Quantifying the effects of parasites on the maintenance of sex

Sang Woo Park, Benjamin M Bolker

January 14, 2018

1 Introduction

The evolution of sexual reproduction still remains unsolved (Otto, 2009). Despite being the dominant mode of reproduction (Vrijenhoek, 1998), sexual reproduction is followed by many costs (Lehtonen et al., 2012). The most often cited is the cost of producing males (Smith, 1978). As males cannot produce offspring, a sexually reproducing population is expected to be outgrown by an asexually reproducing counterpart that grows twice as fast (given that the sexual population produces 50% males and 50% females) – hence the infamous *two-fold cost of sex* (Smith, 1978). However, the two-fold cost of sex relies on the assumption that all else is equal, and this is not necessarily true in nature. So what drives sex to persist?

One explanation for sex is the Red Queen Hypothesis (Bell, 1982). The Red Queen Hypothesis predicts sexually reproducing organisms to overcome the two-fold cost of sex under strong parasite selection by producing genetically diverse offspring that are resistant to infection (Haldane, 1949; Jaenike, 1978; Hamilton, 1980). Hosts with rare genotypes are expected to result in higher fitness and become common eventually. Furthermore, parasites coevolves with the host population to infect the common "rare" genotype [CITE], and negative frequeny depednece drives an oscillation in genotypic frequencies in host and parasite populations (Clarke, 1976; Hamilton, 1980).

Since the Red Queen Hypothesis has been postulated, much of theoretical work has focused on figuring out conditions under which parasites can maintain sexual reproduction.

On the other hand, empiricists have tried to quantify the Red Queen.

While the Red Queen Hypothesis is well supported both empirically and theoretically, there still remains a gap between theory and data. Many models for Red Queen Hypothesis rely on simplifying assumption that are not applicable to natural populations (Ashby and King, 2015). For example,

Recent studies have shown that ecological feedbacks can allow sex to be maintained more easily [CITE]. We find that a simple model can capture the dynamics Here, we try to bridge this gap further.

In this study, we extend the model suggested by Lively (2010b) to include stochasticity and population structure and fit the to data from Dagan et al. (2013); McKone et al. (2016); Vergara et al. (2014) using approximate bayesian computation (ABC). [Something about what we found out from the model fitting] Using posterior distributions obtained from the fit, we perform power analysis to test the idea that infection prevalence is positively correlated with frequency of sexual reproduction Lively (2001). We find that cycling of parasites and hosts can lead to negative correlation between prevalence of infection and sex. Instead, positive correlated predicted by Lively (2001) is driven by populations that do not go through Red Queen dynamics (i.e., having low level of sexual and infection).

2   Methods

2.1   Data

We consider two snail-trematode population in New Zealand (McKone et al., 2016; Vergara et al., 2014) and a similar snail-trematode population in Isral (Dagan et al., 2013). First, the snail-trematode system in New Zealand has been extensively studied under the Red Queen Hypothesis [CITE] so we expect a Red Queen model to fit well. Second, [Something about Israel snails being different]. Finally, by considering both spatial (Dagan et al., 2013; McKone et al., 2016) and spatiotemporal data (Vergara et al., 2014), we can compare how structure of the data may affect model fits. Data for Vergara et al. (2014) and Dagan et al. (2013) were obtained from their Dryad repository [CITE] and data for McKone et al. (2016) was extracted from their figure.

2.2   Model

We model obligate sexual hosts competing with obligate asexual hosts in multiple population by extending the model introduced by Lively (2010b). A model that Lively (2010b) used is essentially an SI model on a discrete time scale with natural mortality and virulence (reduction in offspring production by infected hosts). It is a suitable candidate for this study as it captures essential structures that are present in basic epidemiological and population population dynamics models and is general enough to be applied to broad systems in nature. We do not consider mechanistic details of the snail-trematode system such as life history of trematodes (Vergara et al., 2014) as it would require us to estimate or make assumptions on parameters on mechanisms, both of which are unrealistic.

All hosts are assumed to be diploids with two biallelic loci, and parasites are assumed to be haploids. Let $S_{ij}^k(t)$ and $A_{ij}^k(t)$ be the number of sexual and asexual hosts with genotype ij from population k at generation t. For simplicity, we drop the superscript representing population and write $S_{ij}(t)$ and $A_{ij}(t)$ unless necessary. Following Lively (2010b), the expected number of offsprings (without recombination or outcrossing) produced by a sexual population is given

by

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

where s is the proportion of males produced, $S_{ij,U}$ and $S_{ij,I}$ are the number of uninfected and infected sexual hosts in a population, and W_U and W_I are their corresponding fitnesses. We allow for cost of sex to vary by using a scale parameter, c_b, to the growth term ($2/c_b$ corresponds to two fold cost of sex) (Ashby and King, 2015). Equivalently, S'_{ij} is the amount of genotypic contribution by sexual hosts to the next generation.

Asexual hosts are assumed to be strictly clonal. Then, the expected number of offsprings produced by an asexual population is given by

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

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

We assume that there is mixing between population, determined by rate ϵ_{mix}. Then, the expected number of offspring in the next generation (accounting for contributions from all populations) is given by

$$\mathrm{E}\left(S^k_{ij}(t+1)\right) = f_{\mathrm{sex}}\left((1-\epsilon_{\mathrm{mix}})\left(S^k_{ij}\right)' + \frac{\epsilon_{\mathrm{mix}}}{n_{\mathrm{population}}-1}\sum_{l\neq k}\left(S^l_{ij}\right)'\right),$$

$$\mathrm{E}\left(A^k_{ij}(t+1)\right) = (1-\epsilon_{\mathrm{mix}})\left(A^k_{ij}\right)' + \frac{\epsilon_{\mathrm{mix}}}{n_{\mathrm{population}}-1}\sum_{l\neq k}\left(A^l_{ij}\right)', \tag{3}$$

where $f_{\mathrm{sex}}(x)$ is the function that models sexual reproduction, including recombination at a rate r_{host} and outcrossing, and $n_{\mathrm{population}}$ is the number of populations present in the system. We then take a poisson random variable to simulate process error and allow for stochastic migration to avoid fixation:

$$S^k_{ij}(t+1) \sim \mathrm{Poisson}\left(\lambda = \mathrm{E}\left(S^k_{ij}(t+1)\right)\right) + \mathrm{Bernoulli}\left(p = p_{ij,\mathrm{sex}}\right),$$

$$A^k_{ij}(t+1) \sim \mathrm{Poisson}\left(\lambda = \mathrm{E}\left(A^k_{ij}(t+1)\right)\right) + \mathrm{Bernoulli}\left(p = p_{ij,\mathrm{asex}}\right), \tag{4}$$

where $p_{ij,\mathrm{sex}}$ and $p_{ij,\mathrm{asex}}$ are the probability of a sexual and an asexual host with genotype ij entering the popualation.

Infection is modeled using the matching alleles model (Otto and Michalakis, 1998). We assume that heterozygous individuals are equally susceptible to parasites that match either genotypes

The expected number of infected hosts that carry parasite with genotype i at generation t is given by:

$$I_i(t) = \sum_p 2^{\delta_{ij}}\left(S_{ip,i,I}(t) + A_{ip,i,I}(t)\right), \tag{5}$$

where $S_{ip,i,I}$ and $A_{ip,i,I}$ are the expected number of sexual and asexual hosts infected with genotype i parasite. Following Ashby and King (2015), we assume

3

that mutation can occur in one locus with probability r_{parasite}. We also allow for external migration of an infected host carrying parasite i with probability $p_{i,\text{parasite}}$ to avoid fixation.

The amount of infectious contacts made by infected hosts from a population is given by $\lambda_i^k = \beta^k I_i'^k(t)$, where β^K is the transmission rate at each population, and $I_i'(t)$ is the number of infected hosts accounting for mutation. Since we allow for mixing between population, infected hosts in a population can make contact to susceptible hosts in other populations as well. Then, all infectious contacts made by infected hosts that applies to a host in a single population is

$$\lambda_{i,\text{total}} = (1 - \epsilon_{\text{site}})\lambda_i + \frac{\epsilon_{\text{site}}}{n_{\text{site}} - 1}\sum_{l \neq k}\lambda_i^l \tag{6}$$

Then, the force of infection that a susceptible host experiences is

$$\text{FOI}_{ij}^k = \frac{\lambda_{i,\text{total}} + \lambda_{j,\text{total}}}{2N(t+1)} \tag{7}$$

The probability of infection for a host with genotype ij at site k in the next generation is

$$P_{ij}^k(t+1) = 1 - \exp\left(\text{FOI}_{ij}^k\right). \tag{8}$$

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

$$\begin{aligned} S_{ij,I}(t+1) &\sim \text{Binom}(S_{ij}(t+1), P_{ij}), \\ A_{ij,I}(t+1) &\sim \text{Binom}(A_{ij}(t+1), P_{ij}). \end{aligned} \tag{9}$$

Expected number of individuals infected with haplotype i in the next generation is given by a ratio of λ:

$$\begin{aligned} S_{ij,i,I}(t+1) &= \frac{\lambda_{i,\text{total}}}{\lambda_{i,\text{total}} + \lambda_{j,\text{total}}}S_{ij,I}(t+1) \\ A_{ij,i,I}(t+1) &= \frac{\lambda_{i,\text{total}}}{\lambda_{i,\text{total}} + \lambda_{j,\text{total}}}A_{ij,I}(t+1) \end{aligned} \tag{10}$$

2.3   Simulation design and parameterization

Most Red Queen models have studied a competition between a single asexual genotype and sexual genotypes or have assumed equal genetic diversity of asexual and sexual hosts (see (Ashby and King, 2015)) but neither of these assumptions are realistic. Instead, Ashby and King (2015) adopted a more realistic approach by allowing for stochastic migration of an asexual genotype to a population. Here, we combine these methods. We allow for stochastic migration of migration of asexual hosts with genotypes genotypes but fix number of asexual genotypes (denoted by G_{asex}) that can be present in the population, while the number of sexual genotypes (G_{sex}) that can be present in the population

is equal to size of the genotypic space ($= 10$ in diploid hosts with two biallelic loci).

In the beginning of every simlation, G_{asex} is randomly chosen between 1 and 10 and asexual genotypes that can be introduced to the population are uniformly sampled from the entire genotypic space. To account for differing number of sexual and asexual genotypes, we let

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

$$p_{ij,\text{asex}} = \begin{cases} 1 - (1 - p_{\text{host}})^{1/G_{\text{asex}}} & \text{if } ij \in \{\text{asexual genotypes}\} \\ 0 & \text{otherwise} \end{cases}, \tag{11}$$

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

$$p_{i,\text{parasite}} = 1 - (1 - p_{\text{infected}})^{1/4}, \tag{12}$$

where p_{infected} is the probability that at least one infected host enters the population in a generation.

Each simulation consists of 40 populations. Every population consists of 2000 sexual hosts where 80 of them are infected. They are assumed to be in Hardy-Weinberg equilibrium with ratio between alleles being exactly half. Transmission rate parameter, β, is randomly drawn from a lognormal distribution with parameters β_{meanlog} and β_{sdlog} for each population. Simulation runs for 500 generations without introduction of asexuals. At generation 501, 10 asexual hosts of a single genotype are introduced to each population (note that asexual genotype introduced can vary across population) and simulation runs for 600 generations while allowing for stochastic migration of asexuals as well.

2.4 Approximate Bayesian Computation

We use Approximate Bayesian Computation (ABC) to fit the model (Toni et al., 2009). ABC relies on comparing summary statistics of observed data and those of simulated data and can be particularly useful when the exact likelihood function is not available [CITE?]. We consider mean proportion of infected and sexually reproducing snails in the system and variation in these proportions – measured by coefficient of variation (CV) – across space and time as summary statistics of interest.

For each sample from the posterior, the model is simulated and a subsample from the simulated data is drawn so that the number of simulated population is equal to the number of sites collected in a study. Then, the summary statistics are calculated by considering last 100 generations of simulated data and the parameter sample as well as the simulation generated by the sample is accepted if the distance between simulated and observed data (measureed by the absolute sum of differences in summary statistics) is less than the tolerance value. By simulating greater number of populations than number of sampled sites, we

account for uncertainty in unobserved population as well as interaction between populations.

We use uninformative priors for all parameters except c_b, a scale parameter for the cost of sex. The prior distribution for the scale parameter is chosen so that 95% quantile of cost of sex $(2/c_b)$ is approximately equal to the 95% confidence interval reported by Gibson et al. (2017) (see ??? for the list of parameters). As Dagan et al. (2013) and McKone et al. (2016) report proportion of males, instead of proportion of sexual hosts, we double the proportion of males reported when we fit the model.

To allow for more efficient estimation of the posterior distribution, we use Sequential Monte Carlo Turner and Van Zandt (2012) (see Appendix for details).

2.5 Power analysis

To calculate the power to detect positive correlation between infection prevalence and frequency of sexual hosts, we simulate an observed sample by drawing a random multinomial sample from simulated data. For simplicity, we assume that a year contains of two snail generations [CITE] and that samples are taken within a short period of time. For each simulation from the posterior distributions, we start by setting a reference generation to generatation 1001 and choosing n populations at random from 40 simulated popoulations. For each selected population, hosts are divided into four categories based on their infection status (infected/uninfected) and reproductive mode (asexual/sexual), and mean proportion of hosts in each category is calculated by averaging over two consecutive generations. A multinomial sample of size m is drawn from each selected population independently based on the four categories. Positive correlation between proportion of infected hosts and proportion of sexual hosts is tested using Spearman's rank correlation at 5% significance level. We repeat the process 100 times for each simulation from a posterior distribution.

3 Results

Our model can reproduce infection prevalence and frequency of sexual hosts observed acrosss three snail-trematode systems reasonably well (Fig. 1). However, our model tends to overestimate mean infection prevalence across all three systems. In particular, McKone et al. (2016) reported mean infection prevalence of 5.1% but the mean (95% quantile) estimate of the fitted model is 23.7% (9.1% - 40.3%). The fitted model underestimates mean frequency of sexual hosts in Vergara et al. (2014) population by almost 10% when compared with the mean estimate but the observed mean frequency still lies within a reasonable range of estimated values. We also find that there is large variation in the estimates even though model fitting was performed based on these summary statistics.

To further diagnose the fit, we compare the relationships between infection prevalence and frequency of sexual hosts predicted by the model with those observed across three systems (Fig. 2). Despite being able to reproduce summary
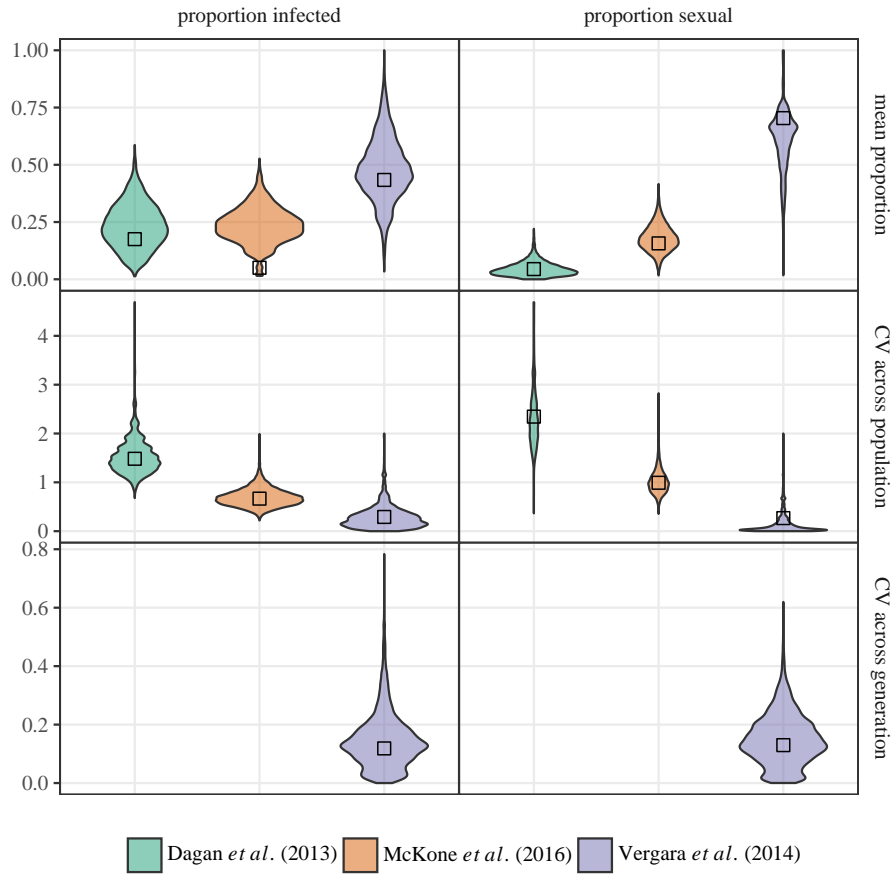
Figure 1: **Summary statistics of the observed data v.s. summary statistics of the simulated data from the posterior samples.** Violin plots show distribution of summary statistics from each simulated set. For each simulated data set accepted along with a posterior parameter sample, simulated populations are sampled at random 100 times so that each sample consists of equal number of populations as number of sites in fitted data. Then, summary statistics are calculated for each sample by taking last 100 generations. Squares represnt observed summary statistics.

statistics reported by Dagan et al. (2013) well, our model is unable to predict qualitative trend bewteen proportion of sexual hosts and proportion of infected hosts (Fig. 2; Dagan et al. (2013)). While both simulated data and observed data mostly consist of asexual populations, our model predicts some amount of sexual reproduction to be maintained when infection prevalence is greather than 20%. On the other hand, the data provided by Dagan et al. (2013) shows that sexual reproduction is only maintained at low levels of infection prevalence and is barely maintained otherwise. Due to such badness of fit, we exclude fits to Dagan et al. (2013) from further discussion (including parameter estimation and power analysis).
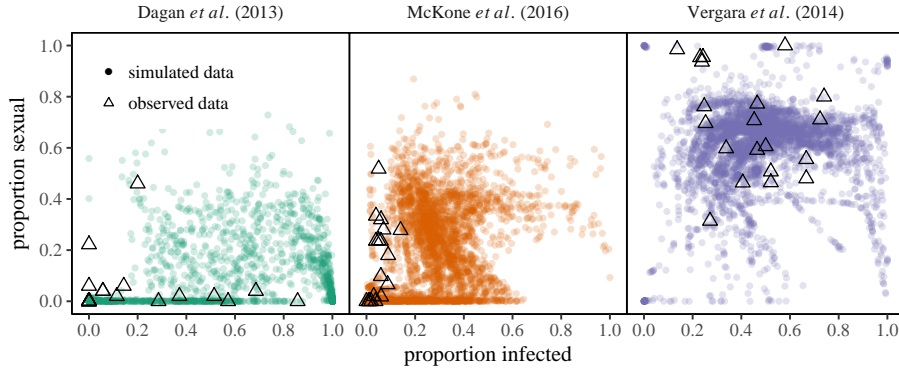

Figure 2: **Comparison in trends of the observed data v.s. simulated data from the posterior samples.** Solid circles represent a population within a simulation. Triangles represent observed data.

The most striking result from Fig. 2 is that there is a region in which proportion of infected hosts increases while proportion of sexual hosts increases (most clearly seen from the fits to McKone et al. (2016) and Vergara et al. (2014)). As infection rate increases, selection for sexual hosts increases but increasing number of resistant offsprings prevents further infection from happening and hence decreases prevalence of infection. This region connects the overlap in prevalence between sexual and asexual populations noted by Lively (2001) We also find that proportion of sexual hosts starts to decrease when infection prevalence is very high. Decrease in fitness of sexual hosts with increase in prevalence was predicted by Ashby and King (2015) but can also be also observed in an earlier work by Lively (2010b) (not discussed in the paper).

The model estimates similar parameters for both snail-trematode systems studied by McKone et al. (2016) and Vergara et al. (2014) (Fig. 3). The most surprising result is that the posterior distributions of the scale parameter, c_b, estimated by the model is much wider than the prior distribution and have noticeably higher mean: (mean (95% credible interval)) 1.19 (0.73-1.79) and 1.29 (0.79-2.01) for McKone et al. (2016) and Vergara et al. (2014), repsectively. Ashby and King (2015) defined c_b as additional costs and benefits of sex, where

$c_b = 1$ corresponds the two fold cost. Under their interpretation, our estimate corresponds to the following mean and 95% CIs for cost of sex: 1.68 (1.12-2.73) and 1.55 (1.00-2.52), respectively.
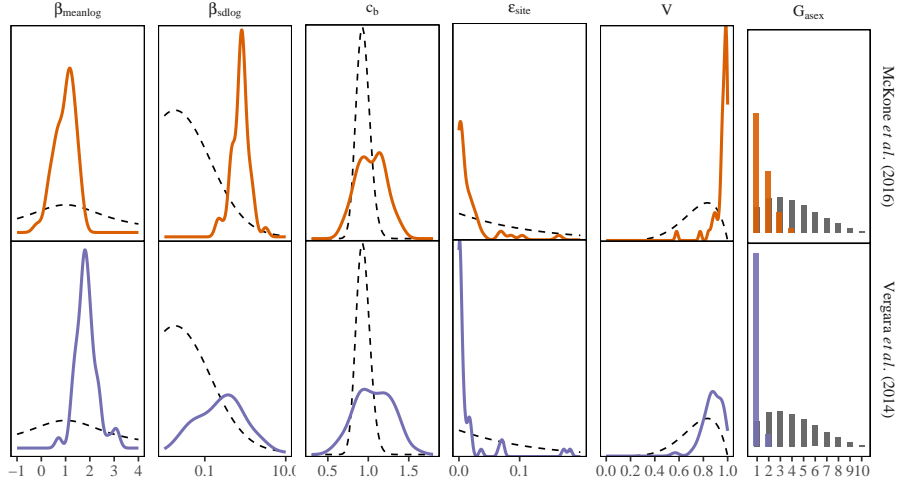

Figure 3: **Parameter estimates from Sequential Monte Carlo Approximate Bayesian Computation.** (dashed lines) prior distribution (solid lines) posterior distribution. 100 posterior samples were obtained from SMC ABC. Prior distributions are specified in [TODO]

The model estimates high virulence overall and almost 100% virulence for the data from McKone et al. (2016): (mean (95% CI)) 88.8% (75.0%-99.2%) and 95.0% (73.4%-100.0%) for Vergara et al. (2014) and McKone et al. (2016), repsectively. In contrast to Lively (2010b), who modeled 1 asexual genotype competing with 9 sexual genotypes, our model estimate shows that higher asexual to sexual genotypic ratio can be supported (G_{asex} panel in Fig. 2). However, it is still necessary for sexual hosts to have higher genetic diversity than asexual hosts.

Finally, a power analysis show that the power for detecting a positive correlation between infection prevalence and frequency of sexual hosts is not very high (Fig. 4). Given that McKone et al. (2016) collected about 100 snails from each of the 18 sites, they would have had 45% power to detect the positive correlation. Surprisingly, the power to detect positive correlation in the system studied by Vergara et al. (2014) is extremely low (less than 10%) and there is much higher (but still low) power to detect a negative correlation. Increasing number of samples per site has small effect on power once the sample size greater than 50. Increasing number of sites leads to greater increase in power but the power appears to saturate as number of sites increases.
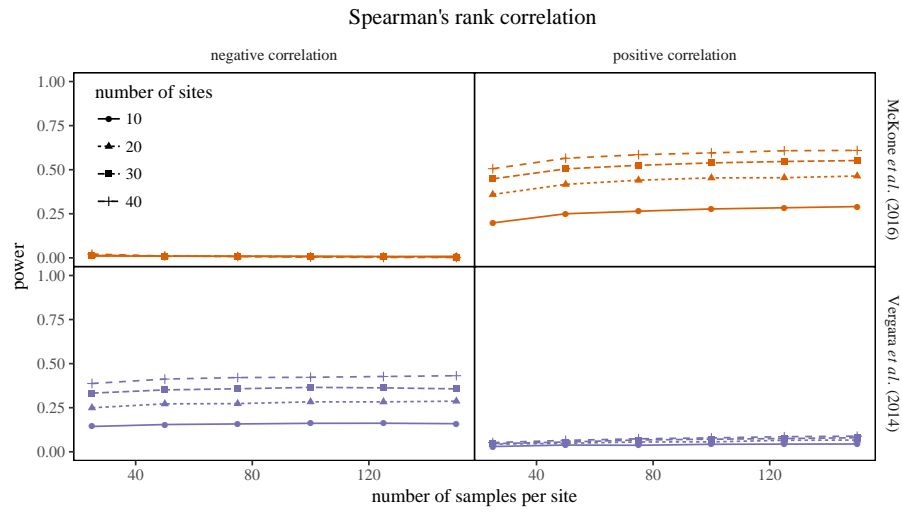
Figure 4: **Power to detect a significant correlation between infection prevalence and frequency of sexual hosts.** Spearman's rank correlation was used to detect correlation between infection prevalence and frequency of sexual hosts in simulated data from the posterior distributions. Detecting the positive correlation predicted by Lively (2001) is difficult. Increainsg number of samples and number of sites increases power but it is more efficient to increase number of sites.

4   Discussion

Our results challenge how the Red Queen Hypothesis for sex has been studied. Previous studies have relied on assumed parameters to understand how parasitism maintains sexual reproduction in host population [CITE]. Here, we show that (1) model parameters are estimable and (2) a simple Red Queen model with population structure can reproduce key summary statistics observed across three different snail populations (Fig. 1). While our model is able to reprouce observed summary statistics, there still remain some discrepancies between model prediction and observed trends in the relationship between infection prevalence and frequency of sexual hosts (Fig. 2). These discrepancies suggest that existing theoretical framework for the Red Queen Hypothesis may not be sufficient to explain sexual reproduction observed in nature. We also find that detecting a positive correlation between infection prevalence and frequency of sexual hosts is difficult (Fig. 4) and hence not a good measure for measuring the effects of the Red Queen.

A model that does not fit well can carry more information than a model that fits well. In order for our model to produce similar summary statistics as those observed in the snail-trematode systems, it requires the scale parameter (c_b) for the cost of sex to have mean greater than 1 (Fig. 3). In other words, a model that assumes two fold cost of sex ($c_b = 1$) would not have fitted well. However, $c_b > 1$ does not necessarily imply that cost of sex is less than two fold. Instead, an estimate of $c_b > 1$ should be interpreted as the amount of compensation required in order for the model to reproduce observed data. Additional compensation can include less than two fold cost of sex or any other mechanisms that promote sex.

A suitable candidate for model compensation is increase in host genetic diversity. Although exact genetic architecture that determines trematode infecion in snails is not known, genetic diversity of snails that have been documented is far greater than what we have assumed (two biallelic loci model yields 10 genotypes) (King et al., 2011; Dagan et al., 2013). In addition, increasing genetic diversity of the model may resolve overestimation of infection prevalence (Fig. 1). Previous studies have shown that moderately high genetic diversity can allow sexual hosts to escape infection more easily and therefore reduces prevalence of infection given equal amount of sexual reproduction (Lively, 2010a; King and Lively, 2012; Ashby and King, 2015). Until recently, many theoretical models have assumed limited genetic diversity. Of the M models reviewed by Ashby and King (2015) (Table 2), N out of M models... [TODO]

While all three studies (Dagan et al., 2013; McKone et al., 2016; Vergara et al., 2014) considered similar snail-trematode systems, the snail populations studied by Dagan et al. (2013) live in intrinsically different environments. Some populations are subject to seasonal flash floods, which can affect reproduction strategies (Ben-Ami and Heller, 2007). As a result, our simple model that does not take any exogeneous factors into account predicts a qualitatively different trend (Fig. 2). We conclude that the Red Queen Hypothesis alone cannot explain maintenance of sexual reproduction in this population. Overall, our results

indicate that more modeling effort is required to understand prevalence of sexual reproduction in nature.

Our power analysis (Fig. 4) appears to contradict the positive correlation predicted by Lively (1992, 2001) and findings of many empirical studies that have confirmed the prediction (Lively, 1987; Lively and Jokela, 2002; Kumpulainen et al., 2004; Vergara et al., 2013; McKone et al., 2016). Since parasitism is the only mechanism that drives sexual reproduction in our model, we expected the power to be high. However. the power analysis predicts almost no power for detecting a positive corrleation in the population studied by Vergara et al. (2014) and higher but still low power for detecting a negative correlation, contradicting an earlier work by Vergara et al. (2013) that reported a positive correlation between between infection prevalence and male frequency in the same population. The key premise behind the postive correlation predicted by Lively (2001) is that there must be large variation in infection prevalence. In particular, comparing Vergara et al. (2013) and Vergara et al. (2014) shows that the positive correlation is driven by differences between three populations from different habitats: populations that do not support the Red Queen Hypothesis (because they live in evolutionary coldspots) act as leverage points. In figure 3 of Kumpulainen et al. (2004), there a clearer distinction between two groups of populations that can support the Red Queen Hypothesis (high infection and high sexual reproduction) and those that cannot support the Red Queen Hypothesis (low infection and low sexual reproduction). Since both data sets provided by Vergara et al. (2014) and McKone et al. (2016) only consists of populations with high and low infection prevalence respectively, there is not enough variance in prevalence to observe the difference between two groups of populations. Therefore, the model predicts low power for detecting the positive correlation.

On the other hand, negative correlation between prevalence of infection and frequency of sexual hosts can be explained by cycling of host and parasite populations. A main component of the Red Queen Hypothesis is that negative frequency dependence drives oscillation in host reproductive strategy [CITE]. When temporal variation is taken into account, association between infection prevalence and frequency of sexual hosts can change depending on what phase each population is going through in its cycle (Fig. 5). The negative correlation in the population does not contradict the positive correlation predicted by Lively (2001) because their prediction did not take temporal variation into account.

Depressing power to detect positive correlation between infection prevalence and frequency of sexual hosts can be concerning. However, as lack of variation explains low power, sampling sites with large variation in infection prevalence will increase power to detect the positive correlation. In the Appendix, we show that doubling the log standard deviation of the transmission rates from the posterior distribution increases power. Alternatively,
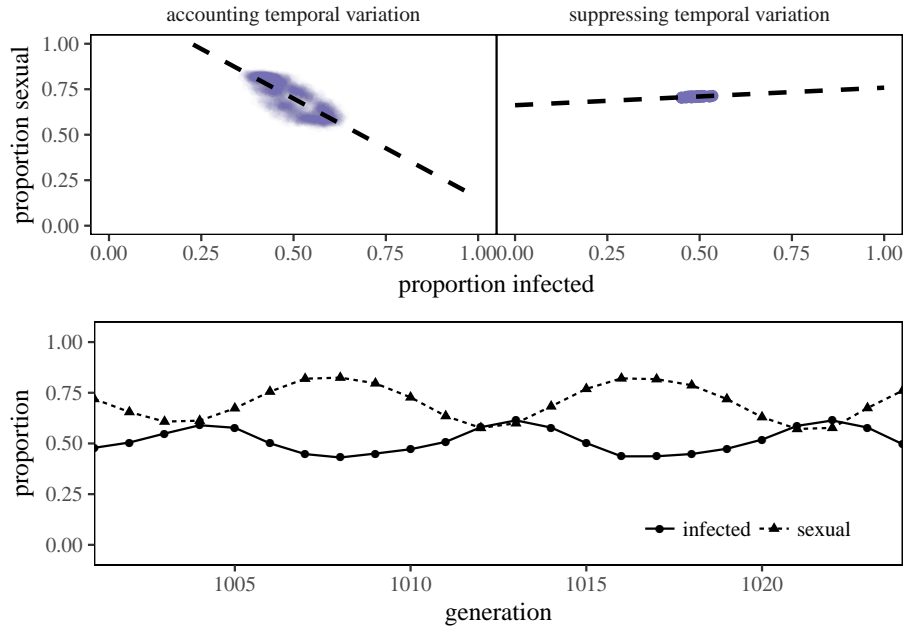
Figure 5: **Simulated data from the posterior distribution fitted to Vergara et al. (2014).** A particular simulation is chosen from the posterior distributions to demonstrate that it is possible to predict opposite relationship when temporal variation is taken into account. (Top left) each point represents proportion of infected hosts and proportion of sexually reproducing hosts of each population at each generation. Last hundred generations are plotted. (Top right) each point represents mean proporiton of infected hosts and mean proportion of sexually reproducing hosts of each population averaged over the last hundred generations. Dashed lines represent least squares fit to all point. (Bottom) A typical host parasite cycle observed in a population from this simulated data.

References

Ashby, B. and K. C. King (2015). Diversity and the maintenance of sex by parasites. *Journal of evolutionary biology 28*(3), 511–520.

Bell, G. (1982). *The Masterpiece of Nature: The Evolution and Genetics of Sexuality.* University of California Press.

Ben-Ami, F. and J. Heller (2007). Temporal patterns of geographic parthenogenesis in a freshwater snail. *Biological journal of the Linnean Society 91*(4), 711–718.

Clarke, B. (1976). The ecological genetics of host-parasite relationships. *Genetic aspects of host-parasite relationships. Blackwell, London*, 87–103.

Dagan, Y., K. Liljeroos, J. Jokela, and F. Ben-Ami (2013). Clonal diversity driven by parasitism in a freshwater snail. *Journal of evolutionary biology 26*(11), 2509–2519.

Gibson, A. K., L. F. Delph, and C. M. Lively (2017). The two-fold cost of sex: Experimental evidence from a natural system. *Evolution Letters 1*(1), 6–15.

Haldane, J. B. S. (1949). Disease and evolution. *La Ricerca Scientific Supplement 19*, 68–76.

Hamilton, W. D. (1980). Sex versus non-sex versus parasite. *Oikos*, 282–290.

Jaenike, J. (1978). An hypothesis to account for the maintenance of sex within populations. *Evol. theory 3*, 191–194.

King, K. and C. M. Lively (2012). Does genetic diversity limit disease spread in natural host populations? *Heredity 109*(4), 199–203.

King, K. C., J. Jokela, and C. M. Lively (2011). Parasites, sex, and clonal diversity in natural snail populations. *Evolution 65*(5), 1474–1481.

Kumpulainen, T., A. Grapputo, J. Mappes, and M. Björklund (2004). Parasites and sexual reproduction in psychid moths. *Evolution 58*(7), 1511–1520.

Lehtonen, J., M. D. Jennions, and H. Kokko (2012). The many costs of sex. *Trends in ecology & evolution 27*(3), 172–178.

Lively, C. M. (1987). Evidence from a new zealand snail for the maintenance of sex by parasitism. *Nature 328*(6130), 519–521.

Lively, C. M. (1992). Parthenogenesis in a freshwater snail: reproductive assurance versus parasitic release. *Evolution 46*(4), 907–913.

Lively, C. M. (2001). Trematode infection and the distribution and dynamics of parthenogenetic snail populations. *Parasitology 123*(07), 19–26.

Lively, C. M. (2010a). The effect of host genetic diversity on disease spread. *The American Naturalist 175*(6), E149–E152.

Lively, C. M. (2010b). An epidemiological model of host–parasite coevolution and sex. *Journal of evolutionary biology 23*(7), 1490–1497.

Lively, C. M. and J. Jokela (2002). Temporal and spatial distributions of parasites and sex in a freshwater snail. *Evolutionary Ecology Research 4*(2), 219–226.

McKone, M. J., A. K. Gibson, D. Cook, L. A. Freymiller, D. Mishkind, A. Quinlan, J. M. York, C. M. Lively, and M. Neiman (2016). Fine-scale association between parasites and sex in *Potamopyrgus antipodarum* within a New Zealand lake. *New Zealand Journal of Ecology 40*(3), 1.

Otto, S. P. (2009). The evolutionary enigma of sex. *The American naturalist 174*(S1), S1–S14.

Otto, S. P. and Y. Michalakis (1998). The evolution of recombination in changing environments. *Trends in ecology & evolution 13*(4), 145–151.

Smith, J. M. (1978). *The Evolution of Sex*, Volume 54. Cambridge Univ Press.

Toni, T., D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf (2009). Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface 6*(31), 187–202.

Turner, B. M. and T. Van Zandt (2012). A tutorial on approximate bayesian computation. *Journal of Mathematical Psychology 56*(2), 69–85.

Vergara, D., J. Jokela, and C. M. Lively (2014). Infection dynamics in coexisting sexual and asexual host populations: support for the Red Queen hypothesis. *The American naturalist 184*(S1), S22–S30.

Vergara, D., C. M. Lively, K. C. King, and J. Jokela (2013). The geographic mosaic of sex and infection in lake populations of a new zealand snail at multiple spatial scales. *The American Naturalist 182*(4), 484–493.

Vrijenhoek, R. C. (1998). Animal clones and diversity. *Bioscience 48*(8), 617–628.