A multilocus perspective on the evolutionary dynamics of multistrain pathogens

David V. McLeod*[1,2,3], Claudia Bank†[2,3,5], and Sylvain Gandon‡[4,5]

[1]Département de mathématiques et statistique, Université de Montréal, Montréal, Canada
[2]Institute of Ecology and Evolution, Universität Bern, Bern, Switzerland
[3]Swiss Institute of Bioinformatics, Lausanne, Switzerland
[4]CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France
[5]equal contribution

January 23, 2024

Abstract

Many human pathogens, including malaria, dengue, influenza, *Streptococcus pneumoniae*, and cytomegalovirus, coexist as multiple genetically distinct strains. Understanding how these multistrain pathogens evolve is of critical importance for forecasting epidemics and predicting the consequences of vaccination. One factor believed to play an important role is naturally acquired immunity. Consequently, a large body of research has sought to predict how acquired immunity molds the genomics of pathogen populations (i.e., what shapes pathogen strain structure). However, the diversity and specificity of existing models has resulted in conflicting evolutionary predictions. This has led to an ongoing debate about which predictions are most broadly applicable. Here, we adopt a multilocus population genetics perspective that unifies the predictions of existing models. We identify three key factors that determine the role of naturally acquired immunity in the evolution of pathogen strain structure: (i) the strength and specificity of immune protections, (ii) the dynamic immunological landscape, and (iii) the number of loci coding for the antigens of the pathogen. Isolating and discussing these three factors clarifies the relationship among previous models of multistrain dynamics, and establishes a solid theoretical foundation for the study of the evolutionary epidemiology of multistrain pathogens.

Keywords: evolutionary epidemiology, multistrain pathogens, adaptive immune response, pathogen strain structure

*david.mcleod@umontreal.ca
†claudia.bank@unibe.ch
‡sylvain.gandon@cefe.cnrs.fr

1 Introduction

Genomic data of many pathogens of importance for human health, including malaria [1], dengue [2], *Streptococcus pneumoniae* [3], and cytomegalovirus [4], reveals species genetically subdivided into distinct genotypes, or strains, characterized by statistical associations between alleles at different loci [5–8]. Disentangling what produces these genomic patterns (henceforth, pathogen strain structure or PSS) is a major focus of evolutionary epidemiology, owing to its relevance for a variety of human health issues, including forecasting epidemics [9] and predicting the epidemiological consequences of vaccination campaigns [10, 11]. One factor believed to play an important role is naturally acquired host immunity, and a large body of work has sought to discern how host immunity affects the dynamics of the genomics of pathogen populations [6, 9, 12–20].

The aspect of naturally acquired immunity that has received the most attention is the role of cross-protection, broadly defined as any immune protection an infection provides against genetically similar strains [e.g., 6, 9, 12, 14–16, 21–25]. Roughly speaking, the predictions are as follows [15, 26]: if cross-protection is weak, there is no PSS; if it is of intermediate strength, evolutionary oscillations are common; if it is strong, a set of strains sharing no antigens in common is over-represented in the population [6, 12]. Yet each of these predictions comes with exceptions. This has led to debate about the generality of the results. For example, it is not well-understood why only some models produce evolutionary oscillations [27], how predictions about PSS scale with an increasing number of immunogenic loci and alleles [18–20], or what patterns of PSS are either possible or likely [28]. One reason for this debate is that models which combine naturally acquired immunity with pathogen genetic diversity require a large number of variables [29]. Consequently, existing work has taken subtly different approaches to simplify the dynamics, with poorly understood consequences [27, 29, 30].

Here we take a population genetics perspective to review, generalize, and provide evolutionary intuition about when, and why, existing predictions hold (or not). We do so in three parts, each of which emphasizes biological intuition [see e.g., 27, 30, for a more quantitative perspective]. In Section 2, we define PSS. In Section 3, we use this definition to identify the evolutionary forces producing PSS. In Section 4, we show how these evolutionary forces are shaped by the epidemiology of naturally acquired immunity. This results in the identification of three crucial factors: the strength and specificity of the immune protections, the dynamic immunological landscape, and the amount of pathogen genetic diversity. We conclude by discussing future directions. Notation and acronyms used are detailed in Table 1.

2 What is pathogen strain structure (PSS)?

Any definition of pathogen strain structure (PSS) requires genomic patterns unexplainable by selection for individual alleles. It is thus insufficient for some strains to be more abundant than others; this could arise solely through variation in allele frequencies. Instead, PSS minimally requires some strains to be *over-represented* relative to their expectation under a random shuffling of alleles, and so identification of PSS necessitates specifying the multilocus pathogen genetics. For a pathogen genome consisting of two diallelic loci, $\{a, A\}$ and $\{b, B\}$,

and so four strains $\{ab, Ab, aB, AB\}$, PSS requires

$$D(t) \equiv p_{ij}(t) - p_i(t)p_j(t) \neq 0, \tag{1}$$

where $p_i(t)$ and $p_{ij}(t)$ denote the frequency of infections by strains carrying allele(s) i and ij, respectively, at time t. In classical population genetics, the quantity $D(t)$ is pairwise linkage disequilibrium (LD) [31]. LD can be generalized to a larger number of loci and alleles [31], and since LD captures any nonrandom associations between alleles, the presence of LD is necessary for some strains to be over-represented. Hence, the minimal definition of PSS is the presence of pairwise (or higher-order) LD.

A more stringent form of PSS, *discordant strain structure* (DSS), was first proposed by Gupta et al. [6]. DSS occurs when a set of strains whose members share no alleles (e.g., $\{ab, AB\}$ and $\{Ab, aB\}$) is over-represented in the population [6, 12, 32]. Generally speaking, the presence of pairwise LD is necessary, but not sufficient, for DSS [32, 33].

Finally, although PSS requires at minimum the presence of LD, in the long term, the presence of LD could occur by LD reaching a stable equilibrium value or by LD undergoing sustained evolutionary oscillations (potentially chaotic). If LD reaches a non-zero stable equilibrium, we will refer to this as *stable PSS*, whereas if LD undergoes sustained oscillations, we will refer to this as *oscillatory PSS*.

3 Which evolutionary forces produce strain structure?

As PSS requires the presence of LD, the dynamics of LD are key to understanding its evolution. To describe these dynamics, we initially focus on two diallelic loci. For multistrain models of adaptive immunity, two timescales often emerge [27, 34]. Over the short timescale, strong negative frequency-dependent selection (NFDS) operates through the additive selection coefficients, $w_A \equiv r_{Ab} - r_{ab}$ and $w_B \equiv r_{aB} - r_{ab}$, which pushes the alleles to an equal frequency (here, r_{ij} is the per-capita growth rate of strain ij; see Box 1). As this occurs, the strains self-organize into two sets, $\{ab, AB\}$ and $\{Ab, aB\}$, so that over the long term, LD evolves according to the equation

$$\frac{\mathrm{d}D}{\mathrm{d}t} \approx w_{AB}\left(\frac{1}{4} + D\right)\left(\frac{1}{4} - D\right) + \Delta_D\rho, \tag{2}$$

where w_{AB} denotes epistasis in fitness between alleles A and B, that is,

$$w_{AB} \equiv r_{AB} + r_{ab} - r_{Ab} - r_{aB},$$

and $\Delta_D\rho$ is the net change in LD due to recombination (Sup. Info. S3). Equation (1) and (2) reveal the evolutionary forces generating and maintaining PSS.

Box 1: A population model of naturally acquired immunity

Models of adaptive immunity largely differ on three types of assumptions (Sup. Info. S2): the type of immune protection considered (e.g., does immunity reduce infection or transmissibility?); how immunity is gained and lost; and whether hosts can be simultaneously infected by multiple strains (co-infection). Typically, the choice of assumptions is made to simplify the mathematics. The model we use here prioritizes conceptual simplicity. We will primarily discuss this model throughout the paper; any required extensions and links to other models can be found in the Supplementary Information.

We consider a pathogen whose genome consists of some number of multiallelic loci; each allele, or antigen, induces a (host) adaptive immune response. Transmission occurs exclusively to uninfected hosts (so we do not permit co-infection) via mass-action with rate constant β, whereas the probability of successful infection, σ_x^g, depends upon the host's immune history, x, and the genotype, g, of the strain. In principle, x could be any combination of antigens and/or strains and could also account for the temporal order of infections (e.g., more recent infections induce a more protective response). However, here we assume x is an unordered set of prior infections. Thus for n pathogen strains, there are 2^n possible infection histories. For a host with immunity to $n = 0, 1, ...$ antigens of strain g but no prior infection by that strain, $\sigma_x^g = \sigma_n$, whereas for a host with prior infection by strain g, $\sigma_x^g = \sigma_h$ (homologous immunity). For example, for two diallelic loci, a host with immune history $x' = \{ab, Ab\}$ has $\sigma_{x'}^{ab} = \sigma_{x'}^{Ab} = \sigma_h$ and $\sigma_{x'}^{aB} = \sigma_{x'}^{AB} = \sigma_1$. For simplicity, we will assume that $\sigma_0 \geq \sigma_1 \geq \cdots \geq \sigma_n \geq \sigma_h$. Hosts clear infections and die from natural and virulence-related mortality at per-capita rates μ, d, and α, respectively. Immunologically naive hosts ($x = \{\varnothing\}$) enter the population at rate b; immunity is life-long and acquired at clearance. Let $S_x(t)$ and $I_x^g(t)$ denote, respectively, the density of uninfected, and infected with strain g, hosts with immune history x at time t. Then

$$\begin{aligned} \frac{\mathrm{d}S_x}{\mathrm{d}t} &= b\mathbf{1}_0(x) - dS_x - \beta \sum_g \sigma_x^g \sum_z I_z^g S_x + \sum_g \sum_{z:g\cup z=x} \mu I_z^g \\ \frac{\mathrm{d}I_x^g}{\mathrm{d}t} &= \beta \sum_z I_z^g \sigma_x^g S_x - (d+\mu+\alpha) I_x^g \end{aligned} \tag{3}$$

where $\mathbf{1}_0(x) = 1$ if $x = \{\varnothing\}$ and 0 otherwise.

The per-capita growth rates of the different strains link the epidemiological and evolutionary dynamics through the additive selection coefficients and epistasis. To extract them from model (3), we focus upon two diallelic loci (i.e., $g = ij$ where $i \in \{a, A\}$ and $j \in \{b, B\}$). Let $\mathcal{A}_{\{i,j\}}(t)$ and $\mathcal{S}_{ij}(t)$ denote the density of uninfected hosts with immunity to *at least* the set of antigens $\{i, j\}$ and prior infection by strain ij, respectively (Fig. 1). For example, uninfected hosts with infection history $\{ab, AB\}$ would contribute to $\mathcal{A}_{\{a,b\}}(t)$, $\mathcal{A}_{\{A,b\}}(t)$, $\mathcal{A}_{\{a,B\}}(t)$ and $\mathcal{A}_{\{A,B\}}(t)$ but only to $\mathcal{S}_{ab}(t)$ and $\mathcal{S}_{AB}(t)$. Using this notation, the per-capita growth rate of strain ij is

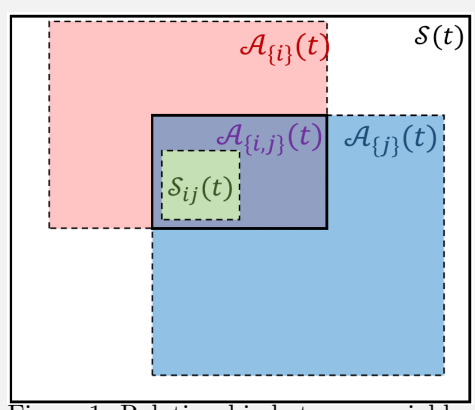

Figure 1: Relationship between variables describing the distribution of immunity.

$$r_{ij} \equiv \sum_x \frac{\mathrm{d}I_x^{ij}/\mathrm{d}t}{\sum_y I_y^{ij}} = \beta\left(\sigma_0 \mathcal{S} - (\sigma_0 - \sigma_1)(\mathcal{A}_{\{i\}} + \mathcal{A}_{\{j\}}) + \sigma_{\mathrm{PI}} \mathcal{A}_{\{i,j\}} - \sigma_{\mathrm{HI}} \mathcal{S}_{ij}\right) - d - \mu - \alpha.$$

where $\sigma_{\mathrm{PI}} \equiv \sigma_0 - 2\sigma_1 + \sigma_2$ and $\sigma_{\mathrm{HI}} \equiv \sigma_2 - \sigma_h$, and $\mathcal{S} = \sum_x S_x$ is the total density of uninfected hosts, irrespective of immune history.

First, from equation (1) the presence of PSS requires both loci to be polymorphic. Maintaining genetic polymorphisms is rarely a problem in models of adaptive immunity as NFDS will tend to maintain allelic diversity. In fact, most models tend to maximize this diversity; for example, in the case of two diallelic loci, allele frequency tends to 1/2. As our interest is what favours or disfavours some strains given allelic variation observable from genomic data, we will not consider further what favours or disfavours allelic diversity [e.g., see 8, 35, 36].

Second, assuming both loci are polymorphic, the long-term presence of PSS requires a force to generate LD. From equation (2), there are two possibilities: epistasis and (biased) recombination. Because virtually all multistrain models implicitly revolve around epistasis [but see 37, 38], we will assume that recombination is weak, $\Delta_D \rho \approx 0$. Then the dynamics of (2) hinge on epistasis, w_{AB}, which indicates whether the combination of alleles A and B are more, $w_{AB} > 0$, or less, $w_{AB} < 0$, favoured than expected.

4 Epistasis in an epidemiological context

How is epistasis in fitness shaped by epidemiology? To answer this requires an epidemiological model. When possible, we will use the model of Box 1. Briefly, in this model, exposure to an allele (or antigen) induces an adaptive immune response that protects against future infection. In particular, the probability of infection of a host with immunity to $n = 0, 1, 2$ antigens of the transmitted strain is σ_n and the probability a host is re-infected by the same strain is σ_h (homologous immunity). Then epistasis is

$$\begin{aligned} w_{AB} = \beta(\overbrace{\sigma_0 - 2\sigma_1 + \sigma_2}^{\sigma_{\mathrm{PI}}})(\overbrace{\mathcal{A}_{\{a,b\}}(t) + \mathcal{A}_{\{A,B\}}(t) - \mathcal{A}_{\{a,b\}}(t) - \mathcal{A}_{\{a,B\}}(t)}^{\mathcal{A}_\Delta(t)}) \\ - \beta(\underbrace{\sigma_2 - \sigma_h}_{\sigma_{\mathrm{HI}}})(\underbrace{\mathcal{S}_{ab}(t) + \mathcal{S}_{AB}(t) - \mathcal{S}_{Ab}(t) - \mathcal{S}_{aB}(t)}_{\mathcal{S}_\Delta(t)}), \end{aligned} \quad (4)$$

where β is a transmissibility rate constant and $\mathcal{A}_{\{i,j\}}(t)$ and $\mathcal{S}_{ij}(t)$ denote the density of uninfected hosts with immunity to *at least* the set of antigens i and j and prior infection by strain ij, respectively, at time t (Box 1). Two factors contribute to equation (4). The first is the strength and specificity of the immune protection, captured by the composite parameters σ_{PI} and σ_{HI} (Fig. 2). The second is the dynamic immunological landscape, captured by the composite variables $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$.

4.1 The specificity and strength of immune protections

The specificity and strength of immune protections partitions existing models into two groups. Hosts acquire *antigen-specific* immunity in the first group [e.g., 18–20, 39]: here, immunity to strains Ab and aB translates to immunity to the set of antigens $\{a, A, b, B\}$ and so full protection against strain AB. In this case, $\sigma_2 = \sigma_h$ and so $\sigma_{\mathrm{HI}} = 0$, that is, homologous immunity is no more protective than full antigenic immunity. Consequently, the evolutionary dynamics hinge on σ_{PI}, which controls whether the rate of increase of antigen-specific cross-protection is increasing ($\sigma_{\mathrm{PI}} < 0$) or decreasing ($\sigma_{\mathrm{PI}} > 0$) as the number of antigens recognized by the immune system of the host increases (Fig. 2). Models of antigen-specific immunity apply to, for example, pathogens that temporally express different antigens to

which the host develops an immune response (e.g., malaria [18–20], pathogenic *Neisseria* [40]).

Hosts acquire *strain-specific* immunity in the second group [e.g., 6, 9, 12, 14–16, 26, 41]: the level of cross-protection against a particular strain is determined by the most antigenically similar strain that has infected the host, but this is less protective than homologous immunity, that is, $\sigma_2 > \sigma_h$ and so $\sigma_{\mathrm{HI}} > 0$. Thus σ_{HI} controls the degree to which homologous immunity is more protective than full antigenic immunity (Fig. 2). For example, although 'strains' might be defined based on the antigenic variation at a few immunodominant epitopes, strains might also vary at other weakly antigenic sites. If this variation is more similar within a strain than between strains, homologous immunity may be more protective than full antigenic immunity.

More generally, if the adaptive immune system recognizes, and targets, the pathogen at the level of its constitutive antigens, rather than as a unit (i.e., strain), then $\sigma_{\mathrm{HI}} = 0$. If instead the pathogen is targeted as a unit and cross-protection arises through single-strain similarity, then $\sigma_{\mathrm{HI}} > 0$. These two perspectives yield very different evolutionary dynamics, as we will see below.

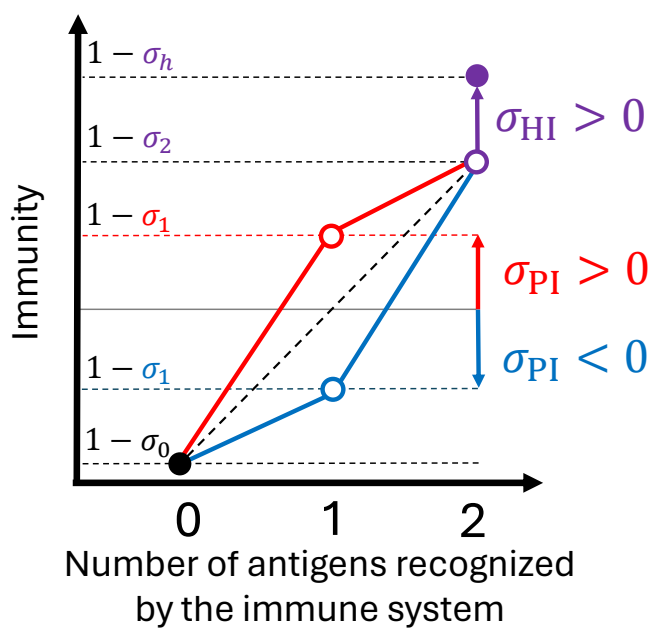

Figure 2: **Schematic of cross-protection**. Specification of cross-protective immunity translates to two composite parameters controlling how cross-protection affects epistasis. The first composite parameter, σ_{PI}, captures the degree to which the increase in cross-protection is increasing ($\sigma_{\mathrm{PI}} > 0$) or decreasing ($\sigma_{\mathrm{PI}} < 0$) as the number of antigens recognized by the immune system increases. Indeed, if we consider the function mapping antigenic immunity to the strength of immune protections, $h : n \mapsto \sigma_n$, then σ_{PI} represents the central difference approximation of $\mathrm{d}^2h/\mathrm{d}n^2$ evaluated at $n = 1$. The second composite parameter, σ_{HI}, captures whether homologous immunity provides equivalent ($\sigma_{\mathrm{HI}} = 0$) or more ($\sigma_{\mathrm{HI}} > 0$) protection than full antigenic immunity.

Rapidly increasing antigen-specific cross-protection produces stable PSS

When full antigenic immunity is as protective as homologous immunity (i.e. $\sigma_2 = \sigma_h$ and $\sigma_{HI} = 0$), epistasis reduces to $w_{AB} = \sigma_{\mathrm{PI}}\mathcal{A}_\Delta(t)$. The variable $\mathcal{A}_\Delta(t)$ indicates whether more uninfected hosts have full antigenic immunity to the set of strains $\{ab, AB\}$ (if positive) or $\{Ab, aB\}$ (if negative). Thus, if one strain set, say $\{ab, AB\}$, is over-represented in the population, our intuitive expectation would be that more hosts have full antigenic immunity to strains belonging to that set. That is, $\mathcal{A}_\Delta(t)$ and $D(t)$ share the same sign. If this holds, there are three possible outcomes (Fig. 3). First, if $\sigma_{\mathrm{PI}} < 0$, all strains stably coexist at equal frequencies. Second, if $\sigma_{\mathrm{PI}} = 0$, epistasis vanishes. Hence, any observed PSS is neutral and contingent upon the initial conditions. Third, if $\sigma_{\mathrm{PI}} > 0$, then stable PSS occurs; in this case because there is a single epidemiological feedback, $\mathcal{A}_\Delta(t)$, one strain set is competitively excluded. Thus, the sign of σ_{PI} drives the formation of PSS by controlling whether competition between strains sharing one antigen is stronger or weaker than expected if the strength of protection offered by the immune response to multiple antigens was additive.

Strongly protective homologous immunity generates oscillatory PSS and maintains strain diversity

Importantly, the sign of σ_{PI} alone cannot sustain evolutionary oscillations (oscillatory PSS), nor can it stably maintain an intermediate level of stable PSS (i.e., $0 < |D(t)| < 1/4$); two common predictions of multistrain models. This is not surprising: when $\sigma_{\mathrm{HI}} = 0$, there is a single epidemiological feedback on epistasis, whether more hosts have *antigenic immunity* to the combination of antigens of strains belonging to one set or the other, $\mathcal{A}_\Delta(t)$. Consequently, there is no reason to expect the sign of epistasis to change based upon the amount of LD. However, when homologous immunity is more protective than full antigenic immunity, $\sigma_{\mathrm{HI}} > 0$. This creates a second epidemiological feedback, whether more hosts have *homologous immunity* to strains belonging to one set or the other, $\mathcal{S}_\Delta(t)$. The second feedback promotes strain diversity because the more abundant strains will have caused more infections and so be at a competitive disadvantage. If $\sigma_{\mathrm{PI}} < 0$, both feedbacks prevent strains from becoming over-represented, leading to a stable coexistence of all strains at an equal frequency (Fig. 3). However, if $\sigma_{\mathrm{PI}} > 0$, in principle promoting PSS, this will be counteracted by the feedback created by $\sigma_{\mathrm{HI}} > 0$ which inhibits PSS (Fig. 3).

The counteracting forces of $\sigma_{\mathrm{PI}} > 0$ and $\sigma_{\mathrm{HI}} > 0$ can produce oscillatory PSS, where the amount of PSS undergoes sustained oscillations (potentially chaotic). Indeed, not only is this the source of the complex dynamics seen in many models [6, 9, 12, 16, 42], but it also explains why oscillations are absent in so-called 'status-based' models but not 'history-based' models [13, 27, 30, 34]. History-based models, such as the model in Box 1, assume that following infection, hosts may be partially protected against genetically similar strains (i.e., have a reduced, but non-zero probability of infection). On the other hand, status-based models assume that following infection, hosts become either completely protected or unprotected against each strain, with the probabilities dependent on immune history (Sup. Info. S2). This is sometimes referred to as *polarized immunity*, as for each strain, the population is divided into either naive hosts or hosts with sterilizing immunity (i.e., if the host recognizes *any* antigens, it cannot be infected so $\sigma_n = 0$ for $n \geq 1$ and $\sigma_h = 0$). Consequently, for status-based models, there is no difference between full antigenic or homologous immunity (so $\sigma_{\mathrm{HI}} = 0$), and so no potential for sustained evolutionary oscillations.

Strain structure depends on the life-history effects of immune protection

Of course, immunity need not only protect against infection. Suppose in the model of Box 1 any of transmissibility, β, clearance rate, μ, and virulence mortality, α, depend upon the immune history of the host. Although in this case epistasis can be written as a time-varying quantity as in equation (4), the effective reproductive number of an infection will depend on its current environment (immune history of the host). Consequently, the per-capita growth rate of strain ij is the weighted average across possible environments, where the weights are the frequency of strain ij infections by environment (Sup. Info. S4). As such, time-varying epistasis is less amenable to interpretation, and so here we focus on epistasis at equilibrium. To do so, suppose homologous immunity is no more protective than full antigenic immunity. Then the population will evolve to a stable equilibrium, whether it be stable PSS or no strain structure. At the stable PSS equilibrium, epistasis shares the same sign as LD if

$$\epsilon_{\mathrm{PI}} \equiv f_0 - 2f_1 + f_2 \tag{5}$$

is positive (Sup. Info. S4). In equation (5), $f_n \equiv \frac{\sigma_n \beta_n}{\mu_n + \alpha_n + d}$ is proportional to the effective reproductive number of an infection in a host with immunity to n of its antigens (Sup. Info. S4). Thus, whether an infection in a host with immunity to one of its antigens has a smaller ($\epsilon_{\mathrm{PI}} > 0$) or larger ($\epsilon_{\mathrm{PI}} < 0$) effective reproductive number than expected determines the evolutionary outcome: if it is smaller than expected, PSS occurs.

Equation (5) reveals that the type of immune protection affects the evolutionary predictions in two ways. First, although increasing the protection offered by immunity to one antigen promotes PSS if immunity reduces susceptibility and/or transmissibility [6, 12, 15, 43] or increases clearance rate [18–20, 39], it inhibits PSS if immunity reduces virulence mortality [44, 45]. Second, if each antigen produces an additive increase in the strength of immunity [8], then if immunity reduces susceptibility to infection or transmissibility, $\epsilon_{\mathrm{PI}} = 0$, whereas if immunity increases clearance rate [a common assumption for malaria models 18–20], $\epsilon_{\mathrm{PI}} > 0$. Thus, immunity that increases the clearance rate is more likely to produce PSS. This is particularly true when pathogens produce chronic infections (larger d) and/or cause host death (larger α).

The derivation of equation (5) assumes that homologous immunity is equivalent to full antigenic immunity. This suggests that the analogous quantity to σ_{HI} is $\epsilon_{\mathrm{HI}} \equiv f_2 - f_h$. Consequently, $\epsilon_{\mathrm{HI}} > 0$ is a necessary condition for evolutionary oscillations and oscillatory PSS.

4.2 The dynamic immunological landscape

In addition to the specificity and strength of immune protections, epistasis also depends upon the dynamic immunological landscape controlled by $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$. When models become more complex than that of Box 1, this landscape plays an increasingly prominent role. However, because these complexities typically alter the dynamics of uninfected hosts, the analysis of equation (4) helps understand the influence of the immunological landscape on the presence of PSS. In particular, this analysis illustrates the impact of re-infections and how immunity is gained and lost.

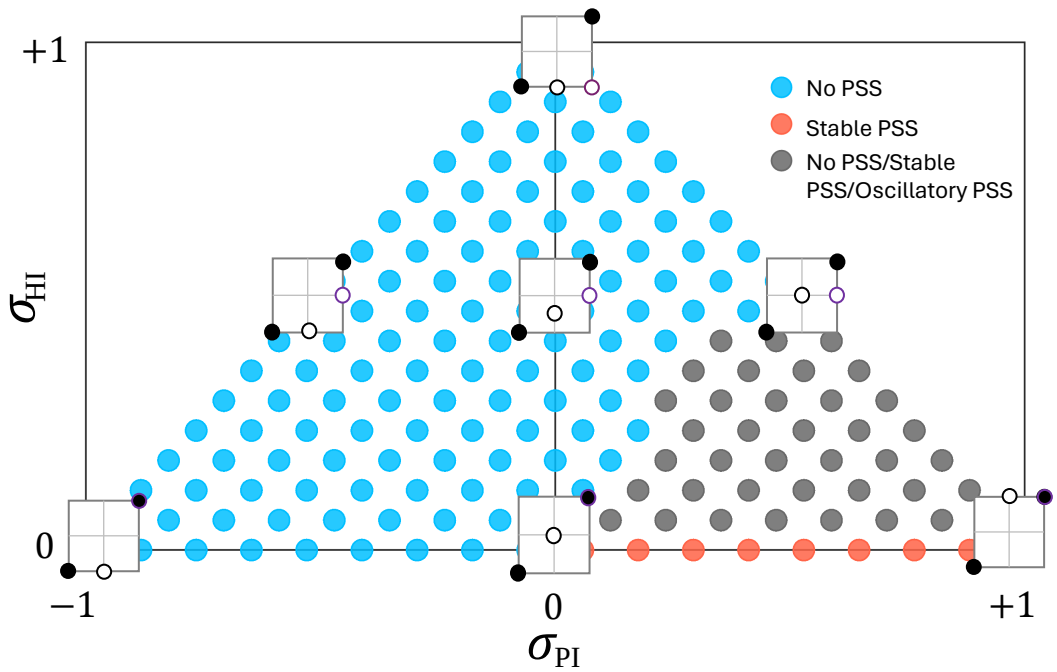

Figure 3: **The strength and specificity of cross-protection determines when PSS can form.** The functional form of cross-protection is controlled by σ_{HI} and σ_{PI} (see Fig. 2); as they vary, this permits the long-term existence of stable or oscillatory PSS or some combination of both. Each circle represents the long-term outcome of 100 simulations of equation (3), where we loop over 10 evenly-spaced values of $R_0 = \beta/(\mu + d) \in [1.1, 8]$ and $\theta = d/(\mu+d+\alpha) \in [0.01, 0.5]$. The insets show how immunity varies in $(\sigma_{\mathrm{PI}}, \sigma_{\mathrm{HI}})$ space using the same format as Figure 2. Note that the values of σ_{PI} and σ_{HI} are constrained due to the assumption that $\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq \sigma_h$. Representative dynamical behaviour through time for the different cases is shown in Figure S1. All simulations use parameters $b = d = 0.01$, $\alpha = 0$, $\sigma_0 = 1$, $\sigma_h = 0$, and $\rho = 10^{-6}$.

The frequency of re-infections shapes the likelihood of evolutionary oscillations

We have established that evolutionary oscillations require homologous immunity to be more protective than full antigenic immunity in order to produce a second epidemiological feedback acting on epistasis ($\mathcal{S}_\Delta(t)$). However, equation (4) illustrates that this condition is not sufficient. In addition, the two epidemiological feedbacks must sufficiently diverge from one another to produce counteracting effects. In particular, there needs to be a difference between the distribution of immunity to different antigens, $\mathcal{A}_\Delta(t)$, and different strains, $\mathcal{S}_\Delta(t)$. It can be shown that only hosts with immunity to the antigens defining a single strain contribute to $\mathcal{A}_\Delta(t)$, whereas $\mathcal{S}_\Delta(t)$ depends upon hosts with immunity to more than one strain (Fig. 4; Sup. Info. S3.2). Consequently, the 'length' of the immune history x, is critical (Fig. 4). Generally speaking, the 'length' of the immune history is simply the number of strains to which a host currently has immunity. However, in the absence of waning immunity, the 'length' of the immune history is the number of prior infections by genetically distinct strains. If hosts have immunity to only one strain, then $\mathcal{A}_\Delta(t) = \mathcal{S}_\Delta(t)$. In this case, equation (4)

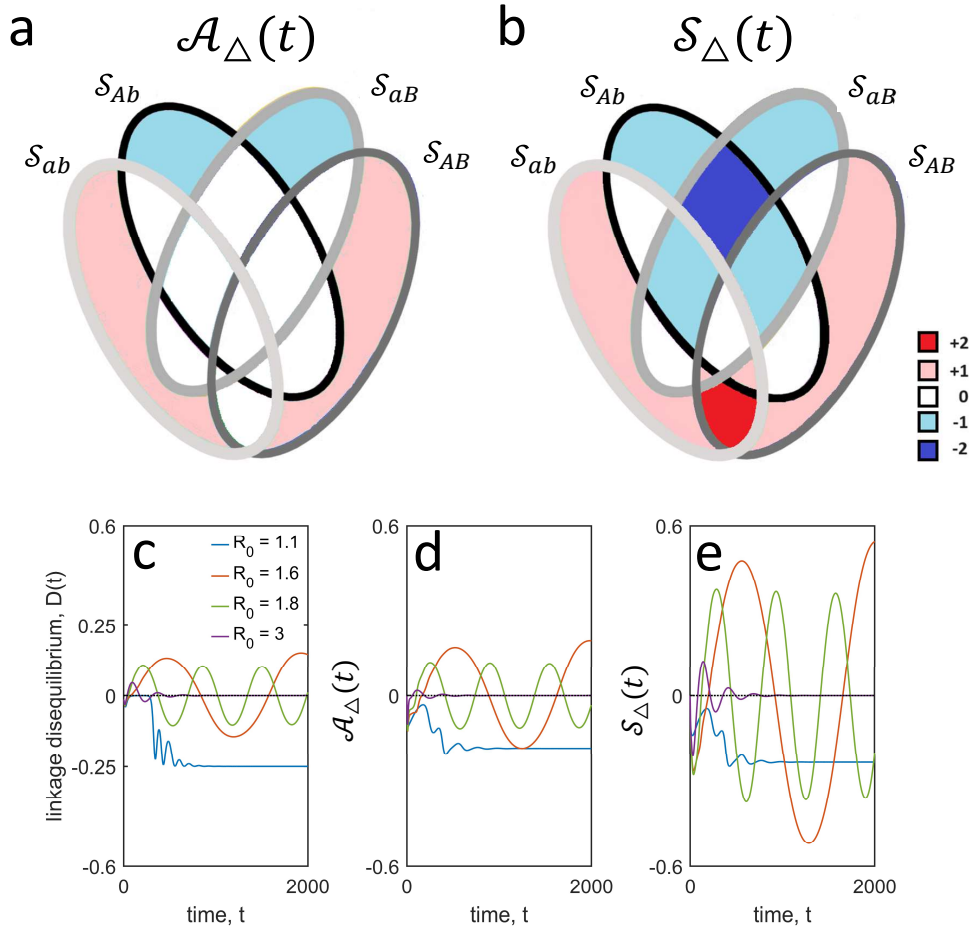

Figure 4: **The dynamic immunological landscape**. Panels **a** and **b** are Venn diagrams of $\mathcal{A}_\Delta(t)$ (**a**) and $\mathcal{S}_\Delta(t)$ (**b**). Each ellipse corresponds to a different $\mathcal{S}_{ij}(t)$, which is the density of hosts with immunity to *at least* strain ij. Then $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$ are computed as the weighted sum of the coloured regions, where the weights are colour-dependent. Although $\mathcal{A}_\Delta(t)$ only involves hosts with homologous immunity to a single strain, $\mathcal{S}_\Delta(t)$ involves hosts with immunity to multiple strains. Panels **c-e** show how the dynamic immunological landscape is shaped by the 'length' of the immune history, which in this case is controlled by the basic reproductive number, R_0. For small R_0 (blue curve), most hosts are infected a single time during their lives and so $\mathcal{A}_\Delta(t) \approx \mathcal{S}_\Delta(t)$. For intermediate R_0 (orange and green), hosts are infected multiple times. Here, $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$ diverge, sustaining evolutionary oscillations. For large R_0 (purple), hosts are always infected multiple times. Here, $\mathcal{A}_\Delta(t) \to 0$ faster than $\mathcal{S}_\Delta(t) \to 0$ and thus no PSS is possible. Parameter values used for the simulations of panels **c-e** are $\beta = 1$, $b = d = 0.01$, $\alpha = 0$, $\sigma_0 = 1$, $\sigma_1 = 0.5$, $\sigma_2 = 0.4$ (i.e. $\sigma_{\mathrm{PI}} = 0.2$), and $\sigma_h = 0.25$ (i.e. $\sigma_{\mathrm{HI}} = 0.15$). μ is chosen so that R_0 is equal to indicated value.

yields that stable PSS occurs only if $\sigma_{\mathrm{PI}} > \sigma_{\mathrm{HI}}$; sustained PSS oscillations are not possible. If, instead, almost all hosts have immunity to more than one strain, then $\mathcal{A}_\Delta(t) \to 0$ much faster than $\mathcal{S}_\Delta(t) \to 0$. Here, neither stable PSS nor oscillatory PSS are possible. Only

between these two extremes can oscillations occur.

Two generic pathogen life-history quantities affect the relationship between $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$. The first is the basic reproductive number, $R_0 \equiv \beta\sigma_0/(\mu+\alpha+d)$ (Fig. 4). When R_0 is small, hosts are unlikely to be infected more than once, so $\mathcal{A}_\Delta(t) \approx \mathcal{S}_\Delta(t)$. When R_0 is large, hosts being infected multiple times is common, so $\mathcal{A}_\Delta(t) \to 0$. Thus, oscillatory PSS is more likely for pathogens with intermediate R_0. Moreover, any PSS is unlikely for pathogens with high R_0. The second quantity is the ratio of the infectious period to host lifespan, which is $\theta \equiv d/(\mu+d+\alpha)$ [27, 46]. This ratio controls whether infections are acute (small θ) or chronic (large θ). The importance of θ hinges upon whether co-infections are possible. If co-infection is not possible, chronic infections reduce the likelihood of hosts being infected more than once, so $\mathcal{A}_\Delta(t) \approx \mathcal{S}_\Delta(t)$. If co-infection is possible, whether an infection is acute or chronic is less important as infected hosts remain susceptible to infection.

Because R_0 and θ are present in virtually every epidemiological model, they will influence the consequences of other factors altering the 'length' of immune memory. For example, if immunity wanes over time [47] the length of immune memory will be shortened by causing the host to 'forget' previous infections. Likewise, if there is short-term strain-transcending immunity [17, 41, 48–52] (analogous to no co-infection), the length of immune memory will alalso be shortened as the likelihood of becoming re-infected is reduced. If R_0 is very high, shortening the immune memory will increase the density of hosts contributing to $\mathcal{A}_\Delta(t)$. This increases the likelihood of both oscillatory and stable PSS. If R_0 is intermediate, the likelihood of oscillations will be reduced. This dependence upon R_0 explains why shortening immune memory through waning immunity [47], conserved immune responses [53], or strain-transcending immunity [41] has been observed to both decrease [47, 53] and increase [41] the likelihood of oscillations.

The impact of cross-protection depends on the gain and loss of immunity

The predicted consequences of the sign of σ_{PI} hinge upon the expectation that the distribution of infections is coupled to the immunological landscape. That is, we assume a positive correlation between $D(t)$ and $\mathcal{A}_\Delta(t)$. The intuition behind this expectation is that the more abundant set of strains will have caused more infections. This tightly couples the sign of $D(t)$ and the immunological landscape, $\mathcal{A}_\Delta(t)$. However, this relationship can be disrupted if a host infected by strain ij does not always simultaneously acquire and/or lose immunity to the combination of antigens $\{i,j\}$. For example, suppose host heterogeneity in antigen-presenting cells means that hosts are more likely to acquire immunity to only a single, randomly chosen antigen of an infecting strain rather than both. Moreover, suppose that strain set $\{ab, AB\}$ is over-represented, that is, $D(t) > 0$. Naive hosts are then most likely to become infected by one of the over-represented strains (say AB), after which they are most likely to become infected by the other over-represented strain (say ab) as it is more abundant. But if each infection provides immunity to a single antigen, following the second infection, the host will have immunity to either the set of antigens $\{A, b\}$ or $\{a, B\}$. Therefore, the host will have antigenic immunity to the under-represented strain set. Consequently, $\mathcal{A}_\Delta(t)$ will have the opposite sign to $D(t)$. Thus, equation (4) yields that PSS can only occur if $\sigma_{\mathrm{PI}} < 0$ (Sup. Info. S5).

Indeed, there are a variety of immunological reasons why hosts may not simultaneously gain and/or lose immunity to the combination of antigens $\{i,j\}$. For example, host variation

in the adaptive immune response might mean that different hosts target and/or lose immunity to different antigens, based upon host genetics (e.g., MHC alleles [8]) or randomly (e.g., which B-cell encounters which epitope first; see [54] and Sup. Info. S5 for a longer discussion). Importantly, this heterogeneity can disrupt the positive correlation between the distribution of infections and the immunological landscape, reversing predictions about the consequences of the sign of σ_{PI}.

4.3 The number of loci and the possible patterns of PSS

Many of the key predictions from our model of two diallelic loci hold as the number of alleles and number of loci increase: increasing the rate at which antigen-specific cross-protection increases with the number of antigens recognized by the host still promotes PSS, homologous immunity being more protective than full antigenic immunity still promotes strain diversity and is necessary for evolutionary oscillations, and oscillations are more likely when hosts tend to be infected by more than one strain. What changes is that as the number of loci increases, the population need not self-organize into sets of strains whose members have no antigens in common (DSS). As a consequence, different types of PSS are possible [9, 16, 28, 41]. This is because LD is no longer restricted to non-random associations spanning two loci. Instead, LD can involve as many loci as are segregating, with each order of LD corresponding to a different type of PSS. For example, triple LD is an over-representation of strains sharing one antigen. More generally, the number of over-represented strains increases with the order of LD. Although the evolutionary dynamics are more complicated, in the long-term at most one order of LD will be nonzero and evolve. This evolution is still dictated by epistasis, whether it be pairwise or higher-order (Sup. Info. S7).

When full antigenic immunity provides as much protection as homologous immunity, each order of epistasis hinges on the rate of increase of antigen-specific cross-protections. For N loci, this rate is captured by the $N-1$ composite parameters (Fig. 5a) ,

$$\sigma_{\mathrm{PI}}^{(n)} \equiv \sigma_{n-1} - 2\sigma_n + \sigma_{n+1}, \quad n = 1, ..., N-1. \tag{6}$$

As before, the sign of $\sigma_{\mathrm{PI}}^{(n)}$ determines whether the rate of increase of the strength of immune protections against n antigens is increasing, $\sigma_{\mathrm{PI}}^{(n)} > 0$, or decreasing, $\sigma_{\mathrm{PI}}^{(n)} < 0$. Thus, these parameters control how the strength of competition between strains changes based on genetic similarity. This, in turn, shapes epistasis (Fig. 5). For example, for three diallelic loci, triple epistasis in fitness, w_{ABC}, can maintain triple LD if $\sigma_{\mathrm{PI}}^{(2)} > \sigma_{\mathrm{PI}}^{(1)}$. On the other hand, pairwise LD between alleles i and j can be maintained if it shares the same sign as $w_{ij} + w_{ABC}/2$, which can be approximated as $\sigma_{\mathrm{PI}}^{(2)} > -\sigma_{\mathrm{PI}}^{(1)}$ (Sup. Info. S7). One important consequence of multiple possible types of PSS is that the conditions permitting different orders of LD to evolve can be simultaneously satisfied. When this occurs, an evolutionary multistability emerges such that the type of PSS present is contingent upon the initial conditions (Fig. 5; [see also 28]).

Given $N > 2$ loci and the possibility of multiple types of PSS, what genomic patterns are expected? This hinges upon the number of antigens sufficient to induce a strongly protective response. For example, if immunity to two antigens is strongly protective, $\sigma_{\mathrm{PI}}^{(1)} \to -1$ and $\sigma_{\mathrm{PI}}^{(2)} \to 1$, triple LD is more likely than pairwise LD (Fig. 5). If instead immunity to one antigen is strongly protective, $\sigma_{\mathrm{PI}}^{(1)} \to 1$ and $\sigma_{\mathrm{PI}}^{(2)} \to 0$, then pairwise LD is more likely

(Fig. 5). This explains why models, which assume immunity to one antigen is strongly cross-protective (i.e., $\sigma_1 = \sigma_2 = \cdots = \sigma_n$, and $\sigma_n \to \sigma_h$), predict DSS or pairwise LD [6, 12].

5 Discussion

Knowledge of when, and why, naturally acquired immunity produces non-random genomic patterns in pathogen populations (pathogen strain structure or PSS) is critical for understanding and predicting pathogen dynamics. Here we review and generalize predictions for the evolution of PSS by focusing on LD and the forces that drive its dynamics. In doing so, we establish the evolutionary connections between the multitude of existing models, providing a unifying perspective for understanding the evolution of PSS. In particular, our analysis illustrates the factors which determine (i) the long-term presence of PSS occurring as a stable equilibrium (stable PSS) and (ii) the long-term presence of PSS occurring as evolutionary oscillations (oscillatory PSS).

The long-term presence of stable PSS hinges upon two factors. First, stable PSS is more likely for pathogens against which prior exposure to *some* antigens of a given strain provides much stronger immune protection than expected ($\sigma_{\mathrm{PI}} > 0$), while prior exposure to *all* antigens of a given strain provides similar protection to homologous immunity ($\sigma_{\mathrm{HI}} \approx 0$). Second, stable PSS is more likely if hosts have immunity to a few strains. This occurs for pathogens with a low-to-intermediate R_0, pathogens causing chronic infections (if co-infection is limited), or when hosts display either waning or strain-transcending immunity. The most likely type of stable PSS, or order of LD, hinges upon the number of antigens required to induce a strongly protective immune response. As the number of antigens needed increases, so will the expected order of LD, and thus the fraction of possible strains present in the population. More broadly, although each type of PSS is in principle detectable from genomic data, it is necessary to test for the correct order of LD. For example, a highly structured population dominated by triple LD would appear unstructured if PSS is assessed based on the presence of pairwise LD. That the shape of the function controlling antigen-specific cross-protections (see Fig. 2 and Fig. 5**a**) can limit the amount of genotypic diversity provides testable predictions. For example, vaccine challenge studies of influenza A virus (IAV) in humans and equine influenza virus (EIV) in horses predict that antigen-specific cross-protection increases roughly linearly with antigenic distance for IAV, but more rapidly for EIV (see Fig. 1a in [55]). Our model would therefore predict that, all else being equal, EIV should be less *genotypically* diverse than IAV. Interestingly, EIV can be roughly divided into three antigenic clusters [56], whereas IAV consists of 11 antigenic clusters [57] (see also [58]). Of course, EIV and IAV differ with respect to other properties (e.g., IAV undergoes more rapid antigenic drift [58]); testing how the shape of the function controlling antigen-specific cross-protections limits genotypic diversity and constrains evolutionary dynamics warrants further work.

Evolutionary oscillations, or oscillatory PSS, have been a focal point of research for multistrain pathogens [e.g., 6, 9, 12, 16, 42]. Our multilocus perspective reveals that oscillatory PSS is most likely when two conditions are met (Fig. 3). First, the gain in protection from immunity to each additional antigen is larger than the additive expectation (σ_{PI} is positive), but only homologous immunity triggers a strongly protective immune response (so σ_{HI} is positive). Second, hosts are likely to be infected by, and so have homologous immunity to,

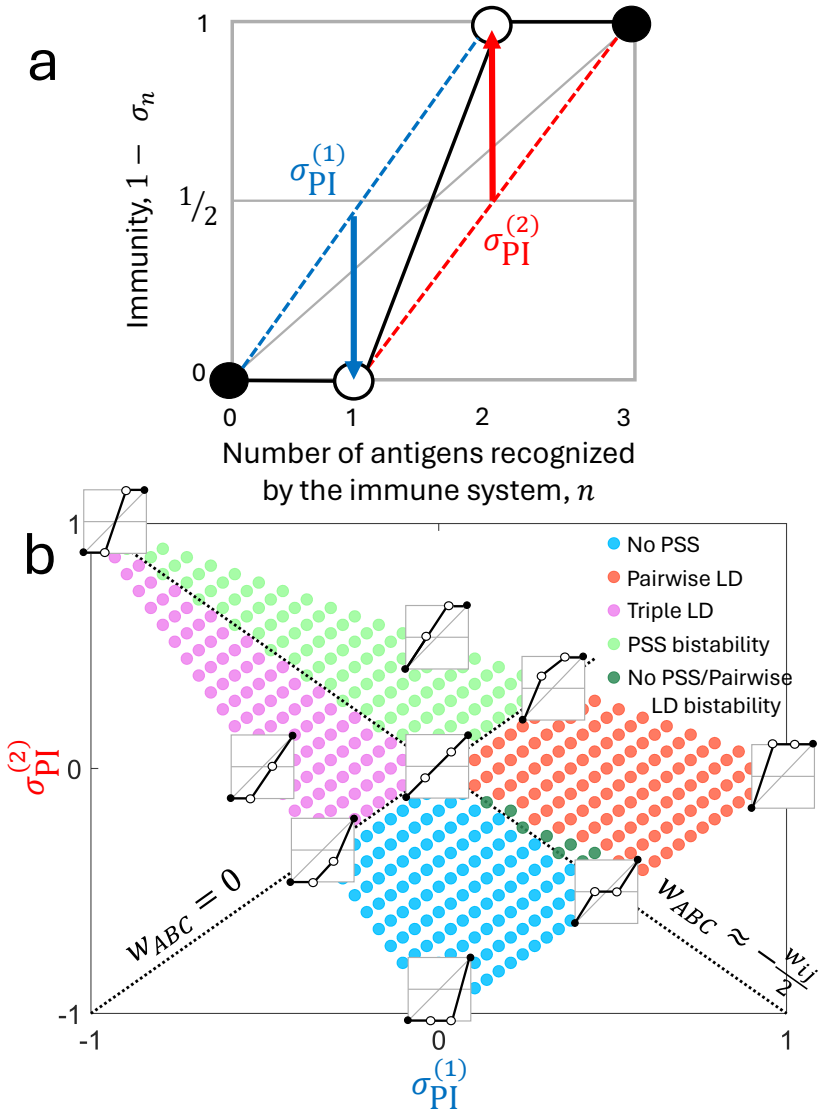

Figure 5: **As the number of loci increases, so do the possible types of PSS**. The number of loci determines the number of parameters capturing the shape of the function that determines how cross-protections increase with the number of antigens recognized by the immune system (panel **a**). As these parameters vary, different long-term outcomes are possible (panel **b**; see Fig. S5 for the representative dynamics through time). When $\sigma_{\mathrm{PI}}^{(2)} > \sigma_{\mathrm{PI}}^{(1)}$, triple LD shares the same sign as w_{ABC} and so can be stably maintained. When $\sigma_{\mathrm{PI}}^{(2)} > -\sigma_{\mathrm{PI}}^{(1)}$, pairwise LD between alleles i and j will generally share the same sign as $w_{ABC} + w_{ij}/2$ (Sup. Info. S7) and so can be stably maintained. As both conditions can be simultaneously met, evolutionary bistability can occur, where the type of PSS that evolves depends upon the initial conditions. Our assumption of $\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq \sigma_3$ constrains the possible values that $\sigma_{\mathrm{PI}}^{(i)}$ can take in panel **b**. Because $\sigma_{\mathrm{HI}} = 0$, when PSS occurs, it is stable. Parameter values used in panel **b**: $\beta = 1$, $d = b = 0.01$, $\alpha = 0$, $\mu = 0.4$, $\rho = 10^{-7}$, $\sigma_0 = 1$, $\sigma_3 = \sigma_h = 0$, whereas σ_1 and σ_2 vary.

multiple strains over the course of their lives. This is most likely for pathogens with either an intermediate R_0, or those with a high R_0 in the presence of waning or strain-transcending immunity (i.e., something shortening the 'length' of immune memory). Whether these conditions are met or not will depend upon the pathogen species. Identifying these key biological quantities may have greater utility than identifying regions of parameter space for a given model that permit oscillations. For example, Recker et al. [9] hypothesized that rather than being a mutation-driven process [13, 14, 50], influenza evolution may be better explained by the periodic recycling (or oscillations) of a finite set of antigens (but see [36]). Although this argument was supported by oscillations occurring over a wide range of parameter space, there is value in considering whether influenza meets the biological conditions for displaying such behaviour.

Co-infection and within-host interactions

The model of Box 1 assumes that transmission only occurs to uninfected hosts. It thus ignores within-host interactions. Multistrain models typically assume that hosts can be simultaneously infected by any number of *different* strains (co-infection), each of which is unaffected by the presence of other co-infecting strains [e.g., 6, 12, 15, 41, 52, 59, 60]. Under these assumptions, and by redefining certain variables in the model of Box 1, all our key predictions hold (Sup. Info. S2.3). However, this model of co-infection is biologically unsatisfying as it maintains high levels of genetic diversity, including otherwise selectively disadvantageous trait variation [e.g., hypervirulence 22]. Although simulations comparing different within-host interactions found limited effect of these interactions [18], this may not hold generally. For example, co-infection in *Streptococcus pneumonia* can lead to the simultaneous clearance of all infecting strains [8, 61, 62]. Since this will disproportionately (negatively) affect the more abundant strains, it will remove PSS. This may partially explain why *S. pneumoniae* genomic data shows low amounts of strain structure [63]. Thus the impact of within-host interactions on PSS warrants further work.

Demographic stochasticity

Our analysis is based on a deterministic model that neglects demographic stochasticity. Previous studies have shown that demographic stochasticity can have important effects on multistrain dynamics. First, demographic stochasticity can produce random oscillations not seen in a deterministic model [18, 64]. Because PSS is inhibited or promoted by epistasis, whenever epistasis is weak, demographic stochasticity can generate LD, cause LD to drift in a quasi-neutral fashion, and/or trigger infrequent oscillations between strains. Indeed, this is likely the source of the quasi-stable PSS observed in some stochastic models, in which the over-represented strain set infrequently changes [18]. For two diallelic loci, epistasis is weak if any of the following hold: $\sigma_{\mathrm{HI}} = 0$ and hosts are likely to be infected multiple times (R_0 is large); $\sigma_{\mathrm{HI}} \approx 0$ and $\sigma_{\mathrm{PI}} \approx 0$; and/or hosts are likely to be infected once and $\sigma_{\mathrm{HI}} \approx \sigma_{\mathrm{PI}}$.

Second, demographic stochasticity is expected to increase the chance loss of strains or antigens [21, 23, 48, 52, 65, 66]. This is especially relevant if the segregating loci have an unequal number of alleles. Here, antigens at the more diverse locus are predicted to be lost in finite populations, particularly when cross-protection is high [23]. This can be explained as follows (see also Sup. Info. S6). If one locus is more antigenically diverse than the others, e.g., the two loci are $\{a, A\}$ and $\{B_1, B_2, B_3\}$, it is not possible for each antigen to be present

at its expected frequency, $p_A(t) = 1/2$ and $p_{B_i}(t) = 1/3$, *and* LD between each $\{A, B_i\}$ pair to be maximal in magnitude. This creates tension between selection on antigens (through the additive selection coefficients) and selection on strains (through epistasis in fitness). This tension is modulated by the relative strength of σ_{PI} and σ_{HI}. As cross-protection increases, selection maintaining antigens at their expected frequencies becomes weaker (the strength of σ_{HI} decreases). At the same time, epistasis becomes stronger, favouring some strains (σ_{PI} becomes increasingly positive). Consequently, at the more diverse loci some antigens are forced to lower frequencies than expected and so are more likely to be stochastically lost. More generally, models incorporating an unequal number of alleles across loci disrupt the symmetry that permits the separation of the short- and long-term timescales as in equation (2). For example, when evolutionary oscillations occur, both allele frequencies and LD will fluctuate (rather than just LD). Understanding the resulting dynamics warrants future work.

The multilocus perspective clarifies the source of PSS and the explanations for antigenic variation

One of the strengths of the multilocus perspective is that it provides an evolutionary lens that clarifies the factors shaping the evolutionary dynamics of multistrain pathogens. For example, that epistasis is the source of PSS runs counter to the view that the source of PSS is NFDS [18–20, 67]. Although NFDS is required to maintain genetic polymorphisms, it need not yield epistasis (e.g., if $\sigma_{\mathrm{PI}} = \sigma_{\mathrm{HI}} = 0$), and it is epistasis that produces strain over-representation. Indeed, past work has compared time series of the genomic patterns produced by models of PSS against a null model in which antigenic variation is neutral [19, 20]. The multilocus perspective suggests an intermediate null model, in which there is NFDS but epistasis is zero (e.g., $\sigma_{\mathrm{PI}} = 0$, $\sigma_{\mathrm{HI}} = 0$). In such a model, NFDS would maintain antigenic variation but nothing would maintain distinct allelic combinations, leading to stochastic fluctuations in LD.

Second, a recent review considered different explanations for the abundance of antigenic variation [8]. In addition to host heterogeneity and functional constraints, the authors discussed three hypotheses. First, each antigen contributes to host immunity additively (so $\sigma_{\mathrm{PI}} = 0$); second, immunity to a single antigen provides strong cross-protection [i.e., $\sigma_1 = \sigma_{i+1}$, $i > 0$; 6, 12]; and third, there is an interaction between antigens at different loci. The multilocus perspective reveals considerable overlap between these hypotheses; for example, although the first and second predict, respectively, no PSS and stable pairwise LD (DSS), the third simply means that there is epistasis (e.g., $\sigma_{\mathrm{PI}} \neq 0$), which can lead to either situation, as well as higher-order LD. Indeed, as the number of antigens required to induce a strongly protective immune response determines the probable order of LD, the likelihood a novel antigen can evade population immunity increases with the order of LD. Thus, we would predict a positive correlation between the order of LD and the amount of allelic variation at each locus.

Multilocus models and vaccination

Understanding how genetically diverse pathogen populations respond to vaccination has been a focus of recent work [68–71]. Our analysis provides new insight into 'evolutionarily optimal' vaccine design, as the type of LD present is generally indicative of how many antigens are required to generate a strongly protective immune response. However, our analysis

has also emphasized some of the challenges associated with predicting the consequences of vaccination. First, the input of vaccinal immunity to the population may disrupt the relationship between strain frequency and population immunity with complex dynamical consequences [e.g., 10, 11, 72, 73]. Second, different conclusions will be reached based upon whether homologous immunity is more protective than full antigenic immunity ($\sigma_{\mathrm{HI}} > 0$), and if so, whether vaccination provides equivalent protection to full antigenic or homologous immunity. Third, the pathogen population response will hinge upon whether evolutionary bistability is possible, as the influx of vaccinal immunity may alter the type of PSS. Fourth, vaccines are often chosen to target a subset of strains, such as those linked to virulence factors [74]. The inclusion of additional pathogen life history variation, such as virulence factors, will induce fitness asymmetries among strains that are difficult to address with existing models. However, by emphasizing key evolutionary quantities, the multilocus perspective provides a useful frame of reference to tackle these questions and should be a starting point for future work.

Conclusion

Understanding the dynamics of multistrain pathogens is important not only to unravel what maintains and structures genetic diversity, but as a question of public health as many key human pathogen species, including malaria, dengue, *S. pneumoniae*, *N. meningitidis* and cytomegalovirus, consist of multiple strains. Indeed, the approach used here may be useful for understanding the ongoing evolutionary dynamics of SARS-CoV-2 as the build-up of natural immunity seems to be a key force. Yet combining pathogen genetic diversity with realistic epidemiological feedbacks remains a daunting task. Existing approaches often obscure the evolutionary predictions. One way forward is adopting a population genetics perspective. As we have shown here, this allows us to disentangle the factors driving the evolutionary dynamics, thus illustrating when, and why, pathogen strain structure occurs.

Acknowledgements

SG thanks Julia Gog and Sunetra Gupta for many inspiring discussions.

Notation	Description
$p_i(t)$, $p_{ij}(t)$	Frequency of infections by strains carrying allele(s) i and ij, respectively, at time t.
$D(t)$	Pairwise linkage disequilibrium (LD) between alleles A and B.
PSS	Pathogen strain structure; indicates the presence of LD of some order.
DSS	Discordant strain structure; is a specific form of PSS in which a set of strains sharing no alleles is over-represented.
w_k	Additive selection coefficient for allele k. Indicates whether allele k is favoured ($w_k > 0$) or disfavoured ($w_k < 0$).
w_{ij}, w_{ijk}	Epistasis between alleles i, j and i, j, k. If positive (resp. negative), the combination of alleles are more (resp. less) favoured than expected.
σ_n, σ_h	Probability of infection of a host with immunity to $n = 0, 1, 2, ...$ antigens, or homologous immunity h, of the transmitted strain
σ_{PI}	Controls the rate of increase of cross-protection as the number of antigens recognized by the immune system increases. If positive (resp. negative), cross protections increase more (resp. less) rapidly than linearly.
σ_{HI}	Controls whether homologous immunity is equivalent ($\sigma_{\mathrm{HI}} = 0$) or more ($\sigma_{\mathrm{HI}} > 0$) protective than full antigenic immunity.
$\mathcal{A}_\Delta(t)$	Indicates whether more uninfected hosts have full antigenic immunity to the set of strains $\{ab, AB\}$ ($\mathcal{A}_\Delta(t) > 0$) or $\{aB, Ab\}$ ($\mathcal{A}_\Delta(t) < 0$).
$\mathcal{S}_\Delta(t)$	Indicates whether more uninfected hosts have homologous immunity to the set of strains $\{ab, AB\}$ ($\mathcal{S}_\Delta(t) > 0$) or $\{aB, Ab\}$ ($\mathcal{S}_\Delta(t) < 0$).
f_n	Proportional to the effective reproductive number of an infection in a host with immunity to n of its antigens.
ϵ_{PI}	Indicates whether an infection in a host with immunity to one of its antigens has a smaller ($\epsilon_{\mathrm{PI}} > 0$) or larger ($\epsilon_{\mathrm{PI}} < 0$) effective reproductive number than expected.
ϵ_{HI}	Indicates whether an infection in a host with full antigenic immunity has an equivalent ($\epsilon_{\mathrm{HI}} = 0$) or larger ($\epsilon_{\mathrm{HI}} > 0$) effective reproductive number than an infection in a host with homologous immunity.
R_0	Basic reproductive number.
θ	Ratio of infectious period to host lifespan. If small, infections are acute; if large, infections are chronic.

Table 1: Key notation used in main text.

References

[1] Gupta, S., Trenholme, K., Anderson, R. M. & Day, K. P., 1994 Antigenic diversity and the transmission dynamics of *Plasmodium falciparum*. *Science* **263**, 961–963.

[2] Holmes, E. C. & Burch, S. S., 2000 The causes and consequences of genetic variation in dengue virus. *Trends Microbiol.* **8**, 74–77.

[3] Donati, C., Hiller, N. L., Tettelin, H., Muzzi, A., Croucher, N. J., Angiuoli, S. V., Oggioni, M., Hotopp, J. C. D., Hu, F. Z., Riley, D. R. *et al.*, 2010 *Genome Biology* **11**, R107.

[4] Lehner, R., Stamminger, T. & Mach, M., 1991 Comparative sequence analysis of human cytomegalovirus strains. *J. Clinical Microbiol.* **29**, 2494–2502.

[5] Maynard Smith, J., Smith, N. H., O'Rourke, M. & Spratt, B. G., 1993 How clonal are bacteria? *Proc. Natl. Acad. Sci.* **90**, 4384–4388.

[6] Gupta, S., Maiden, M. C. J., Feavers, I. M., Nee, S., May, R. M. & Anderson, R. M., 1996 The maintenance of strain structure in populations of recombining infectious agents. *Nature Medicine* **2**, 437–442.

[7] Lipsitch, M. & O'Hagan, J. J., 2007 Patterns of antigenic diversity and the mechanisms that maintain them. *J. R. Soc. Interface* **4**, 787–802.

[8] Georgieva, M., Buckee, C. O. & Lipsitch, M., 2019 Models of immune selection for multi-locus antigenic diversity of pathogens. *Nature Reviews Immunology* **19**, 55–62.

[9] Recker, M., Pybus, O. G., Nee, S. & Gupta, S., 2007 The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc. Natl. Acad. Sci.* **104**, 7711–7716.

[10] Lipsitch, M., 1997 Vaccination against colonizing bacteria with multiple serotypes. *Proc. Natl. Acad. Sci.* **94**, 6571–6576.

[11] Gupta, S., Ferguson, N. M. & Anderson, R. M., 1997 Vaccination and the population structure of antigenically diverse pathogens that exchange genetic material. *Proc. R. Soc. B* **264**, 1435–1443.

[12] Gupta, S., Ferguson, N. & Anderson, R., 1998 Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* **280**, 912–915.

[13] Gog, J. & Grenfell, B. T., 2002 Dynamics and selection of many-strain pathogens. *Proc. Natl. Acad. Sci.* **99**, 17209–17214.

[14] Gomes, M. G. M., Medley, G. F. & Nokes, D. J., 2002 On the determinants of population structure in antigenically diverse pathogens. *Proc. Roy. Soc. B* **269**, 227–233.

[15] Ferguson, N. M. & Andreasen, V., 2002 The influence of different forms of cross-protective immunity on the population dynamics of antigenically diverse pathogens. In *Mathematical Approaches for Emerging and Re-emerging Infectious Diseases: Models, Methods and Theory* (eds. C. Castillo-Chavez, S. Blower, P. van den Driessche, D. Kirschner & A. Yakubu), volume 125, chapter 9, pp. 157–171. Springer.

[16] Recker, M. & Gupta, S., 2005 A model for pathogen population structure with cross-protection depending on the extent of overlap in antigenic variant repertoires. *J. Theor. Biol.* **232**, 363–373.

[17] Cobey, S. & Lipsitch, M., 2012 Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science* **335**, 1376–1380.

[18] Artzy-Randrup, Y., Rorick, M. M., Day, K., Chen, D., Dobson, A. P. & Pascual, M., 2012 Population structuring of multi-copy, antigen-encoding genes in *Plasmodium falciparum*. *eLife* **1**, e00093.

[19] He, Q., Pilosof, S., Tiedje, K. E., Ruybal-Pesantez, S., Artzy-Randrup, Y., Baskerville, E. B., Day, K. P. & Pascual, M., 2018 Networks of genetic similarity reveal non-neutral processes shape strain structure in *Plasmodium falciparum*. *Nat. Comm.* **9**, 1817.

[20] Pilosof, S., He, Q., Tiedje, K. E., Ruybal-Pesantez, S., Day, K. P. & Pascual, M., 2019 Competition for hosts modulates vast antigenic diversity to generate persistent strain structure in *Plasmodium falciparum*. *PLoS Biology* **17**, e3000336.

[21] Buckee, C. O., Koelle, K., Mustard, M. J. & Gupta, S., 2004 The effects of host contact network structure on pathogen diversity and strain structure. *Proc. Natl. Acad. Sci.* **101**, 10839–10844.

[22] Buckee, C. O., Jolley, K. A., Recker, M., Penman, B., Kriz, P., Gupta, S. & Maiden, M. C. J., 2008 Role of selection in the emergence of lineages and the evolution of virulence in *Neisseria meningitidis*. *Proc. Natl. Acad. Sci.* **105**, 15082–15087.

[23] Buckee, C. O., Recker, M., Watkins, E. R. & Gupta, S., 2011 Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations. *Proc. Natl. Acad. Sci.* **108**, 15504–15509.

[24] Adams, B. & Sasaki, A., 2007 Cross-immunity, invasion and coexistence of pathogen strains in epidemiological models with one-dimensional antigenic space. *Math. Biosci.* **210**, 680–699.

[25] Adams, B. & Sasaki, A., 2009 Antigenic distance and cross-immunity, invasibility and coexistence of pathogen strains in an epidemiological model with discrete antigenic space. *Theor. Pop. Biol.* **76**, 157–167.

[26] Ferguson, N. M. & Galvani, A. P., 2003 The impact of antigenic variation on pathogen population structure, fitness and dynamics. In *Antigenic variation* (eds. A. Craig & A. Scherf), chapter 18, pp. 403–432. Academic Press.

[27] Dawes, J. H. P. & Gog, J. R., 2002 The onset of oscillatory dynamics in models of multiple disease strains. *J. Math. Biol.* **45**, 471–510.

[28] Calvez, V., Korobeinikov, A. & Maini, P. K., 2005 Cluster formation for multi-strain infections with cross-immunity. *J. Theor. Biol.* **233**, 75–83.

[29] Wikramaratna, P. S., Kucharski, A., Gupta, S., Andreasen, V., McLean, A. R. & Gog, J. R., 2015 Five challenges in modelling interacting strain dynamics. *Epidemics* **10**, 31–34.

[30] Kucharski, A. J., Andreasen, V. & Gog, J. R., 2016 Capturing the dynamics of pathogens with many strains. *J. Math. Biol.* **72**, 1–24.

[31] Slatkin, M., 2008 Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics* **9**, 477–485.

[32] Gupta, S. & Anderson, R. M., 1999 Population structure of pathogens: the role of immune selection. *Parasitology Today* **15**, 497–501.

[33] Gupta, S. & Maiden, M., 2001 Exploring the evolutionary diversity in pathogen populations. *Trends in Microbiology* **9**, 181–185.

[34] Gog, J. R. & Swinton, J., 2002 A status-based approach to multiple strain dynamics. *J. Math. Biol.* **44**, 169–184.

[35] Lange, A. & Ferguson, N. M., 2009 Antigenic diversity, transmission mechanisms, and the evolution of pathogens. *PLoS Comp. Biol.* **5**, e1000536.

[36] Zinder, D., Bedford, T., Gupta, S. & Pascual, M., 2013 The roles of competition and mutation in shaping antigenic and genetic diversity in influenza. *PLoS Pathogens* **9**, e1003104.

[37] Hastings, I. M. & Wedgwood-Oppenheim, B., 1997 Sex, strains and virulence. *Parasitology Today* **13**, 375–383.

[38] Recker, M., Arinaminpathy, N. & Buckee, C. O., 2008 The effects of a partitioned *var* gene repertoire of *Plasmodium falciparum* on antigenic diversity and the acquisition of clinical immunity. *Malaria Journal* **7**, 18.

[39] Severins, M., Klinkenberg, D. & Heesterbeek, H., 2012 How selection forces dictate the variant surface antigens used by malaria parasites. *J. Roy. Soc. Interface* **9**, 246–260.

[40] Palmer, G. H., Bankhead, T. & Seifert, H. S., 2016 Antigenic variation in bacterial pathogens. *Microbiology Spectrum* **4**, VMBF–0005–2015.

[41] Minayev, P. & Ferguson, N., 2009 Improving the realism of deterministic multi-strain models: implications for modelling influenza A. *J. Roy. Soc. Interface* **6**, 509–518.

[42] Ferguson, N., Anderson, R. & Gupta, S., 1999 The effect of antibody-dependent enhancement of the transmission dynamics and persistence of multiple-strain pathogens. *Proc. Natl. Acad. Sci.* **96**, 790–794.

[43] White, L. J., Cox, M. J. & Medley, G. F., 1998 Cross immunity and vaccination against multiple microparasite strains. *IMA Journal of Math. Applied in Med. Biology* **15**, 211–233.

[44] Adams, B. & Boots, M., 2006 Modelling the relationship between antibody-dependent enhancement and immunological distance with application to dengue. *J. Theor. Biol.* **242**, 337–346.

[45] Kawaguchi, I., Sasaski, A. & Boots, M., 2003 Why are dengue viruse serotypes so distantly related? Enhancement and limiting serotype similarity between dengue virus strains. *Proc. R. Soc. B* **270**, 2241–2247.

[46] Andreasen, V., Lin, J. & Levin, S. A., 1997 The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.* **35**, 825–842.

[47] Watkins, E. R., Grad, Y. H., Gupta, S. & Buckee, C. O., 2014 Contrasting within- and between-host immune selection shapes *Neisseria* opa repertoires. *Sci. Reports* **4**, 6554.

[48] Abu-Raddad, L. J. & Ferguson, N. M., 2004 The impact of cross-immunity, mutation and stochastic extinction on pathogen diversity. *Proc. Roy. Soc. B* **271**, 2431–2438.

[49] Wearing, H. J. & Rohani, P., 2006 Ecological and immunological determinants of dengue epidemics. *Proc. Natl. Acad. Sci.* **103**, 11802–11807.

[50] Ferguson, N. M., Galvani, A. P. & Bush, R. M., 2003 Ecological and immunological determinants of influenza evolution. *Nature* **422**, 428–433.

[51] Nagao, Y. & Koelle, K., 2008 Decreases in dengue transmission may act to increase the incidence of dengue hemorrhagic fever. *Proc. Natl. Acad. Sci.* **105**, 2238–2243.

[52] Minayev, P. & Ferguson, N., 2009 Incorporating demographic stochasticity into multi-strain epidemic models: application to influenza A. *J. R. Soc. Interface* **6**, 989–996.

[53] Gupta, S. & Galvani, A., 1999 The effects of host heterogeneity on pathogen population structure. *Proc. R. Soc. B* **354**, 711–719.

[54] Cobey, S. & Pascual, M., 2011 Consequences of host heterogeneity, epitope immunodominance, and immune breadth for strain competition. *J. Theor. Biol.* **270**, 80–87.

[55] Park, A. W., Daly, J. M., Lewis, N. S., Smith, D. J., Wood, J. L. N. & Grenfell, B. T., 2009 Quantifying the impact of immune escape on transmission dynamics of influenza. *Science* **326**, 726–728.

[56] Lewis, N. S., Daly, J. M., Russell, C. A., Horton, D. L., Skepner, E., Bryant, N. A., Burke, D. F., Rash, A. S., Wood, J. L., Chambers, T. M. *et al.*, 2011 Antigenic and genetic evolution of equine influenza a (h3n8) virus from 1968 to 2007. *J Virol* **85**, 12742–12749.

[57] Smith, D. J., Lapedes, A. S., de Jong, J. C., Bestebroer, T. M., Rimmelzwaan, G. F., Osterhaus, A. D. M. E. & Fouchier, R. A. M., 2004 Mapping the antigenic and genetic evolution of influenza virus. *Scinece* **305**, 371–376.

[58] Whitlock, F., Murcia, P. R. & Newton, J. R., 2022 A review on equine influenza from a human influenza perspective. *Viruses* **14**, 1312.

[59] Kucharski, A. J. & Gog, J. R., 2012 Age profile of immunity to influenza: effect of original antigenic sin. *Theor. Pop. Biol.* **81**, 102–112.

[60] Hamelin, F. M., Allen, L. J. S., Bokil, V. A., Gross, L. J., Hilker, F. M., Jeger, M. J., Manore, C. A., Power, A. G., Rua, M. A. & Cunniffe, N. J., 2019 Coinfections by noninteracting pathogens are not independent and require new tests of interaction. *PLoS Biol.* **17**, e30000551.

[61] Trzcinski, K., Thompson, C. M., Srivastava, A., Basset, A., Malley, R. & Lipsitch, M., 2008 Protection against nasopharyngeal colonization by *Streptococcus pneumoniae* is mediated by antigen-specific CD4+ T cells. *Infect. Immun.* **76**, 2678–2684.

[62] Li, Y., Gierahn, T., Thompson, C. M., Trzcinski, K., Ford, C. B., Croucher, N., Gouveia, P., Flechtner, J. B., Malley, R. & Lipsitch, M., 2012 Distinct effects on diversifying selection by two mechanisms of immunity against *Streptococcus pneumoniae*. *PLoS Pathogens* **8**, e1002989.

[63] Croucher, N. J., Campo, J. J., Le, T. Q., Liang, X., Bentley, S., Hanage, W. P. & Lipsitch, M., 2017 Diverse evolutionary patterns of pneumococcal antigens identified by pangenome-wide immunological screening. *Proc. Natl. Acad. Sci.* pp. E357–E366.

[64] Koelle, K., Pascual, M. & Yunus, M., 2006 Serotype cycles in cholera dynamics. *Proc. Roy. Soc. B* **273**, 2879–2886.

[65] Cummings, D. A. T., Schwartz, I. B., Billings, L. B., Shaw, L. B. & Burke, D. S., 2005 Dynamic effects of antibody-dependent enhancement on the fitness of viruses. *Proc. Natl. Acad. Sci.* **102**, 15259–15264.

[66] Restif, O. & Grenfell, B. T., 2007 Vaccination and the dynamics of immune evasion. *J. R. Soc. Interface* **4**, 143–153.

[67] He, Q., Pilosof, S., Tiedje, K. E., Day, K. P. & Pascual, M., 2021 Frequency-dependent competition between strains imparts persistence to perturbations in a model of *Plasmodium falciparum* malaria transmission. *Frontiers in Ecology and Evolution* **9**, 633263.

[68] Watkins, E. R., Penman, B. S., Lourenco, J., Buckee, C. O., Maiden, M. C. J. & Gupta, S., 2015 Vaccination drives changes in metabolic and virulence profiles of *Streptococcus pneumoniae*. *PLoS Pathogens* **11**, e1005034.

[69] Azarian, T., Martinez, P. P., Arnold, B. J., Qiu, X., Grant, L. R., Corander, J., Fraser, C., Croucher, N. J., Hammitt, L. L., Reid, R. *et al.*, 2020 Frequency-dependent selection can forecast evolution in *Streptococcus pneumonaie*. *PLoS Biology* **18**, e3000878.

[70] Lourenco, J., Daon, Y., Gori, A. & Obolski, U., 2021 Pneumococcal competition modulates antibiotic resistance in the pre-vaccination era: a modelling study. *Vaccines* **9**, 265.

[71] Obolski, U., Lourenco, J., Thompson, C., Thompson, R., Gori, A. & Gupta, S., 2018 Vaccination can drive an increase in frequencies of antibiotic resistance among nonvaccine serotypes of *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci.* **115**, 3102–3107.

[72] McLean, A. R., 1995 Vaccination, evolution and changes in the efficacy of vaccines: a theoretical framework. *Proc. Roy. Soc. B* **261**, 389–393.

[73] Pitzer, V. E., Patel, M. M., Lopman, B. A., Viboud, C., Parashar, U. D. & Grenfell, B. T., 2011 Modeling rotavirus strain dynamics in developed countries to understand the potential impact of vaccination on genotype distributions. *Proc. Natl. Acad. Sci.* **108**, 19353–19358.

[74] Colijn, C., Corander, J. & Croucher, N. J., 2020 Designing ecologically optimized pneumococcal vaccines using population genomics. *Nature Microbiology* **5**, 473–485.

[75] Andreasen, V. & Sasaki, A., 2006 Shaping the phylogenetic tree of influenza by cross-immunity. *Theor. Pop. Biol.* **70**, 164–173.

[76] Kryazhimskiy, S., Dieckmann, U., Levin, S. A. & Dushoff, J., 2007 On state-space reduction in multi-strain pathogen models, with an application to antigenic drift in influenza A. *PLoS Comp. Biol.* **3**, 1513–1525.

[77] Day, T. & Gandon, S., 2012 The evolutionary epidemiology of multilocus drug resistance. *Evolution* **66**, 1582–1597.

[78] McLeod, D. V. & Gandon, S., 2021 Understanding the evolution of multiple drug resistance in structured populations. *eLife* **10**, e65645.

[79] Thomson, G. & Baur, M. P., 1984 Third order linkage disequilibrium. *Tissue Antigens* **24**, 250–255.

[80] Robinson, W. P., Asmussen, M. A. & Thomson, G., 1991 Three-locus systems impose additional constraints on pairwise disequilibria. *Genetics* **129**, 925–930.

[81] Hurford, A., Cownden, D. & Day, T., 2010 Next-generation tools for evolutionary invasion analyses. *J. Roy. Soc. Interface* **7**, 561–571.

[82] Kouyos, R. D., Fouchet, D. & Bonhoeffer, S., 2009 Recombination and drug resistance in HIV: population dynamics and stochasticity. *Epidemics* **1**, 58–69.

[83] Mani, R., St. Onge, R. P., Hartman, J. L., Giaever, G. & Roth, F. P., 2008 Defining genetic interaction. *Proc. Natl. Acad. Sci.* **105**, 3461–3466.

Supplementary Information

Contents

S1 Epidemiological model **S1**
S1.1 Basic model without adaptive immunity S1
S1.2 Specification of adaptive immunity S2
S1.3 Epidemiological dynamics and per-capita growth rates S2

S2 Biological connections between existing models **S3**
S2.1 Immune protections S3
S2.2 Acquisition of immunity S4
S2.3 Co-infection S5

S3 Separation of evolutionary and epidemiological dynamics **S6**
S3.1 The additive selection coefficients and epistasis in fitness S7
S3.2 Relationship between $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$ S8
S3.3 Equations for the evolutionary dynamics S9
S3.4 Fast- and slow-timescales in the evolutionary dynamics S9

S4 Different types of immune protections **S11**
S4.1 Epistasis must share same sign as LD for local asymptotic stability S11
S4.2 Epistasis at equilibrium S12

S5 The link between strain frequencies and the immunological landscape **S13**

S6 Variation in allelic diversity across loci **S15**

S7 More than two loci **S16**
S7.1 Fast- and slow-timescales, and different types of strain structure S19
S7.2 Shape of function controlling antigen-specific cross-protections S20
S7.3 $\sigma_{\mathrm{HI}} > 0$ promotes oscillations and maintains diversity S23

S1 Epidemiological model

S1.1 Basic model without adaptive immunity

In the absence of immunity, the host-pathogen model considered here is straightforward: the pathogen is transmitted to uninfected hosts by mass-action with rate constant β while infections are naturally cleared by the host at a per-capita rate μ. Hosts enter the population at a fixed rate b, and are removed from the population due to natural- and virulence-related mortality (if infected) at per-capita rates d and α, respectively. Hosts cannot be simultaneously infected by multiple strains (no co-infection); we later discuss how co-infection affects the dynamics. Let $S(t)$ and $I(t)$ denote the density of uninfected hosts, and total density of

infected hosts, respectively, at time t. Then the epidemiological dynamics are

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= b - dS - \beta SI + \mu I \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= (\beta S - d - \mu - \alpha)I. \end{aligned} \tag{S1}$$

S1.2 Specification of adaptive immunity

Adaptive immunity is built into the basic model of equations (S1) as follows. We consider a pathogen whose genotype consists of some number of multiallelic loci. The host can develop an adaptive immune response to the alleles (antigens) at each of these loci which can potentially reduce:

(i) The susceptibility to infection, and we let σ_x^g denote the probability of infection of a host with immune history x by strain (genotype) g.

(ii) The transmissibility of an infection, and we let β_x^g denote the transmissibility of an infection by strain g in a host with immune history x.

(iii) The duration of infection, and we let μ_x^g denote the per-capita rate at which a host with immune history x clears an infection by strain g.

(iv) The virulence-related mortality, and we let α_x^g denote the per-capita rate at which a host with immune history x dies due to an infection by strain g.

For each of these protections, we distinguish between immunity to antigens and immunity to previously-infecting strains (homologous immunity). In particular, for a host with immunity to $n = 0, 1, 2, ...$ antigens of strain g, the immune protections are

$$(\sigma_x^g, \beta_x^g, \mu_x^g, \alpha_x^g) = (\sigma_n, \beta_n, \mu_n, \alpha_n),$$

while for a host with homologous immunity to strain g, the immune protections are

$$(\sigma_x^g, \beta_x^g, \mu_x^g, \alpha_x^g) = (\sigma_h, \beta_h, \mu_h, \alpha_h).$$

We assume homologous immunity is at least as strong as full antigenic immunity, and that the host immune response is always the most protective available to the host, based upon its immune history. For example, for a two diallelic loci pathogen genome, $\{a, A\}$ and $\{b, B\}$, then a host with immune history $x' = \{ab, Ab\}$ has immune protections

$$(c_{x'}^{ab},\, c_{x'}^{Ab},\, c_{x'}^{aB},\, c_{x'}^{AB}) = (c_h,\, c_h,\, c_1,\, c_1), \quad c \in \{\sigma, \beta, \mu, \alpha\}.$$

S1.3 Epidemiological dynamics and per-capita growth rates

Let $S_x(t)$ and $I_x^g(t)$ denote the density of uninfected hosts with prior infection by strains belonging to the set x (so x is the immune history), and density of strain g infections in hosts with immune history x, respectively, at time t. Then the epidemiological dynamics are

$$\begin{aligned} \frac{\mathrm{d}S_x}{\mathrm{d}t} &= b\mathbf{1}_0(x) - dS_x - \sum_g \sigma_x^g \sum_z \beta_z^g I_z^g S_x + \Gamma_x \\ \frac{\mathrm{d}I_x^g}{\mathrm{d}t} &= \sum_z \beta_z^g I_z^g \sigma_x^g S_x - (d + \mu_x^g + \alpha_x^g) I_x^g + \Delta_{g,x}\rho \end{aligned} \tag{S2}$$

where $\mathbf{1}_0(x) = 1$ if $x = 0$ and 0 otherwise, $\Delta_{g,x}\rho$ is the net change due to recombination of strain g infections in hosts with immune history x (for now, left arbitrary), and Γ_x controls the rate at which hosts gain (or lose) the immune history x. If immunity is life-long and is updated when infections are cleared,

$$\Gamma_x = \sum_g \sum_{z:g\cup z=x} \mu_z^g I_z^g. \tag{S3}$$

Using this model, we let r_g denote the per-capita growth rate of strain g; this is the growth rate of strain g averaged across the different immune environments x:

$$r_g = \sum_x \frac{\mathrm{d}I_x^g/\mathrm{d}t}{I^g} = \sum_x \sum_z \beta_z^g \frac{I_z^g}{I^g} \sigma_x^g S_x - d - \sum_x (\mu_x^g + \alpha_x^g) \frac{I_x^g}{I^g}, \tag{S4}$$

where $I^g = \sum_x I_x^g$ is the total density of strain g. In Box 1 we present a version of this model where we have restricted immune protections to reducing σ_x^g. That is, immunity only reduces host susceptibility to infection.

S2 Biological connections between existing models

Next, we consider the relationship between model (S2) and existing models. Multistrain models typically diverge from one another with respect to three biological properties:

(i) the immune protections, and how these are affected by antigenic distance;

(ii) the specification of how hosts acquire immunity, that is, the form of Γ_x; and

(iii) whether co-infection is permitted or not.

In what follows we discuss each of these in more detail, and in Table S2 we summarize how the model (S2) relates to existing models. Before doing so, it is important to note that the combination of assumptions employed by previous models are typically made because they reduce the number of equations. Doing so makes the problem more tractable for analysis and simulation. As our purpose here is to understand and, where possible, generalize the biological predictions of these models, we will simply state the assumptions made by each model without showing the mathematical benefits of these assumptions. We refer interested readers to Kucharski & Gog [30], who provide a more detailed treatment of the mathematics behind some of the commonly used models.

S2.1 Immune protections

Multistrain models vary in what type of protection immunity provides, as well as how prior infections combine to determine the strength of immunity.

First, most models assume immunity reduces transmissibility or susceptibility; the choice of which is typically for model-specific mathematical reasons. For example, 'status'-based models often assume reduced susceptibility [13, 27, 34], while 'history'-based models often assume reduced transmissibility [6, 9, 12, 16] (but not always, [14, 46]). More recently, simulation models have assumed immune protections reduce duration of carriage, as this is more

biologically realistic for malaria [18–20]. Finally, some models of antibody-dependent enhancement have allowed immunity to reduce virulence-related mortality [44, 45]. Although immune protections reducing virulence-related mortality can have a qualitatively different effect [e.g., 44], it is generally believed that there is not a qualitative difference between protections reducing susceptibility, transmissibility, or clearance [15, 26]. At least one consequence of the differences in immune protections is that if immunity reduces susceptibility, hosts are less likely to acquire immunity to different strains, provided immunity is only acquired after successful infection [15].

Second, models differ in how multiple prior infections combine to determine the strength of immunity. We will illustrate these differences using the example immune history $x'' = \{Ab, aB\}$, so we are focusing upon a pathogen genome consisting of two diallelic loci. The most common approach is to assume immunity against a particular strain is determined by the *single* most antigenically-similar previously infecting strain [6, 9, 12, 16, 25]. Using the immune history x'',

$$(\sigma_{x''}^{ab}, \sigma_{x''}^{Ab}, \sigma_{x''}^{aB}, \sigma_{x''}^{AB}) = (\sigma_1, \sigma_h, \sigma_h, \sigma_1).$$

Less commonly, some models assume that immunity is the *product* of the individual cross-protections against previously infecting strains [14, 24]. For the immune history x'',

$$(\sigma_{x''}^{ab}, \sigma_{x''}^{Ab}, \sigma_{x''}^{aB}, \sigma_{x''}^{AB}) = \left((\sigma_1)^2, \sigma_h, \sigma_h, (\sigma_1)^2\right).$$

Finally, immunity to each antigen can also be treated independently, and so the strength of immunity is *antigen-specific* and depends upon the total number of antigens to which the host has protection against [18–20]. For the immune history x'',

$$(\sigma_{x''}^{ab}, \sigma_{x''}^{Ab}, \sigma_{x''}^{aB}, \sigma_{x''}^{AB}) = (\sigma_2, \sigma_h, \sigma_h, \sigma_2).$$

Note that although some models restrict immune memory to only the most recent infections [65, 75], we do not consider this further here as this is more similar to a form of waning immunity.

S2.2 Acquisition of immunity

The model of (S2) assumes that after hosts have naturally cleared an infection, they have homologous immunity to the infecting strain, as well as some degree of protection against antigenically-similar strains. Such a framework has been referred to as a 'history-based' model [13, 34], and is a commonly taken approach [6, 12, 14, 46]. An alternative to history-based models are so-called 'status-based' models [13, 27, 34]. Status-based models assume each host is either completely protected, or completely susceptible to a given strain (so immunity reduces susceptibility); there is no 'partial' protection at the level of the individual host. Thus, following clearance of an infection (or at time of infection, depending upon when immunity is acquired), each host becomes either fully-protected or fully-sensitive to each antigenically similar strain. Then the probability of acquiring protection against a particular strain depends upon the antigenic similarity between that strain and the other strains to which the host has protection against [13, 27, 34].

To convert model (S2) to a status-based model, we first restrict immune protections to

reducing susceptibility, such that

$$\sigma_x^g = \begin{cases} 0 & \text{if } g \in x \\ 1 & \text{otherwise} \end{cases}, \quad \text{while} \quad (\beta_x^g, \mu_x^g, \alpha_x^g) = (\beta, \mu, \alpha).$$

Then we suppose that upon clearing an infection by strain g, a host with immune history z becomes a host with immune history x with probability $p_{z\to x}^g$ (so $\sum_x p_{z\to x}^g = 1$), and thus

$$\Gamma_x = \sum_g \sum_z \mu_z^g p_{z\to x}^g I_z^g. \tag{S5}$$

Under these assumptions, model (S2) is a status-based model.

S2.3 Co-infection

The model of (S2) assumes that transmission only occurs to uninfected hosts, and thus hosts cannot be simultaneously infected by different strains (co-infection). This assumption was made primarily for modelling clarity, however, it is biologically reasonable if, for example, the pathogen is not circulating at a high prevalence, infections are acute, and/or infection induces short-term strain-transcending immunity. Yet these conditions need not apply generally, and in fact, co-infection is a common attribute of multistrain models. Virtually all non-simulation models that include co-infection make the following assumptions [6, 12, 15, 41, 52]:

(i) Hosts can be simultaneously infected by any number of *different* strains and these strains do not interact. Thus, transmission and clearance rates are unaffected by the presence, or number, of co-infecting strains, while the immune environment a strain experiences is set at the time of infection and is unaffected by subsequent infections.

(ii) There is no virulence mortality ($\alpha_x^g = 0$) to ensure independence of co-infecting strains [59, 60].

(iii) Homologous immunity is sterilizing and life-long ($\sigma_h = 0$).

These assumptions alter the specification of Γ_x, as well as the biological interpretation of $S_x(t)$ in (S2). In particular, $S_x(t)$ now denotes the density of hosts with immune history x, irrespective of their current infection status, while

$$\Gamma_x = \sum_g \sum_{z_1 : z_1 \cup g = x} \sigma_{z_1}^g S_{z_1} \sum_{z_2} \beta_{z_2}^g I_{z_2}^g. \tag{S6}$$

Importantly, because homologous immunity is sterilizing ($\sigma_h = 0$), and acquired at time of infection, a host that is infected by strain g' cannot be reinfected by strain g'. As a consequence, and because co-infections are assumed to be non-interacting, it is not necessary to keep track of how many (and which) strains currently infect a given host.

Treating co-infecting strains as non-interacting is somewhat biologically unsatisfying, given a host simultaneously infected with n strains is n-fold as infectious as a host infected with one. This means genetic diversity is unreasonably easily maintained. For example, Buckee et al [22] studied the evolution of hypervirulent lineages of *Neisseria meningitidis* using a model in which each strain is defined by its sequence type (ST) and level of virulence

(which causes host mortality). Since hypervirulence shortens the duration of infection, it should be eliminated from the population by selection. But as hosts can be simultaneously infected by any number of genetically distinct strains, each level of virulence creates 'new' strains that can co-infect the same host (even if they share the same ST), thus maintaining variation in virulence. Note that incorporating virulence in models of co-infection is not trivial, and will mean co-infecting strains do interact [59, 60].

Model	Immune protections	Time immunity updated	History or status?	Effect of multiple prior infections	Coinfection	Sign of ϵ_{PI}	Is $\epsilon_{\mathrm{HI}} > 0$?
[6, 12]	$\beta_{n+1} = \beta_0(1-\gamma)$ $\sigma_h = 0$	infection	history	single	yes	positive	yes
[9, 16]	$\beta_1 = \beta_0(1-\gamma)$ $\beta_{n+2} = \beta_0(1-\delta)$ $\sigma_h = 0$	infection	history	single	yes	positive	yes
[13, 27, 34]	$\sigma_{n+1} = 0$ $\sigma_h = 0$	infection	status	single	yes	positive	no
[18–20]	$\mu_n = \mu_0(n+1)$	clearance	history	antigen-specific	yes	positive	no
[25]	$\sigma_n = \left(\frac{N-n}{N}\right)^{\gamma}$ $\sigma_h = 0$	infection	history	single	yes	positive	yes
[12, 21]	$\sigma_n = \left(1 - \left(\frac{n}{N}\right)^{\frac{1}{\gamma}}\right)^{\gamma}$ $\sigma_h = 0$	infection	history	single	yes	positive	yes
[14]	$\sigma_0 = 1$ $\sigma_h = 0$ $\sigma_x^g = \prod_{g_0 \in x} \sigma_{g_0}^g$	infection	history	product	yes	positive	yes
Model (S1)	variable	either	either	variable	no	variable	variable

Table S2: **Summary of key models and their relationship to the framework used here.** Here, γ and δ are positive parameters and N is the total number of loci. We define $\epsilon_{\mathrm{PI}} \equiv f_0 - 2f_1 + f_2$ where $f_i \equiv \sigma_i\beta_i/(\mu_i + \alpha_i + d)$, and $\epsilon_{\mathrm{HI}} \equiv f_2 - f_h$. Note that both ϵ_{PI} and ϵ_{HI} are calculated assuming there are only two diallelic loci. In this case if $\sigma_h = 0$, each of the different ways in which immunity to multiple prior strains combines (single, antigen-specific, and product) is captured by model (S2). There have also been various models whose specification of cross-immunity assumes one-dimensional antigenic space [e.g., 24, 46] and so we do not consider them here.

S3 Separation of evolutionary and epidemiological dynamics

The number of equations associated with model (S2) make it both challenging to analyze and to interpret. Thus a variety of clever mathematical approximations have been employed to simplify the dynamics [e.g., 6, 13, 14, 46, 59, 76]. While these techniques are effective

at reducing the (mathematical) dimensionality of the problem, they struggle to provide an evolutionary context for the predictions [e.g., see discussion in 27]. Thus many important results lack a clear biological interpretation. One way to provide greater clarity on the evolutionary process is to first separate the variables describing the epidemiological dynamics from those describing the evolutionary dynamics [77, 78]; doing so allows identification, and interpretation, of the various evolutionary forces. The evolutionary variables are the allele (antigen) frequencies, and the linkage disequilibrium (LD), or nonrandom associations, between them [31], while the epidemiological variables include the total density of infections, $I(t)$, and the population distribution of immune histories across the susceptible hosts, $S_x(t)$. Here, we focus on extracting the evolutionary variables for a pathogen genome consisting of two diallelic loci; in Sections S6 and S7 we extend this to greater antigenic variation.

S3.1 The additive selection coefficients and epistasis in fitness

For two diallelic loci, there are four possible pathogen genotypes (strains),

$$g \in \{ab, Ab, aB, AB\}.$$

Without loss of generality, take strain ab to be the reference genotype. Then the additive selection coefficients for allele A and B are, respectively,

$$w_A \equiv r_{Ab} - r_{ab} \quad \text{and} \quad w_B \equiv r_{aB} - r_{ab}.$$

The selection coefficients indicate whether allele A (resp. B) is favoured, $w_A > 0$ (resp. $w_B > 0$) or disfavoured, $w_A < 0$ (resp. $w_B < 0$) on the genetic background b (resp. a). We also need to account for whether the fitness of an allele depends upon its genetic background, or *epistasis in fitness*. Epistasis in fitness is defined as

$$w_{AB} \equiv r_{AB} + r_{ab} - r_{Ab} - r_{aB},$$

and indicates whether the combination of alleles A and B are favoured ($w_{AB} > 0$) or disfavoured ($w_{AB} < 0$) more than would be expected based upon the additive selection coefficients. Thus, while the additive selection coefficients can be roughly thought of as controlling selection on individual alleles, epistasis can be thought of as controlling selection on particular combinations of alleles (in the case of two loci, genotypes or strains).

Although we can write down the additive selection coefficients and epistasis for the general model, here we assume immune protections only affect susceptibility. The reason for doing so is because if immune protections only affect susceptibility, the per-capita growth rates of equation (S4) can be written in a particularly simple form. However, in Section S4 we consider other immune protections using an equilibrium analysis. To write down w_A, w_B, and w_{AB}, we define

$$\sigma_{\text{PI}} \equiv \sigma_0 - 2\sigma_1 + \sigma_2 \quad \text{and} \quad \sigma_{\text{HI}} \equiv \sigma_2 - \sigma_h.$$

Conceptually, if we consider the function mapping antigenic immunity to the strength of immune protections, $h : n \mapsto \sigma_n$, then σ_{PI} is the central difference approximation of $\mathrm{d}^2h/\mathrm{d}n^2$ evaluated at $n = 1$. Consequently, σ_{PI} can be thought of as controlling whether antigen-specific cross-protection increases more ($\sigma_{\text{PI}} < 0$) or less ($\sigma_{\text{PI}} > 0$) rapidly than the linear expectation ($\sigma_{\text{PI}} = 0$) as the number of antigens recognized by the immune system increases.

On the other hand, σ_{HI} measures how much more protective homologous immunity is as compared to full antigenic immunity. If $\sigma_{\mathrm{HI}} = 0$, full antigenic immunity is of equal strength to homologous immunity. Using these composite parameters, the per-capita growth rate can be written

$$r_{ij} = \beta(\sigma_0 \mathcal{S} - (\sigma_0 - \sigma_1)(\mathcal{A}_{\{i\}} + \mathcal{A}_{\{j\}}) + \sigma_{\mathrm{PI}} \mathcal{A}_{\{i,j\}} - \sigma_{\mathrm{HI}} \mathcal{S}_{ij}) - d - \mu - \alpha, \tag{S7}$$

where $\mathcal{A}_{\{i\}}(t)$ and $\mathcal{A}_{\{i,j\}}(t)$ are the density of uninfected hosts that have immunity to *at least* antigen i or antigens i and j, $\mathcal{S}(t) = \sum_x S_x(t)$ is the total density of uninfected hosts, and $\mathcal{S}_{ij}(t)$ is the density of uninfected hosts that have homologous immunity (prior infection) to *at least* strain ij.

Then the additive selection coefficients can be written

$$w_A = \beta(\sigma_0 - \sigma_1)(\mathcal{A}_{\{a\}} - \mathcal{A}_{\{A\}}) + \beta\sigma_{\mathrm{PI}}(\mathcal{A}_{\{A,b\}} - \mathcal{A}_{\{a,b\}}) - \beta\sigma_{\mathrm{HI}}(\mathcal{S}_{Ab} - \mathcal{S}_{ab}) \tag{S8}$$

$$w_B = \beta(\sigma_0 - \sigma_1)(\mathcal{A}_{\{b\}} - \mathcal{A}_{\{B\}}) + \beta\sigma_{\mathrm{PI}}(\mathcal{A}_{\{a,B\}} - \mathcal{A}_{\{a,b\}}) - \beta\sigma_{\mathrm{HI}}(\mathcal{S}_{aB} - \mathcal{S}_{ab}) \tag{S9}$$

while epistasis in fitness is

$$w_{AB} = \beta\sigma_{\mathrm{PI}}(\mathcal{A}_{\{A,B\}} + \mathcal{A}_{\{a,b\}} - \mathcal{A}_{\{A,b\}} - \mathcal{A}_{\{a,B\}}) - \beta\sigma_{\mathrm{HI}}(\mathcal{S}_{AB} + \mathcal{S}_{ab} - \mathcal{S}_{Ab} - \mathcal{S}_{aB}). \tag{S10}$$

S3.2 Relationship between $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$

If we define

$$\mathcal{A}_\Delta(t) = \mathcal{A}_{\{A,B\}} + \mathcal{A}_{\{a,b\}} - \mathcal{A}_{\{A,b\}} - \mathcal{A}_{\{a,B\}} \quad \text{and} \quad \mathcal{S}_\Delta(t) = \mathcal{S}_{AB} + \mathcal{S}_{ab} - \mathcal{S}_{Ab} - \mathcal{S}_{aB}, \tag{S11}$$

that is, $\mathcal{A}_\Delta(t)$ and $\mathcal{S}_\Delta(t)$ indicates whether more susceptible hosts have full antigenic-immunity or homologous immunity, respectively, to the set of strains $\{ab, AB\}$ (if positive) or $\{Ab, aB\}$ (if negative), we can rewrite epistasis in fitness (equation (S10)) as

$$w_{AB} = \beta\sigma_{\mathrm{PI}}\mathcal{A}_\Delta(t) - \beta\sigma_{\mathrm{HI}}\mathcal{S}_\Delta(t). \tag{S12}$$

Next, notice that if we expand the sums in $\mathcal{A}_\Delta(t)$, then after some cancellation:

$$\mathcal{A}_\Delta(t) = S_{AB}(t) + S_{ab}(t) - S_{Ab}(t) - S_{aB}(t). \tag{S13}$$

Thus only hosts with immunity to the antigens defining a single strain contribute to $\mathcal{A}_\Delta(t)$. Unlike $\mathcal{A}_\Delta(t)$, however, $\mathcal{S}_\Delta(t)$ depends upon hosts with immunity to more than one strain, since after some cancellation:

$$\mathcal{S}_\Delta(t) = \mathcal{A}_\Delta(t) + S_{ab,AB}(t) - S_{aB,Ab}(t) + \mathcal{S}_{ab,AB}(t) - \mathcal{S}_{aB,Ab}(t) \tag{S14}$$

where $S_{ij,k\ell}(t)$ and $\mathcal{S}_{ij,k\ell}(t)$ indicate the density of uninfected hosts that have homologous immunity to *only*, and *at least*, strains ij and $k\ell$, respectively. Thus as discussed in the main text, in order for $\mathcal{S}_\Delta(t)$ and $\mathcal{A}_\Delta(t)$ to diverge in value, hosts must be infected multiple times during the course of their life (and retain memory of these infections).

S3.3 Equations for the evolutionary dynamics

To then write out the evolutionary dynamics, we first define the evolutionary variables. The allele frequencies are

$$p_A(t) = \frac{I^{Ab}(t) + I^{AB}(t)}{I(t)} \quad \text{and} \quad p_B(t) = \frac{I^{aB}(t) + I^{AB}(t)}{I(t)}, \tag{S15}$$

while the linkage disequilibrium between alleles A and B is

$$D(t) = \frac{I^{AB}(t)}{I(t)} - p_A(t)p_B(t) = \frac{I^{AB}(t)}{I(t)}\frac{I^{ab}(t)}{I(t)} - \frac{I^{Ab}(t)}{I(t)}\frac{I^{aB}(t)}{I(t)}. \tag{S16}$$

Then the evolutionary dynamics can be written

$$\frac{\mathrm{d}p_A}{\mathrm{d}t} = w_A p_A(1-p_A) + w_B D + w_{AB}(1-p_A)p_{AB} + \Delta_{p_B}\rho \tag{S17a}$$

$$\frac{\mathrm{d}p_B}{\mathrm{d}t} = w_B p_B(1-p_B) + w_A D + w_{AB}(1-p_B)p_{AB} + \Delta_{p_B}\rho \tag{S17b}$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \sum_{k=A,B} w_k(1-2p_k)D + w_{AB}p_{AB}(p_{ab} - 2D) + \Delta_D\rho, \tag{S17c}$$

where $\Delta_x\rho$ indicates the effect of recombination on the variable x. Note that $p_{AB} = p_A p_B + D$ and $p_{ab} = (1-p_A)(1-p_B) + D$.

The additive selection coefficients, recombination, and epistasis in fitness need not be constant and instead vary with the epidemiological dynamics. But by writing out the evolutionary dynamics as we have done in (S17), we can identify the evolutionary forces at play and see how they impact the evolutionary process. We can then link each of these forces to the epidemiological model of interest.

S3.4 Fast- and slow-timescales in the evolutionary dynamics

If recombination is unbiased, then it will not affect allele frequencies, $\Delta_{p_B}\rho = \Delta_{p_A}\rho = 0$ and will only act to remove LD from the population, that is, $\Delta_D\rho = -\rho D(t)$ for $\rho \geq 0$. In this case, there are no intrinsic biases favouring one allele over the other at a given loci, and so by symmetry we would expect the population will self-organize so that the per-capita growth rates of the set of strains $\{ab, AB\}$ are similar, and likewise the per-capita growth rates of the set of strains $\{Ab, aB\}$ [27]. Indeed, after some short-term transient dynamics owing to the initial conditions, strong NFDS operating through the additive selection coefficients will push $r_{AB} \to r_{ab}$ and $r_{Ab} \to r_{aB}$. At this point, the antigen frequencies are in quasi-equilibrium, $p_A(t) \approx 1/2$ and $p_B(t) \approx 1/2$, and moreover, $w_A \approx w_B \approx -w_{AB}/2$. Consequently, over the long-term the set of equations (S17) reduces to the single equation

$$\frac{\mathrm{d}D}{\mathrm{d}t} \approx w_{AB}\left(\frac{1}{4} + D\right)\left(\frac{1}{4} - D\right) - \rho D. \tag{S18}$$

Because LD quantifies whether certain strains are more, or less, abundant than predicted based upon allelic frequencies, the presence of LD in the population is necessary for there to be pathogen strain structure or PSS (see the longer discussion in Section 2 of the main text).

Thus over the long timescale, equation (S18) captures the evolutionary dynamics of PSS. Moreover, since $p_{AB}(t) = p_{ab}(t)$ and $p_{Ab}(t) = p_{aB}(t)$, if $D(t) > 0$, then strain set $\{ab, AB\}$ will be over-represented in the population, whereas if $D(t) < 0$, strain set $\{Ab, aB\}$ will be over-represented.

Assuming $\rho = 0$, equation (S18) has four potential equilibria: (1,2) when either strain set has competitively excluded the other ($D(t) = \pm 1/4$); (3) when one set of strains is over-represented, but all strains are present ($w_{AB} = 0$ for $0 < |D| < 1/4$); and (4) when there is no structure ($w_{AB} = 0$ for $D = 0$). Finally, if the sign of epistasis fluctuates, evolutionary oscillations, defined to be any sustained dynamical behaviour of LD (so including cycles and chaos) may occur. In all cases, however, the dynamical behaviour of equation (S18) hinges on epistasis. The long-term behaviour LD, arising from numerical simulation of the full epidemiological model of Box 1 is illustrated in Figure 3, while Figure S1 shows representative examples of the dynamical behaviour.

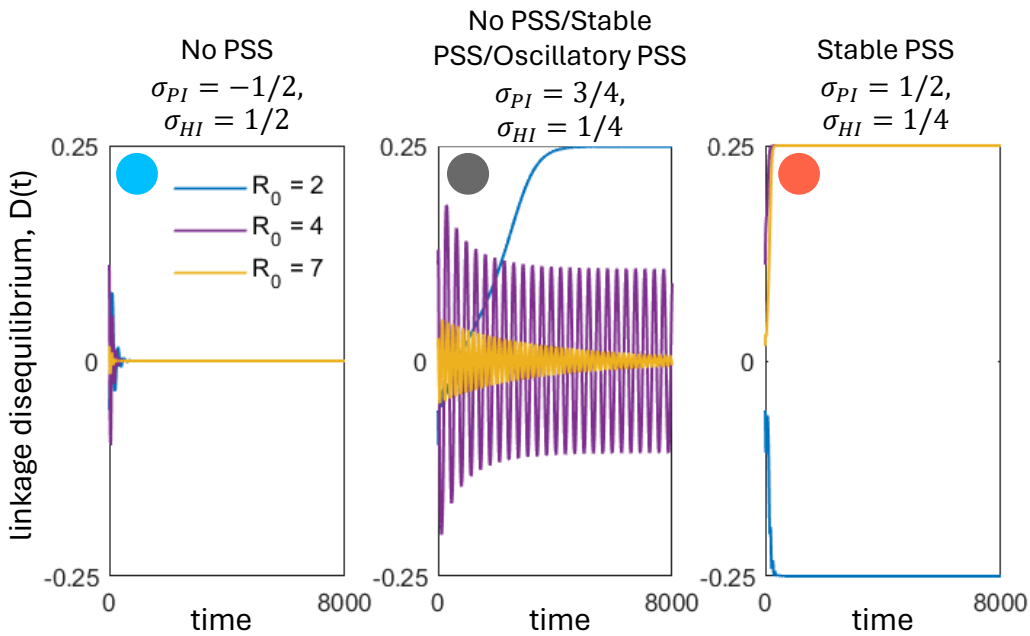

Figure S1: **Dynamical behaviour of LD as $\sigma_{\mathbf{PI}}$, $\sigma_{\mathbf{HI}}$, and R_0 vary** The dynamics of LD depend upon the values of σ_{PI} and σ_{HI}, but also the dynamic immunological landscape (in this case, controlled by R_0). Each panel corresponds to the indicated combination of $\sigma_{\mathrm{PI}}, \sigma_{\mathrm{HI}}$ where the coloured circles indicate the approximate position on Figure 3. The different trajectories correspond to the indicated value of R_0. All simulations used $b = d = 0.01$, $\theta = 0.05$, $\rho = 10^{-5}$, $\sigma_0 = 1$ and $\sigma_h = 0$, while σ_1 and σ_2 vary by panel (μ and β are chosen to satisfy $\theta = 0.05$ and the indicated value of R_0).

Finally, we note that given an observed distribution of allele frequencies, there is a maximum and minimum value LD can possibly take. For example, for two diallelic loci

$$\max(D(t)) = \min\big(p_A(t)(1 - p_B(t)),\, p_B(t)(1 - p_A(t))\big)$$

and

$$\min(D(t)) = \max\big(-p_A(t)p_B(t), -(1 - p_A(t))(1 - p_B(t))\big)$$

[see 31]. Importantly, in order for pairwise LD to achieve this maximum, at least one strain must be absent from the population; in the special case when $p_A(t) = p_B(t)$, two strains are absent. Similar results hold for higher-order LD among more than two loci [79, 80], and when LD reaches either its maximum or minimum, one or more strains are absent from the population. We will use D^* to indicate that the specified order of LD attains either its maximum or minimum, that is, one or more strains are competitively excluded.

S4 Different types of immune protections

Next, we consider the more general case in which immune protections may affect transmissibility, clearance, and virulence mortality, but that homologous immunity is equivalent to full antigenic immunity (so $\sigma_{\text{HI}} = 0$). Although epistasis in fitness can be written out in full, it is more complicated than when immunity only affects susceptibility because if immunity affects transmissibility, clearance, and/or virulence-mortality, the fitness (per-capita growth rate) of an infection will depend upon the immune environment the infection is found in. As a consequence, the per-capita growth rates will depend upon the frequency of strain ij infections found in a particular environment. Thus we instead focus upon deriving epistasis in fitness at equilibrium. Since there is effectively one epidemiological feedback, which is whether more hosts have antigenic immunity to one set of strains or the other (e.g., $\{ab, Ab\}$ or $\{Ab, aB\}$), if recombination is weak then there is no reason to expect epistasis to become zero and/or change sign at some non-zero amount of LD. Consequently, we should expect that if PSS occurs, it will be the largest in magnitude possible given the allele frequencies, that is, $D(t) \to D^*$. Consequently if, for example, $p_A(t) = p_B(t) = 1/2$, then when PSS occurs, $|D(t)| \to 1/4$ (this means two strains will be excluded from the population). It is these two equilibria that we will focus upon here.

S4.1 Epistasis must share same sign as LD for local asymptotic stability

Here we focus upon one of the equilibria $(p_A, p_B, D) = (1/2, 1/2, \pm 1/4)$; our objective is to show that a necessary condition for it to be locally asymptotically stable (LAS) is that epistasis in fitness shares the same sign as LD.

To see this, suppose we take the derivative of the differential equation for LD, $\frac{\mathrm{d}D}{\mathrm{d}t}$, with respect to any of the population variables (arbitrarily indicated as X) under the assumption recombination is negligible ($\rho \approx 0$). Then we obtain

$$\begin{aligned}\frac{\partial}{\partial X}\frac{\mathrm{d}D}{\mathrm{d}t} = D\left((1-2p_A)\frac{\partial w_A}{\partial X} + (1-2p_B)\frac{\partial w_B}{\partial X}\right) \\ + (p_A p_B + D)((1-p_A)(1-p_B) - D)\frac{\partial w_{AB}}{\partial X} + w_A\frac{\partial[1-2p_A]D}{\partial X} \\ + w_B\frac{\partial[1-2p_B]D}{\partial X} + w_{AB}\frac{\partial}{\partial X}[(p_A p_B + D)((1-p_A)(1-p_B) - D)]. \quad \text{(S19)}\end{aligned}$$

Now consider evaluating (S19) when $p_A = p_B = 1/2$ and $D = 1/4$ or $D = -1/4$. In this

case, $r_{AB} = r_{ab}$ and $r_{Ab} = r_{aB}$; thus $w_{AB} = -2w_A = -2w_B$. Therefore

$$\frac{\partial}{\partial X}\frac{\mathrm{d}D}{\mathrm{d}t} = \begin{cases} -2w_{AB}D & \text{if } X = D \\ 0 & \text{otherwise} \end{cases} \tag{S20}$$

Hence if we were to compute the Jacobian of the full system (that combines both the evolutionary and epidemiological variables) evaluated at either of these equilibria, the Jacobian will be block triangular, and one of the entries on the diagonal is given by equation (S20) (when $X = D$). Thus one of the eigenvalues will be $-w_{AB}/2$ (if $D = 1/4$) or $w_{AB}/2$ (if $D = -1/4$). It follows that a necessary condition for PSS to be LAS is that the sign of epistasis matches the sign of LD.

S4.2 Epistasis at equilibrium

Based upon the preceding results, we therefore seek to calculate epistasis at equilibrium. To do so, define

$$f_n \equiv \frac{\sigma_n \beta_n}{\mu_n + \alpha_n + d}; \tag{S21}$$

this can be thought of as proportional to the effective reproductive number of an infection in a host with immunity to n antigens of the infecting strain. This is because, all else being equal, with probability σ_n an infection is established in a host with immunity to n of its antigens; such an infection has an expected lifetime of $(\mu_n + \alpha_n + d)^{-1}$, during which it will produce new infections at a rate proportional to β_n.

Next, at equilibrium, the reproductive number of strain ij [e.g., see 81], which we will denote $\bar{R}_{ij}$, can be written

$$\begin{aligned} \bar{R}_{ij} &= f_0(\bar{\mathcal{S}} - \bar{\mathcal{A}}_{\{i\}} - \bar{\mathcal{A}}_{\{j\}} + \bar{\mathcal{A}}_{\{i,j\}}) + f_1(\bar{\mathcal{A}}_{\{i\}} + \bar{\mathcal{A}}_{\{j\}} - 2\bar{\mathcal{A}}_{\{i,j\}}) + f_2\bar{\mathcal{A}}_{\{i,j\}} \\ &= f_0\bar{\mathcal{S}} - (f_0 - f_1)(\bar{\mathcal{A}}_{\{i\}} + \bar{\mathcal{A}}_{\{j\}}) + (f_0 - 2f_1 + f_2)\bar{\mathcal{A}}_{\{i,j\}}. \end{aligned}$$

Now, the per-capita growth rate, r_{ij} is a measure of fitness in continuous time, whereas the reproductive number, $\bar{R}_{ij}$, is a measure of fitness in discrete time [81]. Although in continuous time, epistasis in fitness is defined as

$$w_{AB} \equiv r_{ab} + r_{AB} - r_{Ab} - r_{aB},$$

in discrete time, epistasis in fitness becomes

$$\bar{w}_{AB} \equiv \bar{R}_{AB}\bar{R}_{ab} - \bar{R}_{Ab}\bar{R}_{aB}$$

[see e.g., 82, 83]. Without loss of generality, focus upon the equilibrium in which the over-represented stain set is $\{ab, AB\}$. At this equilibrium, $\bar{R}_{AB} = \bar{R}_{ab} = 1$; moreover, $\bar{R}_{Ab} = \bar{R}_{aB}$. Therefore, when $(p_A, p_B, D) = (1/2, 1/2, 1/4)$,

$$\bar{w}_{AB} = \bar{R}_{AB}\bar{R}_{ab} - \bar{R}_{Ab}\bar{R}_{aB} = 1 - \bar{R}_{Ab}^2 = (1 - \bar{R}_{Ab})(1 + \bar{R}_{Ab}).$$

Since $\bar{R}_{ij} \geq 0$,

$$\bar{w}_{AB} \propto 2(1 - \bar{R}_{Ab}) = 2(\bar{R}_{ab} - \bar{R}_{Ab}) = 2\left(f_0 - 2f_1 + f_2\right)\left(\bar{\mathcal{A}}_{\{a,b\}} - \bar{\mathcal{A}}_{\{A,b\}}\right) = \epsilon_{\mathrm{PI}}\bar{\mathcal{A}}_{\Delta}, \tag{S22}$$

where proportionality is up to a positive constant, $\epsilon_{\mathrm{PI}} = f_0 - 2f_1 + f_2$, and $\bar{\mathcal{A}}_{\Delta} = \bar{\mathcal{A}}_{\{a,b\}} + \bar{\mathcal{A}}_{\{A,B\}} - \bar{\mathcal{A}}_{\{A,b\}} - \bar{\mathcal{A}}_{\{a,B\}}$ (note that $\bar{\mathcal{A}}_{\{a,b\}} = \bar{\mathcal{A}}_{\{A,B\}}$ and $\bar{\mathcal{A}}_{\{A,b\}} = \bar{\mathcal{A}}_{\{a,B\}}$ at equilibrium). Hence if $\epsilon_{\mathrm{PI}} > 0$, epistasis will share the same sign as LD.

S5 The link between strain frequencies and the immunological landscape

As discussed in the main text, various epidemiological assumptions can disrupt the correlation between which strain set is most abundant in the population, and which strain set has more hosts with immunity to it (so $D(t)$ and $\mathcal{A}_\Delta(t)$, $\mathcal{S}_\Delta(t)$ do not share the same sign).

Here we explore two ways in which this disruption can occur; in both we assume for simplicity that full antigenic immunity is equivalent to homologous immunity, $\sigma_h = \sigma_2$. First, we assume that uninfected hosts lose immunity to each antigen at a per-capita rate of δ. For example, an uninfected host with immunity to antigens $\{a, A, b\}$ becomes immune to antigens $\{a, A\}$ at a per-capita rate δ. Second, we assume that following clearance of infection, hosts gain immunity to a single, randomly selected antigen with probability p, and both antigens with probability $1 - p$.

Since we assume homologous immunity is of the same strength as full antigenic immunity, let us assume that x in this case only tracks the unique antigens to which a host has immunity to (e.g., in the previous formulation, $x = \{ab, Ab\}$, would become $x = \{a, A, b\}$. Then the two assumptions above leave the equations for the I_x^{ij} unchanged, but mean the net rate at which immune history x is gained and lost is

$$\Gamma_x = \frac{1}{2}\sum_g \Big(2\sum_{z:g\cup z=x}(1-p)\mu_z^g I_z^g + \sum_{z:g\cup(z\setminus z_1)=x} p\mu_z^g I_z^g + \sum_{z:g\cup(z\setminus z_2)=x} p\mu_z^g I_z^g\Big) + \delta\Big(\sum_{z:(z\setminus z_i)=x} S_z - \sum_{i\in x} S_x\Big) \tag{S23}$$

where $(z \setminus z_i)$ means the genotype z without the i-th locus; e.g., if $z = ab$, then $(z \setminus z_1) = b$.

When we simulate the dynamics (Fig. S2), we observe two differences.

(1) If immunity wanes in an antigen-by-antigen fashion (i.e., $\delta > 0$), then when σ_{PI} is maximized ($\sigma_1, \sigma_2 \to 0$), strain structure will not form (Fig. S2**a**).

The logic here is as follows. When $\sigma_1, \sigma_2 \to 0$, hosts with immunity to one strain (say AB) can be infected by at most the other strain in the strain set (ab). When this occurs, this will produce hosts with immunity to all four antigens and thus full protection from infection. In order for these hosts to then become susceptible again, they have to lose immunity to one antigen at each locus. Since immunity to each antigen is lost randomly, such hosts are equally likely to produce hosts that are susceptible to one strain set or the other, irrespective of which strain set is more abundant. However, as these hosts become susceptible to infection, the more abundant strain set will infect, and deplete, its available hosts faster. On balance, this will create an excess of hosts with immunity to the less abundant strain set, meaning that $\mathcal{A}_\Delta(t)$ will have the opposite sign to LD. Since $\sigma_{\mathrm{PI}} = \sigma_0 > 0$, epistasis in fitness will also have the opposite sign to population LD. Thus strain structure cannot form, despite σ_{PI} being maximized.

(2) Second, suppose that following infection hosts acquire immunity to both antigens of a strain with probability $1 - p$, or only a single antigen with probability p (each antigen is chosen with probability 1/2). If p is sufficiently large, strain structure occurs only when $\sigma_{\mathrm{PI}} < 0$ (Fig. S2**b**).

The logic here is as follows. Suppose strain set $\{ab, AB\}$ is over-represented in the population ($D(t) > 0$). Therefore, we would expect naive hosts to be more likely to

become infected by either strains ab or AB. Following infection by strain AB (say), the host is then most likely to become infected by strain ab (since it is more abundant). But because each infection can at most provide immunity to a single antigen (since $p \to 1$), when the two randomly selected antigens are at different loci, following a secondary infection the host will then have immunity to either the antigens Ab or aB, and so will have immunity to a strain belonging to the under-represented strain set, $\{Ab, aB\}$. But this will mean that $\mathcal{A}_\Delta(t)$ will have opposite sign to $D(t)$, that is, more hosts will have immunity to the less-abundant strain set, and so if $\sigma_{\mathrm{PI}} < 0$, this provides an advantage for the more abundant strain set and so PSS can occur.

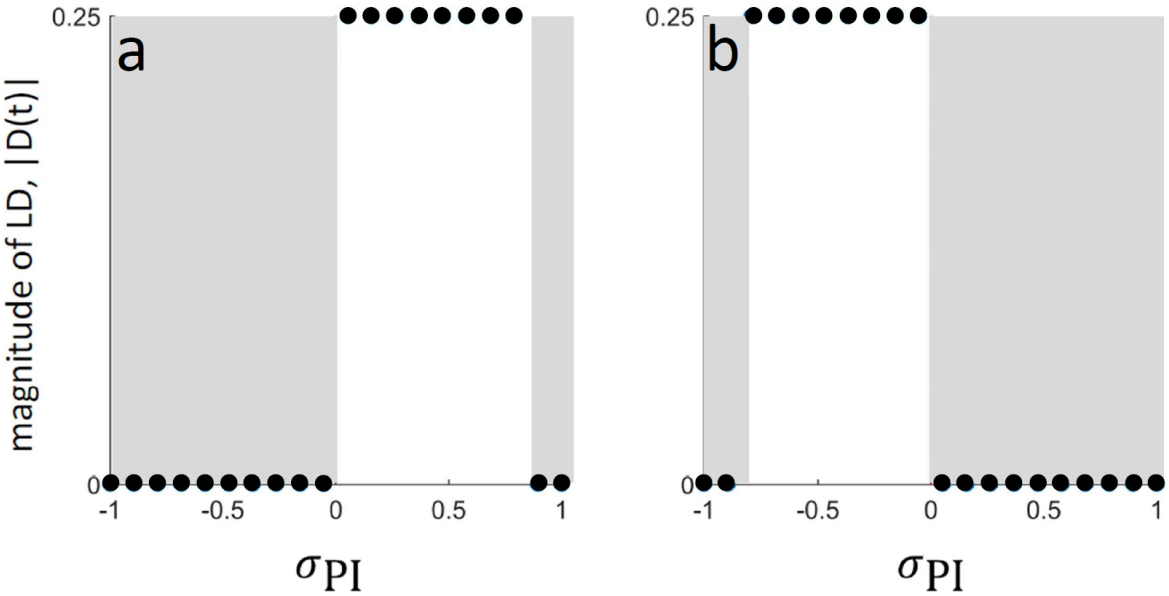

Figure S2: **Disrupting the link between strain frequencies and the immunological landscape**. Here we show two ways in which LD need not share the same sign as $\mathcal{A}_\Delta(t)$. In panel **a**, immunity wanes in an antigen-by-antigen fashion. In panel **b**, hosts acquire immunity to both antigens with probability $1 - p$, and a single, randomly chosen antigen with probability p.

Note that these predictions are sensitive to how immune protections are modelled. In particular, when immunity reduces susceptibility typically only when a host is successfully infected will it potentially gain immunity to the infecting strain [15]. If instead immune protections affect, for example, transmissibility, even an 'unsuccessful' infection (i.e., an infection that is incapable of forward transmission) will update the host's immunity. Generally speaking, these differences are not expected to have a large effect, at least in models with very simple epidemiological assumptions [15]. However, as the underlying epidemiological model becomes more complex, the evolutionary dynamics can become more sensitive to the type of immune protections as the above examples illustrate.

S6 Variation in allelic diversity across loci

The model featured in Section S3 assumes that there are an equal number of alleles at each locus. Here we consider what happens if at one locus there are two alleles $i \in \{A_1, A_2\}$, while at the other there are three, $j \in \{B_1, B_2, B_3\}$, and so six possible strains,

$$ij \in \{A_1B_1, A_1B_2, A_2B_3, A_2B_1, A_2B_2, A_2B_3\}.$$

Without loss of generality, we choose strain A_1B_1 as the reference strain. Then the evolutionary variables are the frequency of alleles A_2 and B_j at time t, which are

$$p_{A_2}(t) = \frac{\sum_j I^{A_2B_j}(t)}{I(t)}, \quad p_{B_j}(t) = \frac{I^{A_1B_j}(t) + I^{A_2B_j}(t)}{I(t)} \; j = 2, 3,$$

as well as the pairwise LD between alleles A_2 and B_j,

$$D_{A_2B_j}(t) = \frac{I^{A_2B_j}(t)}{I(t)} - p_{A_2}(t)p_{B_j}(t), \quad j = 2, 3.$$

Next, the three additive selection coefficients are defined as

$$w_A \equiv r_{A_2B_1} - r_{A_1B_1} \quad \text{and} \quad w_{B_j} \equiv r_{A_1B_j} - r_{A_1B_1}, \; j = 2, 3, \tag{S24}$$

while the two epistasis in fitness terms are

$$w_{A_2B_j} \equiv r_{A_2B_j} + r_{A_1B_1} - r_{A_1B_j} - r_{A_2B_1}, \quad j = 2, 3,$$

that is,

$$\begin{aligned} w_{A_2B_j} = \beta\sigma_{\mathrm{PI}}(\mathcal{A}_{\{A_2,B_j\}} + \mathcal{A}_{\{A_1,B_1\}} - \mathcal{A}_{\{A_2,B_1\}} - \mathcal{A}_{\{A_1,B_j\}}) \\ - \beta\sigma_{\mathrm{HI}}(\mathcal{S}_{A_2B_j} + \mathcal{S}_{A_1B_1} - \mathcal{S}_{A_2B_1} - \mathcal{S}_{A_1B_j}) \end{aligned} \tag{S25}$$

where as before, $\mathcal{A}_{\{i\}}$ and $\mathcal{A}_{\{i,j\}}$ denotes the density of all the uninfected hosts with immunity to at least the antigen i and antigens i and j, respectively, while $\mathcal{Z}_{ij}$ denotes the density of uninfected hosts with homologous immunity to at least strain ij.

It is then straightforward to write out the equations describing the evolutionary dynamics. Here we skip this, and simply interpret the simulated dynamics. To do so, the key is to recognize that additive selection favours the frequency of each allele at a given locus to be equal, that is, $p_{A_2} = 1/2$, $p_{B_j} = 1/3$, while epistatic selection either favours LD to increase or decrease in magnitude, depending upon whether $D_{ij}(t)$ and w_{ij} are of the same or opposite sign, respectively, and an intermediate level of LD only when $w_{ij} = 0$ for some non-zero D_{ij}. In the two-allele, two-locus model of Section S3, both of these could be simultaneously achieved. One consequence of this was that there were two timescales: a short timescale over which allele frequencies were driven to equilibrium $p_k \to 1/2$, and a longer timescale along which LD evolved (see equation (S18)). When there are more alleles at one locus then the other, this need not be true, and thus there is competition between epistatic and allelic selection, which can induce more complex dynamics and long-term outcomes.

To understand this, first consider the case in which $\sigma_{\mathrm{HI}} = 0$. As was true for the model of Section S3, there is a single epidemiological feedback on epistasis and so oscillatory

evolutionary dynamics are not possible and the population will converge to an evolutionary equilibrium. If $\sigma_{\rm PI} < 0$, then epistasis favours linkage equilibrium, $D_{A_2B_i} = 0$, while additive selection favours $p_{A_i} = 1/2$, and $p_{B_j} = 1/3$. Since these can be simultaneously achieved, there is no conflict between epistasis and additive selection (Fig. S3). However, when $\sigma_{\rm PI} > 0$ epistasis favours pairwise LD attaining D^*, but this cannot be achieved for all three $\{A_2, B_j\}$ pairs if the population is in a state favoured by additive selection (i.e., $p_{A_i} = 1/2$, $p_{B_j} = 1/3$). This creates a conflict between epistasis and additive selection, and so the evolutionary outcome depends upon their relative strengths. If $\sigma_{\rm PI}$ is positive but small, epistasis will be relatively weak compared to additive selection. This will cause the allele frequencies to diverge from their expectation, but $D_{A_2B_j}(t) = D^*$ for two of the three $\{A_2, B_j\}$ pairs; the other will be in linkage equilibrium (so two genotypes will have been eliminated from the population). As $\sigma_{\rm PI}$ increases, the allele frequencies diverge further from their unlinked optimum, that is, the distance (measured here using the Euclidean norm)

$$\xi_p = \sqrt{\sum_i (p_{A_i} - 1/2)^2 + \sum_j (p_{B_j} - 1/3)^2}$$

becomes increasingly large (Fig. S3), while the pairwise LD becomes closer to D^*, that is,

$$\xi_D = \sqrt{\sum_j (D_{A_2B_j} - D^*)^2}$$

becomes increasingly small (Fig. S3). Eventually when $\sigma_{\rm PI}$ is sufficiently large, $p_{A_i} = 1/2$, $p_{B_j} = p_{B_k} = 1/4$, $p_{B_\ell} = 1/2$ for $j \neq k \neq \ell = 1, 2, 3$ while $D_{A_2B_j}(t) = D^*$ for each pairwise combination (so three genotypes will have been eliminated from the population) (Fig. S3). Thus the sign of $\sigma_{\rm PI}$ still determines whether PSS forms or not, but now additionally its strength can affect the amount of PSS observed. Importantly, however, even if $\sigma_{\rm PI}$ is maximized, the PSS observed will not be the DSS of Gupta et al [6, 12].

Second, suppose $\sigma_{\rm HI} > 0$. As was the case of the two-locus, two-allele model, evolutionary oscillations are possible. But here there is a important difference: because there is no longer a short- and long-timescale, the allele frequencies will *not* reach an equilibrium. Thus both LD and the allele frequencies will oscillate (Fig. S4).

S7 More than two loci

What happens as the number of loci increases? To answer this, we will suppose there are three diallelic loci, $\{a, A\}$, $\{b, B\}$ and $\{c, C\}$. Without loss of generality, choose strain *abc* as the reference strain. Then the evolutionary dynamics consist of the dynamics of the allele frequencies, $p_k(t)$, $k \in \{A, B, C\}$, e.g.,

$$p_A(t) \equiv \frac{I^{Abc}(t) + I^{ABc}(t) + I^{AbC}(t) + I^{ABC}(t)}{I(t)},$$

three pairwise LD terms, $D_{ij}(t)$, $ij \in \{AB, AC, BC\}$, e.g.,

$$D_{AB}(t) \equiv \frac{I^{ABc}(t) + I^{ABC}(t)}{I(t)} - p_A(t)p_B(t),$$

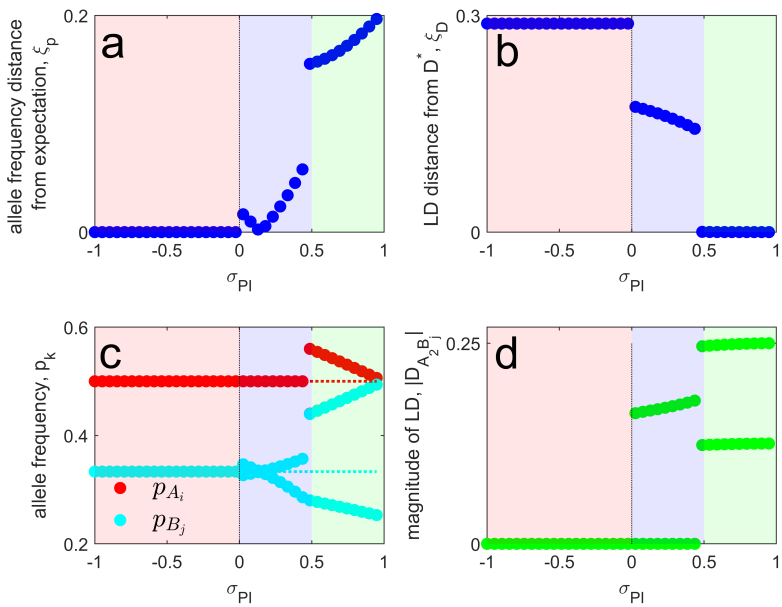

Figure S3: **Consequences of variation in allelic diversity across loci**. An unequal number of alleles across loci can produce conflict between selection on alleles (additive selection coefficients) and on strains (epistasis) because allele frequencies cannot be at their expected values ($p_{A_i} = 1/2$, $p_{B_j} = 1/3$) *and* the LD between each $\{A_2, B_j\}$ pair satisfy $D_{A_2B_j} = D^*$. If $\sigma_{\rm PI} < 0$ (red shaded region), there is no conflict because epistasis does not favour strain structure. If $\sigma_{\rm PI} > 0$ but weak (blue shaded region), although PSS is favoured, selection on alleles is stronger than on strains and so there is no LD between one of the $\{A_2, B_j\}$ pairs. As $\sigma_{\rm PI}$ increases further (green shaded region), so does the strength of epistasis; this pushes all three $D_{A_2B_j} \to D^*$ (panel **b**) and allele frequencies away from their expectation (panel **a**). Parameter values: $\beta = 1$, $\mu = 0.5$, $b = d = 0.01$, $\rho = 10^{-5}$, $\sigma_0 = 1$, $\sigma_2 = \sigma_h = 0$, and $\sigma_1 \in [\sigma_0, \sigma_2]$.

and the triple LD,

$$D_{ABC}(t) \equiv p_{ABC}(t) - p_A(t)p_B(t)p_C(t) - p_A(t)D_{BC}(t) - p_B(t)D_{AC}(t) - p_C(t)D_{AB}(t).$$

Limiting our attention to immune protections that reduce susceptibility, from the model of Section S1, the per-capita growth rate of strain ijk is

$$r_{ijk} = \beta\big(\sigma_0\mathcal{S} - (\sigma_0 - \sigma_1)\sum_{\ell=i,j,k}\mathcal{A}_{\{\ell\}} + \sigma_{\rm PI}^{(1)}\sum_{\ell_1\ell_2\in\{ij,ik,jk\}}\mathcal{A}_{\{\ell_1,\ell_2\}}\big) - \beta\big((\sigma_{\rm PI}^{(1)} - \sigma_{\rm PI}^{(2)})\mathcal{A}_{\{i,j,k\}} + \sigma_{\rm HI}\mathcal{S}_{ijk}\big) - \mu - \alpha - d \quad \text{(S26)}$$

where $\sigma_{\rm HI} \equiv \sigma_3 - \sigma_h$, and

$$\sigma_{\rm PI}^{(n)} \equiv \sigma_{n-1} - 2\sigma_n + \sigma_{n+1}, \quad n = 1, 2.$$

Thus $\sigma_{\rm PI}^{(n)}$ measures whether having antigenic immunity to n antigens of the infecting strain provides more, $\sigma_{\rm PI}^{(n)} > 0$, or less, $\sigma_{\rm PI}^{(n)} < 0$, protection then would be expected based upon the protection offered by having immunity to $n-1$ versus $n+1$ antigens.

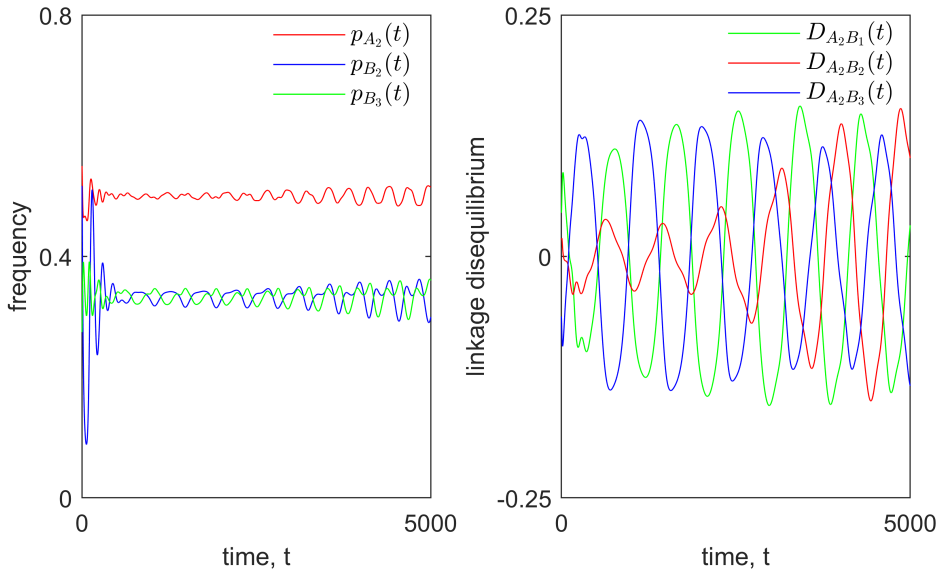

Figure S4: **Conflict between selection on alleles and strains generates more complex evolutionary oscillations**. When there is variation in allelic diversity across loci, there is no longer a fast- and slow-timescale. This is because allele frequencies cannot be at their expectation **and** all pairwise LD terms reach D^*. Consequently, when evolutionary oscillations occur, these involve both allele frequencies and LD. Parameter values: $\beta = 1$, $\mu = 0.4$, $\alpha = 0$, $b = d = 0.01$, $\rho = 0$, $\sigma_0 = 1$, $\sigma_1 = 0.3$, $\sigma_2 = 0.15$, and $\sigma_h = 0$.

Using these per-capita growth rates, the additive selection coefficient for allele A (and similarly for alleles B and C) can be written

$$w_A = r_{Abc} - r_{abc} = -\beta(\sigma_0 - \sigma_1)(\mathcal{A}_{\{A\}} - \mathcal{A}_{\{a\}}) + \beta\sigma_{\mathrm{PI}}^{(1)}(\mathcal{A}_{\{A,b\}} + \mathcal{A}_{\{A,c\}} - \mathcal{A}_{\{a,b\}} - \mathcal{A}_{\{a,c\}}) + \beta(\sigma_{\mathrm{PI}}^{(2)} - \sigma_{\mathrm{PI}}^{(1)})(\mathcal{A}_{\{A,b,c\}} - \mathcal{A}_{\{a,b,c\}}) - \beta\sigma_{\mathrm{HI}}(\mathcal{S}_{Abc} - \mathcal{S}_{abc}). \quad \text{(S27)}$$

Pairwise epistasis for alleles $ij \in \{AB, AC, BC\}$ is defined as before, except we use the allele at the third locus as the reference allele, $\ell \in \{a, b, c\}$; for example,

$$w_{AB} = r_{ABc} + r_{abc} - r_{Abc} - r_{aBc}$$

which using our per-capita growth rates can be written as

$$w_{ij} = \beta\sigma_{\mathrm{PI}}^{(1)}\mathcal{A}_{\Delta}^{(ij)} + \beta(\sigma_{\mathrm{PI}}^{(2)} - \sigma_{\mathrm{PI}}^{(1)})\mathcal{A}_{\Delta}^{(ij)/\ell} - \beta\sigma_{\mathrm{HI}}\mathcal{S}_{\Delta}^{(ij)/\ell}. \quad \text{(S28)}$$

Here $\mathcal{A}_{\Delta}^{(ij)}$ measures whether more hosts have immunity to strains with antigens $\{ij, \hat{i}\hat{j}\}$ or $\{\hat{i}j, i\hat{j}\}$, irrespective of genetic background (note that $\hat{i}$ indicates the other allele at the indicated locus, e.g., if $i = A$, then $\hat{i} = a$). For example,

$$\mathcal{A}_{\Delta}^{(AB)} = \mathcal{A}_{\{A,B\}} + \mathcal{A}_{\{a,b\}} - \mathcal{A}_{\{A,b\}} - \mathcal{A}_{\{a,B\}}. \quad \text{(S29)}$$

On the other hand, $\mathcal{A}_{\Delta}^{(ij)/\ell}$ and $\mathcal{S}_{\Delta}^{(ij)/\ell}$ measures whether more uninfected hosts have immunity to antigens i, j and the reference allele $\ell \in \{a, b, c\}$ at the third locus, or homologous

immunity to the strain whose genotype consists of alleles i, j, ℓ. For example,

$$\begin{aligned}\mathcal{A}_{\Delta}^{(AB)/c} &= \mathcal{A}_{\{A,B,c\}} + \mathcal{A}_{\{a,b,c\}} - \mathcal{A}_{\{A,b,c\}} - \mathcal{A}_{\{a,B,c\}},\\ \mathcal{S}_{\Delta}^{(AB)/c} &= \mathcal{S}_{ABc} + \mathcal{S}_{abc} - \mathcal{S}_{Abc} - \mathcal{S}_{aBc}.\end{aligned}$$

Finally, when there are three loci, we need to also account for triple epistasis in fitness, denoted w_{ABC}. Since by definition, $r_{ABC} = r_{abc} + \sum_k w_k + \sum_{ij} w_{ij} + w_{ABC}$, this is

$$\begin{aligned}w_{ABC} &= r_{ABC} + r_{Abc} + r_{aBc} + r_{abC} - r_{abc} - r_{ABc} - r_{AbC} - r_{aBC}\\ &= \beta(\sigma_{\mathrm{PI}}^{(2)} - \sigma_{\mathrm{PI}}^{(1)})\mathcal{A}_{\Delta}^{(2)} - \beta\sigma_{\mathrm{HI}}\mathcal{S}_{\Delta}^{(2)}\end{aligned}$$

where $\mathcal{X}_{\Delta}^{(2)}$ indicates if more hosts have immunity to the antigens ($\mathcal{X} = \mathcal{A}$) or homologous immunity to the strains ($\mathcal{X} = \mathcal{S}$) belonging to one of the two set of strains whose members share no more than one antigen in common, that is,

$$\{ABC, Abc, aBc, abC\} \quad \text{or} \quad \{abc, ABc, AbC, aBC\}.$$

In particular,

$$\begin{aligned}\mathcal{A}_{\Delta}^{(2)} &= \mathcal{A}_{\{A,B,C\}} + \mathcal{A}_{\{A,b,c\}} + \mathcal{A}_{\{a,B,c\}} + \mathcal{A}_{\{a,b,C\}} - \mathcal{A}_{\{A,B,c\}} - \mathcal{A}_{\{A,b,C\}} - \mathcal{A}_{\{a,B,C\}} - \mathcal{A}_{\{a,b,c\}},\\ \mathcal{S}_{\Delta}^{(2)} &= \mathcal{S}_{ABC} + \mathcal{S}_{Abc} + \mathcal{S}_{aBc} + \mathcal{S}_{abC} - \mathcal{S}_{ABc} - \mathcal{S}_{AbC} - \mathcal{S}_{aBC} - \mathcal{S}_{abc},\end{aligned}$$

We are then in a position to write down the equations describing the evolutionary dynamics. As these are complicated, we will not do this here, but in what follows we use these equations to extract general qualitative predictions. Note that our emphasis is upon biological interpretability, and so we do not attempt to rigorously characterize the dynamics.

S7.1 Fast- and slow-timescales, and different types of strain structure

The first step is to observe that for the three diallelic locus model, there is a fast timescale over which $p_k(t) \to 1/2$, and a slow timescale over which pairwise- and triple-LD (strain structure) evolve. After the fast transient dynamics, on the slow timescale pairwise LD becomes

$$\begin{aligned}D_{AB}(t) &\approx \frac{p_{ABc}(t) + p_{abc}(t) - p_{aBC}(t) - p_{AbC}(t)}{2},\\ D_{AC}(t) &\approx \frac{p_{AbC}(t) + p_{abc}(t) - p_{aBC}(t) - p_{ABc}(t)}{2},\\ D_{BC}(t) &\approx \frac{p_{aBC}(t) + p_{abc}(t) - p_{AbC}(t) - p_{ABc}(t)}{2},\end{aligned} \tag{S30}$$

while triple LD becomes

$$D_{ABC}(t) \approx \frac{p_{ABC}(t) - p_{abc}(t)}{2}. \tag{S31}$$

Inspection of these quantities allows us to identify two distinct ways in which strain structure emerges:

(1) When a set of strains whose members share no more than one allele with any other is over-represented, that is, either strain set

$$\{ABC, Abc, aBc, abC\} \quad \text{or} \quad \{abc, aBC, ABc, AbC\}$$

is over-represented. This type of strain structure is driven by the build-up and maintenance of triple LD, and the removal of pairwise LD. Here, when strain structure is maximized (one set is competitively excluded), $D_{ABC}(t) = \pm 1/8$ while $D_{ij}(t) = 0$.

If we suppose there is no pairwise LD present ($D_{ij}(t) = 0$), then along the slow timescale

$$\frac{\mathrm{d}D_{ABC}}{\mathrm{d}t} \approx w_{ABC}\left(\frac{1}{8} - D_{ABC}\right)\left(\frac{1}{8} + D_{ABC}\right). \tag{S32}$$

By inspection of equation (S32), in order for triple LD to build up and be maintained, triple LD (D_{ABC}) and triple epistasis (w_{ABC}) must share the same sign.

(2) When one of the four sets of strains whose members share no alleles in common is over-represented, that is, one of

$$\{ABC, abc\}, \quad \{Abc, aBC\}, \quad \{aBc, AbC\}, \quad \text{or} \quad \{abC, ABc\}, \tag{S33}$$

is over-represented. This type of strain structure is driven by the build-up and maintenance of pairwise LD, and the removal of triple LD. Here, when strain structure is maximized (three of the sets in (S33) are competitively excluded) then $D_{ij}(t) = \pm 1/4$, and $D_{ABC}(t) = 0$.

If we suppose there is no triple LD present ($D_{ABC}(t) = 0$), then along the slow timescale

$$\frac{\mathrm{d}D_{ij}}{\mathrm{d}t} \approx \left(w_{ij} + \frac{w_{ABC}}{2}\right)\left(\frac{1}{4} - D_{ij}\right)\left(\frac{1}{4} + D_{ij}\right) + \sum_{k\ell}\left(w_{k\ell} + \frac{w_{ABC}}{2}\right)\left(\frac{D_{\hat{k}\hat{\ell}}}{4} - D_{ij}D_{k\ell}\right). \tag{S34}$$

From equation (S34) it is no longer necessary that w_{ij} must share the same sign as D_{ij} in order for pairwise LD to be maintained. Instead, $w_{ij} + w_{ABC}/2$ and D_{ij} need to share the same sign. Note that we can write

$$w_{ij} + \frac{w_{ABC}}{2} = \beta\sigma_{\mathrm{PI}}^{(1)}\mathcal{A}_{\Delta}^{(ij)} + \frac{\beta}{2}(\sigma_{\mathrm{PI}}^{(2)} - \sigma_{\mathrm{PI}}^{(1)})(\mathcal{A}_{\Delta}^{(ij)/\ell} + \mathcal{A}_{\Delta}^{(ij)/\hat{\ell}}) - \frac{\beta}{2}\sigma_{\mathrm{HI}}(\mathcal{S}_{\Delta}^{(ij)/\ell} + \mathcal{S}_{\Delta}^{(ij)/\hat{\ell}}). \tag{S35}$$

S7.2 Shape of function controlling antigen-specific cross-protections

Next, suppose $\sigma_{\mathrm{HI}} = 0$. Here we want to understand how the $\sigma_{\mathrm{PI}}^{(n)}$ affect the evolutionary dynamics. Recall that $\sigma_{\mathrm{PI}}^{(n)}$ measures whether the rate of increase of antigen-specific cross-protection is locally increasing ($\sigma_{\mathrm{PI}}^{(n)} > 0$) or decreasing ($\sigma_{\mathrm{PI}}^{(n)} < 0$) at n.

If we consider $(\sigma_{\mathrm{PI}}^{(1)}, \sigma_{\mathrm{PI}}^{(2)})$ space, there are two curves that divide up the qualitatively different evolutionary behaviour. In what follows we consider these two curves; the results

are presented in the main text Figure 5. Throughout, we assume that the more abundant strain set also has more hosts with antigenic immunity to it, that is, $\mathcal{A}_{\Delta}^{(2)}$ shares the same sign as D_{ABC}, and $\mathcal{A}_{\Delta}^{(ij)}$ shares the same sign as D_{ij}. We will not consider ways in which these relationships could breakdown as in Section S5.

Curve 1: $w_{ABC} = 0$. This occurs when $\sigma_{\mathrm{PI}}^{(2)} = \sigma_{\mathrm{PI}}^{(1)}$, that is, the rate of increase of cross-protection is the same regardless of the number of antigens the host has immunity to. This partitions the parameter space between when triple LD can be maintained and when it cannot. In particular, since triple LD requires the sign of w_{ABC} and D_{ABC} to agree, it can only be maintained if $\sigma_{\mathrm{PI}}^{(2)} > \sigma_{\mathrm{PI}}^{(1)}$.

Curve 2: $w_{ij}+w_{ABC}/2 = 0$. From equation (S34), this curve specifies when pairwise LD can, and cannot, form. Since this curve depends upon the epidemiological dynamics in addition to the parameters $\sigma_{\mathrm{PI}}^{(n)}$, it is not straightforward to precisely specify the curve in $(\sigma_{\mathrm{PI}}^{(1)}, \sigma_{\mathrm{PI}}^{(2)})$ space. However, note that in equation (S35), we can write $\mathcal{A}_{\Delta}^{(ij)} = \mathcal{A}_{\Delta}^{(ij)/r} + \mathcal{A}_{\Delta}^{(ij)/\hat{r}} - \mathcal{A}_{\Delta}^{(ij)/r\hat{r}}$, where

$$\mathcal{A}_{\Delta}^{(ij)/r\hat{r}} = \mathcal{A}_{\{i,j,r,\hat{r}\}} + \mathcal{A}_{\{\hat{i},\hat{j},r,\hat{r}\}} - \mathcal{A}_{\{\hat{i},j,r,\hat{r}\}} - \mathcal{A}_{\{i,\hat{j},r,\hat{r}\}}.$$

Since each of these terms requires hosts be infected by different pathogen strains, if we suppose that the frequency of hosts being infected multiple times is low compared to hosts being infected once, e.g., $\mathcal{A}_{\{i,j,r,\hat{r}\}} \ll \mathcal{A}_{\{i,j,r\}}$ then $\mathcal{A}_{\Delta}^{(ij)/r\hat{r}} \ll \mathcal{S}_{\Delta}^{(ij)/r} + \mathcal{A}_{\Delta}^{(ij)/\hat{r}}$, and so equation (S35) becomes

$$w_{ij} + \frac{w_{ABC}}{2} \approx \frac{\beta}{2}\big(\sigma_{\mathrm{PI}}^{(1)} + \sigma_{\mathrm{PI}}^{(2)}\big)\big(\mathcal{A}_{\Delta}^{(ij)/r} + \mathcal{A}_{\Delta}^{(ij)/\hat{r}}\big), \tag{S36}$$

Hence $w_{ij} + \frac{w_{ABC}}{2} \approx 0$ when $\sigma_{\mathrm{PI}}^{(1)} = -\sigma_{\mathrm{PI}}^{(2)}$, or $\sigma_0 - \sigma_1 - \sigma_2 + \sigma_3 = 0$. Since the maintenance of pairwise LD requires the sign of $w_{ij} + w_{ABC}/2$ and D_{ij} to agree, if we are willing to accept the approximation of (S36), then pairwise LD can be maintained if $\sigma_{\mathrm{PI}}^{(2)} > -\sigma_{\mathrm{PI}}^{(1)}$, otherwise it cannot.

From consideration of these curves and Figure 5, notice that if both

$$\sigma_{\mathrm{PI}}^{(2)} > -\sigma_{\mathrm{PI}}^{(1)} \quad \text{and} \quad \sigma_{\mathrm{PI}}^{(2)} > \sigma_{\mathrm{PI}}^{(1)} \tag{S37}$$

are satisfied, it is possible that pairwise LD can be maintained in the absence of triple LD, but also that triple LD can be maintained in the absence of pairwise LD. This creates the possibility of an evolutionary bistability: depending upon the initial conditions, the population can stably exhibit different types of structure, that is, either $|D_{ABC}|$ is maximized and $D_{ij} = 0$, or $D_{ABC} = 0$ and $|D_{ij}|$ is maximized.

Biologically speaking, from our analysis above and Figure 5 and Figure S5, increasing $\sigma_{\mathrm{PI}}^{(1)}$ and/or $\sigma_{\mathrm{PI}}^{(2)}$, produces strain structure, but depending on which of these increases more rapidly, different types of structure can occur. Increasing $\sigma_{\mathrm{PI}}^{(2)}$ relative to $\sigma_{\mathrm{PI}}^{(1)}$ tends to favour triple LD, and so a population in which there are four strains present, each of which share no more than one antigen with any other. If instead $\sigma_{\mathrm{PI}}^{(1)}$ is increased relative to $\sigma_{\mathrm{PI}}^{(2)}$, this favours pairwise LD, and so a population in which only two strains are present, each of which share no antigens. This is logical: if $\sigma_{\mathrm{PI}}^{(1)}$ is positive and larger than $\sigma_{\mathrm{PI}}^{(2)}$ immune protections

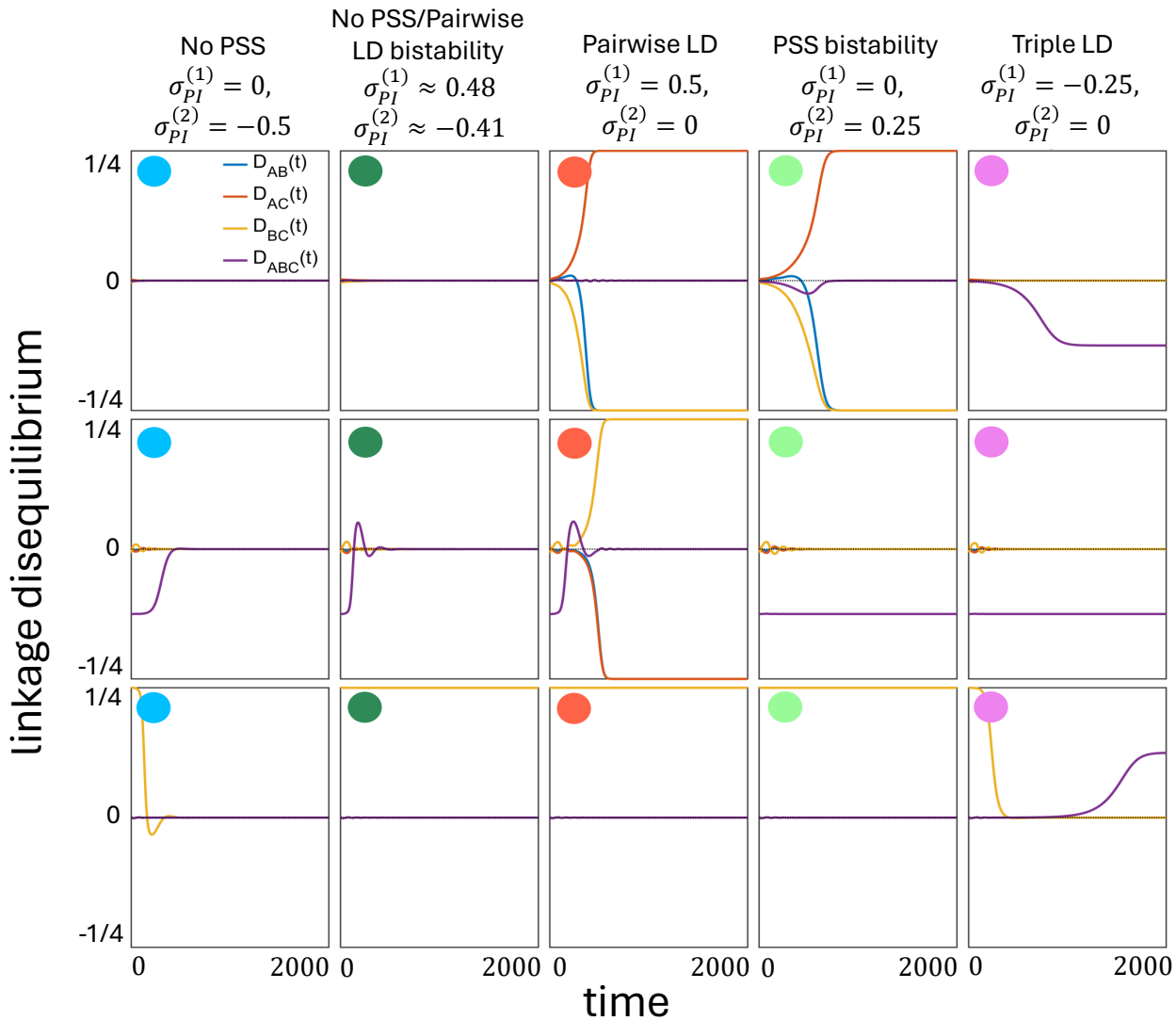

Figure S5: **Dynamical behaviour of LD as $\sigma_{\mathbf{PI}}^{(1)}$, $\sigma_{\mathbf{PI}}^{(2)}$ and the initial conditions vary.** The dynamics of pairwise- and triple LD ($D_{ij}(t)$ and $D_{ABC}(t)$) depend upon the values of $\sigma_{\mathrm{PI}}^{(1)}$ and $\sigma_{\mathrm{PI}}^{(2)}$ (different columns) and the initial conditions (different rows). Each column corresponds to the indicated values of $\sigma_{\mathrm{PI}}^{(1)}$ and $\sigma_{\mathrm{PI}}^{(2)}$, with the coloured circle indicating the approximate position on Figure 5. Each row corresponds to a different initial condition. In the first row, all simulations start near a state of no PSS; in the second row, all simulations start near a state of triple LD PSS; in the third row, all simulations start near a state of pairwise LD PSS. All simulations used $b = d = 0.01$, $\beta = 1$, $\mu = 0.4$, $\rho = 10^{-6}$, $\sigma_0 = 1$ and $\sigma_3 = \sigma_h = 0$, while σ_1 and σ_2 vary as indicated by column.

are much more effective against a single antigen then expected, and so the minimum amount of strain diversity that preserves antigenic variation is favoured. If $\sigma_{\mathrm{PI}}^{(2)}$ is instead large, then immune protections are much more effective when strains share two antigens, and is relatively ineffective against strains sharing a single antigen. Consequently, greater strain

diversity can be maintained in the population.

S7.3 $\sigma_{\mathrm{HI}} > 0$ promotes oscillations and maintains diversity

When $\sigma_{\mathrm{HI}} > 0$, this has a similar effect for three loci as it does when there are only two. In particular, it increases the likelihood that w_{ABC} and $w_{ij} + w_{ABC}/2$ do not share the same sign with D_{ABC} and D_{ij} respectively, and thus acts to break-up strain structure and promote genotypic diversity. Likewise, by introducing a second epidemiological feedback of hosts with homologous immunity, either $\mathcal{S}_{\Delta}^{(ij)/r}$ (for w_{ij}) or $\mathcal{S}_{\Delta}^{(2)}$ (for w_{ABC}), $\sigma_{\mathrm{HI}} > 0$ can produce evolutionary oscillations within either level of strain structure, as the over-represented strain set varies (Fig. S6). Simulations suggest that these oscillations do not occur between different types of structure in the deterministic system, even when bistability is possible, but in the presence of demographic stochasticity it seems likely that this can occur, which would further complicate the dynamics.

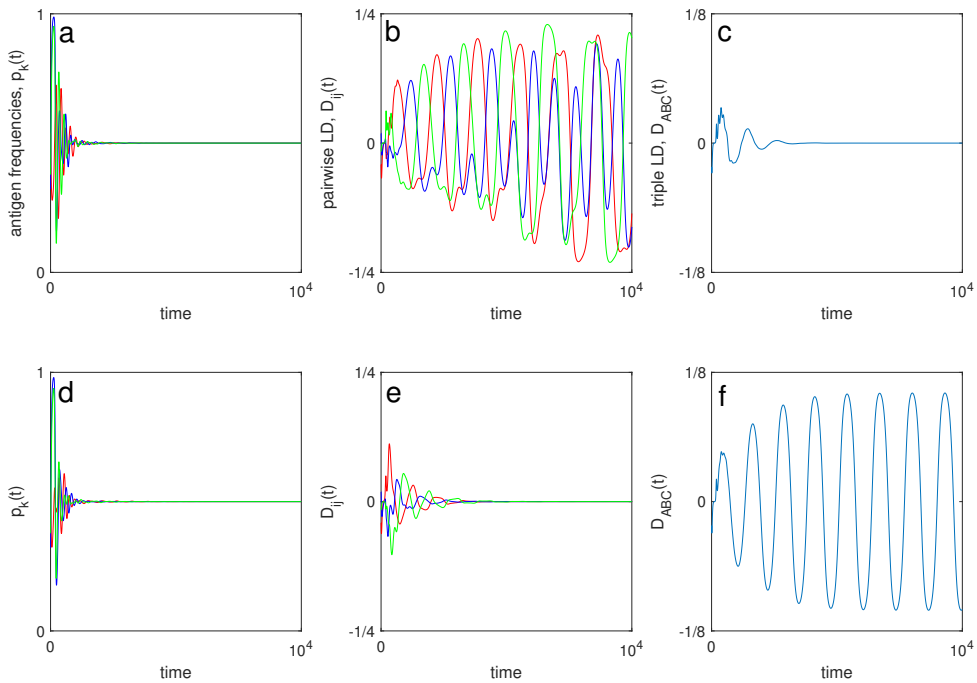

Figure S6: **Evolutionary oscillations for more than two loci.** For more then two loci there is more than one possible type of PSS. Consequently, evolutionary oscillations can occur for any order of LD. All simulations used $b = d = 0.005$, $\beta = 1$, $\mu = 0.5$, $\rho = 10^{-6}$, $\sigma_0 = 1$ and $\sigma_h = 0.05$, while panels **a-c** $\sigma_1 = 0.6$, $\sigma_2 = 0.25$, $\sigma_3 = 0.2$, whereas for panels **d-f**, $\sigma_1 = 0.7$, $\sigma_2 = 0.3$, $\sigma_3 = 0.2$.