PLOS Computational Biology

Multi-strain disease dynamics on metapopulation networks

--Manuscript Draft--

Manuscript Number:	PCOMPBIOL-D-19-01548
Full Title:	Multi-strain disease dynamics on metapopulation networks
Short Title:	Multi-strain disease dynamics on metapopulation networks
Article Type:	Research Article
Keywords:	metapopulation; multi-strain; disease dynamics; movement; network structure; strain structure; simulation
Corresponding Author:	Matthew J. Michalska-Smith University of Chicago Chicago, IL UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Chicago
Corresponding Author's Secondary Institution:	
First Author:	Matthew J. Michalska-Smith
First Author Secondary Information:	
Order of Authors:	Matthew J. Michalska-Smith
	Kimberly L VanderWaal
	Montserrat Torremorell
	Cesar A Corzo
	Meggan E Craft
Order of Authors Secondary Information:	
Abstract:	Many pathogens have clusters of variation in their genotypes that we refer to as strain structure. Importantly, host immunity to one strain is often neither independent from nor equivalent to immunity to related strains. This partial cross-reactive immunity allows repeated infection with (different strains of) the same pathogen and affects disease dynamics across a population, and can influence the effectiveness of intervention strategies. We combine two frameworks well-studied in their own right: multi-strain disease dynamics and metapopulation network structure. We simulate the dynamics of a multi-strain disease on a network of populations connected by movement, and characterize the effects of parametrization and network structures on these dynamics, finding that the movement of (partially) immune individuals tends to have a larger impact than the movement of infectious individuals, dampening infection dynamics in populations further along a chain. Additionally, dynamics propagate from one population to another, even if parameters vary between populations. In addition to providing novel insights into the role of host movement on disease dynamics, this work provides a framework for future predictive modelling of multi-strain diseases across generalized population structures.
Suggested Reviewers:	Sunetra Gupta University of Oxford sunetra.gupta@zoo.ox.ac.uk Creator of multi-strain disease modelling framework used in the work
	Mario Recker University of Exeter mario.recker@zoo.ox.ac.uk expert in disease modelling, including multi-strain disease models

	Marcel Salathé École polytechnique fédérale de Lausanne marcel.salathe@epfl.ch Expert on network epidemiology
	Sweta Bansal University of Georgetown shweta.bansal@georgetown.edu expert on network epidemiology
	James Lloyd-Smith University of California Los Angeles jlloydsmith@ucla.edu
	Nicholas Gotelli University of Vermont Nicholas.Gotelli@uvm.edu
	Shai Pilosof Ben-Gurion University shainova@gmail.com
	Sonia Kéfi Université de Montpellier sonia.kefi@umontpellier.fr
Opposed Reviewers:	
Additional Information:	
Question	**Response**
Financial Disclosure Enter a financial disclosure statement that describes the sources of funding for the work included in this submission. Review the submission guidelines for detailed requirements. View published research articles from *PLOS Computational Biology* for specific examples. This statement is required for submission and **will appear in the published article** if the submission is accepted. Please make sure it is accurate.	MJM, KV, MT, CC, and MEC funded by the University of Minnesota (UMN) College of Veterinary Medicine (CVM) Research Office UMN Ag Experiment Station General Ag Research Funds # MINV-62-057. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Unfunded studies Enter: *The author(s) received no specific funding for this work.* **Funded studies** Enter a statement with the following details: • Initials of the authors who received each award • Grant numbers awarded to each author • The full name of each funder • URL of each funder website • Did the sponsors or funders play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript? • **NO** - Include this sentence at the end of your statement: *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.* • **YES** - Specify the role(s) played. * typeset	
Competing Interests Use the instructions below to enter a competing interest statement for this submission. On behalf of all authors, disclose any competing interests that could be perceived to bias this work—acknowledging all financial support and any other relevant financial or non-financial competing interests. This statement **will appear in the published article** if the submission is accepted. Please make sure it is accurate. View published research articles from *PLOS Computational Biology* for specific examples.	The authors have declared that no competing interests exist.

NO authors have competing interests Enter: *The authors have declared that no competing interests exist.* **Authors with competing interests** Enter competing interest details beginning with this statement: *I have read the journal's policy and the authors of this manuscript have the following competing interests: [insert competing interests here]* * typeset	
Data Availability Authors are required to make all data underlying the findings described fully available, without restriction, and from the time of publication. PLOS allows rare exceptions to address legal and ethical concerns. See the PLOS Data Policy and FAQ for detailed information. A Data Availability Statement describing where the data can be found is required at submission. Your answers to this question constitute the Data Availability Statement and **will be published in the article**, if accepted. **Important:** Stating 'data available on request from the author' is not sufficient. If your data are only available upon request, select 'No' for the first question and explain your exceptional situation in the text box. Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?	Yes - all data are fully available without restriction
Describe where the data may be found in full sentences. If you are copying our sample text, replace any instances of XXX with the appropriate details.	All data is simulated/no empirical data used in this submission. Code to generate is publicly available

• If the data are **held or will be held in a public repository**, include URLs, accession numbers or DOIs. If this information will only be available after acceptance, indicate this by ticking the box below. For example: *All XXX files are available from the XXX database (accession number(s) XXX, XXX.*). • If the data are all contained **within the manuscript and/or Supporting Information files**, enter the following: *All relevant data are within the manuscript and its Supporting Information files.* • If neither of these applies but you are able to provide **details of access elsewhere**, with or without limitations, please do so. For example: *Data cannot be shared publicly because of [XXX]. Data are available from the XXX Institutional Data Access / Ethics Committee (contact via XXX) for researchers who meet the criteria for access to confidential data.* *The data underlying the results presented in the study are available from (include the name of the third party and contact information or URL).* • This text is appropriate if the data are owned by a third party and authors do not have permission to share the data. * typeset	
Additional data availability information:	

Cover Letter

University of Minnesota

Matthew J. Michalska-Smith

Dept. of Veterinary Population Medicine
1988 Fitch Ave
Saint Paul, MN 55108
E-mail: michalsm@umn.edu

September 4, 2019

Dear Editor,

Please find enclosed the manuscript "Multi-strain disease dynamics on metapopulation networks" which we would like to have considered for publication in *PLOS Computational Biology* as a research article in response to the **PLOS Cross-Journal Call for Papers on the Mathematical Modelling of Infectious Disease Dynamics**. In particular, we believe this work directly addresses "The role of population movement on various scales in altering the spread of disease" as well as "Disease modelling and forecasting in both local and global contexts."

Strain structure is ubiquitous among the most widespread diseases in the world today, and has been the subject of extensive epidemiological research over the past decades. Likewise, there has been extensive research on networks of interconnected populations, termed metapopulations, especially within the ecological literature. Surprisingly, however, there have been few attempts to link these two areas of research, despite a need for better understanding and control of pathogen spread.

Here, we use a mathematical model combining multi-strain disease dynamics and meta-population network structure to show that there can be dramatic effects on the temporal dynamics of pathogen prevalence when migration is allowed between populations. We explore several simple network structures and evince some remarkably simple and consistent patterns relating the dynamics of origin and destination populations. We additionally perform simulations to note the effect on larger metapopulation networks. In summary, we find that, when populations are linked, the dynamics of destination populations are significantly influenced by the dynamics of origin populations, with migration having the potential to reduce disease prevalence or even change the dynamical regime in the desitnation population. Generalizing to larger networks, we see a significant effect of network structure on total disease prevalence as well, with more heterogeneity in degree (the number of other populations a given population is connected to) leading to increased prevalence and decreased immunity in the metapopulation as a whole.

This work provides a general framework linking two well-established areas of ecological theory: metapopulations and multi-strain disease dynamics. We believe the insights expounded on in our work will instigate further research into the structures of real-world metapopulations, and, in particular the role of network structures on endemic disease dynamics. This framework will prove particularly useful to researchers of systems with high migration rates, such as in the context of mobile species in fragmented landscapes or the movements of livestock between growing facilities, but we believe the generality of our theoretical approach will prove broadly interesting to the readership of *PLOS Computational Biology*.

Thank you for the attention dedicated to our work.

Sincerely,

Matthew J. Michalska-Smith, Kimberly VanderWaal,
Montserrat Torremorell, Cesar Corzo,
and Meggan E Craft

Multi-strain disease dynamics on metapopulation networks

Matthew Michalska-Smith[1,2], Kimberly VanderWaal[1], Montserrat Torremorell[1], Cesar Corzo[1], and Meggan E Craft[1]

[1]Department of Veterinary Population Medicine, University of Minnesota, St. Paul, MN USA
[2]Department of Plant Pathology, University of Minnesota, St. Paul MN, USA

September 10, 2019

Abstract

Many pathogens have clusters of variation in their genotypes that we refer to as strain structure. Importantly, host immunity to one strain is often neither independent from nor equivalent to immunity to related strains. This partial cross-reactive immunity allows repeated infection with (different strains of) the same pathogen and affects disease dynamics across a population, and can influence the effectiveness of intervention strategies. We combine two frameworks well-studied in their own right: multi-strain disease dynamics and metapopulation network structure. We simulate the dynamics of a multi-strain disease on a network of populations connected by movement, and characterize the effects of parametrization and network structures on these dynamics, finding that the movement of (partially) immune individuals tends to have a larger impact than the movement of infectious individuals, dampening infection dynamics in populations further along a chain. Additionally, dynamics propagate from one population to another, even if parameters vary between populations. In addition to providing novel insights into the role of host movement on disease dynamics, this work provides a framework for future predictive modelling of multi-strain diseases across generalized population structures.

1 Author Summary

Many pathogens have variants that are similar enough that our immune systems recognize them as hostile, but different enough that we are not fully immune to them. This leads to partial immunity against segments of the pathogen population that can have dramatic effects on epidemic size and duration. Also relevant to disease dynamics, many systems can be envisioned as networks of interconnected patches, *e.g.* cities in which there is some migration between cities, but most people remain within a single city. When these two frameworks (multi-strain disease dynamics and meta-population structure) are combined, some surprisingly simple patterns are evinced. First, the movement of immune individuals reduces pathogen prevalence in destination populations. In some cases, migration can even induce a change in the dynamics (*e.g.* steady prevalence through time *vs.* cycles of high and low prevalence), causing populations to look more similar, even when migration rates are low. Finally, while the structure of metapopulation networks can affect prevalence through time, the precise properties governing these effects are not yet known. This study creates a framework for better understanding the interaction between two important factors influencing disease dynamics: the presence of multiple strains and complex metapopulation structure.

2 Introduction

Many of the most impactful infectious diseases that affect humans (influenza, malaria, human papillomavirus, *etc.*), livestock (porcine reproductive and respiratory syndrome, foot-and-mouth disease, *etc.*), and wildlife (anthrax, plague, *etc.*) have clusters in their population-genetic variability that we classify as strains. This variation in pathogen genotype is often associated with differences in phenotype, for example directly affecting the efficacy of host immune defences. While the human immune system is usually capable of preventing re-infection—*i.e.* infection with something to which it has been previously exposed—sufficient, divergent evolution among pathogen strains can reduce the ability of the host to recognize, and thus mount an immunological response to, subsequent exposures. In some cases, this change is not sufficient to completely avoid recognition by the host's

immune system, yielding an immune response that is neither as strong as would be in the case of re-exposure to the same strain, nor as weak as in the case of exposure to a novel pathogen. This partial cross-reactive immunity can likewise lead to reduced transmissibility, affecting disease dynamics across the population.

While the study of multi-strain diseases goes back decades (1; 2), the resulting modelling framework has not yet been generalized to a collection of sub-populations connected through host movement, *i.e.* a metapopulation (but see (3)). Initially introduced through the concepts of island biogeography (4), the network approach of metapopulations can be applied to a variety of systems, including human movement between cities, livestock transport between farms, and populations living in fragmented natural habitats. In each case, there exist relatively high-density areas which are connected to one another through a network of individuals' movement. A metapopulation framework allows the application of network analyses to characterize patterns of connection within the larger system, and can provide unique insights across scales.

Historically, metapopulation studies have been been divided into two main camps: those that model within-population dynamics, and "cell occupancy" models in which only the presence or absence of a given species within a population is recorded (5), with the latter receiving much more theoretical attention. Importantly, this latter case rests on an assumption of temporal separation in which local dynamics occur on a timescale that can be treated as instantaneous relative to that of the between-population dynamics (6). When considering pathogens in systems with relatively high migration rates, however, this assumption rarely holds, and the presence-absence approach can significantly limit model accuracy (7; 8; 9).

The presence of metapopulation structure has been repeatedly associated with increased stability (10; 11; 12). This is due in part to the ability of migration between asynchronous populations to rescue temporarily low density populations from extinction (13). This is particularly relevant when populations are undergoing cyclical or chaotic dynamics, where repeated instances of low density are generally considered to be at greater risk of extinction than a population maintaining steady state dynamics (14; 15).

Here, we build on the strain theory of host-pathogen systems proposed by (16), considering a scenario where a collection of populations undergoing local dynamics are furthermore interconnected through the movement of individuals between populations. We simulate disease dynamics on this system, characterizing the effects of parametrization and network structure on these dynamics. This work is divided into three sections: first, we explore a case of interconnected populations which have been parametrized to display identical dynamics in the absence of host migration. Second, we consider cases where parameters differ between populations. Finally, we explore the role of network structure on disease dynamics in larger networks of connected populations.

3 Results

In the following sections, we provide figures to demonstrate the effect of metapopulation structure on disease dynamics. In these figures, we plot a time series for each of three subsets of the population: those currently infected with a particular strain of the pathogen, those having (complete) specific immunity against the focal strain, and those who have at least partial cross-reactive immunity to the focal strain, due to past exposure to a similar strain (see Methods). We only depict one representative strain in each plot for visual clarity and parametrize the model such that all strains are functionally equivalent (*i.e.* they all have the same transmission and recovery rates).

3.1 Cyclical dynamics are dampened along chains in the metapopulation network

We find that even when all populations share the same parametrizations and initial conditions, that populations further along network chains have reduced proportions of currently infectious individuals and dampened oscillatory dynamics compared to those they would exhibit in isolation (Fig 1). This is due to the movement of (partially) immune individuals between the populations, increasing the proportion of individuals with specific and cross-reactively immunity in populations further along the chain. While infectious individuals move at an equal rate, the proportion of the population that is currently infectious at any given time is much smaller than the proportion with

immunity.

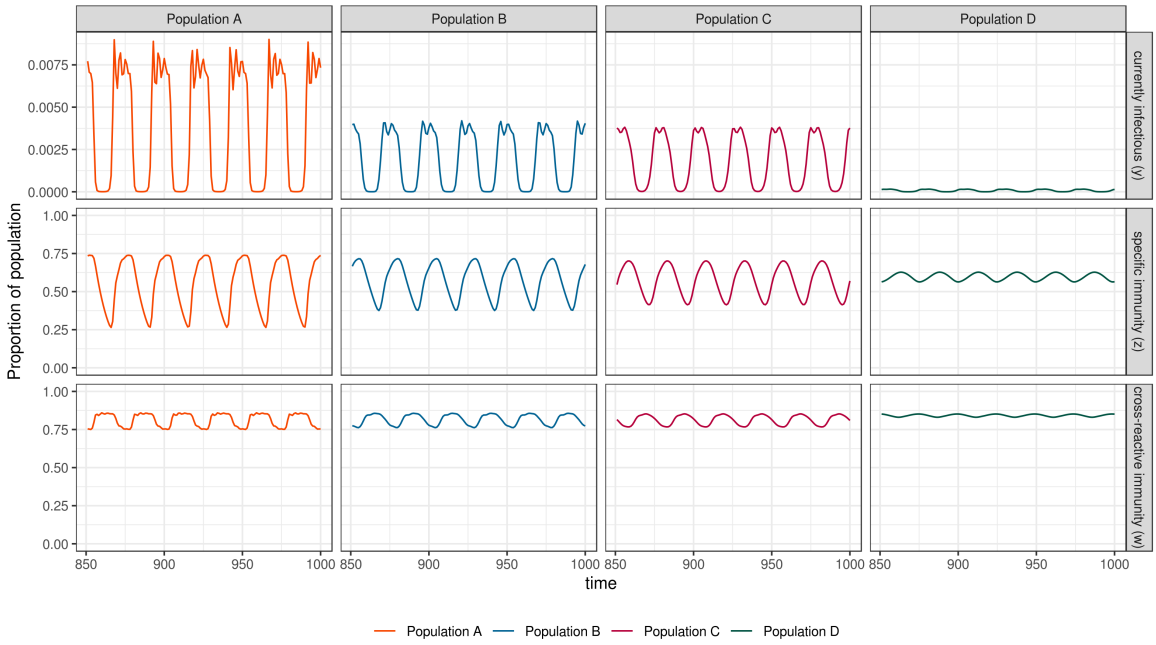

Figure 1: Connecting multiple populations with the same parameters results in reduced pathogen prevalence and dampened cycles in populations further down the chain. Here, populations are connected such that $A \to B \to C \to D$. Each column indicates a population, while each row is one of the three population classes laid out above and in the Methods. The mean level of immunity (both specific (middle row) and cross-reactive (bottom row)) increases in each sequential population, while the mean level of currently infectious individuals (top row) decreases. All populations have parameters $\beta = 40$, $\sigma = 10$, $\mu = 0.05$, $\delta = 0.05$, $\gamma = 0.75$. We use a two-loci, two-allele strain structure, but show only one strain for clarity (but see supporting information Fig S1).

3.2 Dynamics propagate through metapopulation networks

We find that in the case of a simple chain of populations, the dynamics of destination populations can be overridden by the dynamics of origin populations (Fig 2). Interestingly, this is true both of cyclical dynamics overruling stable dynamics and *vice versa*, though the required amount of migration differs according to the origin and destination dynamics (see supporting information Fig S2). This migration can also allow for strain coexistence even in populations where the local parameters would suggest extinction of one or more strains.

The issue of dynamics propagation gets more complicated when there are multiple, varying origin

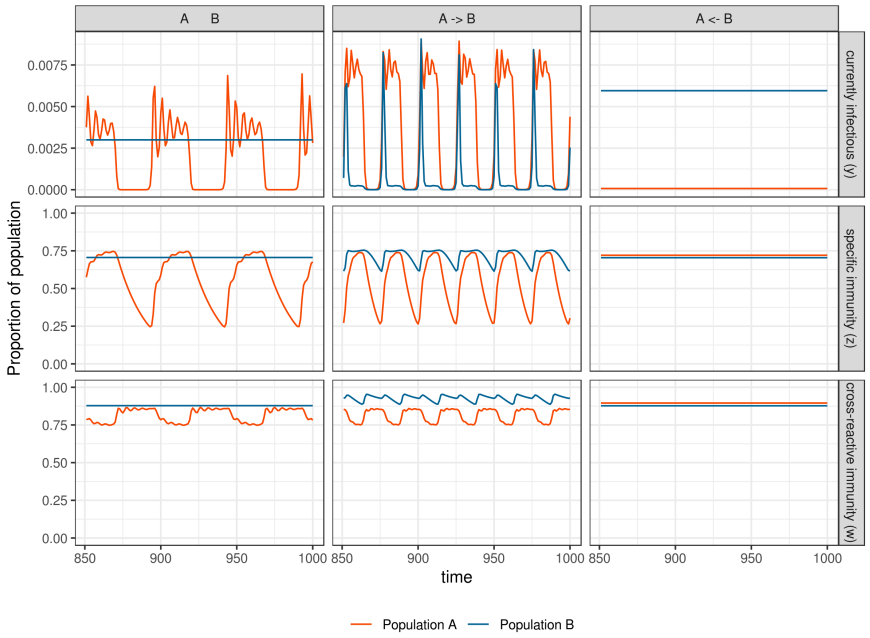

Figure 2: Destination populations tend to inherit origin population dynamics when linking populations with different model parametrizations. As in Fig 1, rows correspond to population classes, but here, columns indicate network structure. While in isolation (left column), population A has cyclical dynamics and population B has steady-state dynamics, when the two populations are linked by migration, the destination population inherits the dynamics of the origin population (center and right columns). This is true regardless of the direction of the movement (depending on the level of migration; see supporting information Fig S2). Populations A and B have parameters $\beta = 40$, $\sigma = 10$, $\mu = 0.05$, $\delta = 0.05$ in common and $\gamma = 0.75, 0.25$, respectively. We use a two-loci, two-allele strain structure, but show only one strain for clarity.

populations for a given destination population. We find that there is a hierarchy of dynamics in their propagation through the network: when there are origin populations with both cyclical and steady state dynamics, the destination population inherits the cyclical dynamics, robustly to imbalance in the relative contributions of the origins. Put another way, if just one of many origin populations (or a small proportion of the total movement) has cyclical dynamics, the destination population will also have cyclical dynamics.

3.3 Degree distribution affects pathogen prevalence and immunity

These simple patterns in the effects of origin population dynamics on those in the destination population have clear implications when pieced together into larger network structures. For instance,

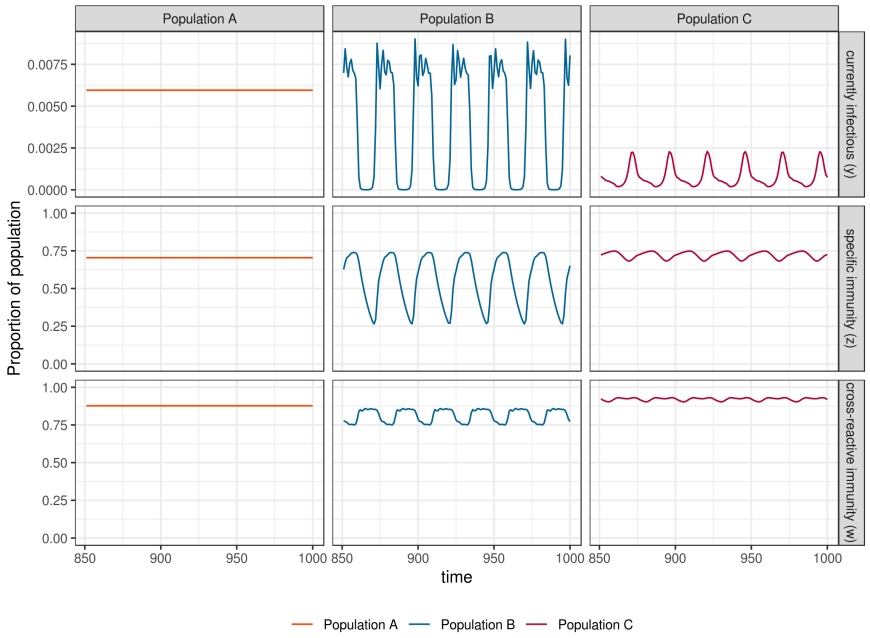

Figure 3: When multiple origin populations differ in their dynamics, the destination population inherits cycles over steady states. As in Fig 1, rows indicate population classes, and columns the component populations. Here, we have populations A and B feeding into population C at the same rate of $\delta = 0.05$. Populations A and C are parametrized to produce steady state dynamics in the absence of migration, with $\beta = 40$, $\sigma = 10$, $\mu = 0.05$, $\gamma = 0.25$. Population B shows cyclical dynamics with $\gamma = 0.75$ and all other parameters the same. Note that, even though the parameters of population C would lead to a steady state in the absence of migration, we see cyclical dynamics being inherited from population B. We use a two-loci, two-allele strain structure, but show only one strain for clarity.

the propagation of immune individuals through the metapopulation suggests that populations further "up the chain" will tend to have higher on-average disease burden and also greater variability. The inheritance of dynamical regimes combined with a hierarchy of dynamics in that inheritence suggests that chaos and cycles should be more common, especially in populations further "down the chain." That is, except in cases where the ultimate origin populations are all disposed toward steady states, in which case the stabilizing effect could overrule downstream local parametrizations, leading to an overall stable system.

In Fig 4, we report the effect of various network structures on three summary statistics of pathogen prevalence (and levels of immunity) using five common network ensembles. Depending on the system being explored, empirical network structures might have elements in common with one or

more of these ensembles, for instance, many social networks are considered to be "small-world" in structure like Barabasi-Albert random graphs, while ecological networks are often commented on for their formation of "modules" or clusters of more densely interacting species as in stochastic block random graphs. Networks were parametrized to have approximately equal connectance and size in order to reduce uninformative variation (see Section 5.3.3). This is because metapopulation size and connectance have known effects on pathogen persistence, independent of further network structure (17; 18; 19).

We find that the network configurations with higher variation in indegree (*i.e.* the number of other populations each population receives migration from) distributions (supporting information Fig S4), such as those found in the tree and Barabasi-Albert networks, tend to have higher levels of infection over time, despite similar levels of immunity as the other three network types. We also see few significant differences in mean time between epidemic peaks across network types. For some parametrizations, such as those in Figure 4, we see slightly lower values for the aforementioned networks with high indegree variance, but this is not consistent across parametrizations (supporting information Fig S5).

4 Discussion

Both metapopulation (20) and strain (16) structure have long been known to be important to disease dynamics and are increasingly being recognized as ubiquitous. Yet the combination of these two areas of theory has been underexplored. We show here that this lacuna can have real consequences for our understanding of disease dynamics in empirical systems.

In probing the relationship between origin and destination dynamics in simple metapopulations, we have demonstrated several patterns that expand our understanding of disease dynamics in these systems. By directly incorporating a movement network into our model framework, we have constructed a very general approach that lends itself to arbitrarily large and complex systems. This is noteworthy, as more and more natural systems are being thought of in terms of networks of interacting components (*e.g.* separate species in ecological communities (21) or host individuals

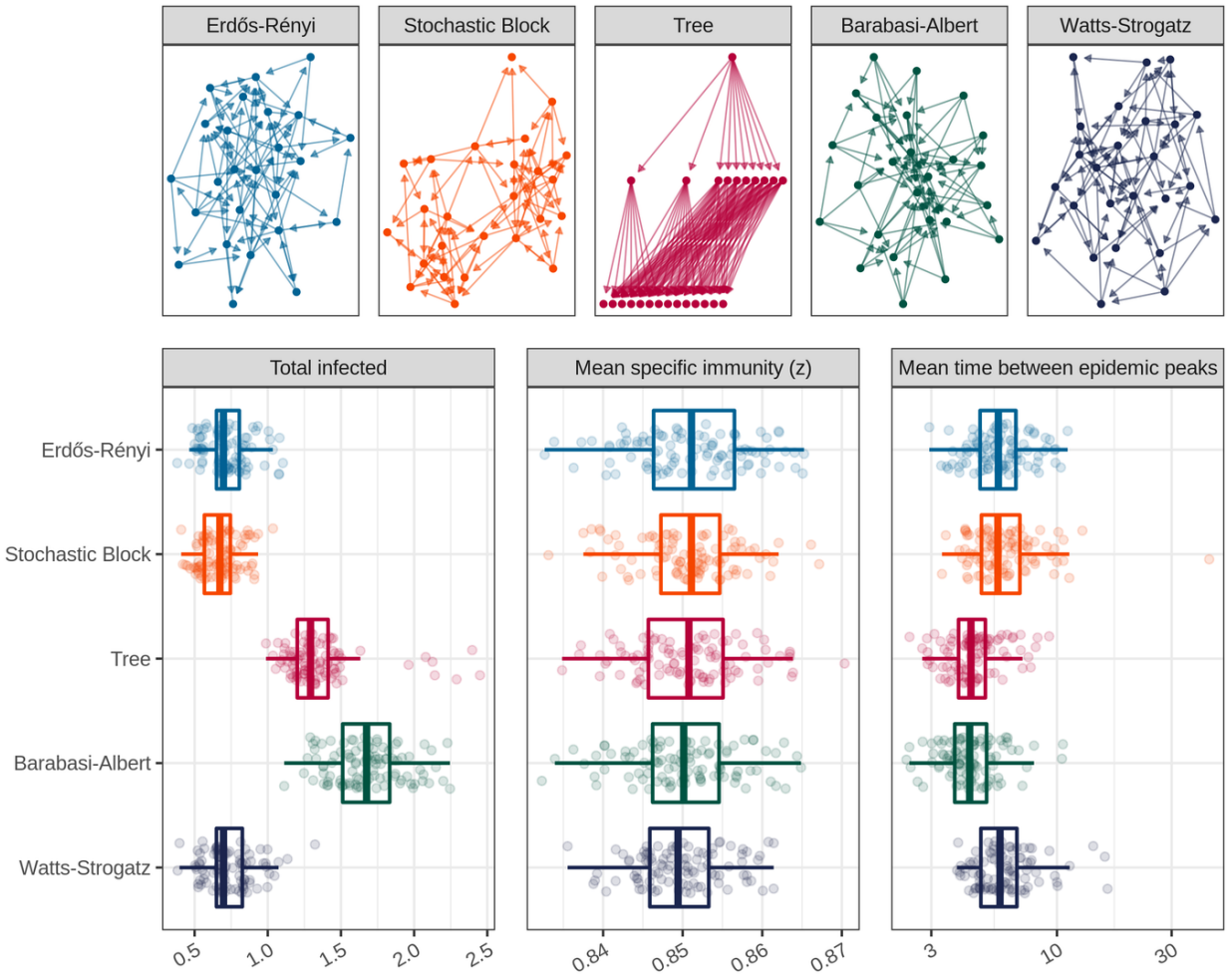

Figure 4: The effect of network structure on pathogen prevalence and levels of immunity through time. In the top row, we depict a representative network from each of the five ensembles. The second row shows the distributions of each of three response variables for prevalence of the pathogen and specific immunity over the course of the simulation. We depict one point for each randomized network structure and box-plots indicating the median and inter-quartile range of of each network-type's distribution. Network generating algorithms were tuned to produce networks of the same size and approximate connectance and model parameters were either the same for all simulations (β=160, σ=40, μ=0.05, and δ=0.05) or randomized between each simulation (initial densities of infectious and immune individuals [0, 1] and γ value [0.05, 0.95] in each population).

exchanging parasites (22)). By adjusting the scale of our metapopulation, we can ask and answer different questions about the forces influencing disease dynamics.

We found that the dynamics of prevalence and immunity among migrationally connected populations are not independent, and that even very small rates of population movement can have profound

effects on a population's disease dynamics: from reducing pathogen prevalence to changing the dynamical regime of destination populations entirely. Our findings regarding the reduction in cycle amplitude (Section 3.1) echo results in dispersal networks in ecology, where population dynamics were dampened following the introduction of migration (23).

Contrary to prior focus in the literature on the role of migrating infectious individuals (24; 25; 20), we found that the migration of immune individuals can be equally (or even more) important. This is noteworthy, as the few previous studies relating multi-strain diseases and metapopulation structure only allow pathogen transmission between populations, not the movement of individuals explicitly (3; 26)—an approach that is more mathematically tractable, but omits the potentially influential transmission of immune individuals.

Finally, we show that larger network structure also has a part to play in disease dynamics, resulting in significant differences in pathogen prevalence across network types. Our results are in agreement with previous results suggesting increased epidemic size in scale-free network structures (such as those found in Barabasi-Albert random graphs) due to the high-degree nodes serving as "super-spreaders" (27; 19). Along these lines, there has been some previous research indicating that node degree (the number of other populations a given population is connected to) is directly related to pathogen prevalence in that focal population ((28), but see (29)), however a complete investigation into network structure at the node-level is beyond the scope of this work. A comprehensive investigation of the role of more complex network structures in disease dynamics, however, remains a topic for further investigation.

In this work, we have utilized a relatively simple model for disease dynamics in an effort to maximize interpretability. Such simplification inevitably comes with a cost, and several of our assumptions can be critiqued as unrealistic. Perhaps foremost is the assumption of continuous movement. While continuous movement might be appropriate for very large populations with frequent, relatively small migrations between them, when any of these three components is not present, we would expect deviation from these predictions. Future work could explore the importance of discrete movement regimes on these patterns. Likewise, in this work we omit strain mutation and recombination (30)

(yet the latter is included in the original framework of (16)). The generation of novel strains is likely important to the global persistence of diseases in humans (31) and animals (32).

This work should not be seen as an attempt at comprehensive categorization of the role of meta-population structure on the dynamics of multi-strain diseases, but rather as an initial step in exploring the complex interplay between the population structure of hosts and strain structure of pathogens. Our results suggest there may be simple rules underlying this relationship, at least for a wide range of parameter values, but it remains to be seen where networks based on empirical data fall in these parameter regimes, as well as how such systems might deviate from theoretical expectations.

5 Methods

5.1 Model framework for one population

We work from a system of ordinary differential equations which delineate a population into classes based on current and past exposure to different strains of a pathogen. Pathogens with strain structure can differ in both the number of strains and the level of cross-reactive immunity afforded by past exposure to similar strains. To model the number of strains, we signify a strain $i = \{x_1, x_2, \ldots, x_n\}$ as a set of n loci, each of which can take on a finite number of alleles. For instance, a pathogen with two loci (a and b) and two alleles at each loci has a total of four potential strains: $\{a_1, b_1\}$, $\{a_1, b_2\}$, $\{a_1, b_1\}$, $\{a_2, b_2\}$. For cross-reactive immunity, we use a parameter γ which indicates the degree of reduced susceptibility a host has to strains that are similar to (*i.e.* strains that share at least one allele with one another) past exposures. Importantly, in this framework, the number of strains is fixed and finite. While strains may go extinct over time, there is no process for the generation of new strains or to re-introduce strains that had previously gone extinct (but see (16)).

The model consists of sets of three nested equations (one set for each strain i): y, z, and w. See (30) for a graphical representation of the model framework. y_i represents the proportion of the population currently infected with strain i (and thus capable of infecting others). Likewise,

z_i represents the proportion of the population that has been exposed to strain i (including those currently infected, *i.e.* y_i). These individuals harbor complete immunity to future infections with strain i. Finally, w_i represents the proportion of the population which has been exposed to any strain j which has at least one allele in common with strain i (including strain i itself), *i.e.* $j \cap i \neq \emptyset$. These individuals have at least partial immunity to strain i. *N.b.* these equations are nested such that any individual in y_i is also in z_i and any individual in z_i is also in w_i, and $y_i \leq z_i \leq w_i$. In traditional Susceptible-Infected, Susceptible-Infected-Recovered, *etc.* single-strain mathematical frameworks: the y class is analogous to I, while w and z are composed of combinations of I and R classes. The susceptible population is not modelled explicitly in this framework.

Explicitly, these three equations (for a given strain i) are:

$$\begin{aligned}
\frac{\mathrm{d}y_i}{\mathrm{d}t} &= \beta\left((1-w_i)+(1-\gamma)(w_i-z_i)\right)y_i - \sigma y_i - \mu y_i \\
\frac{\mathrm{d}z_i}{\mathrm{d}t} &= \beta(1-z_i)y_i - \mu z_i \\
\frac{\mathrm{d}w_i}{\mathrm{d}t} &= \beta(1-w_i)\sum_{j \ni j\cap i\neq\emptyset} y_j - \mu w_i
\end{aligned} \tag{1}$$

As above, we denote strains as subscripts and, in the equation for w_i, we sum over all strains j which share at least one allele with the focal strain i. β, σ, and μ are the infection, recovery, and death rates, respectively. γ (as mentioned above) is an indicator of the level of cross-reactive immunity gained by prior exposure to alleles in the target strain. Note that while we depict only one value per demographic parameter (*i.e.* all strains are functionally equivalent) for clarity of notation, these values could also be written to vary by strain (*i.e.* β_i).

Immunity in this framework is non-waning: exposure to a strain yields consistent protection from future infection over the lifespan of the individual. Moreover, this protection is trichotomous: an individual can either have no protection from a given strain (it has not seen any of the alleles before), complete protection (it has seen this exact combination of alleles before), or a set point in-between according to the parameter γ (it has seen at least one, but not all alleles before). Put another way, we do not distinguish between loci, assuming that sharing an allele at one locus is

functionally identical to sharing an allele at any other locus, or indeed all other loci except one.

5.2 Extensions to more than one population

Following (33), we model movement between populations using a dispersal matrix $\mathbf{\Delta} = \mathbf{A}$ - $\mathbf{E}$, where $\mathbf{A}$ is the weighted adjacency matrix indicating the proportion of individuals moving from from population n (row) to population m (column) and $\mathbf{E}$ is a diagonal matrix representing emigration, where each entry $\mathrm{E}_{kk} = \sum_{k=1}^{n} \mathrm{A}_{kl}$ where n is the number of populations. Thus, the whole system can be depicted by a set of three equations per strain i per population k:

$$\begin{aligned}
\frac{\mathrm{d}y_{i,k}}{\mathrm{d}t} &= \beta\left((1-w_{i,k}) + (1-\gamma)(w_{i,k}-z_{i,k})\right) y_{i,k} - \sigma y_{i,k} - \mu y_{i,k} + \sum_{l} \Delta_{kl} y_{j,l} \\
\frac{\mathrm{d}z_{i,k}}{\mathrm{d}t} &= \beta(1-z_{i,k}) y_{i,k} - \mu z_{i,k} + \sum_{l} \Delta_{kl} z_{j,l} \\
\frac{\mathrm{d}w_{i,k}}{\mathrm{d}t} &= \beta(1-w_{i,k}) \sum_{j \ni j \cap i \neq \emptyset} y_{j,k} - \mu w_{i,k} + \sum_{l} \Delta_{kl} w_{j,l}
\end{aligned} \tag{2}$$

Where each equation from Section 5.1 is now additionally indexed according to population and has an additional term to account for migration between populations. While in principle the elements of $\mathbf{\Delta}$ can take any value [0, 1], signifying a (continuous) movement of between 0 and 100% of individuals, for simplicity we use a constant value δ for the strength of each movement, *i.e.* for each non-zero off-diagonal element of of $\mathbf{\Delta}$. Sensitivity to this value is explored in the supporting information (Fig S2).

This framework can be applied to a metapopulation of arbitrary size and complexity, with the number of equations being linearly related to the number of populations. The dynamics of each population are governed by a set of three equations per pathogen strain, and these equations are interlinked within populations by partial, cross-reactive immunity, and between populations through a movement network. The total number of differential equations for any given system will be three times the number of strains multiplied by the number of populations in the metapopulation.

5.3 Simulation Procedure

All simulations were carried out in Julia (34), with graphics produced using the ggplot2 package (35) in R version 3.6.1 (36). For simplicity of presentation, we fix the values of all variables other than γ (the degree of cross-reactive immunity) and $\mathbf{\Delta}$ (the network of movement information) to be identical for all populations in the metapopulation. γ is varied to demonstrate the variety of dynamics obtainable in this modeling framework (as in (16)), while the $\mathbf{\Delta}$ varies according to the number and interconnectedness of the populations. For the figures of the main text, we utilize a strain structure of two loci, each with two alleles. Sensitivity to these parameter choices is explored in the supporting information (Fig S3).

5.3.1 Populations with identical parametrizations

To assess the effect of migration on population dynamics, we first consider the simplest case of a set of populations sharing the same parametrization. We use a movement network described by a chain of populations, *i.e.* A→B→C→D or

$$\mathbf{\Delta} = \begin{bmatrix} -\delta & \delta & 0 & 0 \\ 0 & -\delta & \delta & 0 \\ 0 & 0 & -\delta & \delta \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $\delta = 0.05$, and ask how the dynamics of populations further down the chain (*i.e.* B, C, D) differ from those of the origin population (*i.e.* A), recalling that, without migration, all populations would have identical dynamics.

5.3.2 Populations with varying parametrizations

We next consider the case where parameters differ between connected populations, we restrict our consideration to a system of two populations, identical in all respects other than the parameter γ, which is set to either induce a steady state of coexisting strains ($\gamma = 0.25$ in population A) or

cyclically coexisting strains ($\gamma = 0.75$ in population B). We then display three potential patterns of connection: no migration (left column), A→ B (middle column), and B→A (right column). Specifically, we set

$$\mathbf{\Delta} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{\Delta} = \begin{bmatrix} -\delta & \delta \\ 0 & 0 \end{bmatrix}, \text{and }, \mathbf{\Delta} = \begin{bmatrix} 0 & 0 \\ \delta & -\delta \end{bmatrix},$$

respectively, again with $\delta = 0.05$.

To address the case of multiple origin populations feeding into a single destination population, we consider a system of three populations: A→C←B, or

$$\mathbf{\Delta} = \begin{bmatrix} -\delta & 0 & \delta \\ 0 & -\delta & \delta \\ 0 & 0 & -\delta \end{bmatrix},$$

where populations A and C have $\gamma = 0.25$, but population B has $\gamma = 0.75$; $\delta = 0.05$ as above.

5.3.3 Larger network structure

Finally, we characterize the role of global network structure through considering the impact of degree distribution on a few summary statistics of overall disease burden: the total proportion infected (area under the currently infectious (*i.e.* y) curve), the mean level of strain-specific immunity (average z value), and the mean time between epidemic peaks (*i.e.* between local maxima in y) over the course of the final 75% of the simulation. We omit the initial period of the simulation to reduce the impact of transient dynamics.

We perform 100 simulations for each of five generic network ensembles each with 25 populations and a connectedness of approximately 0.15. Specifically, we examine Erdős-Rényi (links randomly assigned between populations), stochastic block (a metapopulation consisting of two groups of populations which have high migration within the group, but low migration to populations in the other group), tree-like (where there are many chains of populations and no potential for cycles), Barabasi-

Albert (a scale-free network in which there tends to be a few populations with very many links, and many populations with few links), and Watts-Strogatz (a small-world network structure which is produced by partially re-wiring a spatially connected grid of populations) network structures. To generate these networks, we utilize functions from the tidygraph R package (37), except in the case of the tree and Watts-Strogatz configuration for which we use custom algorithms. In all cases, each migration strength is set to a constant $\delta = 0.05$, only the pattern of connections varies. Each population is assigned a random γ value within [0.05, 0.95]. These results are qualitatively similar if instead every population is assigned the same value of γ.

All code is made available on GitHub: https://git.io/JeqMc.

6 Acknowledgements

We would like to thank José Lourenço for helpful discussions regarding the mathematical framework of this work.

S1 Supporting Information

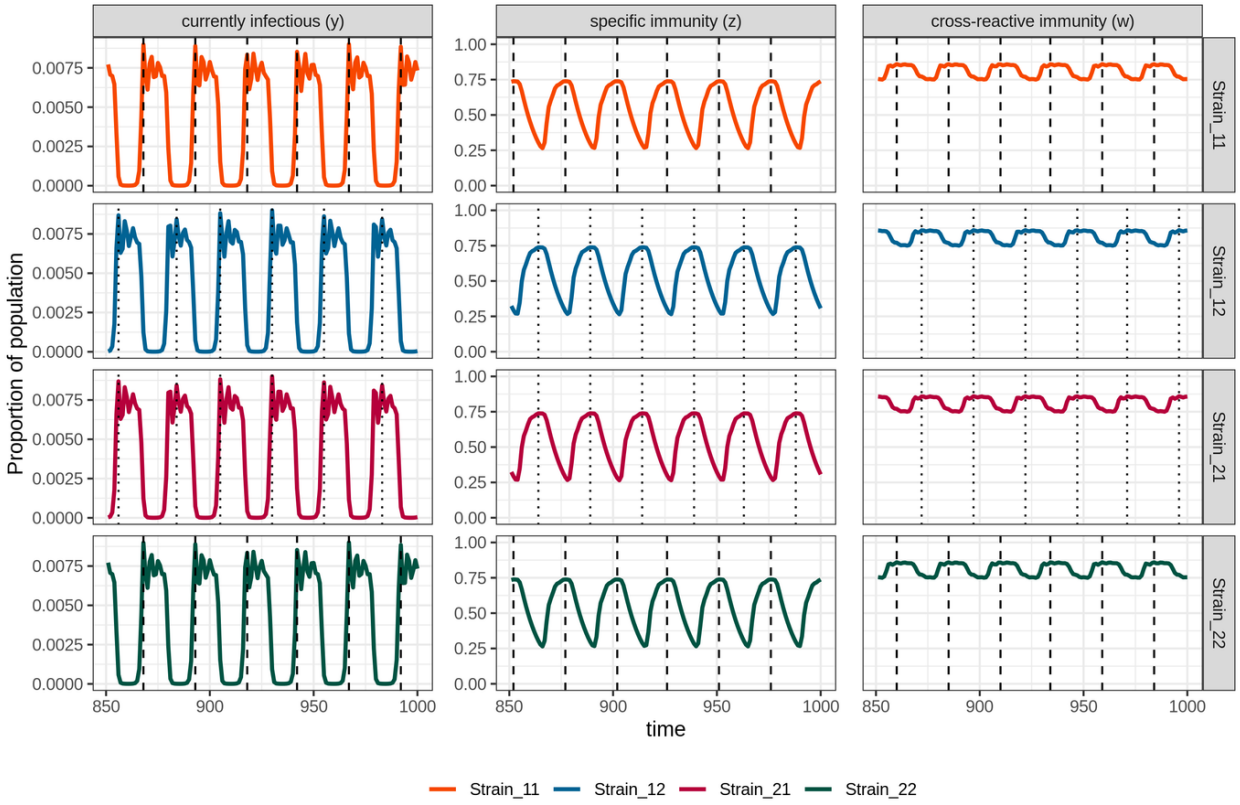

Figure S1: Considering all the dynamics of all four strains from population A in figure 1. Note that lines are colored according to strain rather than population. Strains can be divided into two discordant sets of non-overlapping alleles: {1, 1} and {2, 2}, and {1, 2} and {2, 1}. Each strain of a discordant set behaves identically due to identical parametrization and no interaction between strains that do not share at least one allele, but discordant sets interact with one another due to partial cross-reactive immunity. Thus, when one set is abundant, the other is rare and *vice versa*. We highlight the maximum value of each discordant set's cycle with a vertical in order to facilitate comparisons between strains and sets.

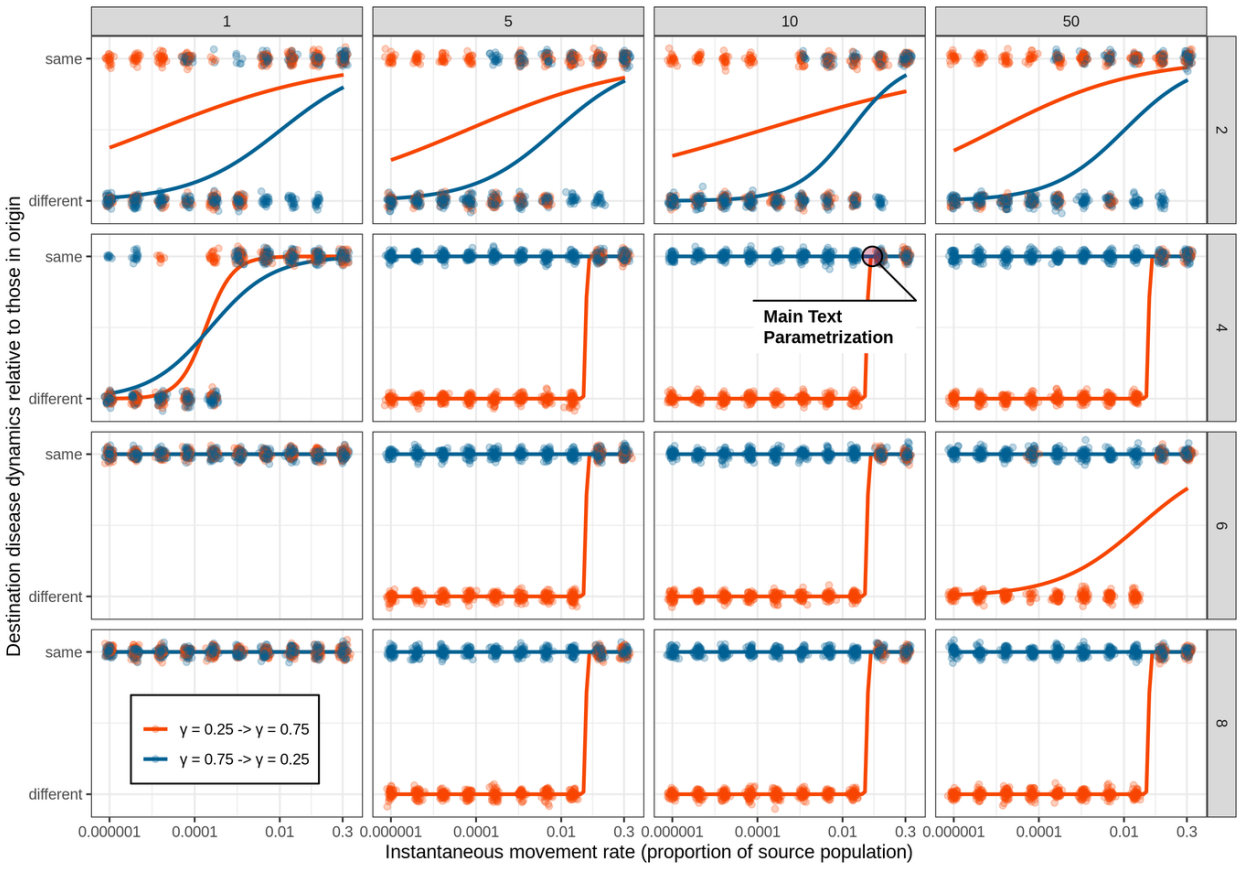

Figure S2: The effect of variable migration rate on transference of dynamical regime from origin to destination in a simple metapopulation of two populations linked by unidirectional movement. Here, each column indicates the value of σ and each row indicates the value of R_0 (these parameters are the same for both populations). We plot whether or not the destination dynamics are the "same" or "different" (vertical axis) for each movement rate ($\delta\cdot$ horizontal axis). We jitter the points slightly for increased visibility and fit a binomial spline to indicate the trend with increasing migration rate. In most cases, a γ value of 0.25 signifies a steady state (in the absence of migration) and a value of 0.75 signifies cyclical dynamics. We see that even with very small rates of migration, a stable population can be converted to a cyclical one (blue points overwhelmingly indicating same dynamics between origin and destination). Yet, it is more difficult to convert a cyclical population to one with steady state dynamics (orange points predominantly indicate a difference between origin and destination dynamics. Remarkably, the transition appears to be sharp in most cases: given a sufficient migration rate, destination dynamics will always be converted to match those of the origin. Note that the specific dynamics of the origin and destination depend upon the parametrization as well as movement rate, so in some cases, such as the panels in the lower left, the origin and destination would have the same dynamics even without movement. In other cases, such as the top row, there are many instances of strain extinction, which is rarely transmitted from origin to destination, though even if extinction dynamics are removed, these parametrizations lead to fewer cases of "successful" transmission of dynamics from origin to destination. We highlight the main-text parametrization of σ=10, R_0=4, δ=0.05.

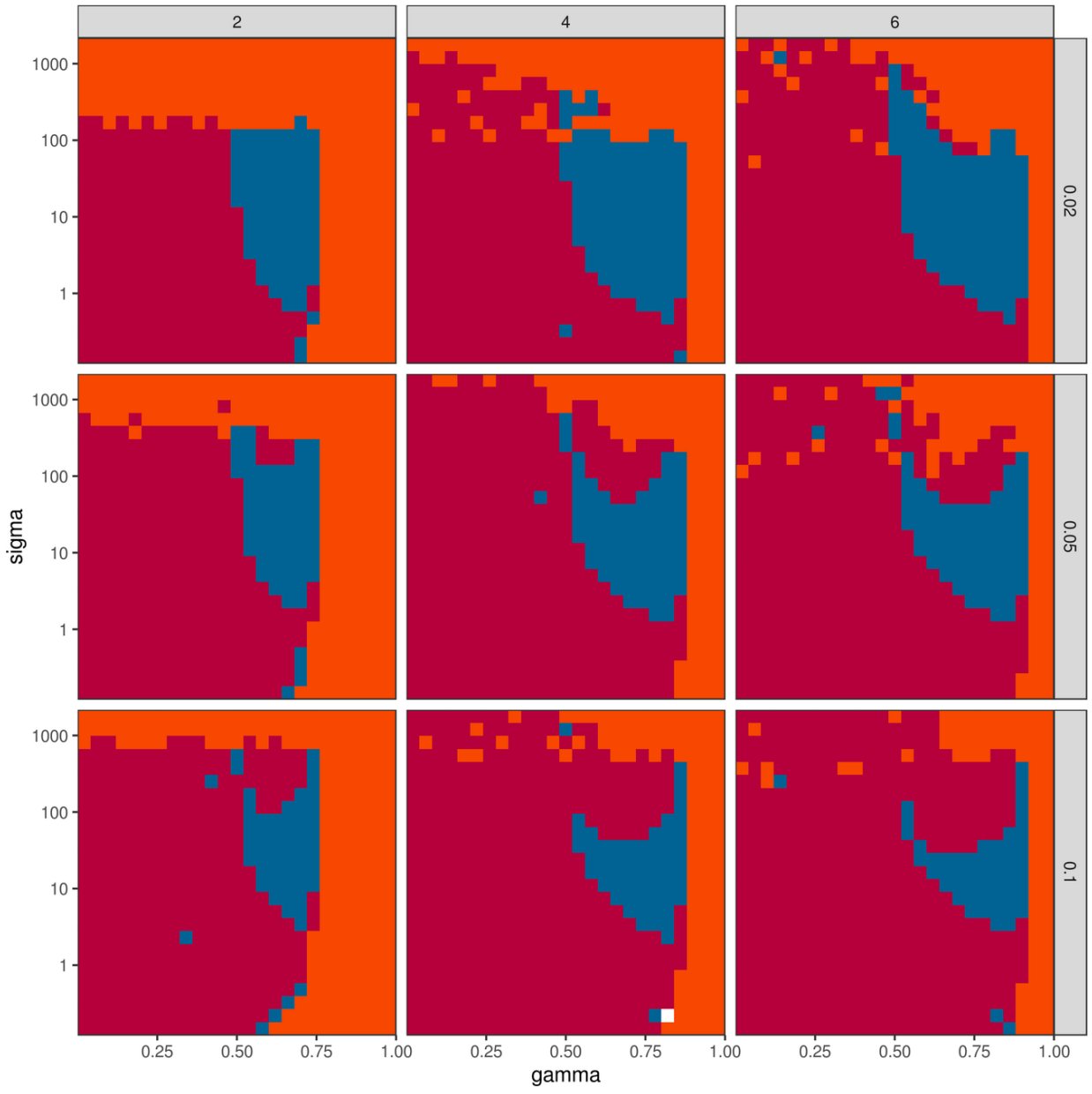

Figure S3: Effect of parametrization on dynamic regime for a population in isolation. Here, columns indicate values of R_0 ($\beta/\sigma\cdot$ main text utilizes a value of 4) and rows indicate values of μ (main text uses a value of 0.05). Depending on the combination of β, σ, and γ, a population can exhibit a range of dynamics including steady states for all strains (red), cyclical or chaotic dynamics for all strains (blue), or partial extinction of some strains (orange). Missing values are due to numerical failure in integration. All simulations here utilize a two-loci, two-allele strain structure. See (16) for a similar figure for alternative strain structures.

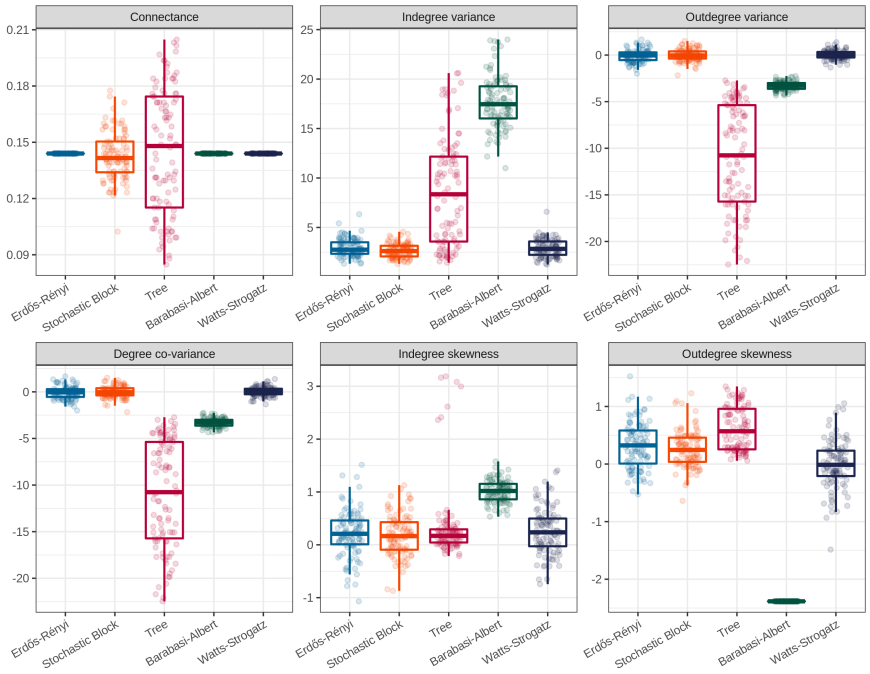

Figure S4: Summary statistics for the degree distributions of each randomized network used for figure 4 in the main text. Networks were constructed to have the same size and approximate connectance, but with the network structure (which populations are connected to which other popuations) otherwise generated according to one of five algorithms: Erdős-Rényi, Barabasi-Albert, and Watts-Strogatz, stochastic block, and tree (see Section 5.3.3 of the main text). Some algorithms allowed perfect matching of connectance (Erdős-Rényi, Barabasi-Albert, and Watts-Strogatz), while others necessitated some minor variation (stochastic block and tree).

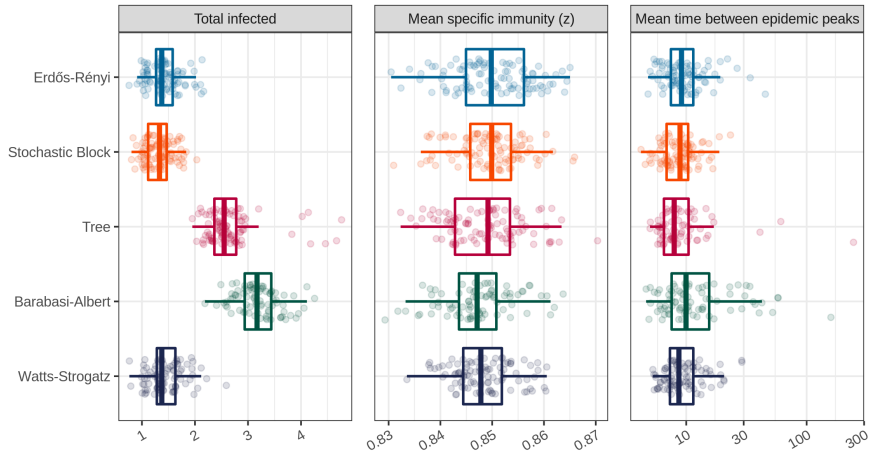

Figure S5: Similar to the lower row of figure 4, but with β and σ equal to 80 and 20, respectively. All other parameters are equal to or set randomly as in figure 4. While the observed differences in total infected are robust, note that here the mean time between epidemic peaks is approximately equal across randomizations.

References

1. Couch RB, Kasel JA. Immunity to influenza in man. Annual review of microbiology. 1983;37(1):529–49.

2. Castillo-Chavez C, Hethcote HW, Andreasen V, Levin SA, Liu WM. Epidemiological models with age structure, proportionate mixing, and cross-immunity.. J Math Biol. 1989;27:233–58.

3. Lourenço J, Recker M. Natural, persistent oscillations in a spatial multi-strain disease system with application to dengue.. PLoS Comput Biol. 2013;9:e1003308.

4. Wilson EO, MacArthur RH. The theory of island biogeography. Princeton University Press; 1967.

5. Taylor AD. Large-scale spatial structure and population dynamics in arthropod predator-prey systems. Annales Zoologici Fennici. 1988;:63–74.

6. Hanski I. A practical model of metapopulation dynamics. The Journal of Animal Ecology. January 1994;63(1):151.

7. Hess G. Disease in metapopulation models: implications for conservation. Ecology. July 1996;77(5):1617–32.

8. Cross PC, Lloyd-Smith JO, Johnson PLF, Getz WM. Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. Ecology Letters. June 2005;8(6):587–95.

9. Lloyd AL, Jansen VAA. Spatiotemporal dynamics of epidemics: synchrony in metapopulation models. Mathematical Biosciences. March 2004;188(1-2):1–16.

10. Ruxton GD. Low levels of immigration between chaotic populations can reduce system extinctions by inducing asynchronous regular cycles. Proceedings of the Royal Society of London Series B: Biological Sciences. May 1994;256(1346):189–93.

11. Earn DJD, Rohani P, Grenfell BT. Persistence chaos and synchrony in ecology and epi-

demiology. Proceedings of the Royal Society of London Series B: Biological Sciences. January 1998;265(1390):7–10.

12. Hanski I. Metapopulation dynamics. In: Metapopulation Biology. Elsevier; 1997. p. 69–91.

13. Brown JH, Kodric-Brown A. Turnover Rates in Insular Biogeography: Effect of Immigration on Extinction. Ecology. March 1977;58(2):445–9.

14. Rosenzweig ML. Paradox of Enrichment: Destabilization of Exploitation Ecosystems in Ecological Time. Science. January 1971;171(3969):385–7.

15. Hilker FM, Schmitz K. Disease-induced stabilization of predator–prey oscillations. Journal of Theoretical Biology. December 2008;255(3):299–306.

16. Gupta S, Ferguson N, Anderson R. Chaos persistence, and evolution of strain structure in antigenically diverse infectious agents. Science. May 1998;280(5365):912–5.

17. Ganesh A, Massoulié L, Towsley D. The effect of network topology on the spread of epidemics. In: Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE; 2005. p. 1455–66.

18. Salathé M, Jones JH. Dynamics and control of diseases in networks with community structure.. PLoS Comput Biol. 2010;6:e1000736.

19. Keeling MJ, Eames KTD. Networks and epidemic models. Journal of The Royal Society Interface. June 2005;2(4):295–307.

20. Grenfell B, Harwood J. (Meta)population dynamics of infectious diseases. Trends in Ecology & Evolution. October 1997;12(10):395–9.

21. Proulx SR, Promislow DE, Phillips PC. Network thinking in ecology and evolution.. Trends Ecol Evol. 2005;20:345–53.

22. Craft ME, Caillaud D. Network models: an underutilized tool in wildlife epidemiology?. Interdiscip Perspect Infect Dis. 2011;2011:676949.

342 23. Holland MD, Hastings A. Strong effect of dispersal network structure on ecological dynamics..
343 Nature. 2008;456:792–4.

344 24. Jesse M, Ezanno P, Davis S, Heesterbeek JAP. A fully coupled mechanistic model for infectious
345 disease dynamics in a metapopulation: Movement and epidemic duration. Journal of Theoretical
346 Biology. September 2008;254(2):331–8.

347 25. North AR, Godfray HCJ. The dynamics of disease in a metapopulation: The role of dispersal
348 range. Journal of Theoretical Biology. April 2017;418:57–65.

349 26. Wikramaratna PS, Pybus OG, Gupta S. Contact between bird species of different lifespans can
350 promote the emergence of highly pathogenic avian influenza strains.. Proc Natl Acad Sci U S A.
351 2014;111:10767–72.

352 27. Shirley MDF, Rushton SP. The impacts of network topology on disease spread. Ecological
353 Complexity. September 2005;2(3):287–99.

354 28. Godfrey SS, Bull CM, James R, Murray K. Network structure and parasite transmission in a
355 group living lizard the gidgee skink, Egernia stokesii. Behavioral Ecology and Sociobiology. April
356 2009;63(7):1045–56.

357 29. VanderWaal KL, Atwill ER, Hooper S, Buckle K, McCowan B. Network structure and preva-
358 lence of Cryptosporidium in Belding's ground squirrels. Behavioral Ecology and Sociobiology. July
359 2013;67(12):1951–9.

360 30. Lourenço J, Wikramaratna PS, Gupta S. MANTIS: an R package that simulates multilocus
361 models of pathogen evolution. BMC Bioinformatics. May 2015;16(1).

362 31. Antia R, Regoes RR, Koella JC, Bergstrom CT. The role of evolution in the emergence of
363 infectious diseases. Nature. December 2003;426(6967):658–61.

364 32. Tompkins DM, Carver S, Jones ME, Krkošek M, Skerratt LF. Emerging infectious diseases of
365 wildlife: a critical perspective.. Trends Parasitol. 2015;31:149–59.

366 33. Xiao Y, Zhou Y, Tang S. Modelling disease spread in dispersal networks at two levels. Mathe-

matical Medicine and Biology. September 2011;28(3):227–44.

34. Bezanson J, Edelman A, Karpinski S, Shah VB. Julia: a fresh approach to numerical computing. SIAM Review. January 2017;59(1):65–98.

35. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016.

36. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2019.

37. Pedersen TL. tidygraph: a tidy API for graph manipulation. 2019.