

RESEARCH ARTICLE

The evolutionary cost of homophily: Social stratification facilitates stable variant coexistence and increased rates of evolution in host-associated pathogens

Shuanger Li^{1‡}, Davorka Gulisija^{2*}, Oana Carja^{1*}

1 Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States of America, **2** Department of Biology, University of New Mexico, Albuquerque, New Mexico, United States of America

‡ Current address: Ecology and Evolution Graduate Program, University of Chicago, Chicago, Illinois, United States of America

* dgulisija@unm.edu (DG); oana@cmu.edu (OC)



OPEN ACCESS

Citation: Li S, Gulisija D, Carja O (2024) The evolutionary cost of homophily: Social stratification facilitates stable variant coexistence and increased rates of evolution in host-associated pathogens. *PLoS Comput Biol* 20(11): e1012619. <https://doi.org/10.1371/journal.pcbi.1012619>

Editor: Roger Dimitri Kouyos, University of Zurich, SWITZERLAND

Received: May 8, 2024

Accepted: November 7, 2024

Published: November 22, 2024

Copyright: © 2024 Li et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All code necessary to reproduce the results presented here are available at https://github.com/shuangerli/virus_polymorphism_homophily.

Funding: This work was supported by the National Institutes of Health (award no. R35GM147445 to OC) and the United States-Israel Binational Science Foundation (award no. 2019266 to OC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Coexistence of multiple strains of a pathogen in a host population can present significant challenges to vaccine development or treatment efficacy. Here we discuss a novel mechanism that can increase rates of long-lived strain polymorphism, rooted in the presence of social structure in a host population. We show that social preference of interaction, in conjunction with differences in immunity between host subgroups, can exert varying selection pressure on pathogen strains, creating a balancing mechanism that supports stable viral coexistence, independent of other known mechanisms. We use population genetic models to study rates of pathogen heterozygosity as a function of population size, host population composition, mutant strain fitness differences and host social preferences of interaction. We also show that even small periodic epochs of host population stratification can lead to elevated strain coexistence. These results are robust to varying social preferences of interaction, overall differences in strain fitnesses, and spatial heterogeneity in host population composition. Our results highlight the role of host population social stratification in increasing rates of pathogen strain diversity, with effects that should be considered when designing policies or treatments with a long-term view of curbing pathogen evolution.

Author summary

Understanding the mechanisms that shape and protect the maintenance of diversity in a pathogen population is of significant importance to curbing their spread in the population. Here we discuss a new diversity-promoting mechanism where, unlike typical balancing mechanisms, the diversity is not buffered from selection in protected life-stages of the pathogen or distinct spatially separated subpopulations, but instead, variant diversity is protected by the social interaction structure of the host population. In turn, increased diversity in the pathogen population increases the rate of their evolution, incurring the

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

health cost to the host population, which we refer to as the evolutionary cost of homophily. Our model can also be interpreted through the lens of contagious social behaviors and showcases how the social architecture of the population can shape the coexistence of a diverse array of cultural variants. For both cultural or biological pathogens, understanding the evolutionary mechanisms that contribute to transient or stable variant diversity is essential for designing responses and policies that prevent increases in strain repertoire of pathogenic variants in the population.

Introduction

From genes to communities, natural systems are characterized by high rates of variant coexistence and understanding the mechanisms that shape and protect the maintenance of diversity in a population is a central problem in biology. Evolutionary mechanisms that are known to promote long-lived, non-neutral strain polymorphism are usually balancing mechanisms, arising from direct negative frequency dependence [1] or from spatially heterogeneous selection pressure [2]. An additional class of mechanisms that have been theoretically shown to promote polymorphism across a wide range of evolutionary scenarios are storage effects, initially recognized in studies of species coexistence [3–6]. Storage effects can promote coexistence when there exist patches of habitat (or specific life stages, or genetic backgrounds) where selection against deleterious variants is diminished, compared to other patches [5, 7–10]. As a result, this effect can store long-lived polymorphism in systems with spatial, life-stage or genetic heterogeneity [7, 8, 11–13].

Here we ask whether patterns of preferential social interaction between individuals can, similarly, generate enough differential in selection pressure to stably increase rates of variant coexistence in a population. Biased preferences of interaction and their effects on population segregation have been widely documented in the social science literature, starting with the ground-breaking studies of Thomas C. Schelling on how micro-motives can shape macro-patterns of population behavior [14–16]. Across many dimensions of phenotypes (including physical, cultural, and attitudinal characteristics), humans exhibit high levels of homophily, the tendency to interact with others of similar type, in social tie formation and patterns of social interaction [17–21, 53]. Moreover, homophily has been shown to significantly shape patterns of pathogen or cultural variant spread. For example, the structure of sexual networks of interaction has been shown to impact the spread of HIV [22] and opinion diffusion through social networks has been shown to influence individuals' decision for vaccination, thus indirectly also shaping contact-dependent vulnerability to pathogens [23]. Recent studies have also investigated the impact of homophily on infection dynamics during the COVID-19 pandemic [24, 25], and highlighted the importance of considering homophily when developing vaccination strategies [26, 27].

In parallel, empirical work has explored how denser patterns of interaction can lead to higher levels of phenotypic and/or immunological similarity between individuals. Local environmental conditions determined by social interactions have been proposed as key determinants of human cellular immune systems, creating different patterns of interaction between different immunological profiles that could significantly impact pathogen spread and diversity [28, 29]. Vulnerabilities to infections correlate with factors like age, ethnicity, exposure histories, and immune system competence [30–38]. Similarly, extensive bacterial strain sharing across human populations, with distinct mother-to-infant, intra-household and intra-

population transmission patterns, points to the important role of the contact network in shaping microbiome diversity and transmission [39–41].

The influence of social patterns of preferential interaction between population groups that are more immunologically and phenotypically similar on the long-term evolutionary dynamics of the pathogen population remains unclear. Here we use a population genetic model to study how social biases of interaction between host population subgroups with varying immune responses or phenotypic characteristics shape rates of pathogen strain polymorphism and long-lived coexistence. We show that, as strains of varying virulence enter the host population, these potential differences in immunity between host subgroups of interaction can promote polymorphism-protecting heterogeneous selective pressure against the pathogen. Therefore, the prevalence of a strain will be increased when circulating within the host group favoring the strain and will be constrained within the social group where the strain is disadvantaged.

This dynamic generates a diversity-promoting mechanism where, unlike typical balancing mechanisms, the diversity is not stored (buffered from selection) in protected life-stages of the pathogen or distinct spatially separated subpopulations, but instead variant diversity is protected by the social interaction structure of the host population. In turn, increased diversity in the pathogen population increases the rate of their evolution, incurring the health cost to the host population, which we refer to as the evolutionary cost of homophily. We show that increased rates of homophily in the host population promote increased rates of heterozygosity and long-lived multi-strain coexistence and study the role of temporal variance in the host social preference of interaction. We find that periodic amplification of existing homophily in the host population, even when short-term, can significantly increase rates of pathogen diversity and polymorphism.

Our model can also be interpreted through the lens of contagious social behaviors, where two cultural variants are maintained in the population due to the preferential interaction between individuals with similar phenotypic characteristics, such as conformist or anti-conformist behaviors [42, 43]. For both cultural or biological pathogens, understanding the evolutionary mechanisms that contribute to transient or stable variant diversity is essential for designing responses and policies that prevent increases in strain repertoire of pathogenic variants in the population.

Model description

To examine the effect of social preferences of interaction on rates of host-associated pathogen diversity and polymorphism, we use a Wright-Fisher model to describe changes in strain frequencies within a host population of fixed size N . The host population consists of two different types of individuals, with the different host immune-phenotypes denoted by S and A , interpreted here as individuals with different sensitivity to pathogen virulence, and broadly referred to as symptomatic (S) and asymptomatic phenotypes (A). These host phenotypic differences correspond to potential differences in immune system composition and effectiveness between people with different genetic backgrounds or in different age groups, which can affect host symptoms and disease progression when infected [30–38]. Since the host-population generation times correspond to an equivalent of many microbial generations, we assume that the proportion of A to S individuals in the host population does not change through the time of pathogen evolutionary dynamics.

Each host individual carries one (and only one) of two different pathogen strains, denoted here by v/V . We assume that these two strains differ in their host-pathogen affinity. For example, two different viral strains could differ in their virulence, with the v strain having a lower virulence than the V strain. Pathogen fitness is determined by the interplay between the

virulence of the strain and the immune vulnerability of the host it infects [44]. We model a scenario with trade-off viral fitness distributions on the two host backgrounds A and S, with optimal virulence smaller on the more sensitive, symptomatic S background than on the asymptomatic A background (Fig 1). It assumes that the benefits of a higher transmission rate can only accrue if the host is healthy enough to interact with other hosts and transmit the strain further. Pathogens with the highest fitness are those with an intermediate level of virulence, striking a balance between within-host production of more transmission forms per unit time and infection strength and length.

This model can also be applied to understand the evolutionary dynamics of two ‘contagious’ cultural traits (for example risk aversion or risk taking behaviors that affect the health of the individual), with different levels of expression on two different individual backgrounds (for example, conformist and anti-conformist hosts or hosts of different status or cultural background [42, 43]), where trait fitness and transmissibility to other hosts depends on the virulence of the trait, in interplay with its host phenotype and sensitivity to the cultural behavior.

We use the following fitness scheme:

	Strain V	Strain ν
Fitness on host A background	$w_{AV} = 1 + s_1$	$w_{A\nu} = 1$
Fitness on host S background	$w_{SV} = 1$	$w_{S\nu} = 1 + s_2$

The more virulent strain V is assumed to have higher fitness on the A-type hosts (determined by positive selection coefficient s_1), while strain ν has higher fitness on the S-type hosts (determined by positive selection coefficient s_2) (Fig 1). For example, one can interpret the model as symptomatic hosts S, infected with a very virulent strain being likely to show symptoms early and reduce social activities, effectively leading to a reduced transmission rate and a selective disadvantage of the V strain on this background. On the other hand, immunocompetent hosts A show fewer symptoms, even with the V strain, which can contribute to higher host transmission rates, thus conferring a selective advantage to the V strain on this host background, as has been described in models of viral-host dynamics [45]. Similarly, for the ν strain, the lower virulence allows for higher transmission on the symptomatic S background. Trade-offs in virulence, transmissibility and pathogen fitness [46–50] are complex and can depend on many parameters. However, the fitness scheme we use is a good baseline model, elegant in

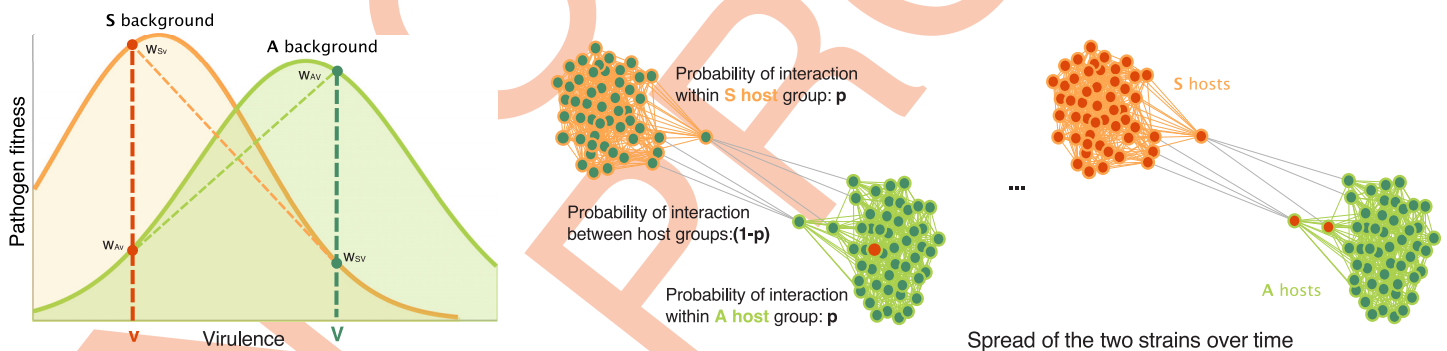


Fig 1. Illustration of the model. Pathogen fitness is determined by the interplay between the virulence of the strain and the immune vulnerability of the host it infects. Left: The more virulent strain V (dark green) is assumed to have higher fitness on the A-type (light green) hosts ($w_{AV} > w_{A\nu}$), while strain ν (dark orange) has higher fitness on the S-type (light orange) hosts ($w_{S\nu} > w_{SV}$). Middle and right: Illustration of host social stratification and implications for pathogen coexistence, over time. Host genotypes, A (light green) and S (light orange), are indicated by the bounding circles, while the pathogen genotypes they carry, ν (dark green) and V (dark orange), are indicated by the color fill. The host population is socially structured, with A and S individuals more likely to interact with other individuals of the same phenotype. At the beginning of the simulation, ν is the resident strain, and a new mutant V enters the population. Through time, the strains may concentrate on their preferred hosts.

<https://doi.org/10.1371/journal.pcbi.1012619.g001>

its simplicity, to study how the composition and interaction patterns in the host population affect pathogen rates of coexistence.

Since our goal is to study the evolution of the pathogen in the population of infected individuals, we assume that all hosts in the population are infected with exactly one of the two strains, the probability of coinfection is considered negligible and, initially, at time $t = 0$, all hosts are assumed to be infected with the wild-type strain v . In the first generation of each simulation run, a new mutation introduces the more virulent strain V in one of the hosts. At each generation, each original host can transmit its pathogen to new hosts, with probability determined by the pathogen strain fitness. Hence, the strain frequencies of the next generation are generated by sampling with replacement proportional to pathogen strain fitnesses and host interaction patterns and we study the change in frequencies of the two pathogen strains in the population. We do not allow for recurrent mutation.

An important aspect of the model is that hosts do not interact randomly, but instead, preferentially with other hosts belonging to the same social strata (henceforth group). That is, we assume that each host belongs to one of the two groups and will interact with a random individual from their own group with probability p , a social preference parameter. In our model, social preference of interaction does not depend on group sizes. This is due to the fact that interactions are probability-based: an individual will get infected with a probability p by someone from their group and $(1 - p)$ by someone from another group (Fig 1). Therefore, if the two social strata are of different sizes, individuals from a smaller group infect more individuals from the larger group and vice versa, which would happen, for example, when an “outsider” individual interacts or moves through a different, more densely populated segment of the population.

In addition, when groups are phenotypically homogenous, the social preference of interaction, or homophily, does not depend on the proportion of the different phenotypes in the population. Note that when the two groups are of equal sizes, a population with randomly interacting individuals would have $p = 0.5$. The setup above directly models the possibility of higher viral transmission opportunities within groups than between groups [15, 16, 19, 51, 52]. In what follows, we make a simplifying assumption that p is equal for the two groups. Most of our simulations also assume that groups are large and phenotypically homogenous (A or S). Thus, p becomes the probability that an individual will interact with an individual of the same phenotype at the population level. In extensions of the model where groups are phenotypically heterogenous (a mixture of A and S), the probability of transmission to a host of the same phenotype in the same group becomes a function of p and the proportion of the phenotype within the group.

Strain coexistence and stable polymorphism in the host population can be quantified by the cumulative expected heterozygosity

$$H = 2\mathbb{E} \left(\sum_{t=1}^{t=T} q_t (1 - q_t) \right), \quad (1)$$

where q_t represents the frequency of strain V at time t and the expectation is over independent simulation runs (ending at times T) [54]. Here H represents the expected sum of heterozygosities over the lifetime of a novel mutant and it can quantify departure from neutrality, independent of population size. When selection is neutral and allele or strain frequencies are only affected by genetic drift, under a randomly mating population, the cumulative expected heterozygosity $H_{neutral}$ has been shown to be equal to two, regardless of the population size [54], and thus $H > 2$ describes an elevated level of polymorphism (i.e. coexistence) relative to that

under neutrality. It can also be shown that the expected heterozygosity in a haploid population under recurrent mutation μ equals $HN\mu$ [54].

Social preference of interaction can however generate population subdivision and inflate the cumulative expected heterozygosity H . We correct for the effects of population stratification by measuring the subdivided cumulative expected heterozygosity,

$$H_L = 2\mathbb{E}\left(\sum_t \sum_i \frac{q_{i,t}(1 - q_{i,t})N_i}{N}\right), \quad (2)$$

where $q_{i,t}$ represents the frequency of strain V in host social interaction group i at time t , and N_i represents the size of the host population of immuno-phenotype i . The subdivided cumulative expected heterozygosity H_L has also been shown to equal two under neutrality, irrespective of the specifics of population subdivision [9]. Therefore, when the levels of polymorphism in the population exceed those under drift, we expect $H_L > 2$ i.e., we observe strain coexistence beyond that expected by chance.

We study the subdivided cumulative expected heterozygosity H_L as a function of the social preference of interaction p , the selection coefficients s_1 and s_2 , the population size N , and the ratio of the two host phenotypes frequencies in the population.

We also extend this baseline model to show its robustness to various relaxations of our assumptions:

Temporally varying preference of social interaction. Social preferences of interaction are known to temporally vary [55]. To relax our requirement of constant social interaction preferences, we consider a model in which the hosts change their interaction preferences periodically. We assume hosts interact with preference p for n_1 virus generations, and interact without preference for n_2 virus generations. Here, n_1 and n_2 need not be equal.

Variance in immune phenotypes within host groups or social strata. To reflect the natural demographic variance in host immuno-phenotype and test the model robustness to variance in immune phenotypes within host subgroups of preferential social interaction, we develop an extension of the model in which we model two randomly interacting population subgroups of fixed population sizes $\frac{1}{2}N$, each containing a mixture of A and S individuals. Therefore, the host phenotype composition is different between the two host subgroups and the probability of interaction and pathogen transmission is now dependent on the A and S phenotype abundances in each subpopulation (Fig A in S1 Text).

Changes in host population size and periodic host bottlenecks. Initially, we assume that the infection rates in the host population are such to generate a large number of infected hosts proxied by a constant population size N . To accommodate for seasonal changes in infection rates, we also model variable N , where the infected population size is initiated at a point in a period of oscillating sizes ranging from $0.05N$ to N to $0.05N$, repeatedly. This generates periods of strong bottlenecks in the affected population.

Implementation and simulation details

We use Monte Carlo simulations [56] to compute the subdivided cumulative expected heterozygosity H_L using an ensemble of at least 5×10^6 independent replicate populations of size $N = 10^5$. In addition to the subdivided cumulative diversity, we record the proportion of simulation runs in which long-lived strain polymorphism is maintained in the population at $100N$ generations. We also record the time of fixation or loss of the two strains, which allows us to quantify the duration of protected polymorphism that perishes. Elevated H_L is compared with values from control simulations either under strain neutrality (all fitness coefficients equal to one) or under a scenario of immune reaction homogeneity across the population.

Simulations are terminated when the new mutant V fixes or goes extinct in the population, or, in the case of stable long-lived polymorphism, the two strains coexist over a range of 10^7 (or $100N$) generations. We assume a host can only be infected by a pathogen strain obtained from one other host at each generation, there is no co-infection, and the effects of lethality or immunity between generations are negligible.

Unless otherwise specified, the population contains equal proportions of A and S hosts. In the case of temporally varying preference of social interaction, we implement periodic changes in the homophily parameter p at deterministic time periods of n_1 and n_2 generations, repeatedly. We implement both symmetric, as well as asymmetric temporal periods, where the duration of the time interval with social preference of interaction can be different from the duration of the time intervals of random interaction and mixing between the hosts.

For the extension of the model to incorporate immune-heterogeneous host subgroups, we simulate $R = 10^7$ independent replicate populations of size $N = 10^5$ (each host group is assumed to have $N = 50000$). We assume two different social subgroups or social strata, each containing a mixture of A and S hosts, with host phenotype composition (i.e. proportions of A and S hosts) different between the two subgroups. In this version of the model, we assume that hosts interact with social preference p with other hosts in their social strata (regardless of host phenotype) and, with probability of interaction $(1 - p)$ with hosts in the other social subgroup (Fig A in [S1 Text](#)). This application could represent different schools, for example, where individuals of different ages are more likely to interact within their school or social strata than between schools, regardless of their immuno-phenotype. Transmission here is now dependent on the A and S abundances in the two population subgroups.

Results

Social preferences of interaction promote long-lived pathogen coexistence

Even in a randomly interacting population, the subdivided cumulative expected heterozygosity H_L is a magnitude higher than drift controls for relatively low values of the selective differential s and further increases with increasing selective differential s ([Fig 2A](#) and [Fig B](#) in [S1 Text](#)).

This is because even though the individuals in the population are randomly interacting, the two host immuno-phenotypes in effect create heterogeneous selection, with each host immune-type providing opposite selection pressure on the virus from the other. This protects pathogen diversity in the population for longer than expected under controls.

As the social stratification of the host population increases ($p > 0.5$), rates of strain heterozygosity raise rapidly, even for small values of social bias of interaction p ([Fig 2A](#) and [Fig C](#) in [S1 Text](#)). This rise in levels of cumulative subdivided heterozygosity H_L is more pronounced as the selective differentials between the two fitness strains become stronger. For large enough selection and social preference of interaction, we observe a long-lived polymorphism that stably persists for at least $100N$ generations under a large set of parameter combinations. We sample frequency trajectories to gain insight into the relative evolutionary dynamics between the two strains and find that their frequencies, f_V and f_v , oscillate inside a small stable frequency range where V preferentially infects A hosts and v preferentially infects S hosts ([Fig 2B](#)) and this coexistence of v and V can be maintained for extended periods of time ([Fig D](#) in [S1 Text](#)).

Interestingly, even though strains are each preferred on a single host immune type and with strong social stratification ($p = 0.95$), we observe high frequencies of the two strains on both A and S individuals. In other words, V strain proliferation on their preferred A hosts, also drive high rates of V on S hosts in a relatively short period of time, effectively maintaining high frequency of a more virulent strain in the more sensitive host group. Our results are robust to

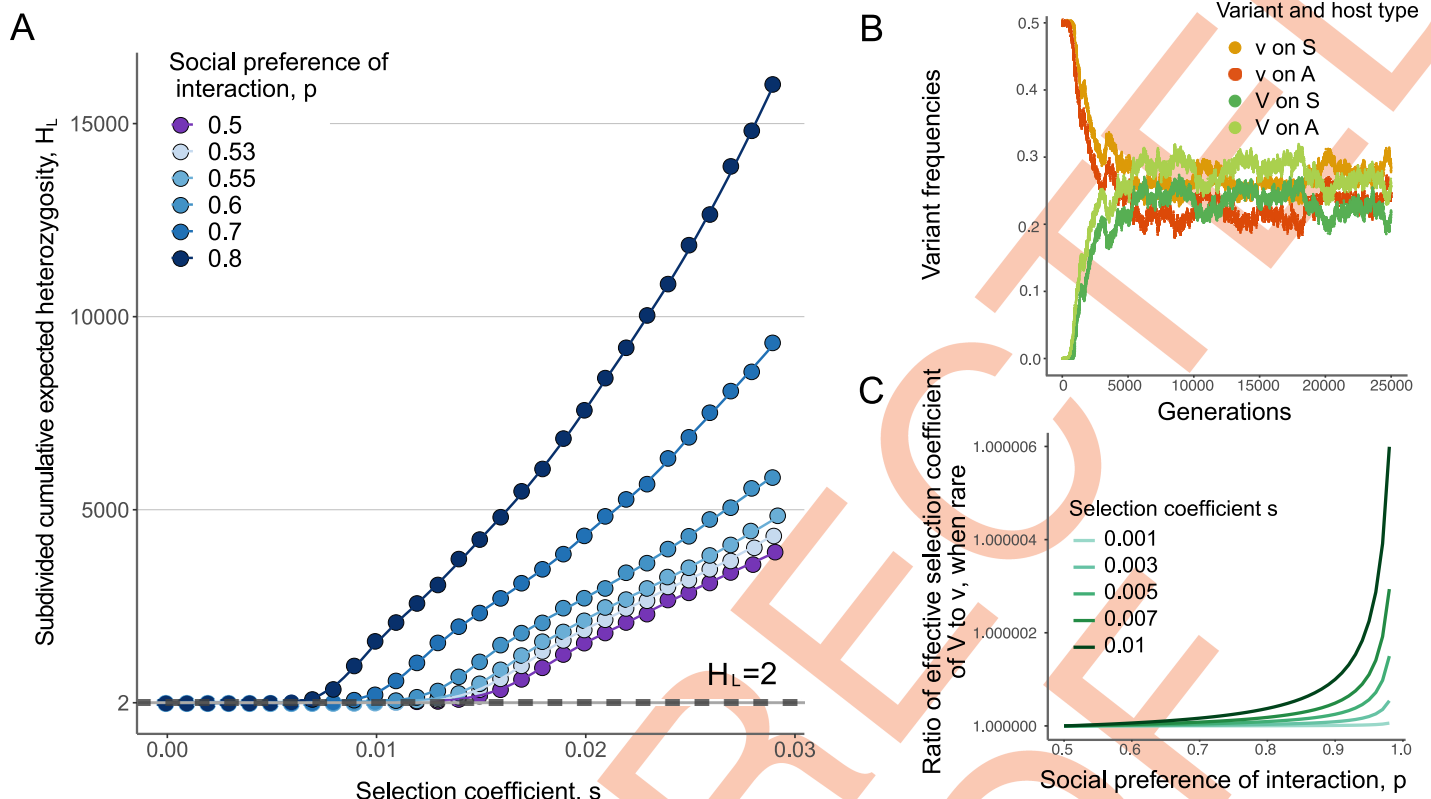


Fig 2. Social preference of interaction promotes increased pathogen strain coexistence. (A) The subdivided cumulative expected heterozygosity, H_L in a population of size $N = 10^5$ as a function of the selection coefficient $s_1 = s_2 = s$. Here, the social preference of interaction p is as presented in the legend. The dots represent ensemble averages across 10^7 replicate Monte Carlo simulations, while the lines represent cubic spline regression. The dotted straight line shows $H_{neutral} = 2$. (B) Population frequencies of each strain-host combination through time for one run of the simulation, with $N = 10^5$, $p = 0.95$, and $s_1 = s_2 = 0.02$ and equal frequencies of A and S individuals. (C) Theoretically-derived relative effective selection difference between the two strains when V is rare in the population, as a function of the strength of selection and social preference. This effective selection difference increases with both the strength of selection s and the social preference of interaction p , showcasing the role of negative frequency dependence in shaping patterns of heterozygosity and polymorphism.

<https://doi.org/10.1371/journal.pcbi.1012619.g002>

changes in population size N and different proportions of A to S host individuals in the population (Figs E and F in S1 Text).

Mechanisms that lead to balancing selection and long-term coexistence can be shown to, directly or indirectly, invoke negative frequency dependence. That is, polymorphism persists if there is a mechanism favoring whichever form is rare. In this model, the increased strain diversity is maintained through indirect negative frequency dependence created by immune-heterogeneity and social stratification patterns of the host population. Each strain experiences heterogeneous selection through the opposite fitness effects that it experiences on the two different host immune-phenotypes. Here, the opposing selection pressures on the two host immuno-phenotypes generate an association between strain and host that is broken as the pathogen is transmitted to another host. Each strain escapes negative selection in a unfavorable host group through transmission to a favorable host (with probability $(1 - p)$) and increases its numbers by within-group transmission on the preferred host background (with probability p) and the combined effect of the two selective forces depends on the frequency of the strain in the population.

To understand the effective negative frequency dependence present in this model, we show that the rare form V is indeed favored in a population. Since the two strains have

symmetrically opposite fitness effects on two immune phenotypes, the derivation also applies to the v strain when rare [57, 58]. Let us assume that the V strain appears on its preferred host phenotype A and it quickly reaches a type of “interaction-selection balance” as it is transmitted between the two host backgrounds A and S . Interaction (host switching, $1 - p$) in this formulation mirrors mutation. At equilibrium, the frequency of V on host phenotype A , x , is given by

$$(w_{SV} - w_{AV})x^2 - (w_{SV} - w_{AV} + (1 - p)(w_{SV} + w_{AV}))x + (1 - p)w_{SV} = 0. \tag{3}$$

This implies that the equilibrium frequency of phenotype V in the A host group is given by f_{AV} :

$$f_{AV} = \frac{w_{AV} - w_{SV} - (1 - p)(w_{SV} + w_{AV})}{2(w_{AV} - w_{SV})} + \frac{\sqrt{4w_{SV}(1 - p)(w_{AV} - w_{SV}) + (w_{SV} - w_{AV} + (1 - p)(w_{AV} + w_{SV}))^2}}{2(w_{AV} - w_{SV})}. \tag{4}$$

The effective selection coefficient of the V strain, when rare, can therefore be written as

$$s_V = w_{SV}(1 - f_{AV}) + w_{AV}f_{AV}. \tag{5}$$

When s_V exceeds the effective selective coefficient of the resident strain v , $s_v = (w_{AV} + w_{SV})/2$ (for equal percentages of A and S host phenotypes), the V mutant strain is preferred, when rare in the population. The effective selection ratio $s_{effective} = s_{V,rare}/s_{v,fix}$ is plotted in Fig 2C and shown to increase with both s and p . Under weak selection, the probability the V strain increases past establishment from a single copy in the population can be approximated by twice the effective selection difference between the strains [59]. A similar argument as above reciprocally holds when the strain v is rare in a population fixed on the V strain.

The effect of temporally changing social preference of interaction on strain coexistence

Population structure in natural host populations is not constant through time, but can dynamically change, with periods of more population mixing [51, 52, 55]. We next ask how elevated rates of polymorphism in the pathogen population change when there are periodic epochs of random and non-random host patterns of interaction. This characteristic also makes our model different from other previous models specifically examining temporally-heterogeneous selection, since changes in social patterns of interaction can act to periodically stratify the population, followed by periods of host random mixing.

We find that social preferences of interaction in the host population need not be constant for elevated pathogen heterozygosity and polymorphism to occur. We consider a periodic temporal regime, with n_1 generations of interaction with fixed preference $p > 0.5$, followed by n_2 generations of random host interaction, $p = 0.5$, which removes population structure. Random host interaction is expected to make intra-strain competition less concentrated in favorable hosts, thereby decreasing heterozygosity. Fig 3A shows the values of heterozygosity H_L , with $n_1 = n_2 = 50$ over the whole range of interaction bias. The levels of strain heterozygosity are smaller when host bias changes through time, with inserted epochs of random interaction, but the temporal periods with stratified interactions are sufficient to maintain balanced patterns of pathogen polymorphism (Fig G in S1 Text).

In fact, even brief periods with non-random host pattern of interaction can promote much higher levels of pathogen coexistence than under random population mixing (Fig 3B). We explore different asymmetric combinations of n_1 and n_2 and show, not surprisingly, that

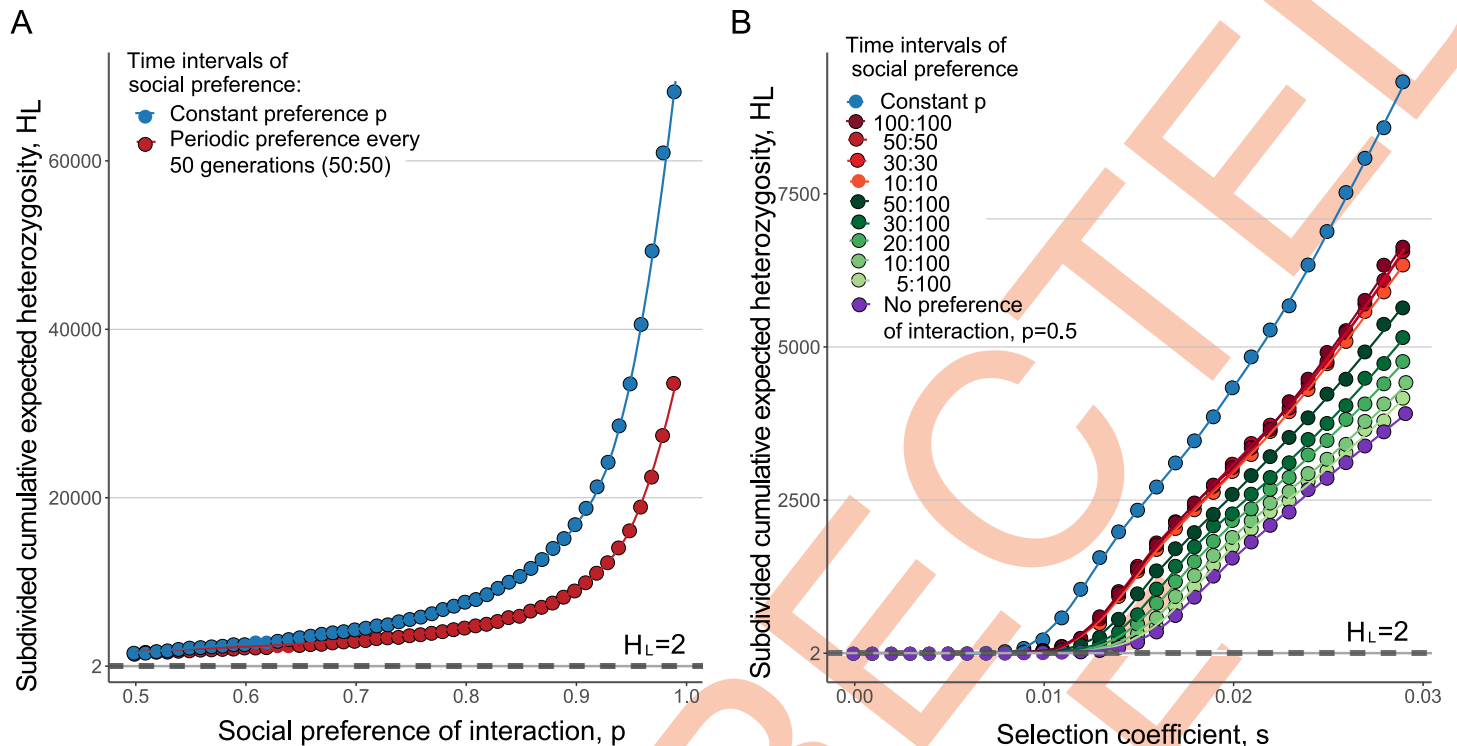


Fig 3. Elevated rates of polymorphism are observed even with periodic interruptions in social patterns of contact. Subdivided cumulative expected heterozygosity, H_L in a population of size $N = 10^5$ and equal frequencies of A and S individuals. The dots represent ensemble averages across 10^7 replicate Monte Carlo simulations of up to 10^7 generations, while the lines represent cubic spline regression. (A) Comparison of constant preference of interaction p with a fluctuating through time regime, where n_1 generations of preference $p > 0.5$ alternate with n_2 generations with no preference of interaction. Here, $n_1 = n_2 = 50$, and selection coefficients $s_1 = s_2 = s = 0.02$. (B) Different colors show different n_1, n_2 combinations, with $p = 0.7$.

<https://doi.org/10.1371/journal.pcbi.1012619.g003>

smaller $\frac{n_1}{n_2}$ ratios lead to lower levels of balanced polymorphism. Even when bias in interaction is only briefly acting in the population (for example, social interaction preference present a tenth of the time duration), the difference in H_L compared to randomly mixing populations increases with increasing selective differentials between the two mutant strains.

The effect of strain fitness differences on strain coexistence

Our results hold in the case of overall fitness differentials between the two strains ($s_d = s_1 - s_2 > 0$). With asymmetric fitness coefficients between the mutants on the two host immunophenotypes, the presence of social preference of interaction maintains elevated levels of strain heterozygosity and is expected to be more evident under strong social preference of interaction p and small fitness differentials s_d . In Fig 4, we show a nonlinear trend in the subdivided cumulative expected heterozygosity H_L . The strain heterozygosity increases with small differences in the selection coefficients, followed by a decrease as the difference in fitness increases. The larger the selective benefit of the wild-type strain on its preferred host type, the more pronounced this nonlinearity and the increase in population heterozygosity (Fig H in S1 Text).

The effect of host phenotypic heterogeneity between preferentially interacting individuals

Factors such as exposure histories, age structures, and ethnicities may create immune heterogeneity within socially isolated host groups or social strata. We next explore the robustness of

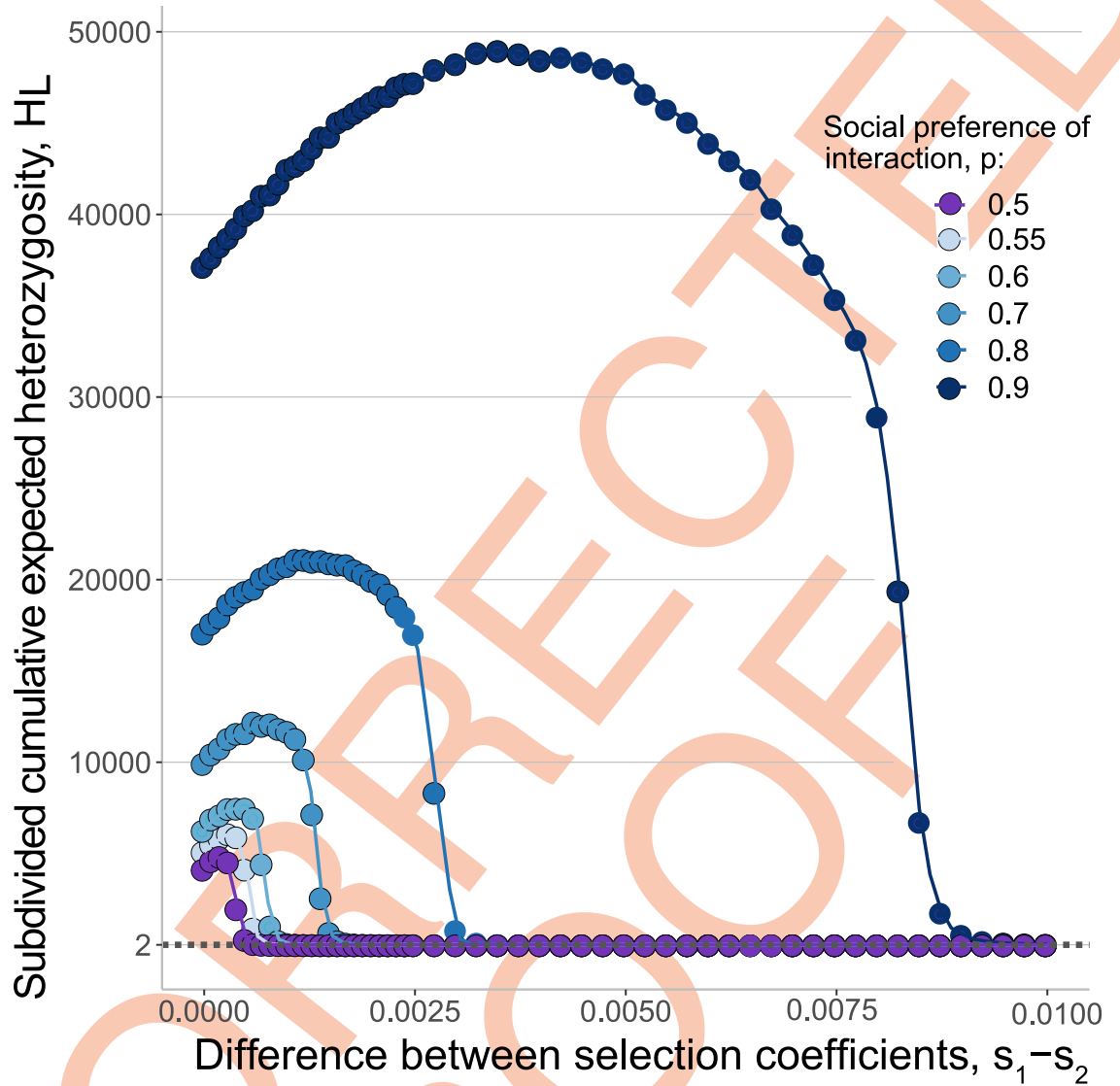


Fig 4. Pathogen strain polymorphism with asymmetric fitness effects between strains. Subdivided cumulative expected heterozygosity, H_L in a population of size $N = 10^5$, equal frequencies of A and S individuals and selection coefficient $s_2 = 0.03$, $0 \leq s_1 - s_2 \leq 0.01$. The dots represent ensemble averages across 10^7 replicate Monte Carlo simulations of up to 10^7 generations, while the lines represent cubic spline regression.

<https://doi.org/10.1371/journal.pcbi.1012619.g004>

the balancing effect arising from our model to immune or phenotypic heterogeneity within host group (A/S). To this end, we split the population into two social subgroups, d_1 and d_2 , each containing a mixture of both A and S individuals, and assume that hosts interact preferentially with another host in their social strata with probability p (regardless of host immunophenotype) and, with probability $(1 - p)$, the interaction occurs between social subgroups (Fig A in S1 Text). We vary the proportion of A phenotypes in each of the two subpopulations, $P_{d1}(A)$ and $P_{d2}(A)$, and assume $P_{d2}(A) = 1 - P_{d1}(A)$.

The strain balancing effect of heterogeneous selection and social distancing is robust to within-group immune diversity and is greater with greater differences in the proportion of A/S within groups (Fig 5). However, coexistence occurs even when the two groups have the same

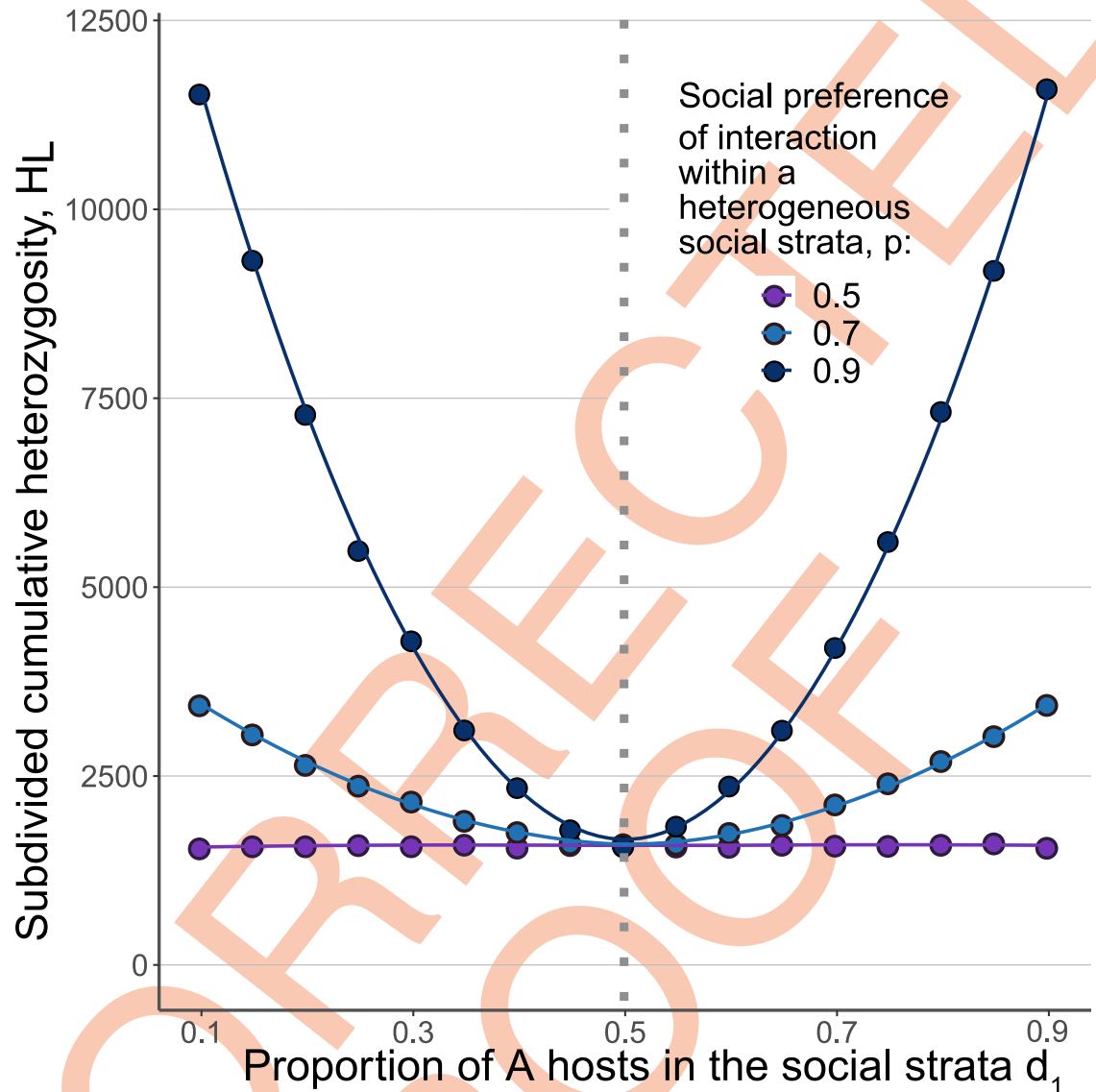


Fig 5. Strain coexistence with heterogeneous host population subgroups. The dots represent ensemble averages across 10^7 replicate Monte Carlo simulations of up to 10^7 generations, while the lines represent cubic spline regression. Here, $N_1 = N_2 = 5 \times 10^4$, $s_1 = s_2 = 0.02$ and random interaction within the subgroups but biased interaction between subgroups, as in legend. The proportion of A immune-phenotypes in group one is given on the x axis.

<https://doi.org/10.1371/journal.pcbi.1012619.g005>

proportions of immuno-phenotypes and are not socially distanced. As previously observed, the effect increases with p .

The effect of changes in host population size and periodic host bottlenecks

Lastly, we find that the reported effect is also robust to changes in the infection rates in the host population. To accommodate for seasonal changes in infection rates, we model severe oscillating bottlenecks in the number of affected individuals and assume an oscillating population size ranging from $0.05N$ to N repeatedly, starting at a random point in this cycle. In Fig I in S1 Text, we show that even frequent periods of bottlenecks in the host population maintain

elevated strain heterozygosity in the viral population and can promote much higher levels of pathogen coexistence than under random population mixing. Overall we find that our results are robust across a wide range of parameters and for a variety of model extensions. It is important to note that here we assume that the resident strain is already circulating in the population at the time an invader strain is introduced. While experiencing small infection rate (bottleneck) could serve as a proxy for a new outbreak, we did not model the outbreak itself and further models could accommodate for a wider set of epidemiological scenarios.

Discussion

As populations continuously adapt from one environment to the other, genetic polymorphism provides a readily available reservoir of adaptive alleles that selection can act upon and thus promotes population persistence [60–63]. Therefore coexistence of multiple pathogen strains in a host population poses significant challenges for vaccine development or treatment outcome. For example, if a vaccine only confers immunity to a subset of viruses circulating in the population, other variants can persist and become reservoirs for viral evolution and further vaccine evasion [63–65]. A central question thus becomes: which mechanisms enable multiple strain coexistence? Here we show how social stratification, combined with the existence of distinct host immuno-phenotypes, can lead to increased levels of pathogen strain or contagious cultural behavior coexistence and highlight the importance of social preference of interaction in promoting long-lived pathogen polymorphism and rates of evolution.

Mechanisms that promote long-term coexistence and maintain balanced polymorphism in a population, can be shown to, directly or indirectly, invoke some kind of negative frequency dependence. Direct negative frequency dependence acts through the interaction of host immunity and virus immunophenotypes, whereby common strains experience more intra-host competition due to prior host exposure to similar strains [1]. This favors the evolution of antigenically diverse strains, as seen in influenza, rotavirus, and HPV [66–68].

Here, we discuss a mechanism that promotes multi-strain pathogen coexistence through indirect negative frequency dependence driven by elevated biases of interaction between individuals in distinct immuno-phenotype groups in the host population. The complex interactions of human society result in heterogeneous contact networks where interactions are common between some individuals in a population and are entirely absent between others [69–71]. This stratification and sparsity of connections in human populations has previously been shown to affect the culture-wide adoptions of behaviors that can have different effects on fitness in different environments or individual phenotypic backgrounds [72]. Similarly, geographic clustering of a host population, which promotes spatio-temporal structures with varying immune selection, has been shown to affect viral diversity and evolution [2, 73].

The mechanism of coexistence in our model is rooted in the fact that different strains of the pathogen experience different selection pressures when infecting hosts of different immune vulnerability, i.e. immuno-phenotypes. Since host interaction patterns can also coincide with host vulnerability, prevalence of a strain is expected to be promoted within the favorable interaction group. Because host populations are immune-heterogeneous and hosts from different groups also interact, strain prevalence is also influenced by inter-strain competitions.

We show that long-lived strain coexistence is promoted by the immune heterogeneity between hosts and the magnitude of host population stratification. Our results are robust to overall fitness differences between the two mutant strains, changes in population size N and different proportions of A to S host individuals in the population, as well as temporal fluctuations in interaction patterns and within-group heterogeneity. We show that temporal fluctuations in the strength of the social preference of host interaction reduce rates of polymorphism,

but stratification nonetheless is the dominant effect, maintaining elevated levels of heterozygosity in the pathogen population.

It is important to note that our mechanism alone is fast enough to create multi-strain coexistence in short time scales (Fig J in [S1 Text](#)). We show that strain coexistence is achieved in a much shorter period of time in a stratified host population compared to a randomly interacting one, and it is maintained for long time scales. These results imply high evolvability in new pathogens with high transmission rates that are constrained between social groups, but allowed to spread within groups.

As a proof-of-concept, our model assumes constant proportions of *A* and *S* hosts in a population. While our robustness analysis shows that the effect we report holds across various proportions of *A* and *S* within preferentially interacting groups, we do not test temporal variation in their ratio that might arise due to immigration of hosts to the population, for example. Likewise, temporal variations in interaction preferences should consider non-deterministic patterns and, to better capture multi-strain coexistence, this model could be extended to multiple strains and multiple host groups. Further extensions of the model could also take into account other types of host-pathogen interactions. For example, different strains of a viral population can infect the same host, resulting in competition or facilitation [74, 75]. Development of acquired immunity should also be considered, especially vulnerability of re-infection [76, 77]. Additionally, as multi-strain coexistence is promoted by multiple mechanisms [1, 2, 66–68, 73], it is important to consider the interplay of these mechanisms when making evolutionary predictions.

Supporting information

S1 Text. Supporting information. This supporting information file contains the ten supplementary figures cited in the paper. (PDF)

Acknowledgments

This research was done using resources provided by the Open Science Grid, which is supported by the National Science Foundation award 1148698, and the U.S. Department of Energy's Office of Science.

Author Contributions

Conceptualization: Oana Carja.

Formal analysis: Shuanger Li, Davorka Gulisija, Oana Carja.

Methodology: Shuanger Li, Davorka Gulisija, Oana Carja.

Supervision: Oana Carja.

Visualization: Oana Carja.

Writing – original draft: Oana Carja.

Writing – review & editing: Shuanger Li, Davorka Gulisija, Oana Carja.

References

1. Cobey S. Pathogen evolution and the immunological niche. *Annals of the New York Academy of Sciences*. 2014; 1320(1):1–15. <https://doi.org/10.1111/nyas.12493> PMID: 25040161

2. Lourenço J, Recker M. Natural, Persistent Oscillations in a Spatial Multi-Strain Disease System with Application to Dengue. *PLOS Computational Biology*. 2013; 9(10):e1003308. <https://doi.org/10.1371/journal.pcbi.1003308> PMID: 24204241
3. Chesson P. Multispecies competition in variable environments. *Theoretical Population Biology*. 1994; 45:227–276. <https://doi.org/10.1006/tpbi.1994.1013>
4. Chesson P. Diversity maintenance by integration of mechanisms over various scales. *Proc 8th International Coral Reef Symposium*. 1997; 1:405–410.
5. Chesson P. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*. 2000; 31(1):343–366. <https://doi.org/10.1146/annurev.ecolsys.31.1.343>
6. Shmida A, Ellner S. Coexistence of plant species with similar niches. *Vegetation*. 1984; 58(1):29–55. <https://doi.org/10.1007/BF00044894>
7. Svardal H, Rueffler C, Hermisson J. Comparing Environmental and Genetic Variance as Adaptive Response to Fluctuating Selection. *Evolution*. 2011; 65(9):2493–2513. <https://doi.org/10.1111/j.1558-5646.2011.01318.x> PMID: 21884052
8. Svardal H, Rueffler C. A general condition for adaptive genetic polymorphism in temporally and spatially heterogeneous environments. *Theoretical Population Biology*. 2015; 99:7–94. <https://doi.org/10.1016/j.tpb.2014.11.002> PMID: 25446960
9. Gulisija D, Kim Y. Emergence of long-term balanced polymorphism under cyclic selection of spatially variable magnitude. *Evolution*. 2015; 69(4):979–992. <https://doi.org/10.1111/evo.12630> PMID: 25707330
10. Gulisija D, Kim Y, Plotkin JB. Phenotypic Plasticity Promotes Balanced Polymorphism in Periodic Environments by a Genomic Storage Effect. *Genetics*. 2016; 202(4):1437–1448. <https://doi.org/10.1534/genetics.115.185702> PMID: 26857626
11. Ellner S, Hairston NG. Role of Overlapping Generations in Maintaining Genetic Variation in a Fluctuating Environment. *The American Naturalist*. 1994; 143(3):403–417. <https://doi.org/10.1086/285610>
12. Ellner S, Sasaki A. Patterns of Genetic Polymorphism Maintained by Fluctuating Selection with Overlapping Generations. *Theoretical Population Biology*. 1996; 50(1):31–65. <https://doi.org/10.1006/tpbi.1996.0022> PMID: 8776837
13. Turelli M, Schemske DW, Bierzychudek P. Stable two-allele polymorphisms maintained by fluctuating fitnesses and seed banks: protecting the blues in *Linanthus parryae*. *Evolution*. 2001; 55(7):1283–1298. <https://doi.org/10.1111/j.0014-3820.2001.tb00651.x> PMID: 11525453
14. Schelling TC. *Micromotives and Macrobehavior*. Norton; 1978.
15. Centola D. An experimental study of homophily in the adoption of health behavior. *Science*. 2011; 334(6060):1269–1272. <https://doi.org/10.1126/science.1207055> PMID: 22144624
16. Flatt JD, Agimi Y, Albert SM. Homophily and health behavior in social networks of older adults. *Family and Community Health*. 2012; 35(4):312–321. <https://doi.org/10.1097/FCH.0b013e3182666650> PMID: 22929377
17. Sampson RJ, Monreoff JD, Gannon-Rowley T. Assessing “Neighborhood Effects”: Social Processes and New Directions in Research. *Annual Review of Sociology*. 2002; 28(1):443–478. <https://doi.org/10.1146/annurev.soc.28.1.443>
18. Ioannides YM, Loury LD. Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature*. 2004; 42(4):105–1093. <https://doi.org/10.1257/0022051043004595>
19. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*. 2001; 27(1):415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
20. Kossinets G. Origins of Homophily in an Evolving Social Network. *American Journal of Sociology*. 2009; 115(2):459–466. <https://doi.org/10.1086/599247>
21. Creanza N, Feldman MW. Complexity in models of cultural niche construction with selection and homophily. *Proceedings of the National Academy of Sciences*. 2014; 111(Supplement 3):10830–10837. <https://doi.org/10.1073/pnas.1400824111> PMID: 25024189
22. Janulis P, Phillips G, Birkett M, Mustanski B. Sexual networks of racially diverse young MSM differ in racial homophily but not concurrency. *Journal of acquired immune deficiency syndromes*. 1999; 77(5):459–466. <https://doi.org/10.1097/QAI.0000000000001620>
23. Zhou X, Coiera E, Tsafnat G, Arachi D, Ong MS, Dunn AG. Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter. *Medinfo*. 2015; p. 761–765. PMID: 26262154
24. Kadelka C, McCombs A. Effect of homophily and correlation of beliefs on COVID-19 and general infectious disease outbreaks. *PloS one*. 2021; 16(12):e0260973. <https://doi.org/10.1371/journal.pone.0260973> PMID: 34855929

25. Hiraoka T, Rizi AK, Kivelä M, Saramäki J. Herd immunity and epidemic size in networks with vaccination homophily. *Physical Review E*. 2022; 105(5):L052301. <https://doi.org/10.1103/PhysRevE.105.L052301> PMID: 35706197
26. Burgio G, Steinegger B, Arenas A. Homophily impacts the success of vaccine roll-outs. *Communications Physics*. 2022; 5(1):70. <https://doi.org/10.1038/s42005-022-00849-8>
27. Kadelka C, Islam MR, McCombs A, Alston J, Morton N. Ethnic homophily affects vaccine prioritization strategies. *Journal of Theoretical Biology*. 2022; 555:111295. <https://doi.org/10.1016/j.jtbi.2022.111295> PMID: 36208667
28. Leschak CJ, Eisenberger NI. Two distinct immune pathways linking social relationships with health: inflammatory and antiviral processes. *Psychosomatic Medicine*. 2019; 81(8):711. <https://doi.org/10.1097/PSY.0000000000000685> PMID: 31600173
29. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. *Cell Research*. 2020; 30(6):492–506. <https://doi.org/10.1038/s41422-020-0332-7> PMID: 32433595
30. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. 2020; 370(6515). <https://doi.org/10.1126/science.abd4570>
31. Bastard P, Rosen LB, Zhang Q, Michailidis E, Hoffmann HH, Zhang Y, et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science*. 2020; 370(6515). <https://doi.org/10.1126/science.abd4585> PMID: 32972996
32. Holmes L, Enwere M, Williams J, Ogundele B, Chavan P, Piccoli T, et al. Black-White Risk Differentials in COVID-19 (SARS-COV2) Transmission, Mortality and Case Fatality in the United States: Translational Epidemiologic Perspective and Challenges. *International Journal of Environmental Research and Public Health*. 2020; 17(12):4322. <https://doi.org/10.3390/ijerph17124322> PMID: 32560363
33. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*. 2020; 323(12):1239. <https://doi.org/10.1001/jama.2020.2648> PMID: 32091533
34. Fuzzele S, Sahay B, Yusufu I, Lee TJ, Sharma A, Kolhe R, et al. COVID-19 Virulence in Aged Patients Might Be Impacted by the Host Cellular MicroRNAs Abundance/Profile. *Aging and Disease*. 2020; 11(3):509–522. <https://doi.org/10.14336/AD.2020.0428>
35. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of COVID-19 disease. *medRxiv*. 2020;.
36. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA*. 2020; <https://doi.org/10.1001/jama.2020.4683> PMID: 32203977
37. Lau MSY, Grenfell B, Nelson K, Lopman B. Characterizing super-spreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA. *PNAS*. 2020; 36(117):22430–22435. <https://doi.org/10.1073/pnas.2011802117>
38. Davies NG, Klepac P, Liu Y, Prem K, Jit M. Age-dependent effects in the transmission and control of COVID-19 epidemics. *medRxiv*. 2020; PMID: 32546824
39. Mueller NT, Bakacs E, Combellick J, Grigoryan Z, Dominguez-Bello MG. The infant microbiome development: mom matters. *Trends in Molecular Medicine*. 2015; 21(2):109–117. <https://doi.org/10.1016/j.molmed.2014.12.002> PMID: 25578246
40. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, et al. Cohabiting family members share microbiota with one another and with their dogs. *Elife*. 2013; 2:e00458. <https://doi.org/10.7554/eLife.00458> PMID: 23599893
41. Brito IL, Gurry T, Zhao S, Huang K, Young SK, Shea TP, et al. Transmission of human-associated microbiota along family and social networks. *Nature Microbiology*. 2019; 4(6):964–971. <https://doi.org/10.1038/s41564-019-0409-6> PMID: 30911128
42. Muthukrishna M, Morgan TJ, Henrich J. The when and who of social learning and conformist transmission. *Evol Hum Behav*. 2016; p. 10–20. <https://doi.org/10.1016/j.evolhumbehav.2015.05.004>
43. Denton KK, Ram Y, Liberman U, Feldman MW. Cultural evolution of conformity and anticonformity. *Proceedings of the National Academy of Sciences*. 2020; 117(24):13603–13614. <https://doi.org/10.1073/pnas.2004102117> PMID: 32461360
44. Fleming-Davies AE, Williams PD, Dhondt AA, Dobson AP, Hochachka WM, Leon AE, et al. Incomplete host immunity favors the evolution of virulence in an emergent pathogen. *Science*. 2018; 359(6379):1030–1033. <https://doi.org/10.1126/science.aao2140> PMID: 29496878
45. McKay B, Ebell M, Dale AP, Shen Y, Handel A. Virulence-mediated infectiousness and activity trade-offs and their impact on transmission potential of influenza patients. *PNAS*. 2020; 287(1927):2020049. <https://doi.org/10.1098/rspb.2020.0496> PMID: 32396798

46. Anderson RM, May RM. Coevolution of hosts and parasites. *Parasitology*. 1982; 84:411–426. <https://doi.org/10.1017/S0031182000055360> PMID: 6755367
47. Jensen KH, Little T, Skorpung A, Ebert D. Empirical Support for Optimal Virulence in a Castrating Parasite. *PLOS Biology*. 2006; 4(7):e197. <https://doi.org/10.1371/journal.pbio.0040197> PMID: 16719563
48. Messenger SL, Molineux IJ, J BJ. Virulence evolution in a virus obeys a trade-off. *Proceedings of the Royal Society of London Series B: Biological Sciences*. 1999; 266(1417):397–404. <https://doi.org/10.1098/rspb.1999.0651>
49. Mackinnon MJ, Gandon S, Read A. Virulence evolution in response to vaccination: The case of malaria. *Vaccine*. 2008; 26:C42–C52. <https://doi.org/10.1016/j.vaccine.2008.04.012>
50. Fraser C, Hollingsworth D, Chapman R, de Wolf F, Hanage WP. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *PNAS*. 2007; 104(44):17441–17446. <https://doi.org/10.1073/pnas.0708559104> PMID: 17954909
51. Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*. 2020; 368(6498):1481–1486. <https://doi.org/10.1126/science.abb8001> PMID: 32350060
52. Singh R, Adhikari R. Age-structured impact of social distancing on the COVID-19 epidemic in India. *arXiv200312005*. 2020;.
53. Kadelka C. Projecting social contact matrices to populations stratified by binary attributes with known homophily. *Mathematical Biosciences and Engineering*. 2022; 20(1):3282–3300. <https://doi.org/10.3934/mbe.2023154> PMID: 36899581
54. Kimura M. The Number of Heterozygous Nucleotide Sites Maintained in a Finite Population Due to Steady Flux of Mutations. *Genetics*. 1969; 61(4):893–903. <https://doi.org/10.1093/genetics/61.4.893> PMID: 5364968
55. Lofgren E, Fefferman NH, Naumov YN, Gorski J, Naumova EN. Influenza Seasonality: Underlying Causes and Modeling Theories. *Journal of Virology*. 2007; 81(11):5429–5436. <https://doi.org/10.1128/JVI.01680-06> PMID: 17182688
56. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Trends in Ecology & Evolution*. 1976; 22:403–434.
57. Carja O, Plotkin JB. The evolutionary advantage of heritable phenotypic heterogeneity. *Scientific Reports*. 2017; 7(1):5090–5100. <https://doi.org/10.1038/s41598-017-05214-2> PMID: 28698577
58. Carja O, Plotkin JB. Evolutionary Rescue Through Partly Heritable Phenotypic Variability. *Genetics*. 2019; 211(3):977–988. <https://doi.org/10.1534/genetics.118.301758> PMID: 30696715
59. Haldane JBS. A Mathematical Theory of Natural and Artificial Selection, Part V: Selection and Mutation. *Mathematical Proceedings of the Cambridge Philosophical Society*. 1927; 23(7):838–844. <https://doi.org/10.1017/S0305004100015644>
60. Lande R, Shannon S. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*. 1996; 50(1):434–437. <https://doi.org/10.2307/2410812> PMID: 28568879
61. Barrett RDH, Schluter D. Adaptation from standing genetic variation. *Trends in Ecology & Evolution*. 2008; 23(1):38–44. <https://doi.org/10.1016/j.tree.2007.09.008> PMID: 18006185
62. Pennings PS. Standing Genetic Variation and the Evolution of Drug Resistance in HIV. *PLOS Computational Biology*. 2012; 8(6):e1002527. <https://doi.org/10.1371/journal.pcbi.1002527> PMID: 22685388
63. Furuse Y, Oshitani H. Mechanisms of replacement of circulating viruses by seasonal and pandemic influenza A viruses. *International Journal of Infectious Diseases*. 2016; 51:6–14. <https://doi.org/10.1016/j.ijid.2016.08.012> PMID: 27569827
64. Yen HL, Webster RG. *Pandemic Influenza as a Current Threat*. Springer; 2009.
65. Carrat F, Flahault A. Influenza vaccine: The challenge of antigenic drift. *Vaccine*. 2007; 25(39):6852–6862. <https://doi.org/10.1016/j.vaccine.2007.07.027> PMID: 17719149
66. Zinder D, Bedford T, Gupta S, Pascual M. The Roles of Competition and Mutation in Shaping Antigenic and Genetic Diversity in Influenza. *PLOS Pathogens*. 2017; 9(1):31003104.
67. Zinder D, Riolo MA, J WR, Pascual M. Role of competition in the strain structure of rotavirus under invasion and reassortment. *bioRxiv*. 2017;.
68. Ranjeva SL, Baskerville EB, Dukic V, Villa LL, Lazcano-Ponce E, Giuliano AR, et al. Recurring infection with ecologically distinct HPV types can explain high prevalence and diversity. *PNAS*. 2017; 114(51):13573–13578. <https://doi.org/10.1073/pnas.1714712114> PMID: 29208707
69. Christakis N, Fowler J. The collective dynamics of smoking in a large social network. *N Engl J Med*. 2008; 358(21):2249–58. <https://doi.org/10.1056/NEJMsa0706154> PMID: 18499567
70. Christakis N, Fowler J. Social contagion theory: examining dynamic social networks and human behavior. *Stat Med*. 2013; 32:556–577. <https://doi.org/10.1002/sim.5408> PMID: 22711416

71. Perc M, Gomez-Gardenes J, Szolnoki A, Floria L, Moreno Y. Evolutionary dynamics of group interactions on structured populations: a review. *J R Soc Interface*. 2013; 10: 20120997. <https://doi.org/10.1098/rsif.2012.0997> PMID: 23303223
72. Carja O, Creanza N. The evolutionary advantage of cultural memory on heterogeneous contact networks. *Theor Popul Biol*. 2019; 129:118–125. <https://doi.org/10.1016/j.tpb.2018.09.006> PMID: 30731105
73. Holmes EC. The phylogeography of human viruses. *Molecular Ecology*. 2004; 13(4):745–756. <https://doi.org/10.1046/j.1365-294X.2003.02051.x> PMID: 15012753
74. Rousseau MC, Pereira JS, Prado JC, Villa LL, Rohan TE, Franco EL Cervical Coinfection with Human Papillomavirus (HPV) Types as a Predictor of Acquisition and Persistence of HPV Infection. *The Journal of Infectious Diseases*. 2001; 184(12):1508–1517. <https://doi.org/10.1086/324579> PMID: 11740725
75. Mavilia MG, Wu GY. HBV-HCV Coinfection: Viral Interactions, Management, and Viral Reactivation. *Journal of Clinical and Translational Hepatology*. 2018; 6(3):296–305. <https://doi.org/10.14218/JCTH.2018.00016> PMID: 30271742
76. Gousseff M, Penot P, Gallay L, Batisse D, Benech N, Bouiller K, et al. Clinical recurrences of COVID-19 symptoms after recovery: Viral relapse, reinfection or inflammatory rebound? *Journal of Infection*. 2020; 81:816–846. <https://doi.org/10.1016/j.jinf.2020.06.073> PMID: 32619697
77. Parry J. Covid-19: Hong Kong scientists report first confirmed case of reinfection. *BMJ*. 2020; 370.