

A simple metapopulation model for Red Queen dynamics: parameter estimation and power analysis

Sang Woo Park and Benjamin M Bolker

October 17, 2019

Abstract

Sexual reproduction persists in nature despite its large cost. The Red Queen Hypothesis postulates that parasite pressure maintains sexual reproduction in the host population by selecting for the ability to produce rare genotypes that are resistant to infection. Mathematical models have been used to lay theoretical foundations for the hypothesis; empirical studies have confirmed these predictions. For example, Lively used a simple host-parasite model to predict that the frequency of sexual hosts should be positively correlated with the prevalence of infection, provided that there is sufficient variance in the risk of infection among populations. Lively *et al.* later confirmed the prediction through numerous field studies of snail-trematode systems in New Zealand. In this study, we fit a simple metapopulation host-parasite coevolution model to three data sets (one from an Israeli population and two from New Zealand populations) by matching the observed prevalence of sexual reproduction and trematode infection among hosts. Using the estimated parameters, we perform a power analysis to test the feasibility of observing the positive correlation predicted by Lively. We discuss anomalies in the data that are poorly explained by the model and provide practical guidance to both modelers and empiricists. Overall, our study suggests that a simple Red Queen model can only partially explain the observed relationships between parasite infection and the maintenance of sexual reproduction.

1 Introduction

Despite being the dominant mode of reproduction in multicellular organisms (???), sexual reproduction entails numerous costs (?). The most commonly mentioned is the cost of producing males (?): as males cannot produce offspring, asexual lineages are expected to outgrow their sexual counterparts. This *two-fold cost of sex* (?) relies on the assumption that everything else is equal. Then, what drives the observed prevalence of sexual reproduction?

One explanation for the persistence of sexual reproduction is the Red Queen Hypothesis (RQH) (?). The RQH predicts that sexually reproducing hosts overcome the cost of sex under strong parasitic pressure by producing genetically diverse offspring that can escape infection (????).

Much of the theoretical literature has focused on determining qualitative conditions under which parasite selection can maintain sexual reproduction in the host population. Here, we describe a few important criteria. First, hosts and parasites must coevolve (?). Host-parasite coevolution creates a time-lagged selective advantage for rare host genotypes, creating an oscillation in genotypic frequencies (???); this process promotes host genetic diversity and not necessarily sex (???). Second, parasites must be highly virulent (?). Although sexual and asexual hosts can coexist at intermediate virulence, sexual reproduction is unable to provide enough advantage to overcome the cost of sex against avirulent parasites (?). Finally, sexual hosts must be genetically more diverse than asexual hosts, as high clonal diversity may diminish the advantage of sexual reproduction (???).

Some theoretical studies have departed from the classical population genetics framework to study the effects of ecological and epidemiological context on the Red Queen dynamics. A few studies have suggested that incorporating more realistic ecological and epidemiological details may assist in supporting sexual reproduction in the host population (??) and maintaining coevolutionary cycles (?). In contrast, ? showed that Red Queen dynamics (i.e., cycles in allele frequencies) can dampen when an explicit epidemiological structure is taken into account with coevolutionary dynamics. Given the wide range of assumptions underlying these eco-evolutionary models, it may be difficult to understand how different ecological details of these models affect the overall maintenance of sexual reproduction; however, it is clear is that these details play important roles in shaping coevolutionary dynamics (???).

Many empirical studies have focused on confirming predictions that stem from the RQH. Typical among them are rapid parasite evolution (?), local adaptation (?????), time-lagged selection (??????), and association between parasite prevalence and host reproductive mode (???). A key example is the snail population in New Zealand that serves as intermediate hosts for trematodes (?). Through decades of work, Lively *et al.* have demonstrated that the population satisfies necessary conditions for the host-parasite coevolutionary dynamics and provides support for the hypothesis (e.g., ???????). While field studies often provide only indirect evidence for the hypothesis, experimental systems can be used to directly test the hypothesis (????).

Even though the Red Queen Hypothesis has gained both theoretical and empirical support, there still remains a gap between theory and data. Many theoretical studies rely on simple models that are not applicable to natural populations and make indirect connections to empirical work. For example, none of the Red Queen models reviewed by ? make statistical connections to empirical data. It is unclear how well these models capture coevolutionary dynamics in natural systems.

Simple models can be still useful for making *qualitative* predictions about

nature. ? initially postulated that infection prevalence should be positively correlated with the frequency of sexual hosts and later formalized the idea using a mathematical model (?). The prediction has since been confirmed in several empirical studies, most of which are based on a snail-trematode system from New Zealand (?????); these New Zealand freshwater snails (*Potamopyrgus antipodarum*), which serve as intermediate hosts for several trematode species (e.g., *Microphallus* sp.), provide key evidence for the Red Queen Hypothesis (?????????). Surprisingly, the predicted correlation was not observed in another snail-trematode system in Israel (?????) even though it appears to demonstrate some key features of the RQH such as virulent parasites (?) and parasite-driven genetic diversity (?).

[SWP: Cite Dagan 2013 somewhere and mention that parasites reproduce parthenogenically but *Microphallus* is not...?]

Here, we try to confront simple models with data. We extend the model used by ? to account for demographic stochasticity and include a simple metapopulation structure. Then, we fit the extended model to data sets from three field studies (???) using Approximate Bayesian Computation (ABC) to estimate biologically realistic parameters, which capture the observed behaviours of real empirical systems, and test whether a simple coevolutionary model can sufficiently explain the data. Using estimated parameters, we test for power (probability of observing a significant effect) to detect a positive correlation between frequency of sexual reproduction and prevalence of infection ?. We provide practical guidance for studying the RQH for modelers and empiricists by discussing discrepancies between a model predictions and the observed data as well as underlying factors that drive the correlation.

2 Methods

2.1 Data

We analyze three data sets from field studies (???) that represent interactions between freshwater snails and sterilizing trematodes. The New Zealand snail populations (??) consist of *Potamopyrgus antipodarum*. These snails serve as an intermediate host for trematodes, among which *Microphallus* sp. is most commonly found (??). The Israeli snail populations (?) consist of *Melanoides tuberculata*, which also serves as an intermediate host for trematodes, including *Centrocestus* sp. (?), *Philophthalmus* sp. (?), and *Transversotrema patialense* (?). Diploid sexual females and polyploid asexual females coexist in both the New Zealand (???) and the Israeli (?) snail populations.

[SWP: CITE] *Microphallus* and *Centrocestus* share roughly similar life cycles. These trematodes reproduce asexually within the snail and sterilize its host. Once infected snails are consumed by waterfowls, trematodes reproduce sexually inside the intestines of ducks and release their eggs through their host's feces. These eggs can either hatch into miracidia that can penetrate snail tissue or be ingested by snails, thus completing their life cycles.

We choose to analyze snail-trematode systems because they have been extensively studied under the context of the Red Queen Hypothesis. The snail-trematode systems in New Zealand exhibit positive correlation between prevalence of infection and frequency of sexual hosts (????); local adaptation (???); negative frequency-dependent selection (????); and periodic selection for sexual reproduction (??). These results are consistent with the host-parasite coevolutionary theory behind the Red Queen Hypothesis. Therefore, we expect a basic Red Queen model to be able to mimic the observed dynamics reasonably well.

We included the Israeli population because there has been a consistent lack of positive correlation between the prevalence of infection and frequency of sexual hosts (????), which sharply contrasts the observations from the New Zealand population. We wanted to test whether a simple host-parasite coevolutionary model can explain the patterns observed in the data.

The data sets include proportions of infected snails, proportions of sexual/asexual snails, and locations of sampling sites. The data set from ? also includes sampling year (for 5 years). Whereas ? report infection status of snails specific to *Microphallus* sp., ? and ? do not distinguish among different species; we use infection by *Microphallus* sp. when we analyze ? but use total infection for two other studies. The data sets collected by ? and ? are obtained from their Dryad repositories (??). The data set collected by ? is extracted from their figure.

2.2 Model

We model obligately sexual hosts competing with obligately asexual hosts in a metapopulation by extending the model introduced by ?. Our model is a discrete time susceptible-infected (SI) model with natural mortality and virulence (defined as a reduction in offspring production among infected hosts). Metapopulation structure is included to model unobserved dynamics among different habitats; equivalently, each subpopulation can be considered as a potential sampling site. A similar metapopulation model was developed by ? in order to study the local adaptation of parasites.

We do not model the life history of snail-trematode interactions explicitly for two reasons. First, many theoretical studies for the RQH (see Introduction) rely on simple models that do not incorporate realistic life cycles of natural systems; we want to test whether these simple models can sufficiently explain the observed patterns in the data. Second, testing a basic model against data can help us identify model assumptions that drive the difference between theoretical predictions and data more clearly (rather than trying to understand how each component of the complex life cycle of trematodes affect the evolutionary outcome).

All hosts are diploids with two biallelic loci controlling resistance; parasites are assumed to be haploids with two biallelic loci. Let $S_{ij}^k(t)$ and $A_{ij}^k(t)$ be the number of sexual and asexual hosts with genotype ij from subpopulation k at generation t , where the subscripts, i and j , represents host haplotypes: $i, j \in \{AB, Ab, aB, ab\}$. For simplicity, we drop the superscript representing

the subpopulation and write $S_{ij}(t)$ and $A_{ij}(t)$; every subpopulation is governed by the same set of equations unless noted otherwise (e.g., when we account for the interaction between populations). Following ?, the expected genotypic contribution (before recombination or outcrossing) by sexual hosts is given by

$$S'_{ij} = c_b(1 - s) \left(W_U S_{ij,U}(t) + W_I S_{ij,I}(t) \right), \quad (1)$$

where s is the proportion of males produced by sexual hosts (assumed to equal to 0.5), and $S_{ij,U}$ and $S_{ij,I}$ are the number of uninfected and infected sexual hosts in a subpopulation. W_U and W_I represent their corresponding fitnesses where virulence is defined as $V = 1 - W_I/W_U$. We allow for the cost of sex to vary by multiplying the growth rate by a scale parameter, c_b , where $2/c_b$ corresponds to a two-fold cost of sex when $c_b = 1$ (?). We discuss this parameter in detail in the Discussion. Recombination and outcrossing are modeled after incorporating genotypic contributions from other populations.

We define

$$W_U = \frac{b_U}{1 + aN(t)}, \quad W_I = \frac{b_I}{1 + aN(t)}$$

where b_U and b_I are the maximum number of offspring produced by uninfected and infected hosts, respectively, and a is the strength of density dependence (??). Following ?, virulence is defined in terms of decrease in offspring production: $V = 1 - b_I/b_U$. While many trematode species (including *Microphallus* sp.) sterilize upon infection (which would correspond to $V = 1$), we allow virulence to vary in order to test whether high virulence is necessary to explain the data.

Asexual hosts are strictly clonal. The expected genotypic contribution by asexual hosts is given by

$$A'_{ij} = W_U A_{ij,U}(t) + W_I A_{ij,I}(t), \quad (2)$$

where $A_{ij,U}$ and $A_{ij,I}$ are the number of uninfected and infected asexual hosts in a subpopulation.

We assume that a proportion ϵ_{mix} of hosts within a subpopulation can move to other subpopulations every generation. The expected number of offspring in the next generation (accounting for contributions from all subpopulations) is given by

$$\begin{aligned} E(S_{ij}^k(t+1)) &= f_{\text{sex}} \left((1 - \epsilon_{\text{mix}}) (S_{ij}^k)^' + \frac{\epsilon_{\text{mix}}}{n_{\text{pop}} - 1} \sum_{h \neq k} (S_{ij}^h)^' \right), \\ E(A_{ij}^k(t+1)) &= (1 - \epsilon_{\text{mix}}) (A_{ij}^k)^' + \frac{\epsilon_{\text{mix}}}{n_{\text{pop}} - 1} \sum_{h \neq k} (A_{ij}^h)^', \end{aligned} \quad (3)$$

where $f_{\text{sex}}(x)$ is the function that models sexual reproduction, including recombination probability r_{host} and outcrossing, and n_{pop} is the number of subpopulations modeled. Here, we assume that migration among subpopulations occur after reproduction and selection. The ordering of such events could affect the

evolutionary outcome of the model (??); for simplicity, we do not explore these directions in this paper.

The total number of sexual and asexual hosts in the next generation is given by Poisson random variables with means specified previously. We also allow for stochastic migration (from outside of the metapopulation) in order to avoid fixation:

$$\begin{aligned} S_{ij}^k(t+1) &\sim \text{Poisson}(\lambda = E(S_{ij}^k(t+1))) + \text{Bernoulli}(p = p_{ij,\text{sex}}), \\ A_{ij}^k(t+1) &\sim \text{Poisson}(\lambda = E(A_{ij}^k(t+1))) + \text{Bernoulli}(p = p_{ij,\text{asex}}), \end{aligned} \quad (4)$$

where $p_{ij,\text{sex}}$ and $p_{ij,\text{asex}}$ are the probabilities of a single sexual or asexual host with genotype ij entering a subpopulation.

Infection is modeled using the matching alleles model (??). We assume that snails are equally susceptible to parasites that match either haplotype. However, parasites must match the host haplotype at both loci in order to successfully infect a host. The expected number of infected hosts that are infected with parasite with genotype i at generation t is given by:

$$I_i(t) = f_{\text{mutation}} \left(\sum_j \left(S_{ij,i,I}(t) + A_{ij,i,I}(t) \right) \right). \quad (5)$$

$S_{ij,i,I}(t)$ and $A_{ij,i,I}(t)$ represent the expected numbers of sexual and asexual hosts that have genotype ij and are infected with parasites that have genotype i . The function f_{mutation} models mutation at a single locus with probability r_{parasite} (?). We further account for stochastic migration with probability $p_{i,\text{parasite}}$ to avoid fixation:

$$I'_i(t) \sim I_i(t) + \text{Bernoulli}(p = p_{ij,\text{parasite}}). \quad (6)$$

The total expected number of infectious contacts made by infected hosts within a population is given by $\lambda_i^k = \beta^k I'_i{}^k(t)$, where β^k is the transmission rate of each population. We model mixing between subpopulations by allowing infected hosts to make contact with susceptible hosts in other populations; we assume that a proportion ϵ_{mix} of the infectious contacts are equally distributed among other subpopulations. The total amount of infectious contact, coming from hosts that carry genotype i parasite, that is received by susceptible hosts in subpopulation k is given by

$$\lambda_{i,\text{total}}^k = (1 - \epsilon_{\text{mix}}) \lambda_i^k + \frac{\epsilon_{\text{mix}}}{n_{\text{pop}} - 1} \sum_{l \neq k} \lambda_i^l \quad (7)$$

Then, the force of infection that a susceptible host with genotype ij experiences in generation $t+1$ is given by

$$\text{FOI}_{ij}^k = \frac{\lambda_{i,\text{total}}^k + \lambda_{j,\text{total}}^k}{2N^k(t+1)}, \quad (8)$$

where $N^k(t+1) = \sum_{i,j} S_{ij}^k(t+1) + A_{ij}^k(t+1)$ is the total number of hosts in generation $t+1$. The probability that a susceptible host with genotype ij in subpopulation k becomes infected in the next generation is given by

$$P_{ij}^k(t+1) = 1 - \exp\left(-\text{FOI}_{ij}^k\right). \quad (9)$$

Finally, the number of infected hosts in the next generation is determined by a binomial process:

$$\begin{aligned} S_{ij,I}^k(t+1) &\sim \text{Binom}\left(S_{ij}^k(t+1), P_{ij}^k(t+1)\right), \\ A_{ij,I}^k(t+1) &\sim \text{Binom}\left(A_{ij}^k(t+1), P_{ij}^k(t+1)\right). \end{aligned} \quad (10)$$

The expected number of infected hosts that have genotype ij and are infected by parasites with genotype i in the next generation is proportional to the amount of infectious contact that was made in the current generation:

$$\begin{aligned} S_{ij,I}^k(t+1) &= \frac{2^{\delta_{ij}} \lambda_{i,\text{total}}^k}{\lambda_{i,\text{total}}^k + \lambda_{j,\text{total}}^k} S_{ij,I}^k(t+1) \\ A_{ij,I}^k(t+1) &= \frac{2^{\delta_{ij}} \lambda_{i,\text{total}}^k}{\lambda_{i,\text{total}}^k + \lambda_{j,\text{total}}^k} A_{ij,I}^k(t+1) \end{aligned} \quad (11)$$

where a δ_{ij} is a Kronecker delta (δ_{ij} equals 1 when $i = j$ and 0 otherwise).

2.3 Simulation design and parameterization

Many Red Queen models have focused on competition between a single asexual genotype and multiple sexual genotypes or have assumed equal genetic diversity between asexual and sexual hosts (see ? for a review of previous Red Queen models) but neither of these assumptions is realistic; asexual hosts should be able to maintain some genetic diversity (e.g., (?)) but are expected to be less diverse than their sexual counterparts. In contrast, ? adopted a more realistic approach by incorporating stochastic external migration of an asexual genotype to a population, which allows for asexual genetic diversity to vary over time.

Here, we try to combine these methods. We allow for stochastic external migration of asexual hosts with different genotypes into the system but assume that the total number of asexual genotypes that can be present in the system (denoted by G_{asex}) is fixed throughout a simulation. Since sexual hosts have limited genetic diversity in our model (diploid hosts with two biallelic loci yields a total of only 10 genotypes), allowing for unlimited migration of asexual hosts will cause the sexual hosts to be easily outcompeted by the asexual hosts. Limiting asexual genetic diversity allows us to account for the intrinsic difference in sexual and asexual diversity, while allowing for some genetic diversity among asexual hosts. At the beginning of each simulation, a pool of G_{asex} asexual genotypes are sampled at random from the entire genotypic space; these are the genotypes that are available for external immigration into subpopulations. Then, we use our model-based comparison with data to estimate G_{asex} to test

whether the data are consistent with the dynamics of a model that allows for more than a single asexual clone.

To account for differing number of sexual and asexual genotypes, we let

$$p_{ij,\text{sex}} = 1 - (1 - p_{\text{host}})^{1/G_{\text{sex}}},$$

$$p_{ij,\text{asex}} = \begin{cases} 1 - (1 - p_{\text{host}})^{1/G_{\text{asex}}} & \text{if } ij \in \{\text{asexual genotypes}\} \\ 0 & \text{otherwise} \end{cases}, \quad (12)$$

where p_{host} is the probability that at least one sexual or asexual host enters a subpopulation in a generation. We scale the probability of an infected host carrying parasite genotype i in a similar way for interpretability:

$$p_{i,\text{parasite}} = 1 - (1 - p_{\text{infected}})^{1/G_{\text{parasite}}}, \quad (13)$$

where p_{infected} is the probability that at least one infected host enters a subpopulation in a generation. The number of parasite genotypes G_{parasite} is equal to 4 (because parasites are assumed to be haploids with two loci).

Each simulation consists of 40 subpopulations. Every subpopulation is initialized with 2000 sexual hosts, of which 80 are infected. They are assumed to be in Hardy-Weinberg equilibrium where the allele frequency in each locus is equal to 0.5. In order to account for variation in infection rates among subpopulations, the transmission rate, β^k , is randomly drawn for each subpopulation from a Gamma distribution with mean β_{mean} and coefficient of variation β_{cv} . The simulation runs for 500 generations without the introduction of asexuals. At generation 501, 10 asexual hosts of a single genotype are introduced to each subpopulation (the asexual genotype introduced can vary across subpopulations) and simulation runs for 600 more generations while allowing for stochastic migration of asexual hosts.

2.4 Approximate Bayesian Computation

We use Approximate Bayesian Computation (ABC) to estimate parameters from data (?). ABC relies on comparing summary statistics of observed data with those of simulated data; it is particularly useful when the exact likelihood function is not available. Each summary statistic (also referred to as a *probe*) represents an aspect of a system (?); probe matching can be more powerful than classical likelihood-based approaches because it allows to evaluate model fits based on certain aspects of a biological system that we are interested in (?). We consider the mean proportion of infected and sexually reproducing snails in the system and variation in these proportions — measured by a coefficient of variation (CV) — across space (subpopulations) and time (generations) as our probes. As ? and ? only reported the proportion of males, the proportion of sexual hosts is assumed to be twice the proportion of males. **[LIVELY: This may not be a safe assumption in *Melaniodes*. There is a positive correlation between male frequency and the confirmed frequency of diploid sexual females, but there is quite a bit of scatter, probably because males are smaller than females.**

They may also have shorter lifespans, but we don't have data on this.] [SWP: I'm not sure how I should address this point.]

CV across space is calculated by first calculating mean proportions by averaging across time (generation) for each site (subpopulation) and then taking the CV of these mean proportions. CV across time (generation) is calculated by first averaging proportions across space (subpopulation) at each generation and then taking the CV. For purely spatial data (? and ?), CV across space is calculated without averaging across time.

Because ABC is a Bayesian method, we must specify prior distributions for all parameters. We use weakly informative priors for all estimated parameters except c_b , a scale parameter for the cost of sex (see Table 1 for prior distributions used and parameters assumed). The prior distribution for the scale parameter is chosen so that 95% prior quantile range of cost of sex ($2/c_b$) is approximately equal to the 95% confidence interval reported by ?. For brevity, all other parameters are assumed to be fixed. While it is a common practice to fix several parameters of a Red Queen model (e.g., ????), failing to explicitly account for uncertainty in all parameters of a model can lead to overconfident conclusions (?).

We use the Population Monte Carlo approach (?), which allows for efficient sampling while ensuring that final result still converges to a correct (approximate) Bayesian posterior. The Population Monte Carlo approach begins with a basic ABC. For each random parameter sample drawn from the prior distribution, the model is simulated and a sample of a subpopulation is drawn from the simulated system, equal to the number of sites collected in a study. Summary statistics are calculated based on the last 100 generations out of 1100 generations; the current parameter set is accepted if the difference between a simulation and the observed data is less than a pre-specified tolerance value (i.e., acceptable amount of difference between the model and the data). We measure the difference between the model and the data by the sum of absolute differences in summary statistics. This process is repeated until 100 parameter sets are accepted.

After the first run ($t = 1$), equal weights ($w_{i,1} = 1/100$) are assigned to each accepted parameter set $\theta_{i,1}$ ($i = 1, 2, \dots, 100$). A weighted random sample (θ^*) is drawn from the accepted parameters of the previous run ($t - 1$) with weights $w_{i,t-1}$. For any run $t > 1$, a parameter sample ($\theta_{i,t}$) is proposed from a multivariate normal distribution with a mean θ^* and a variance covariance matrix that is equal to $\sigma_{t-1}^2 = 2\text{Var}(\theta_{1:N,t-1})$, where $\text{Var}(\theta_{1:N,t-1})$ is the weighted variance covariance matrix of the accepted parameters from the previous run and N is the total number of accepted parameters from the previous run.

The parameter G_{asex} is rounded to the nearest integer and the model is simulated. If a proposed parameter is accepted, the following weight is assigned:

$$w_{i,t} = \frac{\pi(\theta_{i,t})}{\sum_{j=1}^{100} w_{j,t-1} q(\theta_{j,t-1} | \theta_{i,t}, \sigma_{t-1}^2)}$$

where $\pi(\cdot)$ is a prior density and $q(\cdot | \theta_{i,t}, \sigma_{t-1}^2)$ is a multivariate normal den-

Notation	Description	Prior distribution/parameter values	Source
β_{mean}	Mean transmission rate	Gamma($k = 2, \theta = 10$)	Assumption
β_{CV}	CV transmission rate	Gamma($k = 2, \theta = 0.5$)	Assumption
V	Virulence	Beta($\alpha = 6, \beta = 2$)	Assumption
ϵ_{mix}	Mixing proportion	Beta($\alpha = 1, \beta = 9$)	Assumption
G_{asex}	Number of asexual genotypes	$1 + \text{BetaBinomial}(N = 9, p = 3/9, \theta = 5)$	Assumption
c_b	Cost of sex scale	LogNormal($\mu = -0.07, \sigma = 0.09$)	?
s	Proportion of male offspring produced	0.5	Assumption
b_U	Number of offspring produced by an uninfected host	20	?
b_I	Number of offspring produced by an infected host	$(1 - V)b_U$?
a_U	Density dependent effect coefficient of uninfected hosts	0.001	?
a_U	Density dependent effect coefficient of infected hosts	0.001	?
r_{host}	Host recombination probability	0.2	?
r_{parasite}	Parasite mutation probability	0.05	Assumption
p_{host}	Probability that at least one sexual and asexual host enters the population	0.1	Assumption
p_{infected}	Probability that at least one infected host enters the population	0.02	Assumption

Table 1: **Parameter descriptions and values.** Parameters with prior distributions are estimated via Approximate Bayesian Computation (ABC). Parameters k and θ in Gamma distribution represent shape and scale parameters where mean and squared CV are given by $k\theta$ and $1/k$, respectively. α and β in Beta distribution represent shape parameters where mean and squared CV are given by $\alpha/(\alpha + \beta)$ and $\beta/(\alpha^2 + \alpha\beta + \alpha)$. N , p and θ in Beta binomial distributions represent number of trials, probability of success, and overdispersion parameters (?). Parameters μ and σ in a log-normal distribution represent mean and standard deviation on a log scale. All other parameters are fixed throughout simulations.

sity with mean $\theta_{i,t}$ and variance covariance matrix σ_{t-1}^2 . For each run, 100 parameters are accepted and weights are normalized at the end to sum to 1.

For each observed data set, we perform 4 runs with decreasing tolerance. First three tolerance values are chosen in a decreasing sequence to reach the final step quicker. The tolerance value of the final run is chosen so that a parameter set will be accepted if its each simulated summary statistic deviates from the corresponding observed summary statistic by 0.1 units on average. For spatial data (??), four summary statistics are compared: the mean proportion of infected and sexually reproducing hosts and CV in these proportions across populations. Tolerance values of 1.6, 0.8, 0.6 and 0.4 are used for each run. For spatiotemporal data (?), we additionally compare the CV of proportions of infected and sexually reproducing hosts across generations. In this case, larger tolerance values (2.4, 1.2, 0.9 and 0.6) are used for each run to account for a higher number of summary statistics being compared.

All statistical results are weighted by parameter weights of the final run; confidence intervals are obtained by taking weighted quantiles. However, it is always not possible to obtain the exact 95% confidence interval when the smallest (or the largest) value has a weight greater than 0.025. In these cases, we take the lowest (or largest) possible quantile that is closest to the 2.5% (or the 97.5%) quantile. While this procedure results in a slightly narrower confidence interval, the rounding errors in quantiles are negligible. This rule leads to one instance where we use the 2.82% quantile instead of the 2.5% quantile as the lower end of a confidence interval.

2.5 Power analysis

[LIVELY: *I think for the NZ snail, it is probably about one generation per year in the field. It is shorted in the lab, as there are no winters. Also, my guess is that there is some overlap of generations in the field, which can reduced the cost of males per reproductive bout. see Lively, C. M. 2010. Parasite virulence, host life history, and the costs and benefits of sex. Ecology 91: 3-6.]* **[SWP:** *Tried to address this without redoing the analysis*]**]** Using estimated parameters for each data set, we calculate the power to detect a significant positive correlation between infection prevalence and frequency of sexual hosts. For each parameter sample from the final run of the ABC, 10 simulations are run. For each simulation, we take the last two generations — assuming that a year contains two snail generations (estimated generation time is 4-9 months (?)) — from the simulation and choose n subpopulations at random from 40 simulated subpopulations. We might expect New Zealand freshwater snails to have a slightly longer generation time (approximately one year) with overlapping generations (Lively, personal communication); these factors can reduce the cost of males per reproductive time step (?). However, we expect these factors to have little effect on the overall estimates of the power as well as our qualitative conclusions of the analysis.

For each selected subpopulation, hosts are divided into four categories based on their infection status (infected/uninfected) and reproductive mode (asex-

ual/sexual), and the mean proportion of hosts in each category is calculated by averaging over two generations. Independent multinomial samples of size m are drawn from each selected population based on the proportions in every four categories. Correlation between the proportion of infected hosts and the proportion of sexual hosts is tested using the Spearman’s rank correlation at a 5% significance level.

While we do not explicitly try to match observed correlations, this procedure is equivalent to a retrospective power analysis. The main purpose of a power analysis is to design an experiment and determine an appropriate sample size (?). If instead, power is calculated based on an observed effect size after an experiment, the estimated power is likely to be correlated with the observed significance of the result; therefore, interpreting significance (or a lack of significance) of a statistical result based on a retrospective power analysis is inappropriate (??).

Here, we do not try to justify the significant correlation observed by (?) or the nonsignificant correlation observed by (?) based on this power analysis. Instead, we use a power analysis to better understand how a study design or the dynamics of the model might affect the correlation between the prevalence of infection and sexual reproduction while using biologically realistic parameters.

3 Results

[REV: I had difficulty interpreting the figures. I recommend the authors go into more detail in the results section and in figure legends.] [SWP: I tried.]

First, we compare observed summary statistics (calculated from the observed data) with either fitted or predicted summary statistics (Fig. 1). *Fitted* summary statistics (Fig. 1) are values that have been accepted by the ABC procedure as being sufficiently close to observed summary statistics; these can be interpreted as model-based estimates of the summary statistics (and associated confidence intervals) of the study sites. *Predicted* summary statistics, in contrast (Fig. 2), use parameters from the posterior distribution sampled by ABC, but generate new simulations of the dynamics and calculate summary statistics from randomly selected subpopulations. These values represent summary statistics sampled from the full distribution of coevolutionary dynamics that are consistent with the observed dynamics; they represent the expected range of dynamics from randomly sampled host-parasite metapopulations with a given distribution of parameters. Because they include an additional level of dynamical and sampling variability, the predicted summary statistics range much more widely than the fitted summary statistics. The difference between the fitted and predicted summary statistics is analogous to the difference between confidence intervals and prediction intervals from linear regression.

Our simple metapopulation Red Queen model can capture observed variation in infection prevalence and frequency of sexual hosts reasonably well; both temporal and spatial variation (measured by CV across mean proportions) are well-matched by the model. However, as model fitting is performed by mini-

mizing the sum of absolute distance between observed and simulated summary statistics, our method does not guarantee that all summary statistics are equally well-fitted. The model tends to overestimate the mean proportion of infected hosts. The observed mean proportions of infected hosts are 0.175 (?), 0.051 (?), and 0.440 (?), whereas their ABC-based estimates are 0.240 (95% CI: 0.174 - 0.286), 0.313 (95% CI: 0.233 - 0.404), and 0.542 (95% CI: 0.360 - 0.730), respectively. The model underestimates the mean proportion of sexual hosts for ? and ?. Observed mean proportions of sexual hosts are 0.045 and 0.704, respectively, whereas their ABC-based estimates are 0.026 (95% CI: 0.007 - 0.048) and 0.596 (95% CI: 0.441 - 0.679).

To further diagnose the fit, we compare the predicted relationships between the mean proportion of infected hosts and the mean proportion of sexual hosts across subpopulations with the observed data (Fig. 2). This is done by simulating the metapopulation model multiple times from the estimated parameters and calculating the mean proportion of infected and sexual hosts for each subpopulation from each simulation. Then, a two dimensional distribution of these means is compared with the observed proportions. If our model is able to sufficiently capture the observed dynamics, we would expect the observed data points to fall within high density regions. Since we plot the *predicted* densities, they are able to capture the full range of coevolutionary dynamics, which may be only partially represented in the observed data.

Despite its accuracy in reproducing summary statistics reported by ?, our model poorly captures the relationship between the mean proportion of infected hosts and the mean proportion of sexual hosts (Fig. 2; ?). The model predicts sexual reproduction to be well maintained when infection prevalence is high (> 40%) whereas the observed data (?) suggests that sexual reproduction can only be supported when infection prevalence is low (< 20%). ? also found sexually reproducing snails in sites with low infection prevalence (< 20%); we note that infection prevalence may underrepresent the strength of parasite-mediated selection as infected *P. antipodarum* experiences higher risk of predation from ducks to complete the life cycle of *Microphallus* (?). [SWP: added the last sentence as suggested by Lively]

On the other hand, ?'s (?) data set suggests that sexually reproducing subpopulations are likely to have a relatively high prevalence of infection across several years (approximately 20% - 80%). This observation is broadly consistent with our model prediction – most of the data points fall within the range of model predictions (Fig. 2; ?). However, there is one site in which more than 90% of the snails were found to be sexual throughout the study period of 5 years (?); our model fails to maintain such high levels of sexual reproduction.

There is a high posterior density region in which the proportion of infected hosts remains almost constant (around 0.5) among subpopulations while the proportion of sexual hosts can range from 0 to 0.3 (visible in the fits to ?). As transmission rate (β) increases, selection for sexual hosts increases but the increasing number of resistant offspring prevents further infection from occurring and can decrease overall infection prevalence. This trend is consistent with previous results of ? who noted that there is a region in which either sexual

and asexual reproduction can be selected exclusively under the same infection prevalence.

Our model fit to ? predicts that the proportion of sexual hosts will decrease as infection prevalence becomes extremely high (see the shape of the contour lines in Fig. 2: ?). This pattern can be explained by the decrease in fitness of sexual hosts with the increase in infection prevalence, predicted by ?. The same pattern can be seen in earlier work by ?.

Fig. 3 presents parameter estimates. Keeping in mind that we do not obtain good fits to data from ? and ?, we still find that high virulence and a low ratio of asexual to sexual genetic diversity are necessary to explain the observed dynamics. Moreover, we are able to capture observed differences in mean and variation in infection prevalence among studies in our estimates of transmission rate parameters (β_{mean} and β_{CV}).

Our fits to ? suggest that the scale parameter for the cost of sex, c_b , should be higher than our prior assumption based on ? that estimated the cost of sex to be 2.14 (95% CI: 1.81 - 2.55). ? defined c_b as additional costs and benefits of sex, where $c_b = 1$ corresponds to the two fold cost. Under their interpretation, our estimate of c_b corresponds to a slightly lower estimate of the cost of sex: 1.95 (95% CI: 1.68 - 2.4). (We propose an alternate interpretation to this parameter estimate below).

Finally, our power analyses suggest that there is high power to detect a positive correlation between infection prevalence and frequency of sexual hosts in the systems studied by ? and ? (Fig. 4). Such high power predicted for ? is particularly surprising given that they were not able to observe the expected correlation. This discrepancy implies that the snail-trematode system studied by ? show sufficient variation in infection prevalence in order for the correlation to be observed under pure Red queen selection, but other underlying factors that are neglected by our model may have caused the system to deviate from their expected behaviors (predicted by Red Queen models); in particular, only 2 out of 22 populations contains sexual snails. On the other hand, our model predicts low power for detecting the positive correlation for the system studied by ? (Fig. 4).

Overall, our analysis suggests that increasing the number of study sites is a more effective way to increase power than increasing the number of samples per site (Fig. 4). The correlation between the proportion of infected and sexual hosts depends on the spatial distribution of parasite infection; increasing the number of sites can help us better capture this distribution and increase the power to observe the correlation. On the other hand, collecting more samples from a site only gives us a more accurate estimate of the two proportions for that specific site; it does not tell us how these proportions are distributed across space. As a result, the power quickly saturates once we have a sufficient number of samples (> 100) to estimate the two proportions reliably at each site.

While we originally planned to perform power analysis using Spearman's rank correlation, we repeated the analysis using Pearson's correlation after applying arcsine square root transformation (?) to see whether this procedure improves power. Using Pearson correlation (with transformation) gives slightly

higher power to detect the positive correlation between frequency of sexual hosts and infection prevalence (see Appendix).

4 Discussion

[REV: remind the reader of the biologically meaningful parameters.] [SWP: I tried to reword this paragraph] We compared summary statistics (mean and coefficient of variation in proportion of infected and sexual hosts) of model simulations and those of the observed data using Approximate Bayesian Computation (ABC) to infer biologically realistic range of model parameters, which reflects the observed Red Queen dynamics (Fig. 3). Using the estimated parameters, we calculated the power to observe a significant, positive correlation between the proportion of infected hosts and the proportion of sexual hosts (Fig. 4), as predicted by ?. While our simple metapopulation coevolutionary model is able to match the observed prevalence of sexual reproduction and trematode infection across multiple snail populations (Fig. 1), it fails to capture some of the more detailed patterns in the data (Fig. 2); in particular, our model (1) overestimates the required level of infection prevalence to promote sexual reproduction and (2) is unable to match low levels of sexual reproduction observed by ?. These discrepancies between the model predictions and observed data suggest that a simple host-parasite coevolutionary model may not be able to sufficiently explain the maintenance of sexual reproduction observed in nature (Fig. 2). [LIVELY: This is not clear to me. The fit seems good for *Vergara et al.*, but not for the Israeli snail. So it is not clear if you are speaking about the difference between studies, or some kind of average over all studies.] [SWP: Tried to clarify this.]

A model that fits poorly can still provide useful information about a biological system by challenging our model assumptions; for example, lack of fit in dynamical systems (modeled by ordinary differential equations) can suggest misspecification of a model on various levels, including rate equations governing the dynamics of a system and their relevant state variables (?). Our model clearly failed to match the data presented by ? (Fig. 2). The snail-trematode system studied by ? live in qualitatively different environments from the New Zealand system that we considered (??). For example, their habitats are subject to seasonal flash floods, which can act as strong bottlenecks and promote asexual reproduction (?). Because our model did not account for these details, it failed to match the observed relationship between infection prevalence and frequency of sexual reproduction; nonetheless, lack of fit of a simple Red Queen model to this data does not exclude the possibility that sexual reproduction is maintained by parasite selection in this system.

The model fits to ? and ? suggest that the cost of sex can be overcome and sexual reproduction can be maintained only if infection prevalence is much higher than the observed prevalence (Fig. 2). In other words, the benefit of producing offspring with novel genotype is relatively small when infection prevalence is low. In order to support sexual reproduction at lower infection prevalence

lence, the benefit of sex must be greater or other mechanisms must compensate for the difference. As our model relies on a simple structure and strong parametric assumptions, an additional benefit of sex can only be provided by lowering the cost of sex (i.e., increasing the scale parameter, c_b).

The simple structure of the model and limited genetic diversity can explain the discrepancy between model prediction and the observed data by ?. Here, we assumed that host resistance to infection is determined entirely by two biallelic loci, resulting in 10 possible genotypes. However, it is unlikely that such a simple model can capture the genetic interaction between hosts and parasites observed in nature. Although exact genetic architecture that determines trematode infection in snails (e.g., the number and inheritance of loci involved in parasite resistance) is not known, the documented genetic diversity of snails is far greater than what our model assumes (???). Increasing the maximum possible genetic diversity of the model would have allowed sexual hosts to escape infection more easily and maintained sexual reproduction at a lower prevalence of infection (???).

While the positive correlation between frequency of sexual hosts and prevalence of infection can provide evidence for the effect of the parasite on the maintenance of sexual reproduction in the host population, it does not fully capture the coevolutionary process. In particular, the positive correlation represents a contrast between populations that undergo Red Queen dynamics and those that do not (rather than a continuous relationship between the frequency of sexual hosts and prevalence of infection): low-risk populations will be dominated by asexual individuals whereas high-risk populations will consist of both sexual and asexual individuals. ? gave a similar explanation and predicted that large variation in the risk of infection is required to observe the predicted correlation.

The similarity between the correlations predicted by the model and the correlations observed in nature may provide us more confidence that the observed populations provide evidence for the Red Queen Hypothesis (Fig. 5). However, the strength of the correlation is not a direct measure of the strength of selection imposed on the host population. Moreover, when populations are actively coevolving, wide range of correlations can be detected (Fig. 5: ?). While simple statistical summaries, such as a correlation coefficient, are easily accessible, more sophisticated and mechanistic statistical models that yield biologically interpretable parameters may be better suited for understanding the effect of parasites on the maintenance of sexual reproduction in the host population.

Mathematical models have been used extensively to build theoretical foundations for the evolution of sex but only a few models have been confronted with data. The idea of the two-fold cost of sex was only recently tested directly by fitting a theoretical model to observed data (?). Making statistical inference on the observed systems and testing theoretical models against data may provide deeper insight into underpinnings of maintenance of sexual reproduction.

5 Acknowledgements

We thank SHARCnet for providing computational resources, and Curtis Lively, Yannis Michalakis, and an anonymous reviewer for helpful comments.

6 Funding

This work is supported by The Natural Sciences and Engineering Research Council, Undergraduate Student Research Award (to SWP).

Appendix

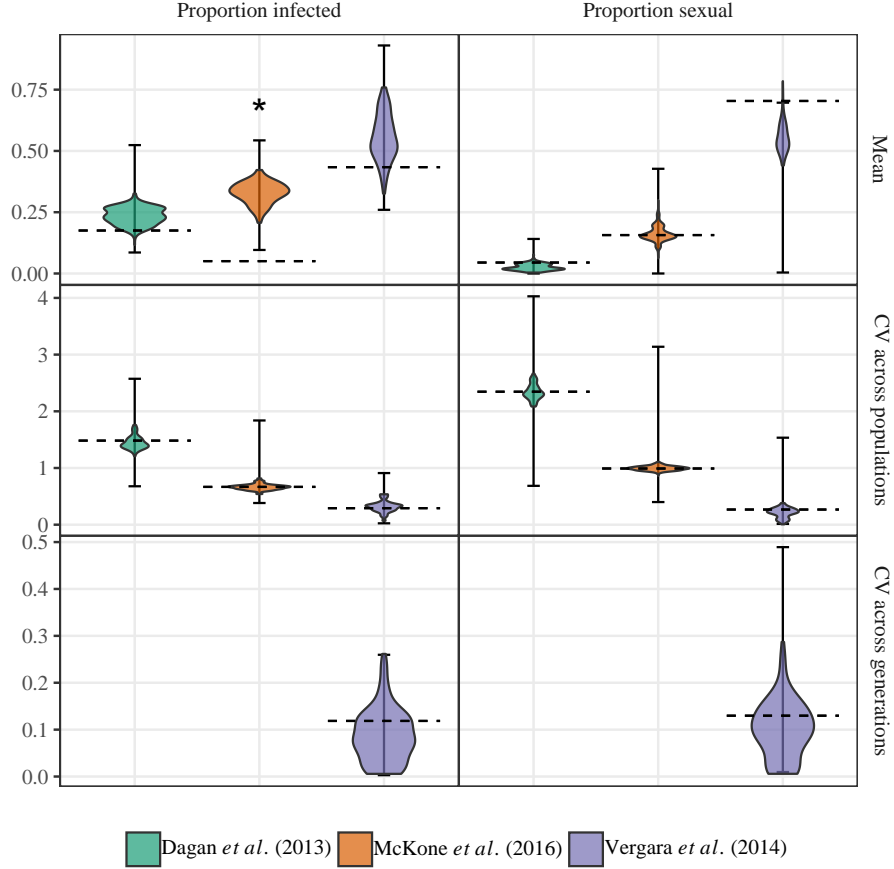


Figure 1: **Summary statistics of the observed data vs. distribution of summary statistics of the simulated data from the posterior samples.** Dotted horizontal lines represent observed summary statistics calculated from the data. Violin plots show the weighted distribution of fitted summary statistics (i.e., summary statistics that were accepted during ABC); they represent the uncertainty in the estimate of the true values. Error bars show 95% weighted quantiles of predicted summary statistics; they are wider than the distribution of the fitted values because they account for the uncertainty in the full distribution of the Red Queen dynamics from the estimated parameter values. The weights correspond to the posterior distribution weights from ABC. For each posterior sample, 10 simulations are run and each simulated system is sampled at random 100 times so that each sample consists of the same number of populations as in the fitted data. All univariate summary statistics are matched reasonably well, except for the mean proportion of infected hosts in ? (indicated by an asterisk).

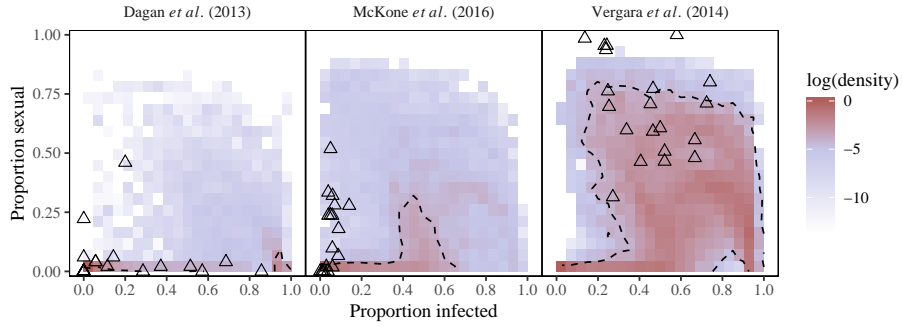


Figure 2: **Predicted relationship between mean infection prevalence and mean proportion of sexual hosts in each population.** For each posterior sample, 10 simulations are run. For each population within a simulation, mean infection prevalence and mean proportion of sexual hosts is calculated by averaging across last two generations (assuming that one year contains two snail generations (?)). Each point, consisting of mean infection prevalence and mean proportion of sexual hosts, is assigned the weight of the parameter used to simulated the population. The density at each cell is calculated as the sum of weights of the points that lie within it. Cell densities are normalized by dividing by the maximum grid densities for each fit. Dashed contours are calculated at $\log(\text{normalized density}) = -4$ to distinguish low- and high-density regions. Open triangles represent the observed data (proportion of sexual hosts is computed by doubling proportion of male hosts). For brevity, we compare data points from ?, which represent 5 years of Red Queen dynamics, with data points from simulations averaged across 2 last generations, which effectively represent 1 year of Red Queen dynamics; because we calculate the densities across 1000 simulations, we expect the distribution of the simulated points to sufficiently capture the range of spatiotemporal dynamics of the Red Queen cycles.

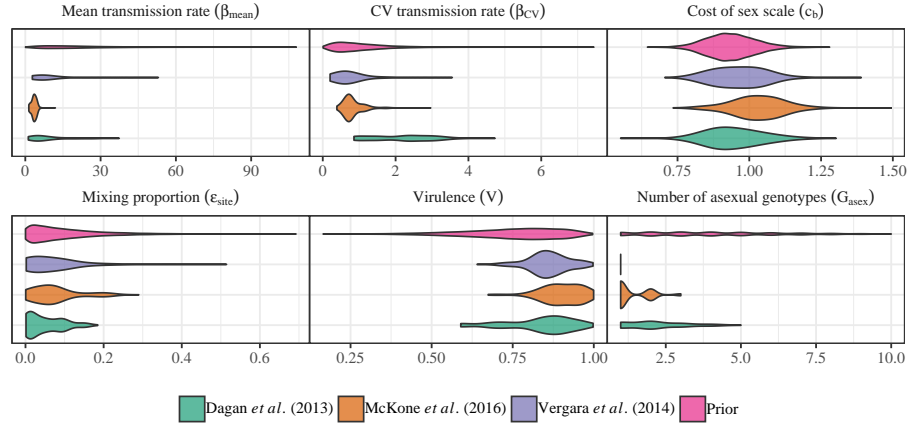


Figure 3: **Parameter estimates from Sequential Monte Carlo Approximate Bayesian Computation.** We estimate mean transmission rate (β_{mean}), CV in transmission rate (β_{cv}), virulence (V), mixing proportion (ϵ_{mean}), number of asexual genotypes (G_{asex}), and a scale parameter for cost of sex (c_b). Violin plots represent weighted distribution of 100 posterior samples obtained from ABC. Violin plots for prior distribution is obtained by drawing 10000 random parameter samples from the prior distribution. G_{asex} is a discrete variable but is drawn on a continuous scale for convenience. See Table 1 for full parameter descriptions and their prior distributions.

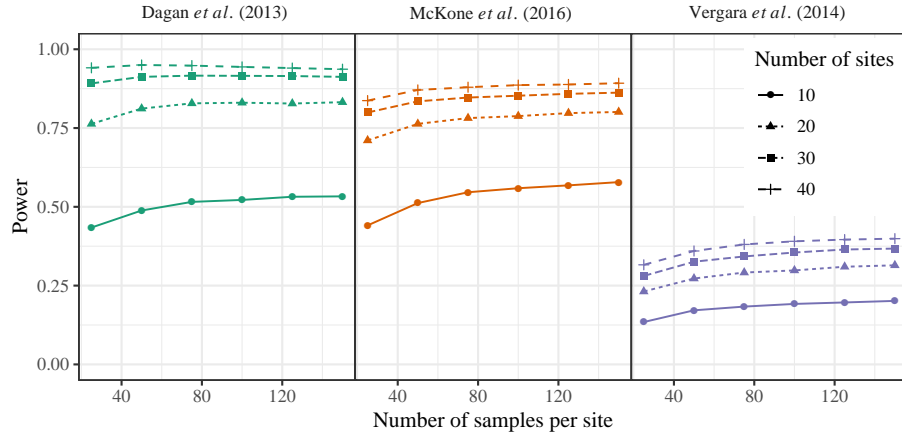


Figure 4: **Power to detect a statistically significant positive correlation between infection prevalence and frequency of sexual hosts.** Spearman's rank correlation was used to test for correlation between infection prevalence and frequency of sexual hosts in simulated data from the posterior distributions.

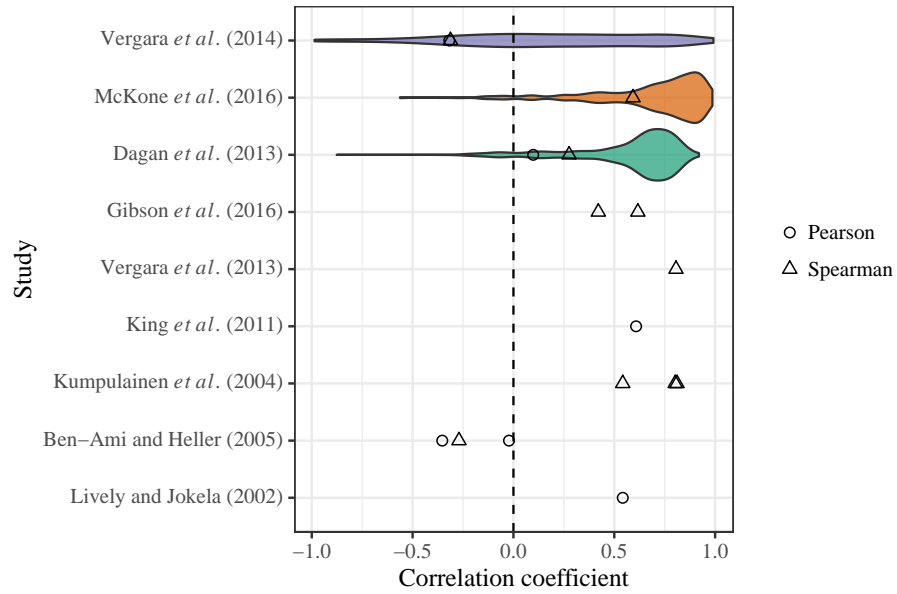


Figure 5: **Predicted vs. observed correlation between infection prevalence and frequency of sexual hosts.** Violin plots show weighted distribution of predicted strength of correlation based on ABC fits. Predicted correlation was measured for each simulation from the posterior by taking into account the last two generations from each simulated population. Triangles and circles represent observed correlation sizes from previous studies. ? and ? do not report the size of the correlation; correlations are re-calculated from their data sets using both Pearson (after applying arcsine square root transformation) and Spearman correlations.

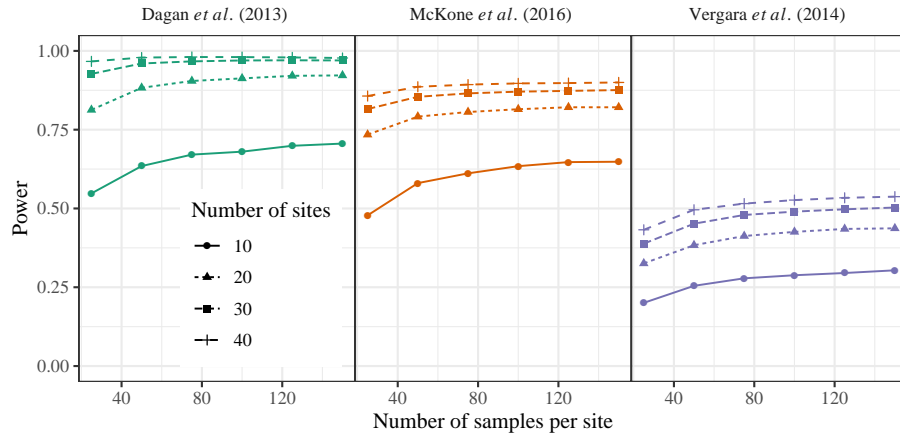


Figure A1: **Power to detect a statistically significant positive correlation between infection prevalence and frequency of sexual hosts.** Pearson correlation was used to test for correlation between square root arcsine transformed infection prevalence and frequency of sexual hosts in simulated data from the posterior distributions.