15 Suppl 4: S41–50. PMID: [11686464](https://pubmed.ncbi.nlm.nih.gov/11686464/)
74. Buve A, Carael M, Hayes RJ, Auvert B, Ferry B, et al. (2001) Multicentre study on factors determining differences in rate of spread of HIV in sub-Saharan Africa: methods and prevalence of HIV infection. *AIDS* 15 Suppl 4: S5–14. PMID: [11686465](https://pubmed.ncbi.nlm.nih.gov/11686465/)
75. Buve A, Lagarde E, Carael M, Rutenberg N, Ferry B, et al. (2001) Interpreting sexual behaviour data: validity issues in the multicentre study on factors determining the differential spread of HIV in four African cities. *AIDS* 15 Suppl 4: S117–126. PMID: [11686460](https://pubmed.ncbi.nlm.nih.gov/11686460/)
76. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, et al. (2005) Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J Infect Dis* 191: 1403–1409. PMID: [15809897](https://pubmed.ncbi.nlm.nih.gov/15809897/)
77. Powers KA, Poole C, Pettifor AE, Cohen MS (2008) Rethinking the heterosexual infectivity of HIV-1: a systematic review and meta-analysis. *Lancet Infect Dis* 8: 553–563. doi: [10.1016/S1473-3099\(08\)70156-7](https://doi.org/10.1016/S1473-3099(08)70156-7) PMID: [18684670](https://pubmed.ncbi.nlm.nih.gov/18684670/)
78. Buchbinder SP, Katz MH, Hessel NA, O'Malley PM, Holmberg SD (1994) Long-term HIV-1 infection without immunologic progression. *AIDS* 8: 1123–1128. PMID: [7986410](https://pubmed.ncbi.nlm.nih.gov/7986410/)
79. CASCADE Collaboration (2000) Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Concerted Action on SeroConversion to AIDS and Death in Europe. *Lancet* 355: 1131–1137. PMID: [10791375](https://pubmed.ncbi.nlm.nih.gov/10791375/)