POPULATION STRUCTURE DOES NOT INCREASE PATHOGEN DIVERSITY IN THEORETICAL BAT POPULATIONS

TIM LUCAS. JANUARY 19, 2016

1. Abstract

1.0.1. *One or two sentences providing a basic introduction to the field.* An increasingly large fraction of emerging diseases come from wild animals and these diseases have a huge impact on human health, healthcare systems and economic development. The chance that a new zoonosis will come from any particular wild host species increases with the diversity of pathogens in that species. However, the factors that control pathogen diversity in wild populations are still unclear.

1.0.2. *Two to three sentences of more detailed background.* Host species traits such as population density, longevity, body size and population structure have been shown to correlate with pathogen diversity. Typically it is assumed that well-connected, unstructured populations promote the invasion of new pathogens and therefore increase pathogen diversity. However, this assumption is largely untested; in particular our mechanistic understanding of how population structure affects pathogen diversity, in the presence of inter-pathogen competition, is poor. Greater understanding is needed to clarify exactly by which mechanisms population structure affects pathogen diversity.

1.0.3. *One sentence clearly stating the general problem (the gap).* It is unknown whether population structure allows invading pathogens to escape from competition by stochastically creating areas of low pathogen prevalence. I hypothesised that low dispersal rates and a low number of connections in a metapopulation network will allow invading pathogens to establish more readily, thus increasing pathogen diversity. I tested these hypothesese using population networks parameterised to mimic wild bat populations as bats have highly varied social structures and have recently been implicated in a number of high profile diseases such as Ebola, SARS, Hendra and Nipah.

1.0.4. *One sentence summarising the main result.* I find that neither population connectedness nor dispersal rate affect the probability that a new pathogen will invade into a population. The assumption that factors causing high R_0 allow new pathogens to invade and therefore increase pathogen diversity is not supported. Instead we find that changes in population structure do not affect the probability of invasion of a new pathogen.

1.0.5. *One or two sentences to put the results into a more general context.* This result implies that population structure does not control pathogen diversity. Instead it seems that local dynamics, that affect the short term branching process at the very start of the epidemic control the chance of pathogen invasion. This also implies that population structure is not a useful proxy for pathogen diversity with respect to zoonotic disease surveillance, for example in bats.

2. Introduction

2.1. General Intro.

2.1.1. *Why is pathogen diversity important?* Over 50% of emerging infectious diseases have an animal source (Jones, Patel, et al. 2008; Smith et al. 2014). Zoonotic pathogens can be highly virulent (Luby et al. 2009; Lefebvre et al. 2014) and potentially have huge public health impacts (Granich et al. 2015), economic costs (Knobler et al. 2004) and slow down international development (World Bank 2014). Therefore understanding and predicting changes in the process of zoonotic spillover is a global health priority (Taylor et al. 2001). The number of pathogen species hosted a wild animal species affects the chance that a disease from that species will infect humans. However, the factors that control the diversity of pathogens in a wild animal population are still unclear (Metcalf et al. 2015); in particular our mechanistic understanding of how population processes inhibit or promote pathogen richness is poor.

2.2. **Specific Intro.**

2.2.1. *We know some factors that correlate with pathogen diversity.* A number of host traits have been shown to correlate with pathogen richness including body size (Kamiya et al. 2014; Arneberg 2002), population density (Nunn, Altizer, et al. 2003; Arneberg 2002) and range size (Bordes and Morand 2011; Kamiya et al. 2014). However, empirical, correlative studies are often contradictory due to small sample sizes, noisy data and because empirical relationships often do not generalise across taxa (though see (Kamiya et al. 2014) for a meta-analysis). The correlation between many traits (e.g. Nunn, Jordán, et al. (2015)) also makes it hard to clearly distinguish which factors are important. Furthermore, knowing which factors correlate with pathogen richness does not tell us how they controls richness. Mechanisms by which a trait could increase pathogen diversity include promoting the evolution of new strains within a species (Buckee et al. 2004), reduction of the rate of pathogen extinction (Rand et al. 1995) and an increased probability of pathogen invasion from other host species (Nunes et al. 2006). These separate mechanisms have not been examined and it is difficult to see how they could be approached through comparative methods. We need explicit mechanistic models in order to tease apart these factors.

2.2.2. *Population structure could be important.* One host-species trait that has been largely understudied with respect to pathogen diversity is population structure.

2.2.3. *Empirical evidence that structure might affect diversity.* Few studies focus on bats, despite their role in recent zoonoses. Maganga *et al.* found that distribution fragmentation predicts viral richness (Maganga et al. 2014), but (Gay et al. 2014) finds the opposite relationship. While the data set in (Gay et al. 2014) is larger, the analysis in (Maganga et al. 2014) is much more focused on fragmentation. Genetic correlates of population structure have also been used. Turmelle *et al.* (Turmelle and Olival 2009), in a small analysis, find that high F_{st} (i.e. a structured population) correlates with high richness. However, they do not account for the widely different spatial scales found in population structure studies, nor do they deal with the differences between F_{ST}, ϕ_{ST} and other measures appropriately.

A number of authors have demonstrated a competitive exclusion principle where only the most optimally spreading disease will survive when infection with one pathogen prevents infection with another pathogen (Bremermann and Thieme 1989; Martcheva and X.-Z. Li 2013; Ackleh and Allen 2003; Ackleh and Salceanu 2014; Turner and Garnett 2002; Castillo-Chavez et al. 1995).

2.2.4. *Network structure has been studied.* Studies of the role of population structure on pathogen diversity have been in very simple systems. These have been so simple that empirical data cannot easily be applied to them to predict pathogen diversity of real wild animal populations. There is a need for models that can be carefully and fully explored, while still capturing the complexities of the real world.

Analytical models of well mixed populations have widely different outcomes: infinite pathogen diversity (May and Nowak 1994; Ackleh and Salceanu 2014) or competitive exclusion have both been predicted (Ackleh and Allen 2003; Bremermann and Thieme 1989; Martcheva and X.-Z. Li 2013; Qiu et al. 2013; Allen et al. 2004). When competitive exclusion occurs, population structure has sometimes been shown to allow coexistence (Qiu et al. 2013; Allen et al. 2004; Nunes et al. 2006).

Competing epidemics, or two pathogens spreading at the same time in a population, is a well studied area (Poletto, Meloni, Colizza, et al. 2013; Poletto, Meloni, Van Metre, et al. 2015; Karrer and Newman 2011). This area is related to the study of pathogen richness in that they indicate that dynamics of multiple pathogens in a population do depend on population structure. However, the results for short term epidemic competition do not directly transfer to the study of long term disease persistence which is the relative time scale when looking at zoonotic potential of a host species.

If competitive mechanisms are strong, pathogens in populations structured such that R_0 will be high will be able to easily out-compete invading pathogens. Only if competitive mechanisms are weak will high R_0 enable the invasion of new pathogens and allow higher pathogen diversity.

Competitive affects are likely to be more or less important depending on the mechanism by which pathogen diversity is created. Competition is likely to be very strong between a newly evolved strain and the existing strain of a pathogen. In this case, there is likely to be strong cross immunity between strains. In the case of pathogens invading from other hosts species, cross immunity is likely to be less strong due to the evolutionary distance between the new and endemic pathogens. However, competition could still occur as disease induced mortality of host individuals reduces host density which promotes pathogen persistence. Finally, it is unclear how competition would affect the persistence of pathogens once they are established, though it seems that cross-immunity can keep a pathogen at a low prevalence in the population (e.g. monkey pox (Rimoin et al. 2010) and monkey malaria (Cox-Singh and Singh 2008)). This would make stochastic extinction more likely.

2.2.5. *Types of population structure.* How structured a population is can be defined in many ways on many scales. The most relevant scale is that of an epidemiological population. This is the population within which a pathogen can spread in an epidemiologically relevant time period (years or decades). It is therefore closely related to a population as defined by population genetics, but with movement defined on a shorter time scale.

The epidemiological contacts within the population can be examined at the individual level (as in contact network epidemiology) or larger scales. I consider the metapopulation network the most appropriate. Ignoring the metapopulation assumes a fully mixed population which is unlikely. Trying to study the contact network relies detailed individual level detail which is not available. Metapopulation models consider a network of small subpopulations. Within subpopulations, epidemiological contacts are fully mixed and relatively fast. Between subpopulations, epidemiological contacts are dependant on an underlying network structure and relatively slow. The network underlying the metapopulation is made up of nodes representing the subpopulations, and edges which represent movement between subpopulations. Animals, and therefore infection, can only move between two subpopulations if they are connected by an edge.

There are two factors that affect how structured a population is, given this model framework. Firstly, dispersal is the rate at which individuals move between subpopulations. Secondly, the metapopulation network structure controls population structure. The simplest measure of how structured the network is the average number of edges each node has. In the extremes, all subpopulations could be either connected to all other subpopulations or only connected to one or two other subpopulations. However,

other measures that take into account second-order structure in the network are also often used.

2.2.6. *Why bats.* Bats have an unusually large variety of social structures. Colony sizes range from 10 to 1 million (Jones, Bielby, et al. 2009). Many bats also have interesting seasonal behaviour such as migration (Richter and Cumming 2008; Fleming et al. 2003), hibernation, birth puleses and swarming behaviours. Furthermore, bats have, over the last decade, become a focus for disease research (Calisher et al. 2006; Hughes et al. 2007). Recently they have been implicated in a number of high profile diseases such as Ebola, SARS, Hendra and Nipah (Calisher et al. 2006; W. Li et al. 2005).

A number of traits have been suggested as predisposing bats towards being reservoirs of zoonotic diseases: high sympatry (Luis et al. 2013), flight (L. Wang et al. 2011) and longevity (L. Wang et al. 2011).

2.3. **The gap.** We have very abstract, simplified models that predict zero or infinite diversity depending on specifics. These cannot be easily applied to real data. They also do not easily predict quantitative or even relative diversity as they often predict either zero or infinite diversity with nothing in between.

We need models that can quantitatively or at least relatively predict diversity in a populations. This requires a middle ground of model complexity.

There are no studies that directly model bat pathogen diversity.

Specifically we use these models to test the affects of population structure on the ability of a new pathogen to invade a population. We test two aspects of structure, dispersal rate and connectedness of the metapopulation network.

2.4. **What I did.** To this end I have run metapopulation, multipathogen, epidemiological simulations based broadly on real-world bat populations. Although still simplified, the model is complex enough that if good measurements of bat populations could be found, simulations of the real world bat population could be run.

I have studied the invasion of new pathogens as a mechanisms for increasing pathogen richness. In particular I have focussed on studying the invasion of a newly evolved pathogen that is therefore identical in epidemiological parameters to the endemic pathogen. Furthermore, this close evolutionary relationship means that cross immunity is strong.

I have studied two metrics for population structure, dispersal rate and metapopulation network topology, to test for effects of population structure on pathogen richness.

2.5. **What I found.** Here I show that, given the assumptions of a metapopulation, population structure does not affect the rate of invasion of new pathogens. This is true whether population structure is altered by changing the dispersal rate or the topology of the metapopulation network.

3. Methods

3.1. Metapopulation model.

3.1.1. *Two pathogen SIR model.* I examine a multipathogen SIR model with vital dynamics. This is a compartment model with individuals being classed as susceptible, infected or recovered with immunity (Figure 1). Susceptible individuals are counted in class S. There are three infected classes, I_1, I_2 and I_{12}, being individuals infected with pathogen 1, pathogen 2 or both respectively. Recovered individuals, R, are immune to both pathogens, even if they have only been infected with one. Furthermore, recovery from one pathogen moves an individual straight into the recovered class, even if the individual is infected with both pathogen. This modelling choice allows the model to be easily expanded to include more than two pathogens. The assumption of immediate recovery from all other diseases is likely to be quite accurate for very closely related

pathogens as is being studied here as once an acquired immune response is activated, all infections are likely to be cleared quickly.

The coinfection rate (the rate at which an infected individual is infected with a second pathogen) is adjusted compared to the infection rate by a factor α (here I used $\alpha = 0.1$ which means coinfection happens at a tenth the rate of first infections). In the application of long term existence of pathogens it is vital to include vital dynamics (births and deaths) as the SIR without vital dynamics has no endemic state. Birth and death rates are assumed to be equal, $b = d$ (see Table 1 for a list of symbols and values used). The time scale of the simulations are set by setting b = 0.05 per year, yielding an average host generation time of 20 years.

3.1.2. *Metapopulation.* The population is divided into a number of subpopulations. This metapopulation is modelled as a network with subpopulations being nodes and dispersal between subpopulations being indicated by edges (Figure 2). Individuals within a subpopulation interact randomly so that the subpopulation is fully mixed. Dispersal between subpopulations occurs at a rate λ. Individuals can only disperse to subpopulations connected to theirs in the network. The rate of dispersal is not affected by the number of edges a subpopulation has (the degree of the subpopulation). So the

TABLE 1. All symbols used.

Symbol	Explanation	Units	Value
S	Number of susceptible individuals		
I_q	N. individuals infectious with diseases q		
I_p^+	Sum of classes infected with pathogen p		
$\bar{n}$	Colony starting size		3,000
β	Transmission rate	Transmission events per year per individual	2, 5, 10
γ	Recovery rate	Recovery events per year per individual.	1
λ	Dispersal	Dispersal events per day per individual	0.001–0.1
b	Birth rate	Births per year per individual	0.05
d	Death rate	Deaths per year per individual	0.05
ρ	No. pathogens		2
p	Pathogen index i.e. $p \in \{1, 2\}$ for pathogens 1 and 2		
q	Disease class i.e., $q \in \{1, 2, 12\}$		
t, t'	Time and time plus waiting time i.e., $t + \delta$	Days	
k_i	Degree of node i		
δ	Waiting time until next event	Days	
α	Cross immunity	Proportion	0.1
n, m	Colony index		
σ	Invading pathogen seed size		10
r_i	The rate that event i occurs.	Days^{-1}	

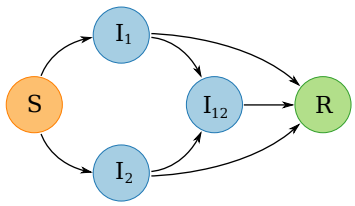


FIGURE 1. Schematic of the SIR model used. Individuals are in one five classes, susceptible (orange, S), infectious with pathogen 1, pathogen 2 or both (blue, I_1, I_2, I_{12}) or recovered and immune from further infection (green, R). Transitions between classes occurs only as indicated by arrows. Note that individuals in I_{12} move into R, not back to I_1 or I_2. That is, recovery from one pathogen causes immediate recovery from the other pathogen.

dispersal rate from a subpopulation m with degree k_m to subpopulation n is $\frac{\lambda}{k_m}$. Note this rate is independent of the degree of subpopulation n.

3.1.3. *Stochastic simulations.* I examine this model using stochastic, continuous time simulations (using the Gillespie algorithm) implemented in R (R Development Core Team 2010). At each step in the simulation we calculate the rate that each possible event might occur. One event is then randomly chosen, weighted by it's rate

$$p(\text{event } i) = \frac{r_i}{\sum_i r_i} \tag{1}$$

where r_i is the rate that event i occurs. Finally, the length of the time step, δ, is drawn from an exponential distribution

$$\delta \sim \text{Exp}\left(\sum_i r_i\right). \tag{2}$$

This means that the total length of each simulation is stochastic. We define the number of events we wish to simulate instead.

We can now write down the rates of all events. I defined I_p^+ to be the sum of all classes that are infectious with pathogen p, for example $I_1^+ = I_1 + I_{12}$. Assuming asexual reproduction, that all classes reproduce at the same rate and that individuals are born into the susceptible class we get

$$P\left(S_{nt'} = S_{nt} + 1\right) = b\left(S_{nt} + \sum_q I_{qnt} + R_{nt}\right) \tag{3}$$

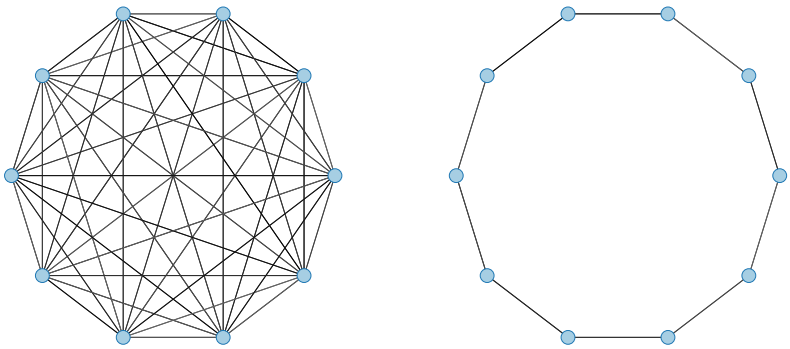
where $P\left(S_{nt'} = S_{nt} + 1\right)$ is the probability that the number of susceptibles in subpopulation n will increase by 1 (a single birth) in the time interval t to t' and $\sum_q I_{qnt}$ is the sum of all infection classes $q \in 1, 2, 12$. The rates of death, given a death rate d are given by

$$P\left(S_{nt'} = S_{nt} - 1\right) = dS_{nt}, \tag{4}$$

$$P\left(I_{qnt'} = I_{qnt} - 1\right) = dI_{qnt}, \tag{5}$$

$$P\left(R_{nt'} = R_{nt} - 1\right) = dR_{nt}. \tag{6}$$

I modelled transmission as density-dependant. This assumption is more suitable than frequency-dependant transmission as I am modelling a disease transmitted by saliva



(A) Fully connected (B) Minimally connected

FIGURE 2. The two network topologies used to test whether network connectedness influences a pathogens ability to invade. A) Dispersal can occur between any subpopulation. B) Animals can only disperse to neighbouring subpopulations. Blue circles are subpopulations of 3,000 individuals. Dispersal only occurs between subpopulations connected by an edge (black line). The dispersal rate is held constant between the two topologies.

or urine in highly dense populations confined to caves, building or potentially a small number of tree roosts. I am notably not modelling an STD as these diseases are not expected to be commonly zoonotic. Infection of a susceptible with either pathogen 1 or 2, $S \to I_p$ where $p \in \{1,2\}$, is therefore given by

$$P\left(I_{pnt'} = I_{pnt} + 1, S_{nt'} = S_{nt} - 1\right) = \beta S_{nt} I^{+}_{pnt}, \tag{7}$$

while coinfection, given a crossimmunity factor α, is given by

$$P\left(I_{12,nt'} = I_{12,nt} + 1,\ I_{1nt'} = I_{1nt} - 1\right) = \alpha\beta I_{1nt} I^{+}_{2nt}, \tag{8}$$

$$P\left(I_{12,nt'} = I_{12,nt} + 1,\ I_{2nt'} = I_{2nt} - 1\right) = \alpha\beta I_{2nt} I^{+}_{1nt}. \tag{9}$$

The probability of migration from colony m (with degree k_m) to colony n, given a dispersal rate λ is given by

$$P\left(S_{nt'} = S_{nt} + 1,\ S_{mt'} = S_{mt} - 1\right) = \frac{\lambda S_{mt}}{k_m - 1}, \tag{10}$$

$$P\left(I_{qnt'} = I_{qnt} + 1,\ I_{qmt'} = I_{qmt} - 1\right) = \frac{\lambda I_{qmt}}{k_m - 1}, \tag{11}$$

$$P\left(R_{nt'} = S_{nt} + 1,\ R_{mt'} = R_{mt} - 1\right) = \frac{\lambda R_{mt}}{k_m - 1}. \tag{12}$$

Finally, recovery from any infectious class occurs at a rate γ

$$P\left(I_{qnt'} = I_{qnt} - 1,\ R_{nt'} = R_{nt} + 1\right) = \gamma I_{qnt}. \tag{13}$$

In each simulation the population was seeded with 10 sets of 200 infected individuals of pathogen 1. There groups were seeded into randomly selected colonies with replacement. For each 200 infected individuals added, 200 susceptible individuals were removed to keep starting colony sizes constant. Pathogen 1 was then allowed to spread and reach equilibrium. After 3×10^5 events, 5 individuals infected with pathogen 2

were added to one colony. Visual inspection of preliminary simulations was used to decide on 3×10^5 as being long enough for the epidemic to reach an equilibrium state. After another 5×10^5 events the invasion of pathgen 2 is considered successful if any individuals are still infected with pathogen 2. Again visual inspection of preliminary simulations was used to determine whether this was long enough to be confident that an invading pathogen.

3.2. **Dispersal.** The values used for the independant variables are chosen to highlight the affects of these variables. Dispersal values are $\lambda = 0.1, 0.01$ and 0.001 dispersals per individual per year. $\lambda = 0.1$ relates to individuals moving between colonies on average twice per lifetime. Therefore exclusively juvenile dispersal would have dispersal rates similar to this. Otherwise it relates to dispersal being a rare event with animals often staying in a colony for many years. $\lambda = 0.01$ relates to 20% of individuals dispersing once in their lifetime. This value is therefore close to male-biased dispersal, with female philopatry. Finally, $\lambda = 0.001$ relates to 2% of individuals dispersing in their lifetime. This therefore relates to a population that does not habitually disperse.

3.3. **Network structure.** The network structure is synthetically created to be either fully or minimally connected (See Figure 2). 10 subpopulations was selected as a trade off between computation time and a network complicated enough that structure might have an effect. This value is artificially small compared to wildlife populations.

3.4. **Parameter selection.** The fixed parameters used are chosen to roughly reflect realistic wild bat populations. The death rate d is set as 0.05 per year giving a generation time of 20 years. The birth rate b is set to be equal to d so that the population size is stable. The recovery rate γ is set to 1 giving a average infection duration of 1 years. This is therefore a long lasting infection but not a chronic infection. It is very difficult to directly estimate infection durations in wild populations. But it seems that these infections might be long lasting ().

Cross immunity is set to 0.1 so that an individual infected with one pathogen is 90% less likely to be infected with another. This is a rather arbitrary value. However, the rationale of the model is that the invading species might be a newly speciated strain of the endemic species. Furthermore, the model assumes complete cross immunity after recovery from infection. Therefore cross immunity to coinfection is likely to be very strong as well.

The population size of each subpopulation is set to 3,000. This is appropriate for many bat species (Jones, Bielby, et al. 2009), especially the large, frugivorous *Pteropodidae* that have been particularly associated with recent zoonotic diseases.

Four values of the transmission rate β are used, 0.1, 0.2, 0.3 and 0.4. All simulations are run under all four transmission rates as this is such a fundamental parameter. Given the recovery, birth and death rates we can calculate an approximation of R_0 that ignores spatial structure. That is, this is R_0 for the local, within-subpopulation dynamics. Furthermore, it is R_0 for the first pathogen; R_0 of the invading pathogen will be lower due to competition. We can calculate that $R_0 \approx \frac{\beta b}{d(d+\gamma)}$. For our four values of $\beta = 0.1, 0.2, 0.3, 0.4$ we therefore get $R_0 \approx 2.9, 4.8, 6.7, 8.6$. These values are very high in part to again find a reasonable trade off between the number of simulations and the reasonableness of the parameters. $R_0 \approx 13.3$ is similar to a highly contagious pathogen such as measles or pertussis.

For the simulations where an invading pathogen is added to the populations the number of invading pathogens added is set to 10. This is a trade off between getting a reasonable proportion of invasions, while still retaining the stochastic nature of invasion.

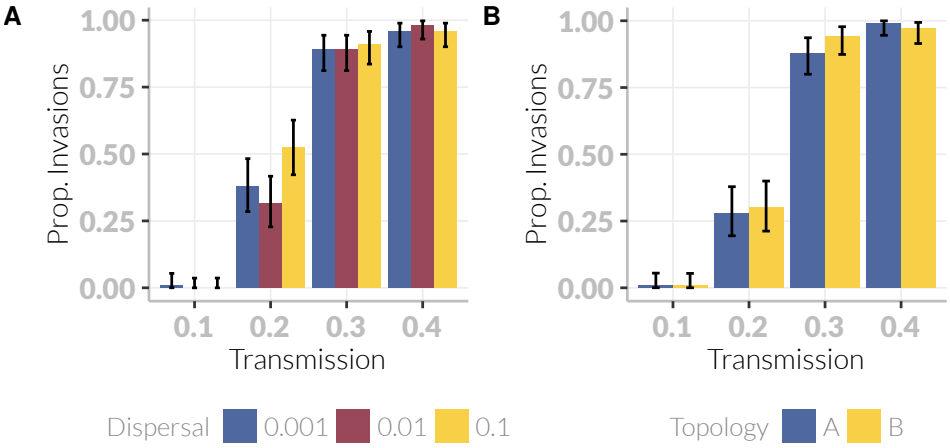


FIGURE 3. The probability of successful invasion. For three different transmission rates, the probability of invasion does not change between different a) dispersal rates or b) network topologies (with network topologies A and B refering to the fully and minimally connected topologies in Figure 2). Error bars are 95% confidence intervals. 100 simulations were run for each treatment. Other parameters are kept constant at: $N = 10$, $b = d = 0.05$, $\gamma = 0.1$, $\alpha = 0.1$. When dispersal is varied, the population structure is fully connected. When population structure is varied, $\lambda = 0.01$.

4. RESULTS

4.1. Invasion.

4.1.1. *Dispersal.* The proportion of invasions was not different across dispersal rates across 1200 simulations run (Figure 3A). This was true at all transmission levels (χ^2 test. $\beta = 0.1$: $\chi^2 = 1.98$, df $= 2$, $p = 0.37$. $\beta = 0.2$: $\chi^2 = 9.44$, df $= 2$, $p = 8.9 \times 10^{-3}$. $\beta = 0.3$: $\chi^2 = 0.29$, df $= 2$, $p = 0.87$, $\beta = 0.4$: $\chi^2 = 0.83$, df $= 2$, $p = 0.66$).

4.1.2. *Network structure.* I ran 800 simulations over 4 transmission values ($\beta = 0.1$, 0.2, 0.3, 0.4). The proportion of invasions was not different between highly connected and largely unconnected metapopulations (Figure 3B). This was true at all transmission levels (χ^2 test. $\beta = 0.1$: $\chi^2 = 5.23 \times 10^{-31}$, df $= 1$, $p = 1$. $\beta = 0.2$: $\chi^2 = 0.02$, df $= 1$, $p = 0.88$. $\beta = 0.3$: $\chi^2 = 1.53$, df $= 1$, $p = 0.22$. $\beta = 0.4$: $\chi^2 = 0.26$, df $= 1$, $p = 0.61$).

4.2. **Transmission.** Inline with theory, increasing the transmission rate increased the probability of invasion (Figure 3). This is true for all three dispersal values (χ^2 test. $\lambda = 0.001$: $\chi^2 = 243.56$, df $= 3$, $p = p < 10^{-5}$. $\lambda = 0.01$: $\chi^2 = 265.4$, df $= 3$, $p < 10^{-5}$. $\lambda = 0.1$: $\chi^2 = 244.95$, df $= 3$, $p < 10^{-5}$) and both network structures (χ^2 test. Fully connected: $\chi^2 = 267.3$, df $= 3$, $p < 10^{-5}$. Minimally connected: $\chi^2 = 267.3$, df $= 3$, $p < 10^{-5}$).

5. DISCUSSION

5.1. **Restate the gap and the main result.** Empirical studies on the role of population structure on pathogen richness are equivocal and cannot examine the specific mechanisms by which pathogen communities are created and maintained. I have used mechanistic, metapopulation models to test whether increased population structure

can promote pathogen richness by facilitating invasion of new pathogens. I find that population structure does not affect the ability of a new pathogen to invade and persist in a population. Instead I find that only transmission rate affects the chance of pathogen invasion.

5.2. **Link results to consequences.** Given that we find no affect of population structure on the ability of a new pathogen to invade, this host-trait is not a useful host trait for predicting the probability that a wild species has many pathogens. This means it cannot be used as a proxy for the probability that a species carries a zoonotic pathogen.

However my simulations do highlight the importance of competition for the spread of a new pathogen. All parameters used correspong to pathogens with $R_0 > 1$. However, the competition with the endemic pathogen means that for some transmission rates the chance of epidemic spread and persistence is close to zero. This has implications for human epidemics as well — if there is strong competiton between a newly evolved strain and an endemicstrain, we ar3 unlikely to see the new strain spread, irregardless of population structure.

5.2.1. *Population structure does not affect pathogen richness.* I have found no evidence that population structure aids the invasion and establishment of newly evolved pathogen species. Instead it seems that invasion relies solely on the local dynamics of the disease. If the transmission rate is high, a pathogen has a good chance of spreading immediately after it's introduction when it is at very density. Once the pathogen has infected a large number of individuals in the local metapopulation, it is bound to spread throughout the metapopulation without going extinct.

This goes against many predictions that increasing R_0 increases pathogen richness. Further work could examine reduced col ony sizes to test when global structure become more important. Measures of population structure should not be used to predict zoonotic potential.

5.2.2. *Dispersal does not affect pathogen richness.* Irrespective of the network topology used, at very high dispersal rates a population will be well mixed.

5.2.3. *Network connectedness does not affect pathogen richness.* This is in direct contrast to (Campos and Gordo 2006). However, the model in (Campos and Gordo 2006) is a contact network, so increasing the connectedness increases the chance of succesful transmission events for the first few transmission generations. This lends support to the idea that I found no affect of connectedness due to the dominance of local dynamics.

5.3. **Discuss assumptions.**

5.3.1. *Complete cross-immunity.* I have assumed that once recovered, individuals are immune to both pathogens. Furthermore, when a coinfected individual recovers from one pathogen, it immediately recovers from the other as well. This is probably a fairly reasonable assumption given that I am modelling a newly evolved strain. However, further work could relax this assumption using a model similar to (Poletto, Meloni, Van Metre, et al. 2015) which contains additional classes for 'infected with pathogen one, immune to pathogen two' and 'infected with pathogen two, immune to pathogen one'. The model here was formulated such that the study of systems with greater than two pathogens is still computationally feasible while a model such as used in (Poletto, Meloni, Van Metre, et al. 2015) contains 3^ρ classes for a system with ρ pathogen species. This quickly becomes computationally restrictive.

5.3.2. *Identical strains.* Many papers on pathogen richness have focussed on the evolution of pathogen traits and have considered a trade off between transmission rate and virulence (Nowak and May 1994; Nowak and May 1994) or infectious period (Poletto, Meloni, Colizza, et al. 2013). However, here I am interested in host traits. Therefore we have assumed that pathogen strains are identical. It is clear however that there are a number of factors that affect pathogen richness and our focus on host population structure does not imply that pathogen traits are not important.

5.4. **Complex social structure and behaviour.** With the models here I have aimed to tread a middle ground between the overly simple models employed in analytical studies (Allen et al. 2004) and the full complexity and variety of true bat social systems (Kerth 2008). Omissions include seasonal migration, maternity roosts, hibernation roosts and swarming sites (Kerth 2008; Fleming et al. 2003; Richter and Cumming 2008; Cryan et al. 2014). While future models might aim to model this complexity more fully the number of parameters that are required to be estimated and varied becomes very large. Furthermore, not all of these social complexities exist in all bat species, so in limiting my analysis to the simpler end of bat social systems it is hoped that the results are more broadly representative of the order.

Furthermore, I have considered a single host species in isolation. In some cases, treating multiple host species as identical could be appropriate. However, more generally, it seems likely that sympatry in bats is epidemiologically important (Brierley et al. 2016; Luis et al. 2013) but was beyond the scope of this study. There is potential for this to be effectively modelled as a multilayered network (W. Wang et al. 2016; Funk and Jansen 2010) and this would be expected to act to reduce population structure.

Finally, many species of bat exhibit strong seasonal birth pulses which are known to affect disease dynamics (Hayman 2015; Peel et al. 2014; Amman et al. 2012). This would be expected to facilitate the invasion of new pathogen species; if a new strain evolved or entered the population by migration during a period of low population immunity, it would have a higher chance of invading and establishing in the population.

5.5. **Conclusions.** In conclusion I have found no evidence that population structure aids the invasion and establishment of newly evolved pathogen species. This suggests that if population structure does have a role in shaping pathogen communities, it is not by this specific mechanism. Practically, this implies that population structure is not a useful metric for predicting pathogen richness in wild bats. Furthermore it implies that population fragmentation by global change will not promote an increase in pathogen richness.

References

Jones, K. E., N. G. Patel, et al. (2008). "Global trends in emerging infectious diseases". In: *Nature* 451.7181, pp. 990–993. DOI: 10.3201/eid0903.020336.

Smith, K. F. et al. (2014). "Global rise in human infectious disease outbreaks". In: *Journal of The Royal Society Interface* 11.101, p. 20140950. DOI: 10.1098/rsif.2014.0950.

Luby, S. P. et al. (2009). "Recurrent zoonotic transmission of Nipah virus into humans, Bangladesh, 2001–2007". In: *Emerging infectious diseases* 15.8, p. 1229. DOI: 10.3201/eid1508.081237.

Lefebvre, A. et al. (2014). "Case fatality rates of Ebola virus diseases: a meta-analysis of World Health Organization data". In: *Médecine et maladies infectieuses* 44.9, pp. 412–416. DOI: 10.1016/j.medmal.2014.08.005.

Granich, R. et al. (2015). "Trends in AIDS deaths, new infections and ART coverage in the top 30 countries with the highest AIDS mortality burden; 1990–2013". In: *PloS one* 10.7, e0131353. DOI: 10.1371/journal.pone.0131353.

Knobler, S. et al. (2004). *Learning from SARS: Preparing for the Next Disease Outbreak: Workshop Summary*. National Academies Press Washington, DC, USA: DOI: 10.1016/j.socscimed.2006.08.004.

World Bank (2014). *Update on the economic impact of the 2014 Ebola epidemic on Liberia, Sierra Leone, and Guinea*. Washington, DC. DOI: 10.1211/pj.2014.20066832. URL: http://documents.worldbank.org/curated/en/2014/12/20454884/update-economic-impact-2014-ebola-epidemic-liberia-sierra-leone-guinea.

Taylor, L. H., S. M. Latham, and E. Mark (2001). "Risk factors for human disease emergence". In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356.1411, pp. 983–989. DOI: 10.1093/acprof:oso/9780198525738.003.0035.

Metcalf, C. et al. (2015). "Five challenges in evolution and infectious diseases". In: *Epidemics* 10, pp. 40–44. DOI: 10.1016/j.epidem.2014.12.003.

Kamiya, T. et al. (2014). "What determines species richness of parasitic organisms? A meta-analysis across animal, plant and fungal hosts". In: *Biological Reviews* 89.1, pp. 123–134. DOI: 10.1111/brv.12046.

Arneberg, P. (2002). "Host population density and body mass as determinants of species richness in parasite communities: comparative analyses of directly transmitted nematodes of mammals". In: *Ecography* 1, pp. 88–94. DOI: 10.1034/j.1600-0587.2002.250110.x.

Nunn, C. L., S. Altizer, et al. (2003). "Comparative tests of parasite species richness in primates". In: *The American Naturalist* 162.5, pp. 597–614. DOI: 10.1086/378721.

Bordes, F. and S. Morand (2011). "The impact of multiple infections on wild animal hosts: a review". In: *Infection ecology & epidemiology* 1. DOI: 10.3402/iee.v1i0.7346.

Nunn, C. L., F. Jordán, et al. (2015). "Infectious disease and group size: more than just a numbers game". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 370.1669, p. 20140111. DOI: 10.1098/rstb.2014.0111.

Buckee, C. O. et al. (2004). "The effects of host contact network structure on pathogen diversity and strain structure". In: *Proceedings of the National Academy of Sciences of the United States of America* 101.29, pp. 10839–10844. DOI: 10.1073/pnas.0402000101.

Rand, D., M. Keeling, and H. Wilson (1995). "Invasion, stability and evolution to criticality in spatially extended, artificial host-pathogen ecologies". In: *Proceedings of the Royal Society of London B: Biological Sciences* 259.1354, pp. 55–63. DOI: 10.1098/rspb.1995.0009.

Nunes, A., M. T. da Gama, and M. Gomes (2006). "Localized contacts between hosts reduce pathogen diversity". In: *Journal of theoretical biology* 241.3, pp. 477–487. DOI: 10.1016/j.jtbi.2005.12.010.

Maganga, G. D. et al. (2014). "Bat distribution size or shape as determinant of viral richness in african bats". In: *PloS one* 9.6, e100172. DOI: 10.1371/journal.pone.0100172.

Gay, N. et al. (2014). "Parasite and viral species richness of Southeast Asian bats: Fragmentation of area distribution matters". In: *International Journal for Parasitology: Parasites and Wildlife* 3.2, pp. 161–170. DOI: 10.1016/j.ijppaw.2014.06.003.

Turmelle, A. S. and K. J. Olival (2009). "Correlates of viral richness in bats (order Chiroptera)". In: *EcoHealth* 6.4, pp. 522–539. DOI: 10.1007/s10393-009-0263-8.

Bremermann, H. J. and H. Thieme (1989). "A competitive exclusion principle for pathogen virulence". In: *Journal of mathematical biology* 27, pp. 179–190. DOI: 10.1007/BF00276102.

Martcheva, M. and X.-Z. Li (2013). "Competitive exclusion in an infection-age structured model with environmental transmission". In: *Journal of Mathematical Analysis and Applications* 408.1, pp. 225–246. DOI: 10.1016/j.jmaa.2013.05.064.

Ackleh, A. S. and L. J. Allen (2003). "Competitive exclusion and coexistence for pathogens in an epidemic model with variable population size". In: *Journal of mathematical biology* 47.2, pp. 153–168. DOI: 10.1007/s00285-003-0207-9.

Ackleh, A. S. and P. L. Salceanu (2014). "Robust uniform persistence and competitive exclusion in a nonautonomous multi-strain SIR epidemic model with disease-induced mortality". In: *Journal of mathematical biology* 1-2, pp. 453–475. DOI: 10.1007/s00285-012-0636-4.

Turner, K. and G. Garnett (2002). "The impact of the phase of an epidemic of sexually transmitted infection on the evolution of the organism". In: *Sexually transmitted infections* 78, pp. i20–i30. DOI: 10.1136/sti.78.suppl_1.i20.

Castillo-Chavez, C., W. Huang, and J. Li (1995). "Dynamics of multiple pathogen strains in heterosexual epidemiological models". In: *Differential Equations and Applications to Biology and Industry. World Scientific Publishing Co., Singapore*, pp. 289–298.

May, R. M. and M. A. Nowak (1994). "Superinfection, metapopulation dynamics, and the evolution of diversity". In: *Journal of Theoretical Biology* 170.1, pp. 95–114. DOI: 10.1006/jtbi.1994.1171.

Qiu, Z. et al. (2013). "The vector–host epidemic model with multiple strains in a patchy environment". In: *Journal of Mathematical Analysis and Applications* 405.1, pp. 12–36. DOI: 10.1016/j.jmaa.2013.03.042.

Allen, L. J., N. Kirupaharan, and S. M. Wilson (2004). "SIS epidemic models with multiple pathogen strains". In: *Journal of Difference Equations and Applications* 10.1, pp. 53–75. DOI: 10.1080/10236190310001603680.

Poletto, C., S. Meloni, V. Colizza, et al. (2013). "Host mobility drives pathogen competition in spatially structured populations". In: *PLoS computational biology* 9.8, e1003169. DOI: 10.1371/journal.pcbi.1003169.

Poletto, C., S. Meloni, A. Van Metre, et al. (2015). "Characterising two-pathogen competition in spatially structured environments". In: *Scientific reports* 5. DOI: 10.1038/srep07895.

Karrer, B. and M. Newman (2011). "Competing epidemics on complex networks". In: *Physical Review E* 84.3, p. 036106. DOI: 10.1103/physreve.84.036106.

Rimoin, A. W. et al. (2010). "Major increase in human monkeypox incidence 30 years after smallpox vaccination campaigns cease in the Democratic Republic of Congo". In: *Proceedings of the National Academy of Sciences* 107.37, pp. 16262–16267. DOI: 10.3410/f.5239959.5179059.

Cox-Singh, J. and B. Singh (2008). "Knowlesi malaria: newly emergent and of public health importance?" In: *Trends in parasitology* 24.9, pp. 406–410. DOI: 10.1016/j.pt.2008.06.001.

Jones, K. E., J. Bielby, et al. (2009). "PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184". In: *Ecology* 90.9, pp. 2648–2648. DOI: 10.1890/08-1494.1.

Richter, H. and G. Cumming (2008). "First application of satellite telemetry to track African straw-coloured fruit bat migration". In: *Journal of Zoology* 275.2, pp. 172–176. DOI: 10.1111/j.1469-7998.2008.00425.x.

Fleming, T. H. et al. (2003). "Ecology of bat migration". In: *Bat ecology*. Ed. by T. Kunz and M. Brock Fenton. Chicago: The University of Chicago Press, pp. 156–208.

Calisher, C. et al. (2006). "Bats: important reservoir hosts of emerging viruses". In: *Clinical Microbiology Reviews* 19.3, pp. 531–545. DOI: 10.1128/CMR.00017-06.

Hughes, J. M. et al. (2007). "Emerging viruses: coming in on a wrinkled wing and a prayer". In: *Clinical Infectious Diseases* 44.5, pp. 711–717. DOI: 10.1086/511078.

Li, W. et al. (2005). "Bats are natural reservoirs of SARS-like coronaviruses". In: *Science* 310.5748, pp. 676–679. DOI: 10.1126/science.1118391.

Luis, A. D. et al. (2013). "A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special?" In: *Proceedings of the Royal Society B: Biological Sciences* 280.1756, p. 20122753. DOI: 10.1098/rspb.2012.2753.

Wang, L., P. Walker, and L. Poon (2011). "Mass extinctions, biodiversity and mitochondrial function: are bats 'special' as reservoirs for emerging viruses?" In: *Current Opinion in Virology*. DOI: 10.1016/j.coviro.2011.10.013.

R Development Core Team (2010). *R: A Language And Environment For Statistical Computing*. R Foundation For Statistical Computing. Vienna, Austria. URL: http://www.R-Project.org.

Campos, P. R. and I. Gordo (2006). "Pathogen genetic variation in small-world host contact structures". In: *Journal of Statistical Mechanics: Theory and Experiment* 2006.12, p. L12003. DOI: 10.1088/1742-5468/2006/12/L12003.

Nowak, M. A. and R. M. May (1994). "Superinfection and the evolution of parasite virulence". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 255.1342, pp. 81–89. DOI: 10.1098/rspb.1994.0012.

Kerth, G. (2008). "Causes and consequences of sociality in bats". In: *Bioscience* 58.8, pp. 737–746. DOI: 10.1641/B580810.

Cryan, P. M., C. A. Stricker, and M. B. Wunder (2014). "Continental-scale, seasonal movements of a heterothermic migratory tree bat". In: *Ecological Applications* 24.4, pp. 602–616. DOI: 10.1890/13-0752.1.

Brierley, L. et al. (2016). "Quantifying Global Drivers of Zoonotic Bat Viruses: A Process-based Perspective". In: *The American Naturalist* 187.

Wang, W. et al. (2016). "Structural diversity effects of multilayer networks on the threshold of interacting epidemics". In: *Physica A: Statistical Mechanics and its Applications* 443, pp. 254–262. DOI: 10.1016/j.physa.2015.09.064.

Funk, S. and V. A. Jansen (2010). "Interacting epidemics on overlay networks". In: *Physical Review E* 81.3, p. 036118.

Hayman, D. T. (2015). "Biannual birth pulses allow filoviruses to persist in bat populations". In: *Proceedings of the Royal Society of London B: Biological Sciences* 282.1803, p. 20142591.

Peel, A. J. et al. (2014). "The effect of seasonal birth pulses on pathogen persistence in wild mammal populations". In: *Proceedings of the Royal Society B: Biological Sciences* 281.1786, p. 20132962. DOI: 10.1098/rspb.2013.2962.

Amman, B. et al. (2012). "Seasonal pulses of Marburg virus circulation in juvenile *Rousettus aegyptiacus* bats coincide with periods of increased risk of human infection". In: *PLoS Pathogens* 8.10, e1002877.