An Eco-Evolutionary Framework to Predict Long-term Strain Diversity and Turnover in Infectious Diseases

Jiawei Liu[1], Qixin He[1*]

[1]Department of Biological Sciences, Purdue University,
915 Mitch Daniels Blvd, West Lafayette, IN 47907,USA

*To whom correspondence should be addressed; E-mail: heqixin@purdue.edu.

Abstract. Strain diversity and regimes of strain dynamics vary significantly among diseases. Understanding the eco-evolutionary forces that shape pathogen strain diversity is central to predicting their response to environmental and demographic changes and informing disease control strategies. However, most existing models focus on short-term invasion processes, which are insufficient for understanding long-term dynamics. Here, we develop Multi-Strain Eco-Evo Dynamics (MultiSEED), a unified theoretical framework that integrates a deterministic n-strain SIRS (Susceptible-Infected-Recovered-Susceptible) model with continuous-time stochastic processes to approximate multi-strain disease dynamics across both ecological and evolutionary scales. MultiSEED accounts for key mechanisms often overlooked in previous models, such as finite strain-specific immunity, mixed infections and within-host competition. Additionally, the model can distinguish between transient and persistent strains and classify strain dynamics into regimes of extinction, replacement, or long-term coexistence. We apply MultiSEED to four human pathogens (influenza A, HIV-1, pneumococcus, and malaria) that differ widely in

life history and transmission ecology. The model accurately reproduces observed levels of strain diversity and lifespan. The model also reveals how interactions among host population size, transmissibility, cross-immunity, and immune loss collectively govern strain diversity, whereas strain innovation rates have a more significant impact on the regimes of strain dynamics. The model further highlights how parameter interactions near the persistence threshold are highly complex, requiring MultiSEED evaluations for regime predictions. By bridging short-term epidemiological processes with long-term evolutionary dynamics, MultiSEED offers a quantitative framework for estimating the emergence, maintenance, and turnover of pathogen strain diversity across a range of infectious diseases and ecological contexts.

The evolution of lymphocytes with diverse antigen receptors marked a key innovation in vertebrate immunity, enabling recognition, memory, and targeted elimination of pathogens (*1*). In jawless and jawed vertebrates, this diversity arises through recombination of leucine-rich repeat sequences or V(D)J gene segments, respectively, generating large repertoires capable of detecting a vast range of antigens with high precision (*2*). In response, pathogens have developed a counter-strategy: antigenic variation, a process that continually alters their surface proteins or polysaccharides to evade immune detection (*3*). This strategy hinders the development of sterilizing immunity, allowing for persistent or recurrent infection (*4, 5*). Given its success, one might expect antigenic variation to be widespread across pathogens, driving high strain diversity in most systems. Yet, this is not what we observe.

Pathogen strain diversity varies widely among diseases, often in ways that defy intuitive expectations. Some pathogens, such as *Streptococcus pneumonia* and *Plasmodium falciparum*, maintain extensive antigenic diversity despite moderate transmissibility (*6, 7*). In contrast, viruses, such as measles and chickenpox, are highly transmissible and mutate rapidly, yet cir-

culate as single dominant strains (*8*). These contrasting patterns raise a fundamental question: what governs the emergence, maintenance, or collapse of strain diversity in host-pathogen systems?

Previous modeling efforts have explored how factors such as strain innovation rates (*9, 10*), mutational patterns (*11–13*), pathogen life history (*14, 15*), population structure (*16*), transmission mechanisms (*17, 18*), and cross-immunity (*19*) shape diversity. What has not been extensively explored is the impact of finite specific immunity and within-host mixed infection. Many infectious diseases do not elicit lifelong immunity to a given strain, with immune protection often waning over time (*5, 20–23*). However, this effect of immune loss has rarely been incorporated into multi-strain modeling frameworks, despite its potential importance for long-term pathogen dynamics (but see (*24*)). Similarly, mixed infections, i.e., simultaneous infection by multiple antigenically distinct strains, are frequently observed in over 50 human diseases and many animal diseases (*25, 26*). The prevalence of different strains is influenced by transmission intensity, clinical risk (*27*), host immunity, and host age (*28*). Importantly, within-host competition resulting from mixed infections (*26*) and its impact on transmissibility and strain persistence have not been captured by multi-strain models.

Existing modeling approaches for antigenic diversity can be broadly grouped into three classes. The first consists of short-term, equation-based deterministic models, often supplemented by stochastic simulations (*14, 29, 30*). These models typically assume a small number of strains, lifelong immunity, and no mixed infections. While useful for exploring outbreak dynamics and short-term competition among strains, they are not designed to capture the long-term processes governing diversity.

The second class comprises long-term stochastic, individual-based simulation models (*7, 9, 10*). These models can incorporate complex immunological and evolutionary mechanisms, including mutation and immune escape, but are generally tailored to specific diseases. Their com-

putational complexity limits analytical tractability and impedes general insights across broader parameter regimes.

A third, more integrative approach is offered by Abu-Raddad & Ferguson (*19*), who developed an equation-based framework linking cross-immunity and mutation to long-term diversity. While analytically elegant, the model omits several critical features. It does not account for mixed infections, assumes lifelong immunity, and overlooks the impact of epidemic burnout on short-term strain persistence (i.e., the loss of newly emergent strains due to post-epidemic troughs in susceptibles and demographic stochasticity) (*31, 32*). Although previous studies have noted the impact of trough size (or overcompensatory effect) on strain persistence (*15, 24*), their quantitative effects on the eco-evolutionary feedback of strain diversity remain poorly understood.

Together, these limitations point to the need for a general, analytically tractable framework that incorporates mixed infections, finite immunity durations, and demographic stochasticity across ecological and evolutionary time scales. We address this gap with **M**ulti-**S**train **E**co-**E**vo **D**ynamics (MultiSEED), a theoretical framework that unifies epidemiological, immunological, and evolutionary processes in multi-strain pathogens. MultiSEED captures both short- and long-term dynamics, accounts for strain interactions within hosts, and enables analytical predictions of key quantities, including the number of circulating strains, the turnover rates of emergent versus persistent variants, and the conditions under which strain coexistence or replacement dominates. We validate this framework using individual-based simulations and compare its predictions to observed diversity patterns in influenza A, HIV-1, *Streptococcus pneumoniae* (hereafter pneumococcus), and *Plasmodium falciparum* (hereafter malaria).

MultiSEED framework

To model the transmission dynamics and long-term persistence of multiple co-circulating strains, MultiSEED integrates a deterministic n-strain SIRS (Susceptible-Infected-Recovered-Susceptible) model with continuous-time stochastic processes at both ecological and evolutionary timescales (Fig. 1, details in SM). The deterministic component extends the classic SIRS model to a multi-strain, status-based formulation, which does not record individual infection histories (Fig. 1A). Instead, it adopts a pathogen-centric perspective, tracking host immune and current infection states with respect to each strain (*33*). We incorporate two forms of immunity that reflect key features of multi-strain pathogen dynamics: (1) strain-specific immunity, which prevents reinfection by the same strain but wanes over time, and (2) lifelong cross-immunity, which reduces susceptibility to other strains following any infection (ρ in Fig. 1A).

For each strain i, hosts are categorized into three compartments: proportions of susceptible (x_i), infected (y_i), and recovered (z_i). The susceptible class x_i is further subdivided into fully naïve hosts ($x_\varnothing$) and partially immune hosts with reduced susceptibility due to prior exposure to other strains or waned specific immunity to strain i ($x_{i\setminus\varnothing}$). This formulation also allows for the convenient incorporation of mixed infections, where a single host can harbor multiple genetically distinct strains of the same pathogen species. As a result, the sum of y_i across strains can exceed 1. Within-host interactions among co-infecting strains take the form of competition for host resources, which reduces per-strain pathogen density and thereby decreases transmissibility. We denote this competition-induced reduction in baseline transmissibility β by β^*, and the corresponding reduction in reproduction number as R_0^*. At the between-host level, co-circulating strains accelerate transitions from $x_\varnothing$ to $x_{i\setminus\varnothing}$ (Fig. 1A), thereby reducing the hosts' overall susceptibility to each strain. The infection recovery rate (γ) and the rate of strain-specific immunity loss (δ) are assumed to be unaffected by inter-strain competition. The

host population size is held constant at demographic equilibrium, with the birth rate equal to the natural death rate (μ), and infection does not contribute to mortality. For analytical and computational tractability, we assume all circulating strains are epidemiologically equivalent and differ only in their antigenic space. This symmetry enables us to derive a single set of equations that characterizes the equilibrium behavior representative of all strains, along with an additional set describing the invasion dynamics of a novel strain introduced into a resident strain population (Eq. S9 and Eq. S10 in SM).

After rescaling time in the model by the expected duration of infection (i.e., $1/(\gamma + \mu)$), the dynamics of each strain are governed by three key parameters: the basic reproduction number (R_0), the relative resource replenishment rate (ε), and cross-strain susceptibility (ρ). The parameter $\varepsilon = \varepsilon_1 + \varepsilon_2$ captures the rate at which the pool of strain-specific susceptible hosts is replenished relative to the pathogen's lifespan. This replenishment occurs through two mechanisms: waning of strain-specific immune memory ($\varepsilon_1 = \delta/(\gamma + \mu)$) and the birth of immunologically naïve individuals ($\varepsilon_2 = \mu/(\gamma + \mu)$). For a detailed derivation and interpretation of ε, see SM section S1.1 and S3.1.

To characterize both the invasion dynamics of novel strains and the demographic stochasticity affecting resident strains, we formulate two classes of stochastic processes—one operating on the ecological timescale and the other on the evolutionary timescale (Fig. 1B). New strains with adequate antigenic distance from existing strains emerge continuously through intrinsic genetic processes, such as mutation, governed by the strain innovation rate ν. However, not all novel strains successfully establish; they must overcome two key stochastic bottlenecks to persist within the system.

The first barrier, which we term the fizzle phase, occurs when a newly formed strain fails to generate enough secondary infections and dies out shortly after introduction (*34*). If a strain escapes this initial extinction risk, it proceeds to its first epidemic wave. However, this expansion

is followed by a sharp depletion of susceptible hosts towards the end of an epidemic, which can lead to a second extinction event, referred to as epidemic burnout (*32*). During this invasion process, resident strains may also face stochastic extinction due to reductions in their infected host populations (SM Section S1.3; Fig. S1).

Strains that survive both the fizzle phase and epidemic burnout are considered successful invaders and transition into long-term circulation. While they may still experience stochastic fluctuations during subsequent oscillatory dynamics, the magnitude of these effects is considerably smaller (*32*). Adapting the approach of Parsons et al. (*32*) for single-strain models, we numerically approximate the probabilities of fizzle and burnout using a combination of deterministic dynamics and time-inhomogeneous stochastic birth–death processes (see SM Section S1.3). For surviving strains that reach endemic levels, we assume quasi-stationary distributions of infected and susceptible hosts for the specific strain. Consequently, the long-term extinction dynamics of these persistent strains can then be approximated by an Ornstein–Uhlenbeck (OU) process (*35, 36*) (SM Section S1.4).

Building on these invasion dynamics and the separation of timescales, we develop a compartmental model to track strain turnover and predict long-term strain diversity. Strains are categorized into two classes: transient strains (n_1), which have survived the fizzle phase but not yet completed their first epidemic, and persistent strains (n_2), which have survived at least one full epidemic wave (Fig. 1C). The total number of co-circulating strains at equilibrium is given by $n = n_1 + n_2$, when the rates of strain invasion and extinction are balanced (SM Section S1.5; Eq. S32). The structure of this model allows us to delineate distinct regimes of strain dynamics based on the ratio of n_1 to n_2, their respective lifespans, and turnover rates.

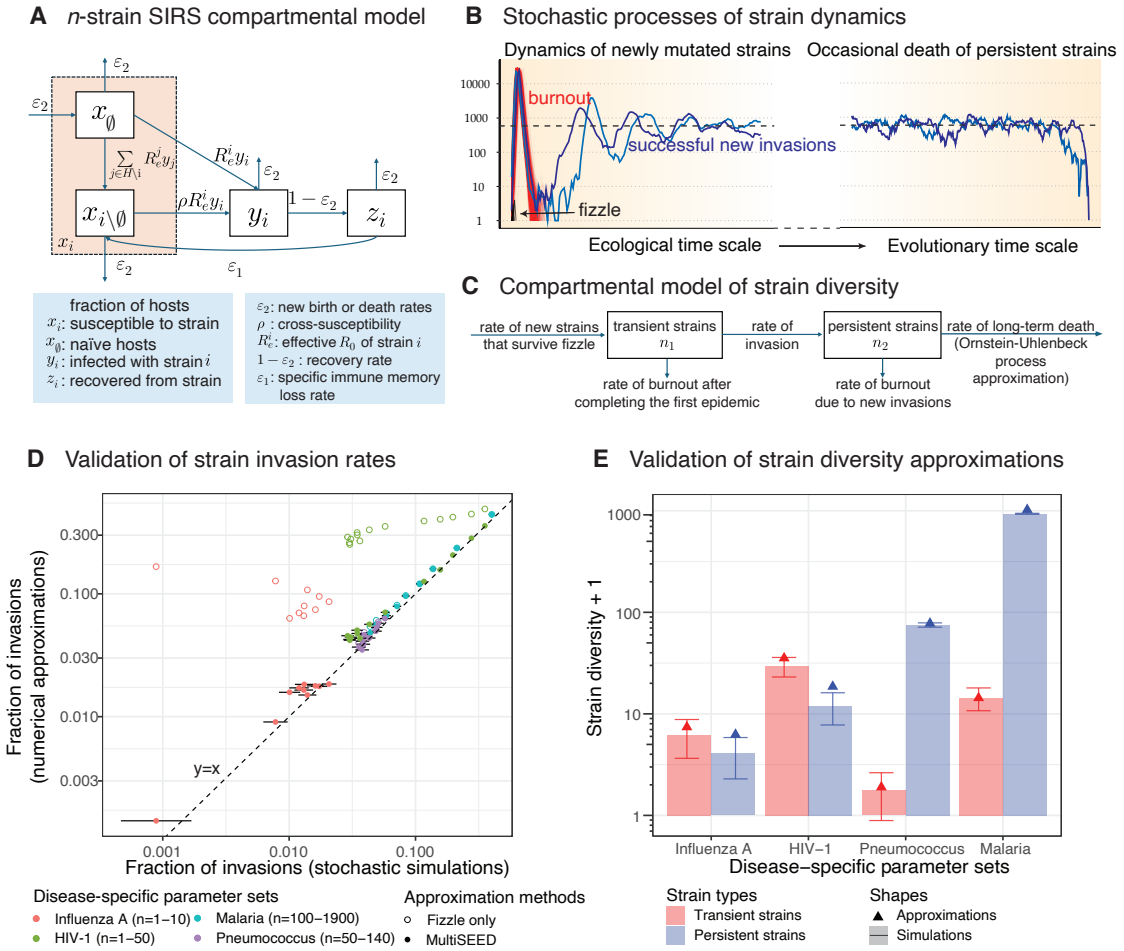

Figure 1: **MultiSEED framework and the validation results.** **(A)** The deterministic n-strain status-based SIRS model. The blue boxes list the definition of host classes per strain and the definition of each parameter. **(B)** The illustration of continuous time stochastic processes in strain dynamics over the ecological (left) and evolutionary (right) time scales. **(C)** The compartmental model of strain dynamics. We categorize strains into transient (n_1) and persistent (n_2) groups and track the rates of fizzle and burnout for transient strains, together with rates of burnout and long-term death for persistent strains, as illustrated in **B**. **(D-E)** Validation of MultiSEED by comparison with the simulation results from an agent-based model, under the same parameter sets for four disease: Influenza A: $\varepsilon = 0.005, \varepsilon_1 = 0.0045, R_0 = 2$; HIV-1: $\varepsilon = 0.022, \varepsilon_1 = 0.002, R_0 = 3$; Pneumococcus: $\varepsilon = 0.013, \varepsilon_1 = 0.0065, R_0 = 1.5$; Malaria: $\varepsilon = 0.035, \varepsilon_1 = 0.026, R_0 = 5$. Other parameters can be found in Table 1. **(D)** Comparison of the invasion probability of new strains into a system with n resident strains between numerical approximations and stochastic simulations; error bars represent 95% binomial proportion confidence interval. **(E)** Comparison of the number of transient and persistent strains at equilibrium between numerical approximations and stochastic simulations; error bars represent standard deviations of simulation runs.

Validation of MultiSEED using an agent-based simulator

We first evaluated whether MultiSEED accurately estimates the invasion probability of a novel strain in the presence of n resident strains by comparing its numerical approximations to results from an agent-based simulation model (adapted from (*7*); see SM Section S1.6). To assess performance across a broad range of epidemiological scenarios, we tested representative parameter sets for four major human pathogens: influenza A, HIV-1, pneumococcus, and malaria, spanning a wide range of assumed resident strain numbers. When both the fizzle phase and epidemic burnout are incorporated as prerequisites for successful invasion, the numerically approximated probabilities of invasion closely match those obtained from stochastic simulations (filled circles in Fig. 1D). In contrast, when only the fizzle phase is considered, an approach taken in previous studies (e.g., (*19*)), the numerical approximations consistently overestimate invasion probabilities, especially for influenza A and HIV-1 (open circles in Fig. 1D). This discrepancy reflects the substantial impact of epidemic burnout in these diseases (Fig. S2).

We then assessed the quality of numerical approximations for long-term strain diversity by MultiSEED in Fig. 1E. The model predictions for the number of transient and persistent strains at equilibrium generally fall within the range of outcomes observed in stochastic simulations, particularly for transient strains, demonstrating strong agreement. However, for the parameter regimes of influenza A and HIV-1, MultiSEED slightly overestimates the number of persistent strains by approximately a factor of two relative to simulation means. This overestimation is primarily due to the known limitation of Ornstein–Uhlenbeck (OU) approximation to estimate the expected lifespan of an endemic strain (T_Q) (see (*35*)). The OU approximation of T_Q captures the correct order of magnitude but becomes less accurate in diseases where many strains stochastically go extinct during successive epidemic troughs before reaching the quasi-stationary endemic state assumed by the OU model.

Despite these deviations in certain parameter regimes, the predicted invasion probabilities and strain diversity estimates from MultiSEED remain within the correct order of magnitude and show strong agreement with results from full agent-based simulations. Thus, MultiSEED provides a reliable framework for estimating long-term strain diversity and turnover, even in systems characterized by complex or volatile dynamics.

Factors affecting strain diversity

To identify disease characteristics that promote or limit long-term strain diversity, we examine a set of key parameters that capture both deterministic and stochastic forces shaping pathogen evolution. In addition to the three core parameters in the deterministic model (i.e., basic reproduction number (R_0), relative resource replenishment rate (ε), and cross-strain susceptibility (ρ)), two additional parameters are required to describe stochastic dynamics: the host population size (N_h) accessible to pathogens and the strain innovation rate (ν). Together, these five high-level parameters represent fundamental aspects of pathogen ecology and evolution, encompassing life history traits, transmission dynamics, host immune responses, and the rate of antigenic diversification. Among these, N_h, R_0, and ε are strongly shaped by host-pathogen interactions, as well as external ecological factors, such as environmental variability, vector and host population structure, and migration. The strain innovation rate ν reflects the intrinsic genetic properties of the pathogens, measuring the mutation and recombination rates weighted by the likelihood that novel strains are antigenically distinct enough to evade existing immunity. While ρ represents immunological effects alone, ε serves as a bridge between immunological and ecological processes, as it is influenced by the rate of immune waning (δ), duration of infection (γ), and host turnover rates (μ). By focusing on this minimal yet biologically meaningful parameter set, we aim to identify general principles governing strain diversity across a wide range of pathogen systems.

We found that the interplay between the relative resource replenishment rate (ε) and the basic reproduction number (R_0) jointly determines the threshold for long-term disease persistence (i.e., the equilibrium strain diversity $\hat{n} >= 1$). This threshold emerges from the inclusion of epidemic burnout in the model dynamics, making the classical condition $R_0 > 1$ insufficient for sustained transmission over evolutionary timescales (Fig. 2A; see also (*32*)). In the lower-left region of Fig. 2A (gray), strain diversity drops to zero, indicating complete failure of strains to persist when either ε or R_0 is too low. In these scenarios, old strains may go extinct stochastically in troughs and new strains may emerge via introduction, mutation or recombination but are quickly lost due to stochastic extinction or epidemic burnout. This may represent pathogens with short infectious periods and/or long-lasting immunity, such as those causing acute infections or resulting from zoonotic spillover. Toward the upper-right region of the parameter space, strain diversity rises rapidly in magnitude as both ε and R_0 reach high values, a rare combination in natural diseases. Instead, most pathogens employ life history in intermediate regimes, where the interaction between R_0 and ε becomes the dominant force shaping strain diversity.

Among all parameters, ε exhibits the most dramatic effect on diversity, effectively acting as a switch. Below a critical threshold ($\varepsilon < 0.003$, under the tested parameter space), the system fails to support persistent strains even when the basic reproduction number (R_0) is over 20. Above the persistence threshold, strain diversity increases sharply with ε and quickly approaches an asymptotic maximum (Fig. 2B-C, Fig S3A, Fig. S4A). Biologically, ε can be increased by higher host turnover (via birth, death, or immigration), faster waning of strain-specific immunity, or longer infection durations. Thus, even modest changes in these underlying processes can profoundly impact a pathogen's ability to sustain strain diversity over time.

Once the persistence threshold is crossed, strain diversity $\hat{n}$ increases rapidly with both R_0 and the host population size (N_h) (Fig. 2B-E, Fig. S3B-D, Fig. S4B-C). For R_0, this relationship exhibits a two-phase pattern: near the threshold, a twofold increase in R_0 can produce up to a

three-order-of-magnitude increase in $\hat{n}$, whereas further from the threshold, increases in $\hat{n}$ scale more gradually with R_0, sometimes reaching plateau status under large N_h and ν. For N_h, the diminishing returns in diversity growth rate are less dramatic, but still display a more rapid increase near the persistence threshold.

Compared with the critical influence of N_h, the effect of the strain innovation rate (ν) is more nuanced (Fig. 2F, Fig. S4D). Diversity increases approximately linearly with ν, but the slope of this relationship flattens as R_0 and N_h increase. This indicates that, under conditions favorable to transmission and persistence, strain diversity becomes less dependent on innovation rate. This contrasts with classical neutral models in population genetics, where $N_h \times \nu$ is the primary determinant of diversity. In our framework, N_h exerts a stronger and nonlinear influence on $\hat{n}$ (see also (*19*)), while a magnitude increase in ν alone does not result in a magnitude increase in $\hat{n}$. This is because large N_h not only increases the supply of new strains but also reduces burnout risk and extends the average lifespan (T_Q) of persistent strains after successful invasion. In contrast, ν affects only the rate of strain emergence, not their survival.

While the strain innovation rate ν has a limited impact on equilibrium diversity in well-established systems, it plays a more prominent role in new diseases when the host-pathogen association was recently established and the systems are far from equilibrium. As shown in the shaded regions of Fig. 2C, increases in ν can significantly boost observed diversity near the persistence threshold. Therefore, when the host-pathogen association is recent and strain accumulation is ongoing, higher ν accelerates the system's approach to equilibrium.

The cross-strain susceptibility parameter ρ has a modest influence on the equilibrium value of $\hat{n}$ except for extreme levels of cross-immunity or close to the persistence threshold. It also has a very minor effect on the persistence threshold along the R_0 or ε axis (Fig. S3A-B).

In conclusion, we have found that ε, R_0 and N_h have more impact on strain diversity than ν or ρ. Our findings suggest that interactive factors, including pathogen life history and host-

pathogen interactions, contribute significantly more to the accumulation of strain diversity than intrinsic pathogen genetic factors and cross-reactivity between antigens.

Factors influencing regimes of strain dynamics

To understand long-term strain dynamics, it is crucial not only to quantify expected diversity, but also to characterize the composition of strains (i.e., whether transient or persistent strains dominate) and their turnover rates, which determine how frequently strains are replaced and over what timescales. MultiSEED enables estimation of both the number and lifespan of transient and persistent strains. These quantities indicate whether the system is dominated by the rapid turnover of a few strains, the coexistence of multiple long-lived strains, or a combination of both. We thus classified disease systems into three strain dynamic regimes based on these properties: 1) Extinction, when no strain can persist over evolutionary timescales; 2) Replacement, when transient and persistent strains occur at similar frequencies and lifespans, indicating ongoing turnover; 3) Coexistence, when persistent strains dominate in both number and longevity. To differentiate between the replacement and coexistence regimes, we used two criteria: 1) the proportion of transient strains, $n_1/(n_1 + n_2)$, and 2) the ratio between the epidemic burnout time and the expected lifespan of persistent strains, $T_e/(T_e + T_Q)$ (Fig. S5 B-C). Low values for both quantities signal a long-term coexistence regime.

The regimes of strain dynamics exhibit complex dependencies on model parameters, and these parameters exert different effects on strain diversity versus strain dynamics. Notably, the two parameters that define the strain persistence threshold, ε and R_0, do not always influence strain regimes in a monotonic manner (Fig. 2G–J; see also Fig. S4A–B for additional parameter combinations). In general, increasing ε or R_0 beyond the persistence threshold promotes long-term strain coexistence. However, when either parameter is small or persistence is close to the threshold, the system tends to favor strain replacement.

Among all parameters, the strain innovation rate ν has a particularly strong and consistent influence on dynamic regimes. Increases in ν drive a near-linear shift toward strain replacement, regardless of other parameter values (Fig. 2K–L). In contrast, host population size N_h, while critical for determining the persistence threshold and setting the upper bound of strain diversity (Fig. 2E), has only a modest effect on regime classification. Larger N_h promotes strain replacement, probably due to its dual role as both a resource base for sustaining existing strains and a reservoir for seeding new invasions (Fig. S4C). Cross-strain susceptibility ρ tends to promote strain coexistence, but its effect becomes more ambiguous near the persistence threshold (Fig. S4E), where parameter interactions become highly nonlinear. Together, these results underscore that the regimes of strain dynamics are shaped by non-additive, sometimes non-monotonic interactions among parameters, especially near the persistence threshold. As such, regime classification under specific parameter combinations must be determined numerically through MultiSEED simulations, rather than inferred solely from theoretical expectations.

Strain diversity and regimes of strain dynamics in empirical diseases

To evaluate MultiSEED's ability to capture realistic strain dynamics, we assessed its predictions of strain composition, lifespan distributions and turnover regimes in four human infectious diseases with well-documented variation in strain diversity: influenza A, HIV-1, pneumococcus, and malaria (see Table 1 for parameter ranges and SM Section S1.8 for data sources and derivation methods). Given the strong influence of the resource replenishment rate (ε) and basic reproduction number (R_0) on predicted diversity and the relatively high uncertainty in their empirical estimates, we explored a range of (ε, R_0) combinations for each disease, while keeping other parameters (N_h, ν, ρ) fixed. The diseases show drastically different combinations of predicted numbers of transient strains ($\hat{n}_1$) and persistent strains ($\hat{n}_2$), in contrast to very similar

trends in disease prevalence among diseases across the parameter space (Fig. 3A, Fig. S5A). Focusing on the empirically plausible regions of (ε, R_0) denoted by white rectangles, we found that observed diversity falls within the model's predicted range for all four diseases (Fig. 3B).

Table 1: Empirical parameter ranges for the four focal human diseases: influenza A, HIV-1, pneumococcus and malaria, along with corresponding literature sources in brackets.

Disease	N_h	ν	ε	R_0	ρ	#Gens/year	#Years
Influenza A	10^6 (*37*)	10^{-4} (*38–44*)	$0.003 - 0.006$ (*23, 41–43*)	$1.2 - 1.85$ (*45*)	0.5 (*46*)	52 (*41–43*)	1000 (*47*)
HIV-1	10^4 (*48*)	10^{-3} (*49–56*)	$0.013 - 0.038$ (*49, 57–59*)	$1.6 - 3.1$ (*60*)	0.5 (*61*)	1 (*49*)	100 (*62, 63*)
Pneumococcus	10^6 (*64*)	10^{-6} (*65–70*)	$0.0065 - 0.023$ (*65, 66, 71–73*)	$1.23 - 2$ (*74–76*)	0.7 (*77–80*)	3 (*65, 66*)	100000 (*70*)
Malaria	10^5 (*81*)	10^{-5} (*7, 82–85*)	$0.033 - 0.059$ (*20, 82–84, 86–88*)	≥ 1 (*89*)	0.7 (*90*)	2.5 (*82–84*)	10000 (*91, 92*)

For influenza A, MultiSEED predicts either disease extinction or circulation of a small number of strains (between 1 and 10) at higher values of ε and R_0, with both transient and persistent strains occurring at low levels (Fig. 3A). This aligns with observed strain diversity of 1–3 co-circulating lineages (*44*). These results suggest that influenza A typically experiences local extinction dynamics, but under conditions of elevated transmission (e.g., due to increased host density or higher ε via host birth or migration), multiple co-circulating strains can still emerge and persist.

For HIV-1, the predicted equilibrium diversity is roughly an order of magnitude higher than for influenza, with both transient and persistent strains ranging between 10^1 and 10^2. However, given HIV-1's relatively recent emergence in humans ($\approx$ 100 years ago) (*62, 63*), the virus is unlikely to have reached equilibrium. We therefore simulated 100-year stochastic trajectories using the compartmental Markov process based on Eq. S32 for the empirical parameter ranges (see SM Section S1.7). This produced a narrower and lower strain diversity distribution than the equilibrium estimate, more consistent with current observations: approximately 13 known types, comprising four major groups and nine subtypes within group M (*93*), as shown by the blue histograms in Fig. 3B.

For pneumococcus, MultiSEED predicts a consistently low number of transient strains (≈ 1)

but a wide range of persistent strain diversity (from 10^1 to 10^3), driven primarily by small changes in R_0. This prediction encompasses the roughly 90 known immunologically distinct serotypes, defined by variation in the polysaccharide capsule (*6*).

For malaria, we explored a broad range of R_0 values associated with vector-borne transmission, which vary significantly with mosquito behavior and ecology. Consequently, predicted strain diversity spans multiple orders of magnitude, with persistent strains dominating the overall diversity. These predictions are consistent with the large variability in malaria parasite diversity observed across biogeographic regions with differing transmission intensity (*94*).

When comparing strain turnover rates, we found that while the expected lifespans of transient and persistent strains are comparable in influenza A and HIV-1, they differ by over two orders of magnitude in pneumococcus and malaria, with persistent strains lasting up to 10,000 years (Fig. 4A). Based on strain composition and their lifespan, we classified the regimes of strain dynamics for the four diseases. We tested cutoff thresholds from 0.1 to 0.4 to delineate these regimes and present results using a threshold of 1/3 in Fig. 4B (see alternative cutoffs in Fig. S6).

Interestingly, beyond the persistence threshold, the (N_h, ν, ρ) sets of pneumococcus and malaria always place them in the coexistence regime regardless of ε and R_0 (Fig. 4B). While regime classifications may shift across broad parameter regions of influenza A and HIV-1, they are always assigned strain replacement regime within their empirical parameter ranges, highlighted by white rectangles. Outside of the empirical region, if these two diseases have increased ε and R_0, the system generally shifts toward coexistence, except for a small band of coexistence regime at low R_0 and high ε (Fig. S6B).

The contrast of strain dynamics between replacement and coexistence regimes is also evident in stochastic simulation results (Fig. 4C). In influenza A, newly emerged strains often drive brief epidemics before going extinct, with rapid replacement. Conversely, pneumococ-

cus exhibits long-term coexistence of multiple strains at endemic equilibrium, punctuated by occasional replacement events.

The predicted regimes of strain dynamics align well with the observed strain turnover rates: strains in influenza A typically last 2–5 years and exhibit constant replacement of dominant lineages (*44*). For HIV-1, transient and persistent strains are predicted to have similar strain lifespans of approximately 100 years at equilibrium. When inspecting the results over a short coevolutionary timescale, we observe a greater disparity in lifespan between transient and persistent strains, indicating a regime that lies between replacement and coexistence. This pattern is reinforced by the long epidemic cycle of HIV-1 (i.e., $T_e \approx$ decades to centuries), and by the fact that predicted diversity is substantially higher than that of influenza A. Nevertheless, because most strains are still in the early phases of spread with respect to T_e, their frequencies fluctuate more dramatically than would be expected under true long-term coexistence. These characteristics are consistent with observed phylodynamic patterns: high population-level diversity is maintained, while within-host evolution is continuous and often marked by reversions to ancestral strains (*95*), leading to relatively stable strain composition. Additionally, there is substantial regional variation in dominant subtypes across the globe (*96*), further supporting the picture of a dynamic yet structured strain landscape.

For pneumococcus and malaria, MultiSEED predicts both high strain diversity and extremely long strain persistence times. Given their long histories of transmission in human populations for at least 10,000 years, equilibrium-based estimates are appropriate for characterizing their long-term strain dynamics. The predicted equilibrium regimes match empirical observations: pneumococcal serotype diversity is high, stable over time and geography, and shaped by frequency-dependent selection driven by host immunity (*6*). For malaria, although recombination continually reshuffles strain compositions, empirical evidence has shown the long persistence of antigenic alleles (*94, 97*), genes (*98*), and modules of strains (*99*), supporting the

coexistence regime identified by MultiSEED. For *var* genes, encoding the major variant surface antigens, trans-specific polymorphisms have been maintained between gorilla-associated *Plasmodium praefalciparum* and *P. falciparum*, predating *P. falciparum*'s transfer to humans ≈50,000 years ago (*94,97*). Analyses of other major antigens also reveal high within-population antigenic diversity, high divergence among haplotypes, and little geographic structure (*98*). In addition, clusters of malaria strains have been shown to persist over extended periods, despite gradual shifts in gene composition due to recombination (*99*). Therefore, while MultiSEED does not account for sexual recombination, its predictions of strain coexistence regime is robust to this limitation.

Summary and Perspectives

In summary, we present a novel numerical framework for predicting long-term strain diversity and regimes of strain dynamics. Unlike previous multi-strain models that focus primarily on short-term invasion processes, MultiSEED enables large-scale analytical exploration across broad parameter spaces, capturing dynamics at both ecological and evolutionary timescales. By applying MultiSEED to four human diseases (i.e., influenza A, HIV-1, pneumococcus, and malaria), we demonstrate that the model accurately reproduces observed patterns of strain diversity, persistence, and dynamic regimes, despite the wide variation in life history traits and transmission modes across these pathogens.

MultiSEED advances prior approaches by incorporating several key characteristics of pathogen evolution: finite strain-specific immunity, which allows for waning immune memory as a source of susceptible replenishment; mixed infections, enabling co-infection by multiple antigenically distinct strains, which results in within-host competition that modulates strain transmissibility; and a distinction between transient and persistent strains due to epidemic burnout, capturing different lifespans and turnover dynamics. These additions allow MultiSEED to make realistic

long-term predictions of strain dynamics that can be directly compared to empirical data.

One of the most significant conceptual contributions of this framework is the introduction of a composite parameter, ε, which represents the relative rate of resource replenishment for susceptible hosts. This parameter accounts for both immunity loss and demographic turnover relative to infection recovery rate, and emerges as a key determinant of disease persistence. Previous studies have shown that in the absence of sufficient host replenishment (e.g., susceptible individuals are replenished solely through new births in the population), long-term persistence is unlikely for most human diseases, particularly for short-duration infections (*32*). Our results expand on this by demonstrating that finite strain-specific immunity significantly boosts ε and promotes high strain diversity, as seen in pneumococcus and malaria. Additionally, high strain innovation rates (ν) or host population size (N_h) can compensate for low ε, promoting persistence in diseases like influenza A with otherwise very short infection durations. The persistence filter effect of ε also impacts the classic concept of critical community sizes (CCS) for disease persistence. Previous studies derived CCS from the estimated long-term lifespans of strains (i.e., T_Q in our model) (*19, 35, 100*). Our study suggests that the required CCS would be higher, as it needs to survive the burnout stage in addition to enduring long-term stochastic loss.

MultiSEED confirms that increases in any of the five key parameters lead to greater long-term strain diversity (*11, 19, 101, 102*), while also quantifying the relative importance of each parameter in shaping that diversity. In general, we show that host population size (N_h) and R_0 have a greater influence on equilibrium diversity. When close to the persistence threshold with small R_0, parameters such as cross-immunity (1-ρ) and innovation rate (ν) become more critical, where a small reduction in cross-immunity can lead to large shifts in diversity magnitudes.

The framework enables systematic analysis of how different parameters influence the regimes of strain dynamics. However, these effects are often non-monotonic and do not follow the more predictable patterns in strain diversity. Instead, the dynamic regime of a disease emerges from

complex interactions among multiple parameters, especially near the persistence threshold. This complexity is exemplified by comparing the dynamics of influenza A and pneumococcus. Despite similar values of R_0, N_h, and overlapping ranges of ε, pneumococcus exhibits stable coexistence, while influenza A displays frequent strain replacement. This divergence is largely driven by lower cross-immunity and lower innovation rate of pneumococcus, which permit long-term persistence of strains (Table 1). These comparisons highlight how specific parameter combinations co-determine the qualitative nature of strain dynamics and illustrate the utility of MultiSEED in exploring a broad array of empirical disease systems.

The framework also provides a powerful tool for studying how pathogen diversity may respond to environmental and demographic changes. Global shifts such as increasing human mobility, urbanization, or climate-induced habitat change can affect key epidemiological parameters like R_0, N_h, and ε, with potentially large effects on strain diversity and turnover. For example, greater connectivity among human populations may increase effective transmissibility and host replenishment, leading to increased strain diversity. Meanwhile, public health interventions, such as vaccination campaigns, will reduce the susceptible pool and the immune loss rate, thereby lowering the likelihood of diverse strain coexistence. MultiSEED can be applied to evaluate how such global trends influence pathogen evolution and to inform strategies for disease control in an increasingly interconnected world.

While MultiSEED captures a wide range of multi-strain dynamics, it also makes several simplifying assumptions. First, it assumes that all strains are epidemiologically equivalent, with equal transmissibility, recovery rates, and immune waning. In reality, fitness differences among strains can also impact their invasion probabilities and persistence (*24*). However, recent simulation work suggests that under strong immune-mediated competition, functional fitness across strains tends to converge over evolutionary timescales while immunological niche differences are widened (*103*). We thus predict that the equal fitness assumption is more valid

in high-diversity systems. In contrast, in transmission-limited or low-diversity scenarios, deviations from this assumption may lead MultiSEED to overestimate the number of coexisting strains. A natural extension would be to allow for two or more classes of strains with differing functional fitness. Second, the current model assumes a single host species, whereas many pathogens infect multiple hosts. A larger host range will increase ε and N_h, promoting persistent strain diversity in density-dependent transmission, or buffer transmission due to a reduction in R_0 (*104*). Future work could incorporate multi-host systems with varying susceptibilities and transmission roles. Third, the SIRS framework may not apply to all pathogens. For instance, tuberculosis involves complex latency dynamics (*105*), while dengue virus features immune enhancement upon reinfection (*106*); however, these interactions are not captured in the current model. Thus, the deterministic model in MultiSEED shall be adapted for these specific transmission systems before its application. Additionally, we have explored our parameter space relevant to human diseases. Applying MultiSEED to plant, animal, or vector-borne pathogens would require adapting assumptions about host demography, immune architecture, and transmission biology to reflect system-specific constraints and mechanisms. Finally, it is important to clarify that in this study, strains refer to immunologically distinct phenotypes. Other studies have focused on the fine-scale mutational pathways and molecular evolution underlying phenotypic diversification (*13, 107*). While such approaches yield insight into how specific antigenic types arise, our focus is on the effective number of distinct immunological entities circulating in the population over time. MultiSEED thus provides a coarse-grained, population-level perspective that complements mutation-level analyses.

In conclusion, MultiSEED represents a significant step forward in understanding the eco-evolutionary drivers of pathogen diversity and strain dynamics. The model not only reproduces empirical observations in multiple human diseases, but also provides a flexible framework for evaluating how pathogen diversity may respond to global change, demographic shifts, and in-

tervention strategies. Future developments include accommodating fitness variation of strains, multi-host dynamics, and investigating different types of transmission dynamics.

Acknowledgments

This research was partly supported by Ralph W. and Grace M. Showalter Award and discretionary funds from the Mary J. Elmore New Frontiers Professorship at Purdue University. We appreciate the support of Rosen Center for Advanced Computing at Purdue University and Purdue Institute of Inflammation, Immunology and Infectious Disease.

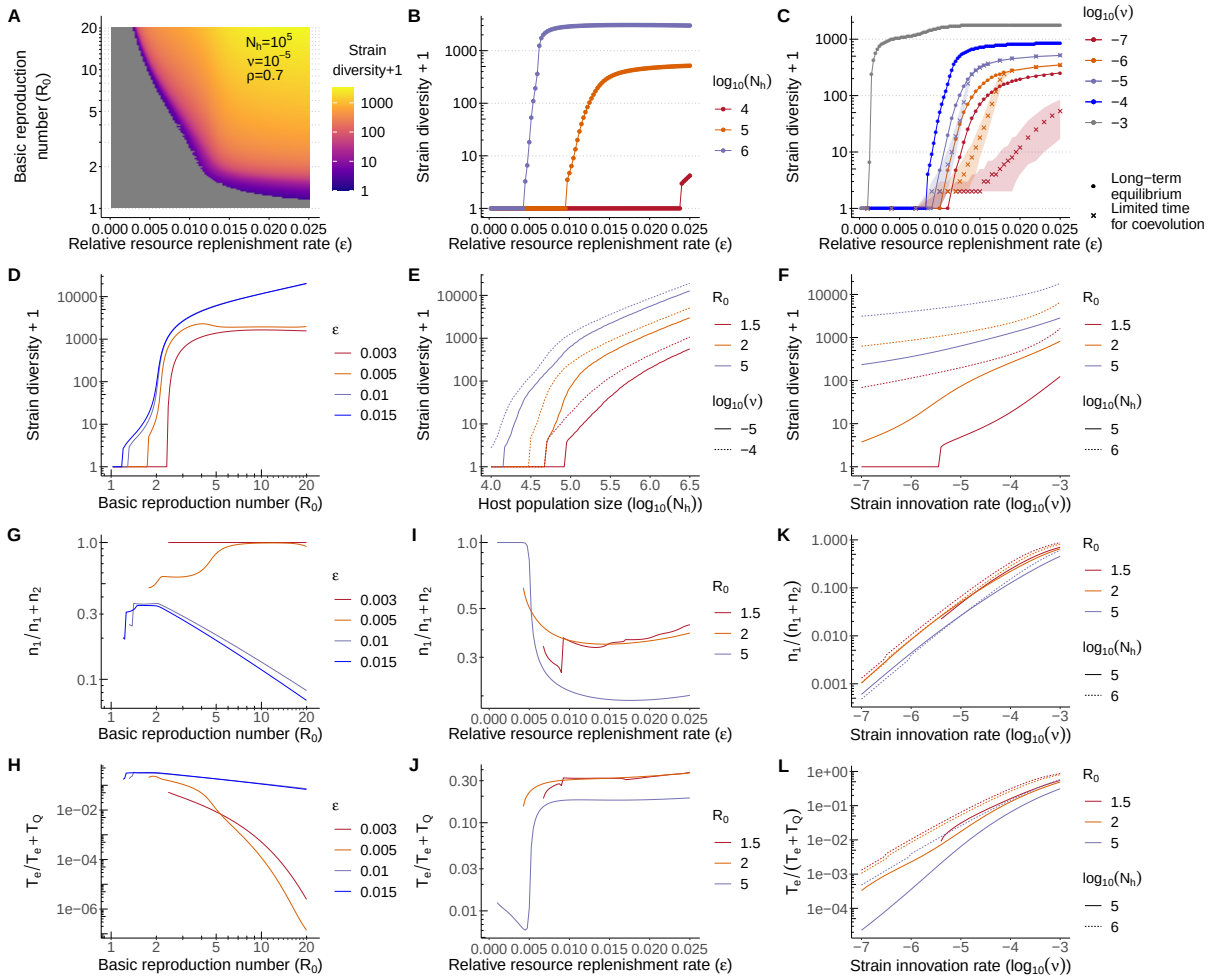

Figure 2: **Effects of key parameters on long-term strain diversity and regimes of strain dynamics predicted by MultiSEED. (A)** Heatmap of strain diversity showing the interplay of relative resource replenishment rate from hosts (ε) and the basic reproduction number (R_0), with all other parameters held constant. Grid cells with zero strain diversity are shown in gray. **(B-C)** Effects of ε on strain diversity, across different magnitudes of host population size N_h and strain innovation rate ν. In **C**, we examine the effect of limited time for strain diversification by running stochastic compartmental model simulations for 80,000 generations, with 1,000 replicates per ε value (for computational efficiency, only 100 replicates were run for $\varepsilon > 0.014$ when $\nu = 10^{-5}$). Crosses and shades represent the median and 5th-95th percentiles across replicates. **(D)** Effect of R_0 on strain diversity. Other parameters: $N_h = 10^6$, $\nu = 10^{-4}$, and $\rho = 0.5$, corresponding to the parameter range for influenza A in Table 1. **(E-F)** Effects of N_h and ν on strain diversity, with fixed parameter values $\varepsilon = 0.015$ and $\rho = 0.7$. **(G-L)** Parameter effects on the two criteria used to classify regimes of strain dynamics: Effects of R_0 **(G-H)** and ε **(I-J)**, using the same parameter settings as in **D**; Effects of ν **(K-L)**, with fixed parameter values $\varepsilon = 0.015$ and $\rho = 0.7$.

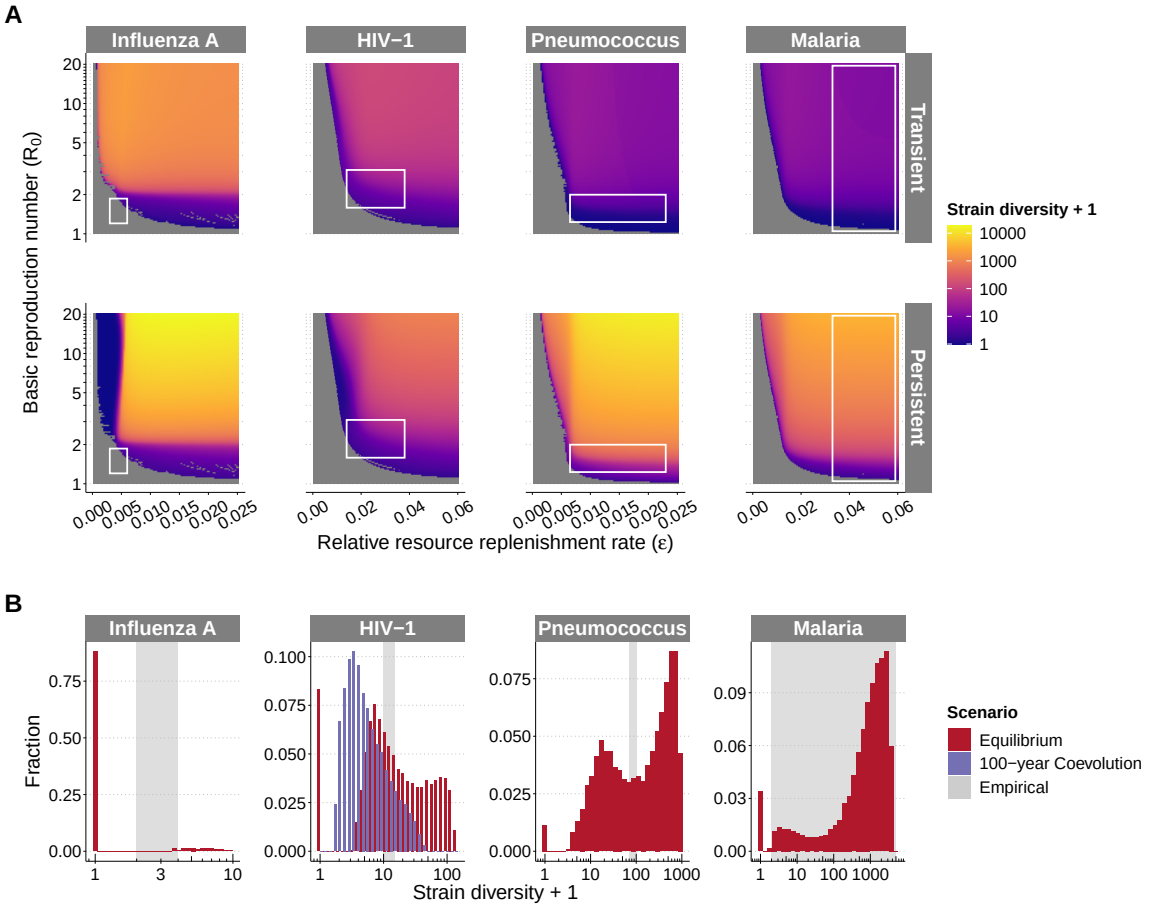

Figure 3: **Predicted strain diversity of four human diseases using MultiSEED. (A)** Predicted numbers of transient and persistent strains at long-term equilibrium with varying ε and R_0. Grid cells with zero strain diversity are shown in gray. White rectangles indicate the empirically estimated ranges of ε and R_0 for each disease (see Table 1 for details). **(B)** Comparison between predicted (colored) and empirical (gray) strain diversity. Histograms in red show the distribution of predicted total strain diversity within the empirical parameter ranges (white rectangles in A). For HIV-1, we further calculate the expected total strain diversity after 100 years of coevolution. Specifically, mean diversity was obtained from over 1,000 compartmental Markov simulations per grid cell and shown as purple histograms.

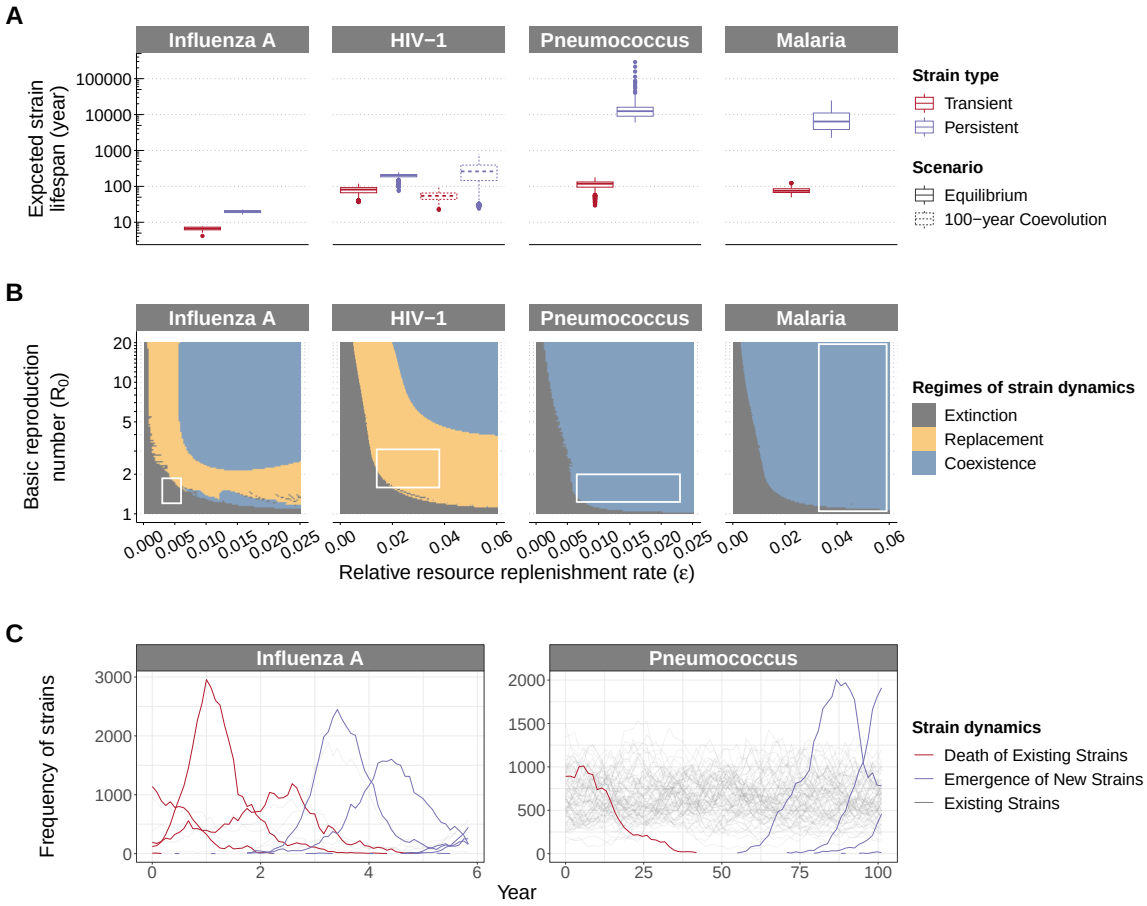

Figure 4: **Predicted strain lifespans and regimes of strain dynamics for four human diseases.** **(A)** Strain lifespan estimated by MultiSEED. The lifespan of a transient strain is defined as the mean duration of its initial epidemic (T_e; see SM). The lifespan of a persistent strain is calculated based on its probabilities of short-term burnout and long-term persistence, using the expression $q_k^{(r)}T_e + (1 - q_k^{(r)})(T_Q)$ (see SM section S1.5). Box plots summarize strain lifespans across all grid cells within the empirical parameter ranges, excluding cells with zero strain diversity. For HIV-1, dashed box plots represent estimates under a 100-year host-pathogen coevolution scenario (see Fig. 3B). **(B)** Regimes of strain dynamics predicted by MultiSEED. Extinction occurs when both $n_1 = n_2 = 0$ at equilibrium. Replacement and coexistence are distinguished using two ratios: the proportion of transient strains, $n_1/(n_1 + n_2)$, and the relative duration of the initial epidemic, $T_e/(T_e + T_Q)$. When both ratios are below 1/3, strains are considered to stably coexist; otherwise, they fall under the replacement regime. White rectangles denote empirical parameter ranges. **(C)** Representative strain dynamics for influenza A and pneumococcus using agent-based stochastic simulations, illustrating typical patterns of the replacement and coexistence regimes, respectively. See Fig. 1 legend for the parameter sets used in the simulations.

Supplementary Materials

Jiawei Liu[1], Qixin He[1*]
[1]Department of Biological Sciences, Purdue University,
915 Mitch Daniels Blvd, West Lafayette, IN 47907,USA

[*]To whom correspondence should be addressed; E-mail: heqixin@purdue.edu.

1 Material and Methods

In the following sections, we introduce the construction of the MultiSEED framework. We start by extending a single-strain SIRS model to a status-based n-strain SIRS model, followed by a detailed description of the system's stochastic dynamics at ecological and evolutionary time scales. We then introduce the numerical methods for obtaining long-term strain diversity. Results from the numerical solutions are verified with simulations using an agent-based stochastic model. Parameter ranges of interest are explored using the empirical ranges obtained from common infectious diseases.

1.1 Single-strain SIRS dynamics

For the SIRS model of a disease with vital dynamics, we use S, I, R to denote the number of susceptible, infected, and recovered people in the total population. Birth and death rates of the hosts are equal, such that the total host population size, $N_h = S(t) + I(t) + R(t)$, is assumed to be constant. Then the deterministic mean-field ODEs are,

$$\begin{cases} \dfrac{dS}{dt} &= \mu N_h - \mu S - \beta IS/N_h + \delta R \\ \dfrac{dI}{dt} &= \beta IS/N_h - \gamma I - \mu I \\ \dfrac{dR}{dt} &= \gamma I - \delta R - \mu R \end{cases} \tag{S1}$$

where μ is the *per capita* natural birth or death rate of the host population, β is the transmission rate, γ is the recovery rate from the infection, and δ is the immunity loss rate of the recovered class.

If we use X, Y, and Z to denote the fractions of susceptible, infected and recovered (i.e., $\dfrac{S}{N_h}$, $\dfrac{I}{N_h}$, and $\dfrac{R}{N_h}$), we can rewrite Eq. S1 to be,

$$\begin{cases} \dfrac{dX}{dt} &= \mu(1-X) - \beta XY + \delta Z \\ \dfrac{dY}{dt} &= (\beta X - \gamma - \mu)Y \\ \dfrac{dZ}{dt} &= \gamma Y - (\delta + \mu)Z \end{cases} \tag{S2}$$

Given that $\dfrac{dX}{dt} + \dfrac{dY}{dt} + \dfrac{dZ}{dt} = 0$ and $X + Y + Z = 1$ being constant, we may study the first two equations to describe the system,

$$\begin{cases} \dfrac{dX}{dt} &= \mu(1-X) - \beta XY + \delta(1-X-Y) \\ \dfrac{dY}{dt} &= (\beta X - \gamma - \mu)Y \end{cases} \tag{S3}$$

The trajectories of the deterministic SIRS model converge to a globally asymptotically stable equilibrium point, which is determined by the basic reproduction number, $R_0 = \dfrac{\beta}{\gamma + \mu}$. When $R_0 <= 1$, the system will be disease-free, whereas $R_0 > 1$, the system goes to an endemic equilibrium,

$$(\hat{x}\ ,\ \hat{y}) = (\frac{1}{R_0}\ ,\ \frac{\mu + \delta}{\mu + \gamma + \delta}(1 - \frac{1}{R_0})) \tag{S4}$$

If we reparameterize the system and rescale time using the expected duration of an individual infection, we define a new time variable $\tau = t/(\gamma + \mu)$, Eq S3 becomes,

$$\begin{cases} \dfrac{dX}{d\tau} &= \varepsilon_2(1-X) - R_0 XY + \varepsilon_1(1-X-Y) \ = \ \varepsilon(1-X) - R_0 XY - \varepsilon_1 Y \\ \dfrac{dY}{d\tau} &= (R_0 X - 1)Y \end{cases} \tag{S5}$$

where $\varepsilon_1 = \dfrac{\delta}{\gamma + \mu}$ represents the rescaled immune loss rate in the units of expected infectious period $(1/(\gamma+\mu))$, $\varepsilon_2 = \dfrac{\mu}{\gamma + \mu}$ represents the rescaled birth and death rates of hosts in the units of expected infectious period, and $\varepsilon = \varepsilon_1 + \varepsilon_2$.

1.2 Multi-strain SIRS dynamics

1.2.1 status-based n-strain SIRS model

Many infectious diseases consist of multiple co-circulating strains, and hosts may be infected by different strains either simultaneously or sequentially. In this study, we extend the classic single-strain SIRS framework to model the dynamics of multiple strains. We assume there are currently n equivalent strains in circulation, denoted as $H = \{1, 2, ..., n\}$, where all strains share identical epidemiological properties, including R_0 and immune loss rate (δ). For simplicity, we initially do not incorporate cross-immunity between strains, but we extend the model in the next section.

Unlike earlier multi-strain models that restrict hosts to infection by only one strain at a time (such as (*19*)), we allow for simultaneous co-infection by multiple strains, reflecting the common occurrence of mixed infections observed in empirical studies (*25*). In our model, strains compete within co-infected hosts by sharing transmission potential: if k different strains co-infect the same host, the transmissibility of each strain is reduced to β/k. We also assume that there is no trade-off between transmissibility and virulence, so the recovery rate γ of each strain within the host remains unchanged from that of single-strain infections.

We do not track the full infection history of individual hosts. Instead, we record their current immunity and infection status. This allows us to continue using Eq. S5 to describe the transmission dynamics of individual strains, with the exception that the transmission rate β and the corresponding R_0^* of strains must be adjusted to account for co-infection, where multiple strains share the same host. Let $x_i(\tau)$, $y_i(\tau)$, and $z_i(\tau)$ denote the proportions of hosts at time τ who are susceptible to, infected by, and recovered from strain i, respectively. By definition, these proportions satisfy $x_i(\tau)+y_i(\tau)+z_i(\tau) = 1$ for each strain i. However, since hosts can be co-infected with multiple strains, the total infected proportion across all strains, $y_T = \sum_{i=1}^{n} y_i$ can exceed 1. Assuming no host age structure and that co-infections are randomly distributed

among hosts, we define the overall disease prevalence (i.e., the probability that a host is infected with at least one strain) as $PR(\tau) = 1 - \prod_{i=1}^{n}(1 - y_i(\tau))$.

Assume that $\mathcal{F}_m^k$ represents the kth random subset of m strains out of $H \backslash i$, the total number of such subsets is $C(n-1, m)$. Then, the reduced transmissibility of strain i because of co-infection with other strains can be calculated as follows,

$$\begin{aligned}\beta_e^i(t) &= \sum_{m=0}^{n-1} (\text{fraction of } \beta|(m+1) \text{ infections}) \cdot P\,(\text{Host having } m \text{ infections in addition to strain } i) \\ &= \sum_{m=0}^{n-1} \frac{\beta}{m+1} \sum_{k=1}^{C(n-1,m)} \left(\prod_{j \subset \mathcal{F}_m^k} y_j(t) \prod_{l \not\subset \mathcal{F}_m^k} (1 - y_l(t)) \right)\end{aligned} \tag{S6}$$

Therefore, we can modify Eq. S5 to describe the dynamics of a particular strain within n equivalent circulating strains to be,

$$\begin{cases} \dfrac{dx_i}{d\tau} &= \varepsilon(1 - x_i) - R_0^{*i} x_i y_i - \varepsilon_1 y_i \\ \dfrac{dy_i}{d\tau} &= (R_0^{*i} x_i - 1) y_i \\ R_0^{*i}(\tau) &= R_0 \displaystyle\sum_{m=0}^{n-1} \frac{1}{m+1} \sum_{k=1}^{C(n-1,m)} \left(\prod_{j \subset \mathcal{F}_m^k} y_j(\tau) \prod_{l \not\subset \mathcal{F}_m^k} (1 - y_l(\tau)) \right) \end{cases} \tag{S7}$$

In quasi-stationary state, each strain is roughly of equal frequency so that $\hat{y_i} = \hat{y_T}/n$. We can then approximate Eq. S6 as follows,

$$\begin{aligned}\hat{R}_e(n, \hat{y_i}) &\approx R_0 \sum_{m=0}^{n-1} \frac{1}{m+1} \binom{n-1}{m} (\hat{y_i})^m (1 - \hat{y_i})^{n-1-m} \\ &= \frac{R_0}{n\hat{y_i}} \sum_{k=1}^{n} \binom{n}{k} (\hat{y_i})^k (1 - \hat{y_i})^{n-k} \\ &= \frac{R_0}{n\hat{y_i}} (1 - (1 - \hat{y_i})^n)\end{aligned} \tag{S8}$$

We can thus solve for the equilibrium values $(\hat{x}_i, \hat{y}_i)$ and $\hat{R}_e(n, \hat{y}_i)$ at the quasi-stationary states.

1.2.2 Incorporating generalized cross-immunity

In the previous section, we assumed that there was no cross-immunity among strains. However, in reality, prior infections often confer partial protection against subsequent heterologous infections. To capture this, we now consider an extended version of n-strain models that incorporates general cross-immunity. We adopt the definition of cross-immunity from (*108*) and generalize it to accommodate multiple strains. For tractability, we assume that hosts with any prior infection history gain a constant level of cross-protection to all strains in the form of reduced susceptibility. Hosts do not lose generalized immunity. While hosts may lose strain-specific immunity and re-enter the susceptible class for that particular strain, they still maintain the generalized protection, modeled as a reduced susceptibility. Accordingly, we modify Eq. S7 for n equivalent strains as follows:

$$\begin{cases} \dfrac{dx_\varnothing}{d\tau} &= \varepsilon_2(1-x_\varnothing) - x_\varnothing \Sigma_{i=1}^{n} R_0^{*i} y_i \\ \dfrac{dx_i}{d\tau} &= \varepsilon(1-x_i) - \varepsilon_1 y_i - R_0^{*i}\left(x_\varnothing + \rho(x_i - x_\varnothing)\right) y_i \\ \dfrac{dy_i}{d\tau} &= R_0^{*i}\left(x_\varnothing + \rho(x_i - x_\varnothing)\right) y_i - y_i \\ R_0^{*i}(\tau) &= R_0 \displaystyle\sum_{m=0}^{n-1} \frac{1}{m+1} \sum_{k=1}^{C(n-1,m)} \left(\prod_{j \subset \mathcal{F}_m^k} y_j(\tau) \prod_{l \not\subset \mathcal{F}_m^k} (1 - y_l(\tau)) \right) \end{cases} \tag{S9}$$

Here, the definitions of x_i, y_i, and R_0^{*i} remain consistent with those in Eq. S7. We specifically track the dynamics of $x_\varnothing$, the proportion of naïve hosts who have never been infected by any strain and are therefore fully susceptible to all strains, without any cross-protection. $x_\varnothing$ will increase only through host birth, and will decrease from host deaths and infection by any strain. The parameter ρ quantifies cross-susceptibility of non-naïve hosts who are still susceptible to

strain i, $x_{i\setminus\varnothing}$. These hosts include the ones infected by strains other than i, or those who have lost specific immunity to strain i. Thus, $x_{i\setminus\varnothing} = x_i - x_\varnothing$. A value of $\rho = 1$ corresponds to no cross-immunity, whereas $\rho = 0$ indicates complete cross-protection. We again use Eq. S8 and Eq. S9 to solve for the equilibrium values $(\hat{x}_\varnothing, \hat{x}_i, \hat{y}_i)$ at the quasi-stationary state.

1.2.3 Deterministic dynamics of a new-strain invasion into n existing strains at equilibrium

We have established the equilibrium dynamics of n strains. To explore the invasion potential by a novel strain, we extend our ODEs to include the dynamics of a single invading strain, ω, introduced into a parasite population where the n resident strains are at equilibrium. We first present the deterministic ODEs governing this invasion process. In the next sections, we incorporate stochastic dynamics to estimate the probability of successful invasion.

Let x_r and y_r represent the susceptible and infected compartments associated with each of the n resident strains, respectively, and x_ω and y_ω denote the corresponding compartments for the invading strain ω. The following system of ODEs describes the deterministic dynamics of this invasion scenario.

$$\begin{cases} \dfrac{dx_\varnothing}{d\tau} &= \varepsilon_2(1 - x_\varnothing) - x_\varnothing(R_0^{*\omega} y_\omega + nR_0^{*r} y_r) \\ \dfrac{dx_r}{d\tau} &= \varepsilon(1 - x_r) - \varepsilon_1 y_r - R_0^{*r}\left(x_\varnothing + \rho(x_r - x_\varnothing)\right) y_r \\ \dfrac{dy_r}{d\tau} &= R_0^{*r} y_r \left(x_\varnothing + \rho(x_r - x_\varnothing)\right) - y_r \\ \dfrac{dx_\omega}{d\tau} &= \varepsilon(1 - x_\omega) - \varepsilon_1 y_\omega - R_0^{*\omega}\left(x_\varnothing + \rho(x_\omega - x_\varnothing)\right) y_\omega \\ \dfrac{dy_\omega}{d\tau} &= R_0^{*\omega} y_\omega \left(x_\varnothing + \rho(x_\omega - x_\varnothing)\right) - y_\omega \end{cases} \tag{S10}$$

Here $x_\varnothing$ is still the fraction of susceptible hosts with no previous infection, while $R_0^{*\omega}$ and R_0^{*r} refer to the effective reproductive number of the invading strain and the resident strains, respectively. Similar to Eq. S8, the reduced transmissibility of each resident strain is due to

coinfection with other resident strains and/or the new invading strain. Thus,

$$
\begin{aligned}
R_0^{*r}(n,\tau) &\approx R_0 \sum_{i_1=0}^{n-1}\sum_{i_2=0}^{1} \frac{1}{i_1+i_2+1}\begin{pmatrix} n-1 \\ i_1 \end{pmatrix} y_r(\tau)^{i_1}(1-y_r(\tau))^{n-1-i_1}\begin{pmatrix} 1 \\ i_2 \end{pmatrix} y_\omega(\tau)^{i_2}(1-y_\omega(\tau))^{1-i_2} \\
&= R_0\,(1-y_\omega(\tau))\sum_{i_1=0}^{n-1}\frac{1}{i_1+1}\begin{pmatrix} n-1 \\ i_1 \end{pmatrix} y_r(\tau)^{i_1}(1-y_r(\tau))^{n-1-i_1} + \\
&\quad R_0 y_\omega(\tau)\sum_{i_1=0}^{n-1}\frac{1}{i_1+2}\begin{pmatrix} n-1 \\ i_1 \end{pmatrix} y_r(\tau)^{i_1}(1-y_r(\tau))^{n-1-i_1}
\end{aligned}
\tag{S11}
$$

The second term of Eq. S11 is equivalent to solving $E[\frac{1}{X+2}]$ given X is a random variable following a binomial distribution $B(n-1, y_r(\tau))$. Using generating functions, we get a closed form solution for $R_0^{*r}(n,\tau)$,

$$
\begin{aligned}
R_0^{*r}(n,\tau) &\approx R_0 \frac{(1-y_\omega(\tau))}{n y_r(\tau)}\left(1-(1-y_r(\tau))^n\right) + \\
&\quad R_0 \frac{y_\omega(\tau)}{y_r(\tau)^2}\left(\frac{1}{n+1} - \frac{(1-y_r(\tau))}{n} - \frac{(1-y_r(\tau))^{n+1}}{n+1} + \frac{(1-y_r(\tau))^{n+1}}{n}\right)
\end{aligned}
\tag{S12}
$$

The reduced transmissibility of the new invading strain, $R_0^{*\omega}$, on the other hand, is due to the competition with the n resident strains. It can be derived in a similar way to Eq. S8:

$$
\begin{aligned}
R_0^{*\omega}(n,\tau) &\approx R_0 \sum_{i=0}^{n}\frac{1}{i+1}\begin{pmatrix} n \\ i \end{pmatrix}(y_r(\tau))^i\,(1-y_r(\tau))^{n-i} \\
&= R_0 \frac{1}{(n+1)y_r(\tau)}\left(1-(1-y_r(\tau))^{(n+1)}\right)
\end{aligned}
\tag{S13}
$$

1.3 Continuous time stochastic processes at the ecological time scale: rates of strain emergence and strain dynamics

To predict the number of strains at the long-term equilibrium, the next step is to quantify the dynamics of strain emergence and extinction. A newly formed strain has to survive two phases

of bottlenecks to become a persistent strain. Strains are mostly prone to extinction during the initial fizzle phase, when a small number of infections may stochastically go extinct. Once the strain accumulates enough copies and reaches the epidemic stage, it does not go extinct until the susceptibles are quickly exhausted and the fraction of infected individuals again decreases sharply, a phenomenon termed epidemic burnout (see trends of light red lines between τ_1 and τ_2 in Fig. S1). Meanwhile, a new epidemic also affects the dynamics of n resident strains, as the fractions of resident strains, y_r, decrease, resulting in resident strain burnout (see trends of light blue lines between τ_3 and τ_4 in Fig. S1).

We build on the boundary layer framework for SIR models developed in (*31*) and (*32*) to estimate short-term extinction probabilities—specifically, the fizzle and burnout rates—following the introduction of a new strain in an n-strain SIRS model. Using the deterministic ODE trajectories described in Section 1.2.3, we approximate the epidemic phase until the prevalence of the invading strain ω and any resident strain r, denoted y_ω and y_r, fall below a threshold $\hat{y}_{n+1}$. This threshold defines a boundary layer below which stochastic effects dominate. Within this stochastic regime, we apply Kendall's q function (*109*), which models a time-inhomogeneous birth-death process, to estimate the extinction probabilities for strains whose prevalence has dropped below $\hat{y}_{n+1}$.

Following the initial epidemic wave, surviving strains may undergo additional cycles of resurgence and decline near $\hat{y}_{n+1}$. However, (*32*) showed that the probability of extinction after surviving the first trough is negligible compared to extinction during the initial wave. Therefore, we restrict our analysis to the extinction probabilities resulting from the first epidemic wave to approximate the short-term fizzle and burnout rates for both invading and resident strains.

According to Kendall (*109*), in a time-inhomogeneous birth-death process with per capita birth rate of $b(\tau)$ and death rate of $d(\tau)$, the probability of the dynamics starting from one individual is extinct by the time τ, $q(\tau)$ is,

$$q(\tau) = \left(1 + \frac{1}{\int_0^\tau e^{-\int_0^t [b(s)-d(s)]ds} d(t)dt}\right)^{-1} \tag{S14}$$

Fizzle stage. Assume that currently the system has n strains at the endemic stage. Recall the dynamics of y_ω is,

$$\frac{dy_\omega}{d\tau} = \left(\underbrace{R_0^{*\omega}\left(x_\varnothing + \rho(x_\omega - x_\varnothing)\right)}_{b(\tau)} - \underbrace{1}_{d(\tau)}\right) y_\omega \tag{S15}$$

The per capita birth rate of ω is $b(\tau) = R_0^{*\omega}\left(x_\varnothing + \rho(x_\omega - x_\varnothing)\right)$, and the per capita death rate is $d(\tau) = 1$. At the beginning of strain emergence (i.e., when τ is small), $x_w \approx 1$, $y_r \approx \hat{y_r}$, $x_\varnothing \approx \hat{x}_\varnothing$ at n-strain equilibrium. Thus $R_0^{*\omega}$ at $\tau = 0$ can be considered constant at the fizzle stage as well. Therefore, the probability of the new strain ω escaping the fizzle stage is,

$$\begin{aligned}
p_1^{(\omega)} &= 1 - q_1(\infty) \\
&= \left(1 + \int_0^\infty e^{-\int_0^t [b(s)-d(s)]ds} d(t)dt\right)^{-1} \\
&= \left(1 + \int_0^\infty e^{(1-b_0)t} dt\right)^{-1} = 1 - \frac{1}{b_0} \\
&= 1 - \frac{1}{R_0^{*\omega}(n,0)\left(\hat{x}_\varnothing + \rho(1-\hat{x}_\varnothing)\right)} \\
&= 1 - \frac{(n+1)\hat{y}_r}{R_0\left(\hat{x}_\varnothing + \rho(1-\hat{x}_\varnothing)\right)\left(1 - (1-\hat{y}_r)^{(n+1)}\right)}
\end{aligned} \tag{S16}$$

For the standard single strain SIR model, $b_0 = R_0$, and $p_1^{(\omega)}$ is equivalent to the classical expression for the establishment probability of a new disease $1 - (1/R_0)^k$, where k is $I(0)$, the initial infectious individuals (*34*).

Burnout phase. Once a newly formed strain survives the fizzle phase, it enters its epidemic stage and reaches a peak of infection. We use the numerical solutions of the deterministic ODEs

in Eq. S10 to approximate the trajectory of strain ω (Fig. S1). When the fraction of infected declines and re-enters the boundary layer (i.e., drops below the threshold $\hat{y}_{n+1}$), we again use Kendall's q (Eq. S14) to estimate burnout probability between τ_1 and τ_2 for the invading strain, ω, and between τ_3 and τ_4 for the resident strains r (Fig. S1).

Within the boundary layer, the per capita death rate remains constant at 1, while the per capita birth rate, $b(\tau)$ varies dynamically due to rapid changes in $x_\varnothing$, x_r,x_ω, $R_0^{*\omega}$, R_0^{*r}. We numerically solve the ODEs to obtain the deterministic time series for these variables and calculate the corresponding $b(\tau)$ using Eq. S12 and Eq. S13, which serve as inputs to Kendall's q function. The critical times τ_1, τ_2,τ_3 and τ_4 are determined by when the prevalence trajectories intersect with the boundary threshold $\hat{y}_{n+1}$ (Fig. S1). We then numerically integrate Kendall's q to approximate the per infection burnout probability for the invading and resident strains respectively. Upon entering the boundary layer, the strains have $k = N_h\hat{y}_{n+1}$ individuals. Thus the final burnout probability is given by $q_k = q_1^{N_h\hat{y}_{n+1}}$.

In the special case where no resident strains are present, the invading strain is the sole circulating strain. The ODE system simplifies to Eq. S17:

$$\begin{cases} \dfrac{dx_\varnothing}{d\tau} &= \varepsilon_2(1-x_\varnothing) - x_\varnothing R_0 y_\omega \\ \dfrac{dx_\omega}{d\tau} &= \varepsilon(1-x_\omega) - \varepsilon_1 y_\omega - R_0 y_\omega\left(x_\varnothing + \rho(x_\omega - x_\varnothing)\right) \\ \dfrac{dy_\omega}{d\tau} &= R_0 y_\omega\left(x_\varnothing + \rho(x_\omega - x_\varnothing)\right) - y_\omega \end{cases} \tag{S17}$$

We follow the same steps as mentioned above to calculate the fizzle and burnout probabilities of this new strain after its first epidemic. In this scenario, there is no risk of burnout for the non-existent established strains.

Another key quantity we aim to estimate is the mean duration of the initial epidemic, T_e, for invading strains that go extinct during the burnout phase. This metric helps quantify the sojourn time of newly introduced strains. As explained earlier, we use the deterministic processes to

approximate the epidemic phase from $\tau = 0$ to τ_1 (Fig. S1). Between τ_1 to τ_2, we apply Kendall's q-function to estimate the mean time to extinction, τ_e. To further characterize strain dynamics between τ_1 and τ_2, we first calculate the mean duration of strain presence—accounting for both extinct and surviving strains—using the survival probability function $q(\tau)$. Let $q_k(\tau)$ represent the probability that a strain with k initial copies has gone extinct by time τ. Then, the expected mean duration of presence across this interval is:

$$E(T_m) = \int_0^{\tau_2-\tau_1} (1 - q_k(t))\, dt, \tag{S18}$$

where $q_k(\tau) = q_1(\tau)^k = q_1(\tau)^{N_h \hat{y_n}}$. Note that this integral does not have a closed-form solution but can be approximated numerically using ODE solvers. For strains that survive the period, the mean duration is simply $\tau_2 - \tau_1$. Combining the contributions from both extinct and surviving strains, we obtain $E(T_m) = \tau_e \cdot q_k(\tau_2 - \tau_1) + (\tau_2 - \tau_1) \cdot (1 - q_k(\tau_2 - \tau_1))$. Solving for τ_e, we rearrange the expression:

$$\tau_e = \frac{E(T_m) - (\tau_2 - \tau_1)(1 - q_k(\tau_2 - \tau_1))}{q_k(\tau_2 - \tau_1)} \tag{S19}$$

Thus, the total expected duration of the first epidemic is given by,

$$T_e = \tau_1 + \tau_e \tag{S20}$$

1.4 Continuous time stochastic processes at the evolutionary time scale: time to extinction of an endemic strain

If an infectious disease strain survives the initial troughs and stays around the endemic equilibrium $(\hat{x}, \hat{y})$, then how long does it take for it to go extinct stochastically? Due to negative frequency-dependent selection, the prevalence of the strain will be maintained at the endemic level, appearing to be in a stable equilibrium before eventually becoming extinct. The sojourn time depends on $\hat{y}$ as well as the intensity of competition against other co-circulating

strains. We write out the Kolmogorov forward (aka Fokker-Planck) equation for the n-strain SIRS model and derive the joint quasi-stationary distributions of I_i and S_i at time τ, $p_{mn}(\tau) = P\left(S_i(\tau) = m, I_i(\tau) = n\right)$. We then use a bivariate Ornstein-Uhlenbeck process to approximate the distribution. The expected sojourn time to extinction is thus the inverse of the probability of having one infected host in the quasi-stationary state times the duration of an individual infection (Eq. 2.8 in (*35*)).

Table S1: Transition rates for the stochastic reparameterized single-strain SIRS model: Eq. S5

Event	Type of transition	Transition rate $g_j(x, y)$
New birth or immunity loss	$(S, I) \to (S + 1, I)$	$N_h g_1(x, y) = \varepsilon_1(N_h - S - I) + \varepsilon_2 N_h$
Death of susceptibles	$(S, I) \to (S - 1, I)$	$N_h g_2(x, y) = \varepsilon_2 S$
Death or recovery of infected	$(S, I) \to (S, I - 1)$	$N_h g_3(x, y) = I$
Infection of susceptibles	$(S, I) \to (S - 1, I + 1)$	$N_h g_4(x, y) = R_0 IS/N_h$

First, we write out all the transition rates for the stochastic version of the one-strain SIRS model. We can then calculate mean, variance, and covariance of $x(\tau)$ and $y(\tau)$ considering a bivariate diffusion approximation. The approach of approximating the mean sojourn time is then extended to the n-strain scenario.

The changes in the state variables x and y during the time interval from τ to $\tau + \Delta\tau$ are denoted Δx and Δy. We derive the mean and the covariance of Δx and Δy. The mean is,

$$\begin{aligned} E\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} &= \begin{pmatrix} g_1(x,y) - g_2(x,y) - g_4(x,y) \\ g_4(x,y) - g_3(x,y) \end{pmatrix} \Delta\tau + o(\Delta\tau) \\ &= \begin{pmatrix} \varepsilon(1-x) - R_0 xy - \varepsilon_1 y \\ (R_0 x - 1)y \end{pmatrix} \Delta\tau + o(\Delta\tau) \\ &= b(x,y)\Delta\tau + o(\Delta\tau) \end{aligned} \tag{S21}$$

The Jacobian matrix of the vector $b(x, y)$ with respect to x and y is denoted $B(x, y)$:

$$B(x,y) = \begin{pmatrix} \dfrac{\partial b(x,y)_1}{\partial x} & \dfrac{\partial b(x,y)_1}{\partial y} \\ \dfrac{\partial b(x,y)_2}{\partial x} & \dfrac{\partial b(x,y)_2}{\partial y} \end{pmatrix} = \begin{pmatrix} -\varepsilon - R_0 y & -R_0 x - \varepsilon_1 \\ R_0 y & R_0 x - 1 \end{pmatrix} \tag{S22}$$

When evaluated at the endemic equilibrium point $(\hat{x}, \hat{y})$, we get,

$$B(\hat{x}, \hat{y}) = \begin{pmatrix} -\varepsilon - \frac{\varepsilon(R_0-1)}{\varepsilon_1+1} & -(1+\varepsilon_1) \\ \frac{\varepsilon(R_0-1)}{1+\varepsilon_1} & 0 \end{pmatrix}. \tag{S23}$$

We then determine the covariance matrix of the vector of changes in the state variables during the time interval $(\tau, \tau + \Delta\tau)$,

$$\begin{aligned} S(x,y) &= \begin{pmatrix} 1^2\left(g_1(x,y)\right) + (-1)^2\left(g_2(x,y) + g_4(x,y)\right) & 1(-1)g_4(x,y) \\ 1(-1)g_4(x,y) & 1^2\left(g_4(x,y)\right) + (-1)^2\left(g_3(x,y)\right) \end{pmatrix} \\ &= \begin{pmatrix} R_0 xy - \varepsilon_1(y+2x) + \varepsilon(x+1) & -R_0 xy \\ -R_0 xy & R_0 xy + y \end{pmatrix} \end{aligned} \tag{S24}$$

When evaluated at the endemic equilibrium point $(\hat{x}, \hat{y})$, we get,

$$S(\hat{x}, \hat{y}) = \frac{\varepsilon}{R_0(1+\varepsilon_1)} \begin{pmatrix} 2(R_0 - \varepsilon_1(1-\varepsilon_2)/\varepsilon) & -(R_0-1) \\ -(R_0-1) & 2(R_0-1) \end{pmatrix} \tag{S25}$$

For large N_h, the process $N_h^{1/2}\{\mathbf{x}(t) - \hat{\mathbf{x}}\}$ is approximated by a bivariate Orstein-Uhlenback process with local drift matrix $B(\hat{x}, \hat{y})$ and local covariance matrix $S(\hat{x}, \hat{y})$. Its stationary distribution approximates the quasi-stationary distribution, which is approximately a bivariate normal distribution with mean 0 and covariance matrix Σ, where $B\Sigma + \Sigma B^T = -S$ (*36*). Σ is thus,

$$\Sigma = \begin{pmatrix} \dfrac{1}{R_0} + \dfrac{1+\varepsilon_1}{\varepsilon R_0(R_0+\varepsilon_1)} & -\dfrac{1}{R_0} \\ -\dfrac{1}{R_0} & \dfrac{(R_0-1)(1+\varepsilon_1) + \varepsilon(R_0+\varepsilon_1)^2}{(\varepsilon_1+1)^2 R_0(\varepsilon_1+R_0)} \end{pmatrix} \tag{S26}$$

Therefore, we can approximate the marginal distribution of infected individuals I with the normal distribution, $N\left(u_I, \sigma_I\right) = (\hat{y}N_h, \sqrt{\Sigma_{yy}N_h})$. Therefore,

$$u_I = N_h \frac{\varepsilon(R_0 - 1)}{(1+\varepsilon_1)R_0} \tag{S27}$$

$$\sigma_I = \sqrt{N_h \frac{(R_0-1)(1+\varepsilon_1) + \varepsilon(R_0+\varepsilon_1)^2}{(\varepsilon_1+1)^2 R_0(\varepsilon_1+R_0)}} \tag{S28}$$

When ε is small, that is, the immune loss and birth rates are much smaller than the recovery rate of the disease,

$$\tilde{\sigma}_I = \frac{\sqrt{N_h(R_0-1)}}{R_0} \tag{S29}$$

We use $f_{\cdot 1}$ to denote the conditional probability of $I = 1$ given I being positive. That is, $f_{\cdot 1} = \frac{p_{\cdot 1}}{1-p_{\cdot 0}} = P\left(I = 1 | I \neq 0\right)$. $q_{\cdot 1}$ can thus be approximated using the probability density function of normal distribution rescaled by truncating the cumulative density at 0.5,

$$f_{\cdot 1} \approx \frac{1}{\sigma_I} \frac{\varphi\left((1-u_I)/\sigma_I\right)}{1 - \Phi\left((0.5-u_I)/\sigma_I\right)} = \frac{1}{\sigma_I} \frac{\varphi\left((1-u_I)/\sigma_I\right)}{\Phi\left((u_I-0.5)/\sigma_I\right)} \tag{S30}$$

Here φ and Φ are the PDF and CDF of the standard normal distribution. Then using Eq. 2.8 in (*35*), we get the approximation for the mean time to long-term stochastic extinction in the unit of the duration of infection,

$$E(T_Q) \approx \frac{1}{f_{\cdot 1}} = \frac{\sigma_I\ \Phi\left((u_I-0.5)/\sigma_I\right)}{\varphi\left((1-u_I)/\sigma_I\right)} \tag{S31}$$

We extend the derivation of T_Q of one strain SIRS model to the n-strain model by modifying Eq. S27 and Eq. S28. Specifically, R_0 in these two equations needs to be replaced by $\hat{R}_0^{*i}\left(\hat{x}_\varnothing/\hat{x}_i + \rho(1-\hat{x}_\varnothing/\hat{x}_i)\right)$ to reflect the competition among strains and the reduced susceptibility from generalized immunity.

1.5 Long-term strain diversity at equilibrium

In previous sections, we described the ecological and evolutionary dynamics of strains, from their emergence to extinction. Strains accumulate through the continual invasion of new mutants and are lost either during early burnout phases or through long-term stochastic extinction. To quantify strain diversity at equilibrium, we model two categories of strains: transient strains (n_1) that recently emerged and have survived the initial fizzle phase but have not yet completed their first epidemic wave, and persistent strains (n_2) that have survived at least one epidemic wave and whose long-term persistence is governed by a stochastic extinction timescale T_Q. Consequently, total strain diversity is $n = n_1 + n_2$.

Birth and Death Dynamics of Transient Strains. Let N_h be the host population size and $n = n_1 + n_2$ be the total number of strains at time τ. Then the overall prevalence is given by $1 - (1 - \hat{y})^n$. Assuming a strain innovation rate ν, the birth rate of transient strains is $N_h \left(1 - (1 - \hat{y})^n\right) \nu p_1^{(\omega)}$, which represents the rate at which new strains are introduced and survive the initial fizzle phase. For the existing transient strains in the system, we will quantify their durations of the first epidemic to calculate the exit rate of the transient class, $\frac{n_1}{T_e}$, where T_e is the mean duration of the first epidemic, as calculated by Eq. S20.

Transition from transient to persistent strains and Extinction of Persistent Strains. Transient strains become persistent at the rate of $\frac{n_1}{T_e}(1 - q_k^{(\omega)})$ where $q_k^{(\omega)}$ is the probability that the strain goes extinct during the epidemic burnout phase. Persistent strains are lost through two processes at two timescales: 1) burnout during new strain invasions, with probability $q_k^{(r)}$ over the interval T_e, 2) Long-term stochastic extinction, over a time window $T_e + E(T_Q)$, with probability $1 - q_k^{(r)}$. The combined death rate of persistent strains is a harmonic mean of the rates of these two processes.

ODE System for Strain Dynamics. We capture these dynamics with the following ODEs:

$$
\begin{cases}
\dfrac{dn_1}{d\tau} \approx N_p \nu p_1^{(\omega)} - \dfrac{n_1}{T_e} \\
\dfrac{dn_2}{d\tau} \approx \dfrac{n_1}{T_e}(1 - q_k^{(\omega)}) - \dfrac{n_2}{q_k^{(r)} T_e + (1 - q_k^{(r)})(E(T_Q))} \\
N_p = N_h \left(1 - (1 - \hat{y}_i)^{n1+n2}\right)
\end{cases}
\tag{S32}
$$

Transient strains (n_1) contribute to persistent strains (n_2) only if they survive the initial epidemic. The total number of strains at equilibrium is solved by setting $\frac{dn_1}{d\tau} = 0$ and $\frac{dn_2}{d\tau} = 0$. To compute all quantities in the ODEs, we compute equilibrium values of $(\hat{x}_\varnothing, \hat{x}_i, \hat{y}_i, \hat{R}_e^i)$ for systems with n and $n-1$ strains using Eq. S9. We then numerically evaluate strain invasion trajectories using Eq. S10 to compute the fizzle probability ($p_1^{(\omega)}$), the burnout probabilities ($q_k^{(\omega)}, q_k^{(r)}$), and the mean epidemic duration (T_e). Note that when evaluating birth rates of n_1, $p_1^{(\omega)}$ is calculated for one invading strain against n residents. In contrast, T_e, $q_k^{(\omega)}$, and $q_k^{(r)}$ are based on the behavior of a single strain invading a background of $n-1$ strains for the exit rates of existing transient strains.

Numerical solutions. We use the `nleqslv` function in the `R` package `nleqslv` (*110*) to solve for the equilibrium values $(\hat{n}_1, \hat{n}_2)$, applying the Broyden method (*111*) and trust-region strategies (*112*). Providing the initial conditions of a guess on (n_1, n_2), together with the parameter array (N_h, R_0, ε, ε_1, ν, ρ), we aim to solve for the pair $(\hat{n}_1, \hat{n}_2)$ at the stable equilibrium, with where both eigenvalues from the Jacobian matrix have negative real parts. To ensure stability and accuracy, we test multiple initial conditions and accept solutions with the minimum sum of absolute values $|\frac{dn_1}{d\tau}| + |\frac{dn_2}{d\tau}|$, at least satisfying $|log_{10}(N_p \nu p_1^{(\omega)}) - log_{10}(\frac{n_1}{T_e})| < 10^{-2}$ and $|log_{10}(\frac{n_1}{T_e}(1 - q_k^{(\omega)}) - log_{10}(\frac{n_2}{q_k^{(r)} T_e + (1 - q_k^{(r)})(T_e + E(T_Q))})| < 10^{-2}$. Only when We prioritize solutions where the Jacobian matrix has negative eigenvalues (indicating local stability). If no stable solution is found, we return the one with the minimum residual $|\frac{dn_1}{d\tau}| + |\frac{dn_2}{d\tau}|$. If no solution meets the tolerance criteria, we conclude that no strains can persist, and set $n = 0$.

1.6 Stochastic compartmental model of strain dynamics using Markov chain methods

In addition to deriving the equilibrium value of $\hat{n}$ using Eq. S32, we also formulate a stochastic version of the model to simulate the temporal trajectories of strain dynamics using a continuous-time Markov process. This approach is based on four types of events derived from the deterministic equation: the birth and death of both transient (n_1) and persistent (n_2) strains. The rates of these events are as follows: emergence of new transient strains: $r_{1,b} = N_p \nu p_1^{(\omega)}$, burnout of transient strains: $r_{1,d} = \frac{n_1}{T_e} q_k^{(\omega)}$, transition of transient to persistent strains: $r_{2,b} = \frac{n_1}{T_e}(1 - q_k^{(\omega)})$, and the extinction of persistent strains: $r_{2,d} = \frac{n_2}{q_k^{(r)} T_e + (1-q_k^{(r)}) E(T_Q)}$. Given an initial state defined by n_1 and n_2, the Markov process proceeds by generating events based on these rates. The time to the next event is sampled from an exponential distribution with rate $r_t = r_{1,b} + r_{1,d} + r_{2,b} + r_{2,d}$. The type of event is then selected probabilistically, weighted by its relative rate. After each event, the state variables n_1 and n_2 are updated accordingly, along with the simulation clock. The process terminates either when a predefined time limit is reached or when all strains go extinct (i.e., $n_1 = n_2 = 0$). For the Markov simulations shown in Figs. 2–4 of the main text, we initialized the system with one persistent strain ($n_1 = 0, n_2 = 1$) and ran 1,000 independent simulations for each parameter set to estimate the expected strain diversity over the specified co-evolutionary time period. This Markov-based stochastic compartmental model can similarly be used to track the evolution of strain diversity over time from arbitrary initial conditions.

1.7 Agent-based model of multi-strain SIRS dynamics

We developed an agent-based stochastic model (ABM) to simulate the eco-evolutionary dynamics of multi-strain transmissions (Github link). Results from the stochastic simulations were used to verify our numerical approximations described in previous sections. The model tracks each infection within each host explicitly. Each infection is associated with a specific

strain. Multiple infections of heterogeneous strains within the same host are allowed, and the transmissibility (β_e) per infection will be reduced to $1/m$ if the host has m concurrent infections. After recovering from an infection by strain i ($1/\gamma$), the host gains specific immunity against strain i. This specific immune memory is lost at the rate of δ. Hosts who have been infected or have cleared infections gain a general protection against the disease, acquiring a reduced susceptibility towards future infections from heterogeneous strains (ρ). The rate of newborns in the population and the death rates of hosts are kept constant (μ). Infections do not cause host deaths. New strains are innovated from existing infections during their generation time, which includes the incubation period and infectious period. Note that ν is in the unit of $\tau = t/(\gamma + \mu)$, so the strain innovation rate in the simulation is scaled to be $\nu * (\gamma + \mu)$ per infected host. The copy number of each strain is tracked throughout the simulation. Origin and death times of strains are also recorded. Simulations are initiated with N_h naïve hosts, characterized by an exponentially distributed age profile with a mean life expectancy of $1/\mu$. When a host reaches its life expectancy and is removed from the simulation, a new naïve host is born. Strain frequencies, host prevalence, and immune profiles are sampled every 30 days from the transmission dynamics. Simulations are stopped when host prevalence and strain diversity have reached a steady state. Simulations may also terminate if the disease is cleared and there remain no infections in the system.

The ABM was developed to be as close as possible to the equation-based multi-strain model. However, the age structure of disease prevalence and immune profile, which naturally emerges from the ABM, is not accounted for in the ODE systems. The discrepancy will generate slight differences between the results from ABM and the numerical methods described above.

1.8 Empirical ranges of parameters in common infectious diseases

In this section, we present empirical estimates for the key parameters used in our model for several well-studied infectious diseases. The host population size N_h is estimated based on typical host behavior and the transmission scale of each pathogen. For each disease, we estimate values from the literature for the following parameters: the strain innovation rate (ν), cross-immunity susceptibility (ρ), the basic reproduction number (R_0), and parameters used to calculate the relative resource replenishment rate ε. As introduced in Section 1.1, we define $\varepsilon = \dfrac{\delta + \mu}{\gamma + \mu}$, where δ is the immunity loss rate, γ is the recovery rate, and $\mu = 0.02\,\text{year}^{-1}$ is the host birth/death rate, assuming an average human lifespan of 50 years. When estimating γ, we account for the full generation time of the infection, which includes both incubation and infectious periods, as consistent with generation interval of the SEIR model framework (see Eq. 3.5 in (*113*)). For the strain innovation rate ν, we estimate it by calculating the observed genetic divergence between known strains and dividing by the expected time to accumulate such divergence, based on known mutation or recombination rates. Lastly, we investigate the origin time of the disease to estimate the length of its co-evolutionary history with the host. This period is critical for determining whether the currently observed strain diversity reflects a long-term evolutionary equilibrium.

1.8.1 Influenza A

Influenza A is highly transmissible in large populations. It has been suggested that large host populations enable the antigenic drift in influenza A and thus the lineage turnover (*37*). Therefore, We assume a relatively large host population size of $N_h = 10^6$. The basic reproduction number R_0 of influenza A is often estimated to be between 1.2 and 1.85, excluding the estimates for the initial 1918 epidemic and in confined settings, where extreme values were reported (*45*). To estimate the strain innovation rate, we focus on the HA1 receptor-binding domain, which

consists of 329 amino acids (*38*). The per-nucleotide mutation rate is approximately 2×10^{-5} per cellular infection (*39*), with roughly two viral replication cycles per day (*40*), yielding a daily rate of 4×10^{-5}. Given that the incubation and infectious periods each span 2–4 days (*41–43*), we estimate a generation time of 7 days, resulting in a per-generation mutation rate of $4 \times 10^{-5} \times 7$ per nucleotide. Since strains with deleterious mutations are subject to strong purifying selection, the strain innovation speed should be scaled by the fraction of beneficial mutations, which is estimated to be $5/57 = 0.088$ according to (*39*). The mean pairwise amino acid divergence among influenza A strains is about 13 residues (*44*), suggesting that a new strain requires a 3.95% change in amino acid sequence $(13/329)$. Therefore, the strain innovation rate is $\nu = \frac{1}{0.0395 \cdot L/(4 \times 10^{-5} \cdot 7 \cdot L \cdot 0.088)} \approx 6.2 \times 10^{-4}$, where L is the nucleotide length of the HA1 domain. We thus take the magnitude of 10^{-4} as the strain innovation rates per infection lifespan. We use the half-life of immune memory to calculate the strain-specific immune loss rate, assuming immunity over time follows an exponential decay. Protection against influenza A strains wanes by half in 3.5 to 7 years (*23*), which represents the median duration of immune memory. Therefore, we parameterize the strain-specific immunity to last from 3.5 to 7 years. Given these values, the replenishment ratio ε ranges from approximately 0.003 to 0.006 ($\varepsilon \approx \frac{1/3.5+0.02}{1/(7/365)+0.02} \approx 0.006, \quad \varepsilon \approx \frac{1/7+0.02}{1/(7/365)+0.02} \approx 0.003$). For cross-immunity, vaccine efficacy studies report that both live-attenuated and trivalent inactivated vaccines offer approximately 50% protection against mismatched strains (*46*). We therefore assume cross-susceptibility $\rho = 0.5$, indicating that prior infection reduces susceptibility to heterologous strains by half. Finally, the most recent common ancestor of influenza A hemagglutinin dates to around 1056 AD (*47*), suggesting that the disease has been co-evolving with humans for around 1000 years.

1.8.2 Pneumococcus (*Streptococcus pneumoniae*)

The second example is the bacterial pathogen *Streptococcus pneumoniae*, which has more than 90 different serotypes detected based on the composition of its capsule polysaccharide (*6*). Since bacteria can be easily transmitted in large urban populations at the magnitude of one million (*64*), we set $N_h = 10^6$. The basic reproduction number of *Streptococcus pneumoniae* has mostly been estimated to be no greater than 2 (*74, 75, 114*). In one specific study covering daycare in multiple countries, the basic reproduction number varies from 1.23 to 4.77 (*76*). However, the focal populations only include young children, and the ones with $R_0 > 2$ have quite small sample sizes, and the only population with sample size larger than 1000 has $R_0 = 1.23$. Therefore, we assume the range of R_0 for this bacterial pathogen is $1.23 - 2$. The carriers of this bacteria in their nose can spread the infection without showing symptoms, and the carriage duration may vary from 6 weeks to almost 5 months, according to one study (*65*), with a median of 132 days reported in another study (*66*). Thus, we assume that the generation period is approximately 120 days, equivalent to three generations of infections per year. For pneumococcus, the main source of genetic changes is recombination rather than spontaneous point mutations (*115, 116*), we will use the recombination rates to estimate ν. The capsule polysaccharide locus (cps), whose genotypes define serotypes, recombines at a rate of approximately 0.005 per year (*67*). The length of cps is 10-30 kilobases (*68*), so that the recombination rate is $0.005/20000 = 2.5 \times 10^{-7}$ per base per year when assuming the length is 20kb. Accordingly, the number of recombination events in a generation period of 120 days over the antigen coding region is $2.5 \times 10^{-7} \times L/365 \times 120$, where L is the length of the antigen coding region. The genetic difference between cps serotypes can be as small as 3% (*69, 70*). Therefore, the number of events required for forming a new strain is around $0.03 \times L$, suggesting that $\nu = \dfrac{1}{0.03 \times L/(2.5 \times 10^{-7} \times L/365 \times 120)} \approx 2.7 \times 10^{-6}$. Thus, we use 10^{-6} as the magnitude of strain innovation rate for pneumococcus. Serotype-specific immune memory wanes

with age, and older individuals are a high-risk group for reinfection; however, they still retain certain levels of B cells, sometimes even into their 100s (*71, 72*). Therefore, the lower bound of ε is estimated as $\frac{0.02}{1/(120/365)+0.02} \approx 0.0065$ assuming lifelong immunity. Another study indicates that the IgG antibody level can drop from 78% to 58% within 6 years after vaccination in elderly individuals (*73*). If we assume an exponential decay of immune memory over these 6 years (18 generations as there are 3 generations per year), we can calculate the decay rate by solving the equation $58 = 78 \times e^{-18\lambda}$. Since $\lambda \approx 0.016$, the mean lifetime of antibodies can be approximated as $1/\lambda \approx 61$ generations in this scenario, which is roughly 20 years. Thus, we estimate the upper bound of ε assuming the waning immunity over 20 year, so that $\varepsilon = \frac{0.05+0.02}{1/(120/365)+0.02} \approx 0.023$. Studies on different conjugate vaccines have commonly indicated that the efficacy against all pneumococcal pneumonia is around 30% (*77–79*),while serotype-independent protection in children is around 40% (*80*). These results suggest that we can roughly estimate that $\rho = 0.7$. The common ancestors of *Streptococcus pneumoniae* and its close relatives date back as far as a million years ago (*70*). It had a tiny effective population size until the recent expansion of the human population in the last 120,000 years. Therefore, we assume that the circulating time of *Streptococcus pneumoniae* in human populations is around 100,000 years.

1.8.3 HIV-1

HIV-1 is transmitted through body fluids, indicating that its spread is highly dependent on individual host behaviors and is typically constrained to small regional populations. Such diseases can transmit within a population with around 10,000 hosts (*48*). Accordingly, we set the host population size to $N_h = 10^4$. The infectious period of HIV-1 is approximately one year (*49*). Both the recombination and mutation rates of HIV-1 fall within the range of 1.4×10^{-5} to 2×10^{-4} per nucleotide per viral replication cycle (*50–53*). Given the short half-life of infected

cells (i.e., about two days (*54*)), we estimate that the infection undergoes approximately 90 replication cycles per year. The key antigenic region of the HIV-1 genome is the Gag domain, which spans roughly 1500 nucleotides. One study suggests that 5% of HIV-1 mutations are beneficial (*55*), while another reports that subtypes differ by 17% in amino acid sequence within the Gag region (*56*). Assuming that 70-75% single-nucleotide mutations are non-synonymous, 17% amino acid changes map to around 20% nucleotide divergence among strains. We estimate the strain innovation rate ν as follows: assuming $0.05 \times 1 \times 10^{-4} \times 1500 \div 4 \times 365$ beneficial recombination events per generation and 0.2×1500 genetic changes are required to produce a new strain, $\nu = \frac{1}{(0.2\times1500)/(0.05\times1\times10^{-4}\times1500/4\times365)} \approx 0.002$. This places the magnitude of ν around 10^{-3} per generation period. Estimating the strain-specific immune loss rate δ is challenging. In untreated individuals, antibody levels tend to stabilize long-term (*57*), and superinfections typically occur only when encountering genetically distinct strains (*58*). Thus, we assume $\delta \approx 0$, representing nearly lifelong strain-specific immunity (*59*). Using this assumption, we estimate the bounds of the replenishment rate ε under different host mortality rates: for a 70-year lifespan ($\mu = 1/70$), $\varepsilon = \frac{1/70}{1+1/70} \approx 0.014$, whereas for a 25-year lifespan ($\mu = 1/25$), $\varepsilon = \frac{1/25}{1+1/25} \approx 0.038$. The basic reproduction number R_0 for HIV-1 is estimated between 1 and 3, with a 95% HPD interval of [1.6,3.1] when assuming general epidemiological parameters in ten subepidemics from one study (*60*). Regarding cross-immunity, a study reports a hazard ratio of 0.47 for superinfection compared to initial infections (*61*), suggesting partial protection. We therefore assume $\rho = 0.5$. Regarding the co-evolutionary history, HIV-1 has been present in humans for roughly 100 years. The earliest known human case was identified in 1959, while the most recent common ancestor of the HIV-1 M group is estimated to date back to around 1930 (*62, 63*).

1.8.4 Malaria

Since falciparum malaria transmission is mostly endemic, and we set the resource host population size N_h to 10^5, according to the reported population size in small to intermediate cities in Africa (*81*). The generation period of malaria varies significantly in treated and untreated patients, with a recovery time range of approximately 20 to 200 days (*82–84*). Here, we use 150 days, based on the duration of infection in naïve hosts. The antigenic diversity of *var* genes in the malaria parasite is largely generated through ectopic recombination (*117*); therefore, we use the recombination rate (around 1.8×10^{-7} per gene per day (*7*)) to estimate the innovation rate of new strains. The pairwise type sharing of *var* repertoires between malaria parasites tends to be kept at very low levels (mostly below 20%) to favor the persistence of parasites (*85*); thus, we assume that completely different *var* gene repertoires are required to form a new strain. In one generation of infection, assuming there are 60 *var* genes on one strain (*117*), we expect $n_r = 1.8 \times 10^{-7} \times 150 \times 60$ recombination events, and the number of generations for recombination to occur on all genes is $60/n_r$. The strain innovation rate is calculated as follows, $\nu = \dfrac{1}{60/n_r} = \dfrac{1}{1/(1.8 \times 10^{-7} \times 150)} = 2.7 \times 10^{-5}$ per generation. Thus, the magnitude of ν for malaria is around 10^{-5}. The immune memory to one antigen can last for at least 3 years, and sometimes more than 8 years or even 16 years (*20, 86–88*). We assume that the immunity to a strain will be lost to enable another infection by itself when the relatively long-lasting immune memory of antigens is lost, so here if we have $\delta = 1/8$ then $\varepsilon = \dfrac{1/8 + 0.02}{1/(150/365) + 0.02} \approx 0.059$, and $\varepsilon \approx 0.0336$ when $\delta = 1/16$. The basic reproduction number for malaria is difficult to estimate, as it varies widely from 1 to 3000 (*89*), because malaria is a vector-borne disease, and transmissibility depends on the host-vector behavior and density ratio. For cross-immunity among strains, the vaccine efficacy study on malaria strains indicated that the efficacy against unmatched strains is about 0.334 (*90*), suggesting that $\rho \approx 0.7$. The malaria parasite *Plasmodium falciparum* was transmitted to humans from apes, and genetic diversity has accumulated

within the past 10,000 years (*91, 92*). Here, we use 10,000 years as the co-evolutionary length between malaria parasites and humans.

2 Supplemental Figures

3 Supplemental Text

3.1 Negligible effects of the ratio of ε_1 to ε

In our main text, when we explore the determinants of strain diversity, we regard ε as a single parameter, rather than studying its two components, ε_1 and ε_2, separately. However, as indicated in the ODE systems S7 and S9, sometimes ε_1 and ε_2 are involved separately in the compartmental model. Therefore, when $\varepsilon = \varepsilon_1 + \varepsilon_2$ is constant, it is possible that the relative difference between ε_1 and ε_2 can have some influence on the strain diversity, especially when general protection from cross-immunity is considered (ε_2 is in equation set S9 but not in S7). However, such influence is usually negligible, as we will show here in this section.

When the ratio varies from 0.1 to 0.9, the effect of this ratio is generally negligible when the values of ε and R_0 are intermediate or high (Fig. S7). This ratio mainly affects strain diversity under small values of ε or R_0 near the persistence threshold. The lower this ratio is, which represents a higher replenishment from host births, the smaller threshold value of ε or R_0 is required for the accumulation of strain diversity. This is because the replenishment of susceptible hosts (ε_2) is more beneficial for the survival of new pathogen strains as it supplies naïve hosts, compared to the waning strain-specific immunity (ε_1) that supplies hosts with protection from generalized immunity. Overall, the differences in strain diversity under different $\varepsilon_1/\varepsilon$ ratios are generally minor and less than one magnitude. Therefore, we simplify ε_1 and ε_2 as one parameter ε for factor exploration in the main text.

3.2 The number of possible equilibria

Sometimes, multiple equilibria can exist in the transmission system with positive feedback between genetic diversity and force of infection (*118*). When we obtain numerical approximations of strain diversity using MultiSEED, we observe scenarios with two possible equilibrium states, where the alternative equilibrium often comes with $\hat{n}_1 < 1$ and $\hat{n}_2 < 10$, which probably represent the unstable equilibrium state that occurs when strains are first accumulating. We systematically examined the parameter space to identify regions where multiple equilibria may be found. We focus on the boundary region of persistence threshold set by R_0 and ε, and fixed the combinations of N_h, ν and ρ for the four focal human diseases. We use several initial conditions of different magnitudes in finding the equilibrium values of $(\hat{n}_1, \hat{n}_2)$. As Fig. S8 shows, multiple equilibrium states exist close to the persistence threshold with low ε. These low equilibria often have one or both values of n_1 and n_2 less than 1, and sometimes one to two magnitudes smaller than those of the corresponding high equilibria. For influenza A and HIV-1, which have a regime of strain replacement, the existence of two equilibria is relatively frequent over their empirical parameter space, particularly when the strains start to accumulate. For pneumococcus and malaria in regimes of coexistence, the existence of two equilibria is rare over their empirical parameter space, validating the stable co-circulation of multiple strains in such diseases. We then test the stability of these equilibria by evaluating the eigenvalues of the Jacobian matrices in finding $(\hat{n}_1, \hat{n}_2)$. The eigenvalues from the numerical solver, `nleqslv`, can be affected by the initial conditions and are thus not 100% reliable. Nevertheless, we still find that the low equilibria are more frequently unstable compared to the high equilibria.

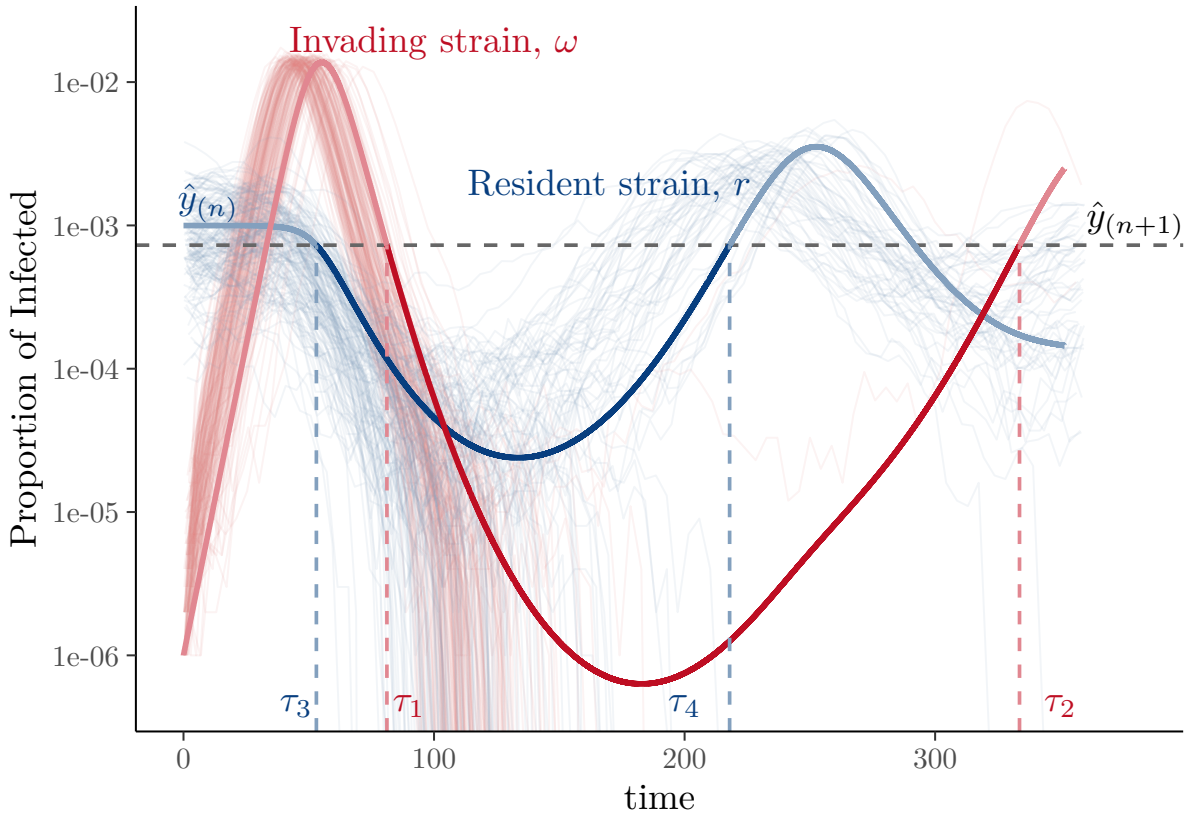

Figure S1: Illustration of deterministic and stochastic dynamics during a new strain invasion. Thick lines show the deterministic trajectories of the proportion of hosts infected with the invading strain ω (red) and a resident strain r (blue), obtained by numerically solving Eq.S10, assuming one resident strain ($n = 1$), a host population size of $N_h = 10^6$, with parameters $\rho = 0.5$, $\varepsilon_1 = 0.0045$, and $\varepsilon_2 = 0.0005$. To capture stochastic effects, 1,000 agent-based simulations were run using the same parameter set, modeling the introduction of a single mutant strain into a system at endemic equilibrium (see Section 1.7 for model details). Translucent lines represent the 162 stochastic trajectories in which the invading strain survived the initial fizzle phase. Of these, only 2 strains successfully established, while 160 went extinct during the subsequent burnout phase. In addition, 112 resident strains were driven to extinction due to the epidemic triggered by the invading strain ω. These outcomes closely match the numerical approximations derived using the methods in Section 1.3: fizzle probability of the invading strain is $p_1^{(\omega)} = 0.833$; burnout probabilities of the invading strain and the resident strain are $q_k^{(\omega)} = 0.992$ and $q_k^{(r)} = 0.728$, respectively.

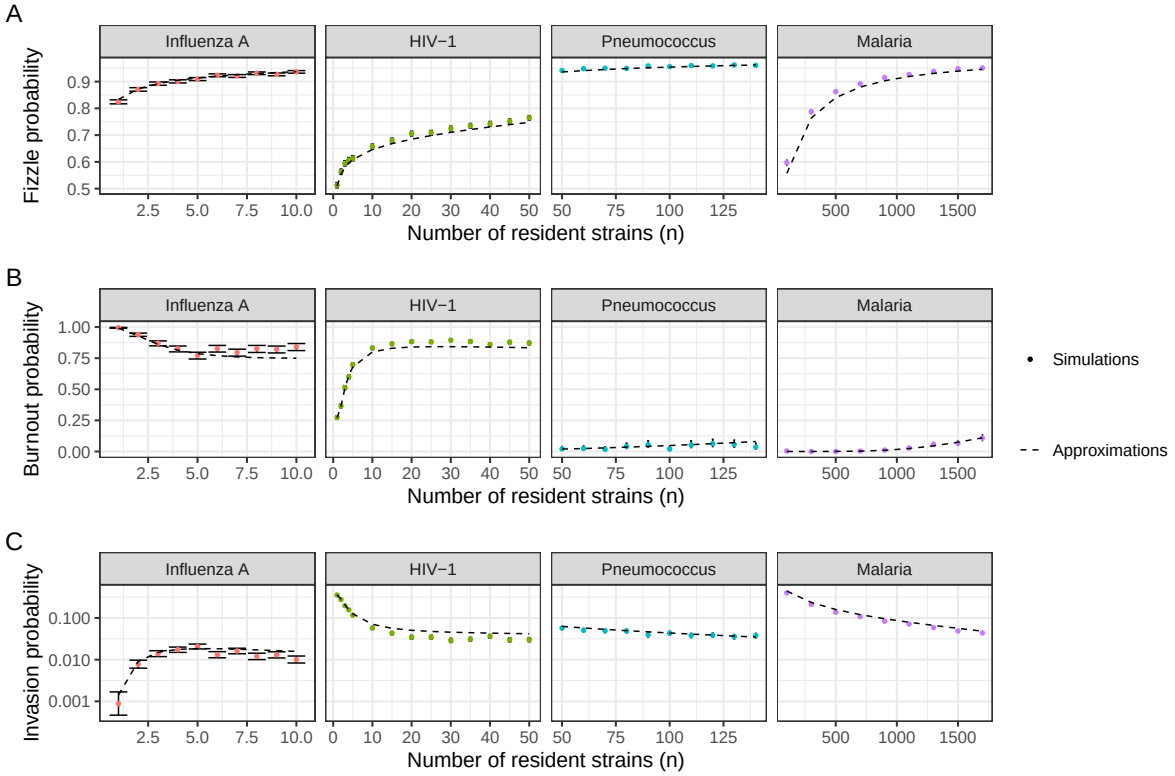

Figure S2: **Comparison of predicted probabilities of stochastic processes at the ecological timescale between MultiSEED and agent-based stochastic simulations. (A)** Fizzle probability, **(B)** burnout probability, and **(C)** invasion probability of a new strain are shown across four human diseases with different numbers of resident strains. Filled circles represent results from agent-based stochastic simulations, and the error bars represent 95% binomial proportion confidence interval. The parameter sets are identical to those described in Fig. 1D-E in the main text.

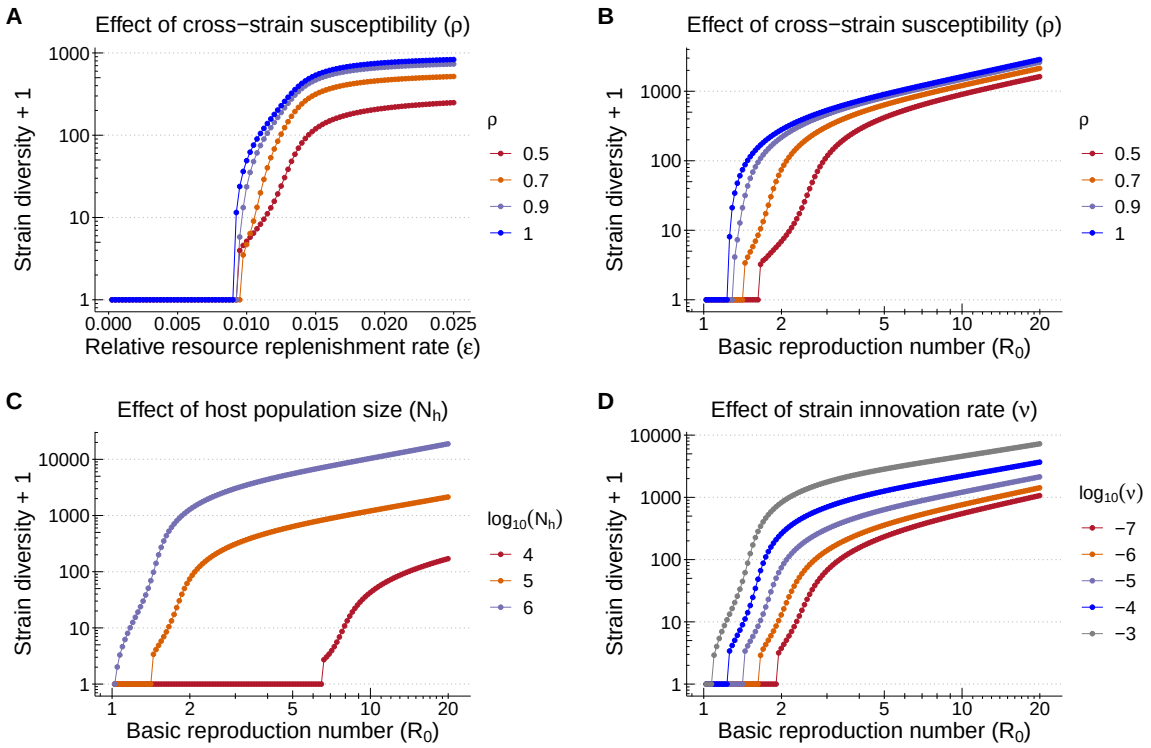

Figure S3: **Effects of parameters on strain diversity and the persistence threshold along the R_0 or ε axis**, as a supplementary for Fig. 2B-C in the main text. **(A-B)** The effects of cross-strain susceptibility (ρ) along the ε and R_0 axis. **(C-D)** The effects of host population size (N_h) and strain innovation rate (ν) along the R_0 axis. The default parameter values are $N_h = 10^5$, $\nu = 10^{-5}$, $\rho = 0.7$, $\varepsilon = 0.015$ and $R_0 = 3$. In each plot, one parameter on the x axis and one parameter indicated by the legend are changed, while others are fixed.

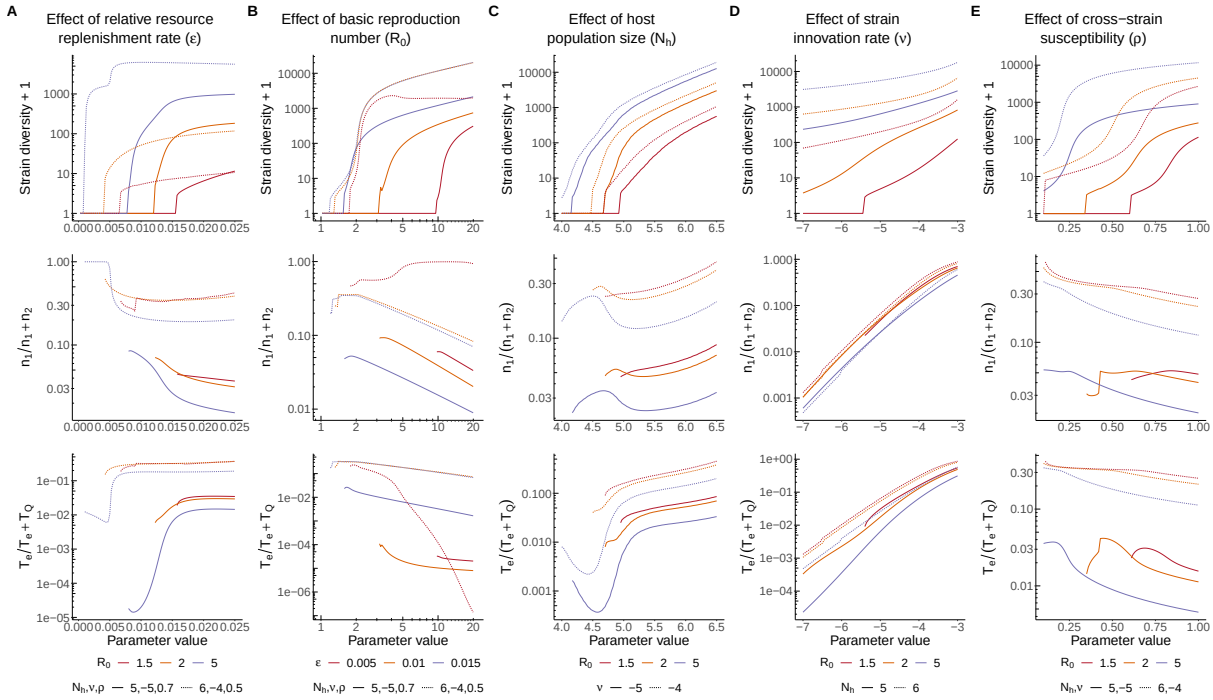

Figure S4: **Effects of different parameters on long-term strain diversity and regimes of strain dynamics.** Plot rows one, two and three correspond to the responses of strain diversity and the two criteria for classifying regimes of strain dynamics given varying parameters. Parameters N_h, ν are always at log_{10} scale. **(A)** The effects of the relative resource replenishment rate ε. Two parameter combinations are tested here: $N_h = 10^5, \nu = 10^{-5}, \rho = 0.7, \varepsilon_1/\varepsilon = 0.75$, representing the broad parameter region of malaria, and $N_h = 10^6, \nu = 10^{-4}, \rho = 0.5, \varepsilon_1/\varepsilon = 0.9$, representing the broad parameter region of influenza A. **(B)** The effects of basic reproduction number R_0. The same parameter combinations in **A** are used. Here $\varepsilon = 0.003$ is not included because for the parameter combination $N_h = 10^5, \nu = 10^{-5}, \rho = 0.7$, strain diversity is always 0 when $\varepsilon = 0.003$. **(C)** The effects of host population size N_h, with fixed parameter values $\varepsilon = 0.015$ and $\rho = 0.7$. **(D)** The effects of strain innovation rate ν, with $\varepsilon = 0.015$ and $\rho = 0.7$. **(E)** The effects of cross-strain susceptibility ρ, with $\varepsilon = 0.015$.

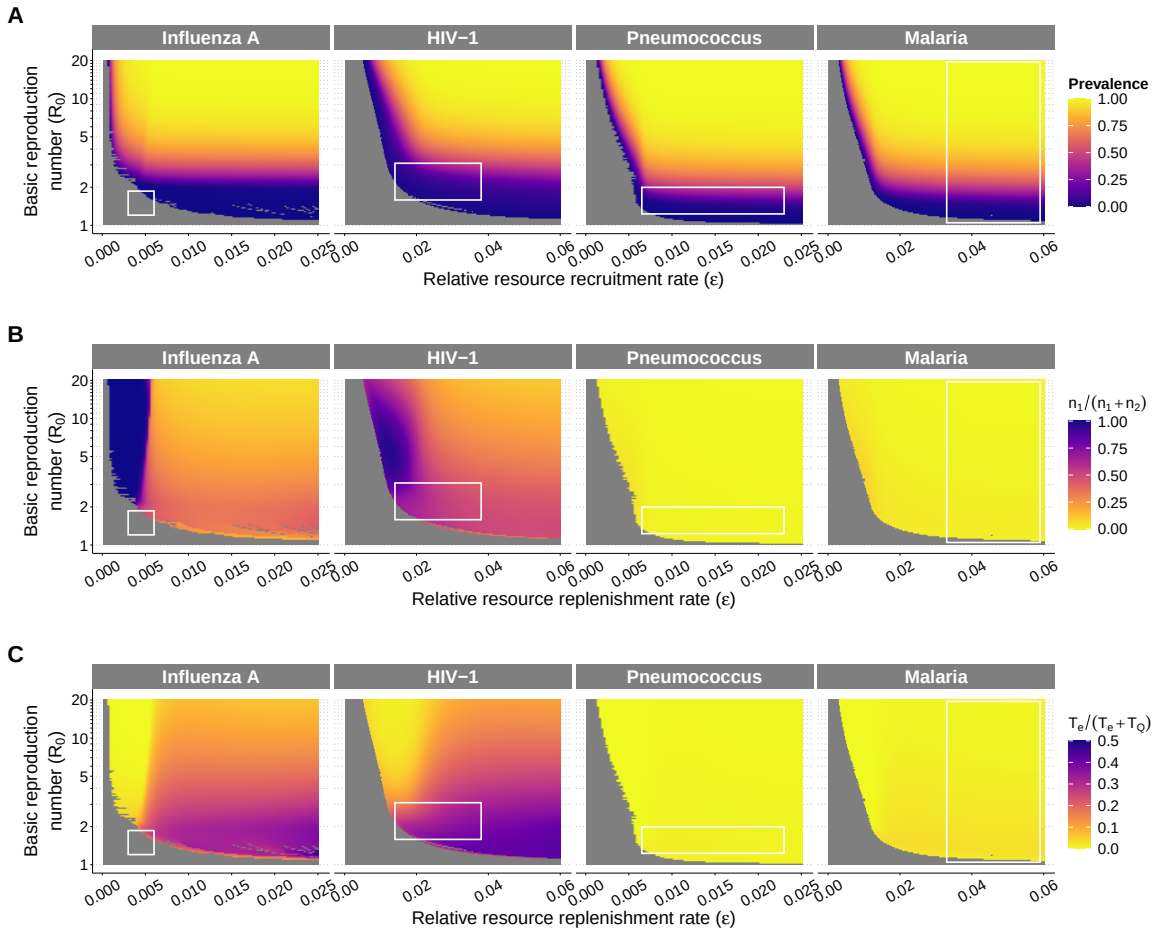

Figure S5: **Predicted prevalence and the two quantities used in defining the replacement and coexistence regimes of the four focal diseases.** The empirical ranges of ε and R_0 are indicated by the white rectangles, identical to Fig. 3 and 4 in the main text. **(A)** Prevalence. (B) The ratio of transient strains in all strains: $n_1/(n_1+n_2)$. **(C)** The ratio between the duration of the first epidemic and the expected extinction time at the evolutionary time scale, $T_e/(T_e+T_Q)$. In Fig. 4B in the main text, the regime of coexistence is defined when both quantities are below 1/3 in (B) and (C). Whenever either quantity is no less than 1/3 outside of the extinction regime, we define the dynamics to be replacement.

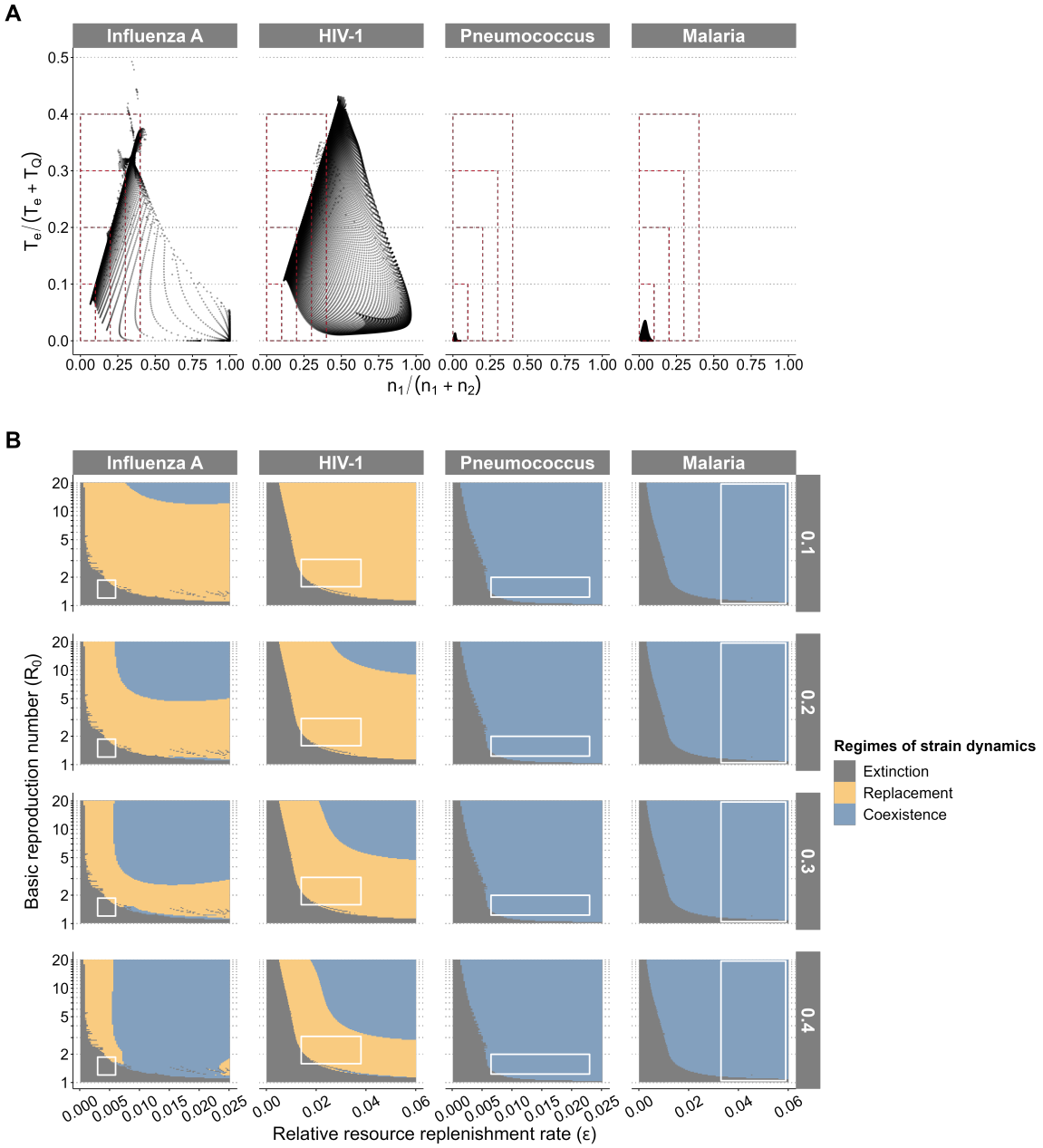

Figure S6: **Effects of the cutoff values in regime delineation.** **(A)** The correlation between the two quantities defining regimes of strain dynamics mentioned in Fig. S5: $n_1/(n_1 + n_2)$ and $T_e/(T_e + T_Q)$. Red rectangles in dashed lines indicate the four cutoff values considered in defining the regime. Points within the rectangles are considered coexistence; otherwise, replacement occurs. For the broad parameter spaces of N_h, ν and ρ in influenza A and HIV-1, the regimes are mostly replacement, but sometimes coexistence depending on the cutoff values. For the parameter spaces of pneumococcus and malaria, the regime is always coexistence, as points accumulate at the bottom-left corner with extremely small values of both quantities. **(B)** The regimes of strain dynamics when using different cutoff values of the two quantities. Coexistence is suggested when both quantities are below the cutoff value.

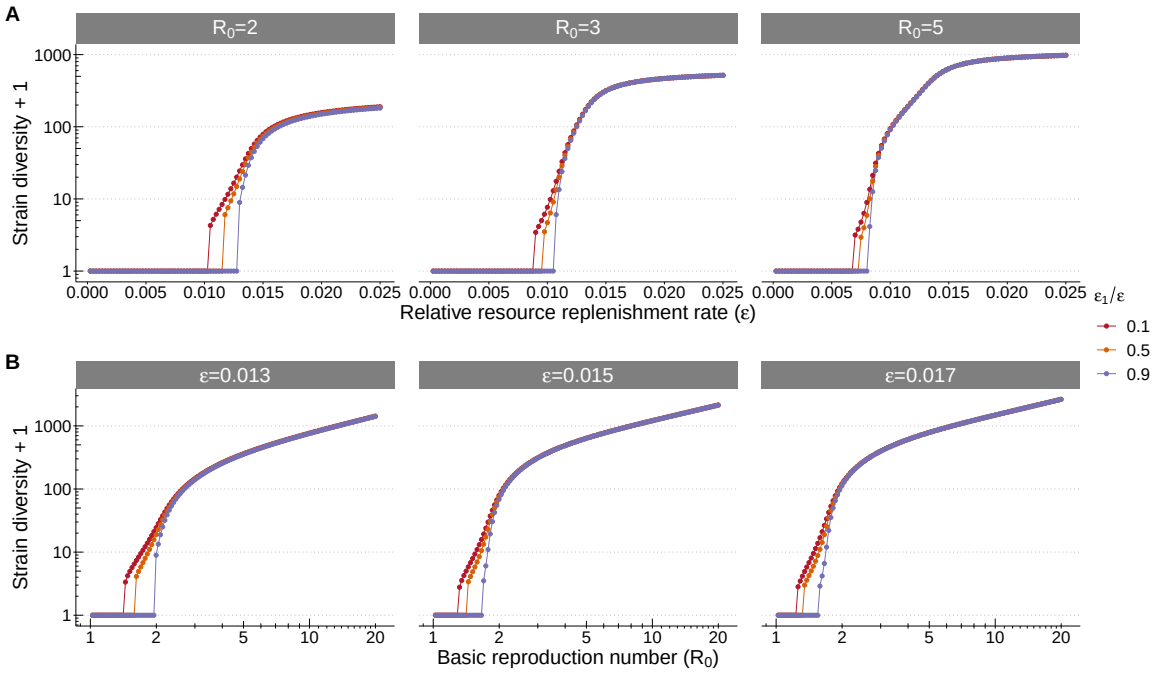

Figure S7: **The effect of $\frac{\varepsilon_1}{\varepsilon}$ on long-term strain diversity**, under different values of R_0 and ε. The values of other parameters are $N_h = 10^5$, $\nu = 10^{-5}$ and $\rho = 0.7$.

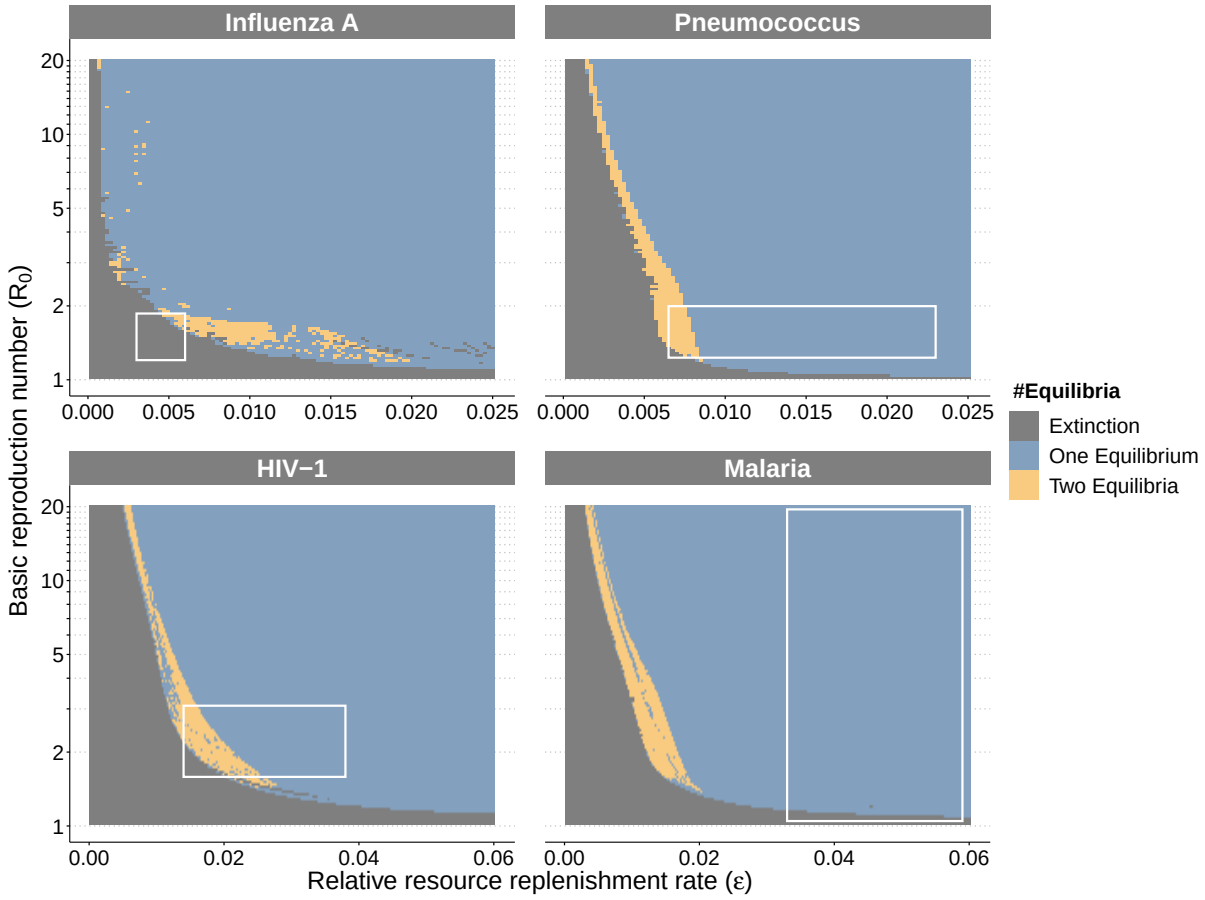

Figure S8: **Estimated number of equilibria using MultiSEED.** Each subplot shows the number of equilibrium under different combinations of ε and R_0 for the four human diseases. The empirical ranges of ε and R_0 are indicated by the white rectangles, identical to Figs. 3 and 4 in the main text. We delineate the boundary region to search for alternative equilibria as follows: For each value of R_0, we first locate the smallest value of ε that leads to a positive strain diversity (ε_s). Then we examine the range from ε_s to (ε_s+0.01) or the upper bound of ε. To find the number of equilibria, we search for multiple initial conditions in each grid cell, and keep all the solutions found for $(\hat{n}_1, \hat{n}_2)$, as described in Section 1.5. The solutions of equilibria are considered identical when the values of $\hat{n}_1$ or $\hat{n}_2$ differ by less than 0.01 on the log scale.

References

1. T. Boehm, C. C. Bleul, *Nature immunology* **8**, 131 (2007).
2. G. W. Litman, J. P. Rast, S. D. Fugmann, *Nature Reviews Immunology* **10**, 543 (2010).
3. K. W. Deitsch, S. A. Lukehart, J. R. Stringer, *Nature Reviews Microbiology* **7**, 493 (2009).
4. K. Marsh, *Parasitology* **104**, S53 (1992).
5. M. Meijers, D. Ruchnewitz, J. Eberhardt, M. Łuksza, M. Lässig, *Cell* **186**, 5151 (2023).
6. K. A. Geno, *et al.*, *Clinical microbiology reviews* **28**, 871 (2015).
7. Q. He, *et al.*, *Nature communications* **9**, 1817 (2018).
8. B. T. Grenfell, O. N. Bjørnstad, J. Kappey, *Nature* **414**, 716 (2001).
9. D. Zinder, *et al.*, *BMC Evol. Biol.* **14**, 272 (2014).
10. T. Bedford, A. Rambaut, M. Pascual, *BMC Biol.* **10**, 38 (2012).
11. J. R. Gog, B. T. Grenfell, *Proceedings of the National Academy of Sciences* **99**, 17209 (2002).
12. K. Koelle, D. A. Rasmussen, *Elife* **4**, e07361 (2015).
13. I. M. Rouzine, G. Rozhnova, *PLoS pathogens* **14**, e1007291 (2018).
14. A. J. Kucharski, J. R. Gog, *Theor. Popul. Biol.* **81**, 102 (2012).
15. O. Restif, B. T. Grenfell, *Proceedings of the Royal Society B: Biological Sciences* **273**, 409 (2006).

16. C. O. Buckee, K. Koelle, M. J. Mustard, S. Gupta, *Proceedings of the National Academy of Sciences* **101**, 10839 (2004).

17. A. Lange, N. M. Ferguson, *PLoS Computational Biology* **5**, e1000536 (2009).

18. S. Gupta, K. Trenholme, R. M. Anderson, K. P. Day, *Science* **263**, 961 (1994).

19. L. J. Abu-Raddad, N. M. Ferguson, *Proceedings of the Royal Society of London. Series B: Biological Sciences* **271**, 2431 (2004). Publisher: Royal Society.

20. J. Wipasa, *et al.*, *PLoS pathogens* **6**, e1000770 (2010).

21. M. OhAinle, *et al.*, *Science translational medicine* **3**, 114ra128 (2011).

22. C. K. Mugyenyi, *et al.*, *The Journal of infectious diseases* **216**, 887 (2017).

23. S. Ranjeva, *et al.*, *Nature Communications* **10**, 1660 (2019).

24. S. W. Park, S. Cobey, C. J. E. Metcalf, J. M. Levine, B. T. Grenfell, *Science* **386**, 175 (2024).

25. O. Balmer, M. Tanner, *The Lancet Infectious Diseases* **11**, 868 (2011).

26. S. Alizon, J. C. De Roode, Y. Michalakis, *Ecology letters* **16**, 556 (2013).

27. A. E. Barry, *et al.*, *Mol. Cell. Proteomics* **10**, M111.008326 (2011).

28. L. Konaté, *et al.*, *Trans. R. Soc. Trop. Med. Hyg.* **93 Suppl 1**, 21 (1999).

29. N. Ferguson, V. Andreasen, *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods, and Theory*, The IMA volumes in mathematics and its applications (Springer New York, New York, NY, 2002), pp. 157–169.

30. P. Minayev, N. Ferguson, *J. R. Soc. Interface* **6**, 989 (2009).

31. T. L. Parsons, D. J. D. Earn, Analytical approximations for the phase plane trajectories of the SIR model with vital dynamics (2023).

32. T. L. Parsons, B. M. Bolker, J. Dushoff, D. J. D. Earn, *Proceedings of the National Academy of Sciences* **121**, e2313708120 (2024). Publisher: Proceedings of the National Academy of Sciences.

33. J. Gog, J. Swinton, *Journal of mathematical biology* **44**, 169 (2002).

34. M. S. Bartlett, *Proceedings of the third Berkeley symposium on mathematical statistics and probability* (Berkeley, 1956), vol. 4, p. 109.

35. I. Nåsell, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**, 309 (1999). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9868.00178.

36. A. D. Barbour, *Advances in applied probability* **8**, 296 (1976).

37. N. M. Ferguson, A. P. Galvani, R. M. Bush, *Nature* **422**, 428 (2003).

38. W. M. Jou, *et al.*, *Cell* **19**, 683 (1980).

39. E. Visher, S. E. Whitefield, J. T. McCrone, W. Fitzsimmons, A. S. Lauring, *PLoS pathogens* **12**, e1005856 (2016).

40. P. Baccam, C. Beauchemin, C. A. Macken, F. G. Hayden, A. S. Perelson, *Journal of virology* **80**, 7590 (2006).

41. J. J. Cannell, M. Zasloff, C. F. Garland, R. Scragg, E. Giovannucci, *Virology journal* **5**, 1 (2008).

42. R. Casagrandi, L. Bolzoni, S. A. Levin, V. Andreasen, *Mathematical biosciences* **200**, 152 (2006).

1042 43. M. Samsuzzoha, M. Singh, D. Lucy, *Applied Mathematics and Computation* **220**, 616
1043 (2013).

1044 44. D. J. Smith, *et al.*, *Science* **305**, 371 (2004).

1045 45. M. Biggerstaff, S. Cauchemez, C. Reed, M. Gambhir, L. Finelli, *BMC infectious diseases*
1046 **14**, 1 (2014).

1047 46. A. C. Tricco, *et al.*, *BMC medicine* **11**, 1 (2013).

1048 47. D. Gatherer, *Indian journal of microbiology* **49**, 352 (2009).

1049 48. G. P. Garnett, R. M. Anderson, *Journal of Infectious Diseases* **174**, S150 (1996).

1050 49. T. D. Hollingsworth, R. M. Anderson, C. Fraser, *The Journal of infectious diseases* **198**,
1051 687 (2008).

1052 50. J. Zhuang, *et al.*, *Journal of virology* **76**, 11273 (2002).

1053 51. M. Zhang, *et al.*, *Retrovirology* **7**, 1 (2010).

1054 52. H. Song, *et al.*, *Nature communications* **9**, 1928 (2018).

1055 53. J. M. Rawson, S. R. Landman, C. S. Reilly, L. M. Mansky, *Retrovirology* **12**, 1 (2015).

1056 54. X. Wei, *et al.*, *Nature* **373**, 117 (1995).

1057 55. E. Bons, F. Bertels, R. R. Regoes, *Virus evolution* **4**, vey029 (2018).

1058 56. B. Korber, *et al.*, *British medical bulletin* **58**, 19 (2001).

1059 57. S. M. Keating, *et al.*, *The Journal of infectious diseases* **216**, 72 (2017).

58. D. M. Smith, D. D. Richman, S. J. Little, *The Journal of Infectious Diseases* **192**, 438 (2005).

59. O. Metadilogkul, V. Jirathitikal, A. S. Bourinbaiar, *BioMed Research International* **2009**, 934579 (2009).

60. T. Stadler, *et al.*, *Molecular biology and evolution* **29**, 347 (2012).

61. K. Ronen, *et al.*, *PLoS Pathogens* **9**, e1003593 (2013).

62. B. Korber, *et al.*, *science* **288**, 1789 (2000).

63. M. Worobey, *et al.*, *Nature* **455**, 661 (2008).

64. A. P. d. O. Menezes, *et al.*, *Vaccine* **34**, 791 (2016).

65. K. L. Sleeman, *et al.*, *The Journal of infectious diseases* **194**, 682 (2006).

66. C. Murad, *et al.*, *International Journal of Infectious Diseases* **86**, 73 (2019).

67. R. J. Mostowy, *et al.*, *Molecular biology and evolution* **34**, 2537 (2017).

68. K. L. Wyres, *et al.*, *The Journal of infectious diseases* **207**, 439 (2013).

69. A. Mavroidi, *et al.*, *Journal of bacteriology* **189**, 7841 (2007).

70. M. Kilian, *et al.*, *PloS one* **3**, e2683 (2008).

71. E. Ramos-Sevillano, G. Ercoli, J. S. Brown, *Frontiers in immunology* **10**, 358 (2019).

72. Y. Shi, *et al.*, *The Journal of Immunology* **175**, 3262 (2005).

73. D. M. Musher, *et al.*, *Clinical Infectious Diseases* **17**, 66 (1993).

74. M. Andersson, *et al.*, *Statistics in Medicine* **24**, 3593 (2005).

75. M. Nurhonen, A. C. Cheng, K. Auranen, *PloS one* **8**, e56079 (2013).

76. E. Gjini, *Scientific Reports* **7**, 3049 (2017).

77. M. J. Bonten, *et al.*, *New England Journal of Medicine* **372**, 1114 (2015).

78. M. Suzuki, *et al.*, *The Lancet infectious diseases* **17**, 313 (2017).

79. K. F. Nielsen, *et al.*, *Emerging Infectious Diseases* **30**, 1164 (2024).

80. S. M. Granat, *et al.*, *The Journal of infectious diseases* **200**, 99 (2009).

81. S. Pauleit, *et al.*, *Futur City* **4** (2015).

82. W. Gu, *et al.*, *Acta tropica* **86**, 71 (2003).

83. S. Mandal, R. R. Sarkar, S. Sinha, *Malaria journal* **10**, 1 (2011).

84. J. H. Huber, G. L. Johnston, B. Greenhouse, D. L. Smith, T. A. Perkins, *Malaria journal* **15**, 1 (2016).

85. Q. He, S. Pilosof, K. E. Tiedje, K. P. Day, M. Pascual, *Frontiers in ecology and evolution* **9**, 633263 (2021).

86. J. Langhorne, F. M. Ndungu, A.-M. Sponaas, K. Marsh, *Nature immunology* **9**, 725 (2008).

87. D. L. Doolan, C. Dobaño, J. K. Baird, *Clinical microbiology reviews* **22**, 13 (2009).

88. A. Ly, D. S. Hansen, *Frontiers in immunology* **10**, 559 (2019).

89. D. L. Smith, F. E. McKenzie, R. W. Snow, S. I. Hay, *PLoS biology* **5**, e42 (2007).

90. D. E. Neafsey, *et al.*, *New England Journal of Medicine* **373**, 2025 (2015).

91. D. E. Loy, *et al.*, *International journal for parasitology* **47**, 87 (2017).

92. P. M. Sharp, L. J. Plenderleith, B. H. Hahn, *Annual review of microbiology* **74**, 39 (2020).

93. F. E. McCutchan, *Journal of medical virology* **78**, S7 (2006).

94. T. D. Otto, *et al.*, *Wellcome Open Research* **4**, 193 (2019).

95. F. Zanini, *et al.*, *Elife* **4**, e11282 (2015).

96. N. Bbosa, P. Kaleebu, D. Ssemwanga, *Current Opinion in HIV and AIDS* **14**, 153 (2019).

97. D. B. Larremore, *et al.*, *Nature communications* **6**, 8368 (2015).

98. I. I. Ciubotariu, *et al.*, *EBioMedicine* **106** (2024).

99. S. Pilosof, *et al.*, *PLoS biology* **17**, e3000336 (2019).

100. M. S. Bartlett, *Journal of the Royal Statistical Society Series A: Statistics in Society* **123**, 37 (1960).

101. M. Lipsitch, J. J. O'Hagan, *Journal of the Royal Society Interface* **4**, 787 (2007).

102. S. Gupta, N. Ferguson, R. Anderson, *Science* **280**, 912 (1998).

103. Q. Zhan, M. Pascual, Q. He, *bioRxiv* pp. 2025–06 (2025).

104. A. Dobson, *the american naturalist* **164**, S64 (2004).

105. J. P. Aparicio, C. Castillo Chavez, *Mathematical Biosciences And Engineering* **6**, 209 (2009).

106. L. C. Katzelnick, *et al.*, *Science* **358**, 929 (2017).

107. K. Koelle, S. Cobey, B. Grenfell, M. Pascual, *Science* **314**, 1898 (2006).

108. V. Andreasen, J. Lin, S. A. Levin, *Journal of mathematical biology* **35**, 825 (1997).

109. D. G. Kendall, *The annals of mathematical statistics* pp. 1–15 (1948).

110. B. Hasselman, M. B. Hasselman, *R package version* **3** (2018).

111. L. Reichel, W. B. Gragg, *ACM Transactions on Mathematical Software (TOMS)* **16**, 369 (1990).

112. J. E. Dennis Jr, R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations* (SIAM, 1996).

113. D. Champredon, J. Dushoff, D. J. Earn, *SIAM Journal on Applied Mathematics* **78**, 3258 (2018).

114. F. Hoti, P. Erästö, T. Leino, K. Auranen, *BMC infectious diseases* **9**, 1 (2009).

115. E. J. Feil, J. M. Smith, M. C. Enright, B. G. Spratt, *Genetics* **154**, 1439 (2000).

116. C. Chaguza, J. E. Cornick, D. B. Everett, *Computational and structural biotechnology journal* **13**, 241 (2015).

117. T. S. Rask, D. A. Hansen, T. G. Theander, A. Gorm Pedersen, T. Lavstsen, *PLoS computational biology* **6**, e1000933 (2010).

118. A. M. de Roos, Q. He, M. Pascual, *Proceedings of the National Academy of Sciences* **120**, e2218499120 (2023).