Population structure promotes the evolution of costly sex in artificial gene networks

Abstract

We build on previous observations that Hill-Robertson interference generates an advantage of sex that, in structured populations, can be large enough to explain the evolutionary maintenance of costly sex. We employed a gene network model that explicitly incorporates interactions between genes. Mutations in the gene networks have variable effects that depend on the genetic background in which they appear. Consequently, our simulations include two costs of sex—recombination and migration loads—that were missing from previous studies of the evolution of costly sex. Our results suggest a critical role for population structure that lies in its ability to align the long- and short-term advantages of sex. We show that the addition of population structure favored the evolution of sex by disproportionately decreasing the equilibrium mean fitness of asexual populations, primarily by increasing the strength of Muller's Ratchet. Population structure also increased the ability of the short-term advantage of sex to counter the primary limit to the evolution of sex in the gene network model—recombination load. On the other hand, highly structured populations experienced migration load in the form of Dobzhansky–Muller incompatibilities, decreasing the effective rate of migration between demes and, consequently, accelerating the accumulation of drift load in the *sexual* populations.

Keywords: Recombination; Hill-Robertson; Genetic architecture; Population structure; Cost of sex

Introduction

Sexual reproduction among eukaryotes is nearly ubiquitous (Bell 1982; Maynard Smith 1978; Vrijenhoek 1998) despite its substantial costs. Foremost among these is the risk that successful, coevolved genetic interactions will be destroyed by recombination. The production of males introduces a further cost, which may approach twofold in anisogamous species (Lehtonen et al. 2012; Gibson et al. 2017). Mating can also be a costly endeavor. In many species, an individual must expend time and resources to locate a mate, risking disease and predation in the process. The widespread prevalence of sex indicates it must convey extraordinary benefits, but the conditions that allow these benefits to counter its significant costs remain controversial.

One benefit of sex lies in its ability to increase the efficiency of selection by breaking up deleterious genetic associations (Burt 2000; Otto and Lenormand 2002). Such deleterious genetic associations can accumulate in finite populations through the combined activities of selection and drift, producing a phenomenon known as Hill-Robertson interference (Hill and Robertson 1966; Felsenstein 1974; Comeron et al. 2008). Hill-Robertson interference impairs evolvability and, in its strongest form, leads to the irreversible accumulation of deleterious mutations via a mechanism known as Muller's Ratchet (Muller 1964; Felsenstein 1974; Haigh 1978). Recombination counters Hill-Robertson interference, generating an advantage of sex. Prior theoretical work on panmictic populations demonstrated that Hill-Robertson interference, particularly in the form of Muller's Ratchet, facilitates the spread of modifiers that increase the recombination rate, that is, the *origin* of sex (Otto and Barton 2001; Iles et al. 2003; Barton and Otto 2005; Keightley and Otto 2006; Gordo and Campos 2008; Hartfield et al. 2010). However, the advantage of sex in panmictic populations was insufficient to allow the evolution of *costly* sex, except when the modifier led to only a small increase in the probability that sex occurs (Keightley and Otto 2006; Hartfield et al. 2010). The advantage of sex was stronger in structured populations, where costly sex was able both to evolve (Martin et al. 2006) and to resist invasion by

asexual mutants (Peck et al. 1999; Salathé et al. 2006; Hartfield et al. 2012). This advantage of sex arising from population structure is expected to operate broadly because all natural populations show at least some structure.

The extent to which these findings can be generalized to real organisms, however, is unclear, as most of these models employed a simple genetic architecture, with limited to no interaction between genes (but see Peck et al. 1999). This precluded an important cost of sex—recombination load, which results from the disruption of coadapted gene complexes. This distinction proved vital in previous work that investigated the evolution of sex in an artificial gene network model that explicitly incorporated interactions between genes (Azevedo et al. 2006; Lohaus et al. 2010; Whitlock et al. 2016). As in real biological systems, mutations in the gene networks have variable effects that depend on the genetic background in which they appear. In panmictic populations of artificial gene networks, the interactions between genes evolved to produce a long-term (equilibrium fitness) advantage for sex that was strongest in small populations ($N < 10^2$). By contrast, sex exhibited a short-term (invasibility) advantage only in the largest populations ($N > 10^3$). Recombination load prevented both the origin and maintenance of sex in the small populations in which sex produced a large fitness advantage. As a result of the mismatch between the population sizes that produced the strongest short- and long-term advantages of sex, recombination load prevented the evolution of costly sex entirely (Whitlock et al. 2016).

We now ask whether population structure can facilitate the evolution of costly sex in the artificial gene network model. Population structure is expected to promote the evolution of sex because it introduces two processes that have a disproportionate negative impact on asexual populations. First, population structure strengthens the influence of Hill-Robertson interference (Martin et al. 2006; Hartfield et al. 2012). Second, it increases the time to fixation of invading asexual mutants, allowing more time for the operation of Hill-Robertson interference and, in particular, Muller's Ratchet (Peck et al. 1999; Salathé

et al. 2006). Because the short- and long-term advantages of sex depend, respectively, on the first (Gordo and Campos 2008) and second (Hartfield et al. 2012; Whitlock et al. 2016) of these effects, population structure has the potential to simultaneously increase both the short- and long-term advantages of sex, a possibility that has not been explored either theoretically or through simulation.

We note, however, that the complex genetic architecture of the gene network model, when combined with population structure, imposes another cost of sex: migration load. If we allow natural selection to operate in a spatially homogeneous manner, migrants will be equally well adapted to any deme, regardless of their sexual strategy. The offspring of sexual migrants will, nonetheless, experience a reduction in fitness (i.e., migration load) if they carry alleles that are incompatible with alleles in resident genotypes. Such Dobzhansky-Muller incompatilities (Dobzhansky 1937; Muller 1942) are common in nature, segregate within species (Corbett-Detig et al. 2013), and have been observed to evolve in a similar model (Palmer and Feldman 2009). Migration load can therefore be understood as an extension of recombination load, in which recombination between residents of different demes disrupts coadapted gene complexes and produces low-fitness hybrids. The impact of migration load on the benefits of sex imparted by population structure is unknown.

Here we evaluate the extent to which population structure supports the origin and maintenance of costly sex in a model with complex, evolving genetic architecture. We found that the relationship between population structure and the advantage of sex is non-monotonic. Small amounts of population structure did simultaneously increase both the short- and long-term advantages of sex by disproportionately increasing both the strength of Muller's Ratchet in asexual populations and the time to fixation of invading asexual or sexual mutants. As a result, population structure enabled both the origin and maintenance of sex in the face of substantial costs. However, these advantages of sex were maximized in moderately structured populations. Highly structured sexual populations

accumulated Dobzhansky–Muller incompatibilities that effectively prevented gene flow and, in some cases, disproportionately accelerated the accumulation of drift load (an analog of Muller's Ratchet) in *sexual* populations.

Methods

The gene network model used here is based on a model introduced by Wagner (1994, 1996).

Genotype

A haploid genotype is modeled as a network of n genes, each encoding a transcription factor that can, potentially, regulate its own expression or the expression of other genes. The gene network is represented by an $n \times n$ matrix, $\mathbf{R}$, where $r_{ij} \in \mathbf{R}$ is the regulatory effect of the product of gene j on gene i.

Phenotype

The expression pattern of an individual is represented by the vector $\mathbf{S}$, where $s_i \in \mathbf{S}$ is the expression state of gene $i = 1, 2, \ldots, n$. Expression states are discrete: a gene is either on $(s_i = +1)$ or off $(s_i = -1)$.

The expression pattern of an individual at time t is given by the system of difference equations

$$s_i(t+1) = f\left[\sum_{j=1}^{n} r_{ij} s_j(t)\right] \quad , \tag{1}$$

where f is a step function that determines how the input from the gene network controls

the expression of the target gene:

$$f(x) = \begin{cases} +1 & \text{if } x \geq 0 \\ -1 & \text{otherwise} \end{cases} .$$

Starting from an initial expression pattern $\mathbf{S}(0)$ at time $t = 0$, gene expression changes according to Equation 1 and is judged to reach a steady state if the following criterion is met: $\mathbf{S}(t) = \mathbf{S}(t-1)$. If a genotype does not achieve a gene expression steady state within $t \leq 100$ time steps, it is considered inviable ($W = 0$, see next section). If a genotype achieves a gene expression steady state within $t \leq 100$ time steps, it is considered viable ($W > 0$), and the steady state gene expression pattern $\hat{\mathbf{S}}$ is its *phenotype*.

Fitness

The fitness of a viable genotype is given by

$$W = \exp\left[-\frac{\delta(\hat{\mathbf{S}}, \dot{\mathbf{S}})}{\sigma}\right] , \tag{2}$$

where $\delta(\mathbf{S}, \mathbf{S}') = \sum_{i=1}^{n}(s_i - s_i')^2/(4n)$ measures the difference between expression patterns $\mathbf{S}$ and $\mathbf{S}'$, $\hat{\mathbf{S}}$ is the phenotype corresponding to the genotype, $\dot{\mathbf{S}}$ is the optimal phenotype, and $\sigma > 0$ is inversely related to the strength of stabilizing selection.

Offspring of asexual reproduction have fitness $W_{\text{asex}} = W$, whereas offspring of sexual reproduction have fitness $W_{\text{sex}} = W/C$, where $C \geq 1$ is the cost of sex. For example, $C = 1$ and $C = 2$ represent no cost and a two-fold cost, respectively.

Random genotype

A random genotype is created by generating a random gene network and a random initial gene expression pattern. A random gene network is obtained by randomly filling an $\mathbf{R}$

matrix with $(1-\gamma)n^2$ zeros and γn^2 random numbers from a standard normal distribution, where γ is the connectivity density of the network. A random initial gene expression pattern is generated by filling the n entries of $\mathbf{S}(0)$ with either -1 or $+1$ with equal probability.

Evolution

Evolution is simulated using an individual-based, Wright-Fisher model with non-overlapping generations. Populations maintain a constant total size, N, and are subdivided into D demes of equal size, $N_d = N/D$, arranged in a ring. In each generation, individuals undergo a selection-reproduction-mutation-migration life cycle.

Initialization: Simulations are initiated by producing N clones of a single viable randomly generated genotype. The phenotype of this founding individual is set as the population's optimal phenotype. The optimum remains constant over the course of the simulation.

Selection: Every generation, in each deme, the individuals that will reproduce are chosen at random, with replacement, with probability proportional to their fitness (Equation 2).

Reproduction: If the parent reproduces asexually, it generates a clone of itself; if a parent reproduces sexually, it mates with another sexually reproducing parent and produces one haploid recombinant offspring. The recombinant $\mathbf{R}$ matrix is generated by copying rows from the $\mathbf{R}$ matrices of the parents with equal probability. This is equivalent to free recombination between regulatory regions and no recombination within regulatory regions.

Mutation: Each individual offspring acquires a random number of mutations drawn from a Poisson distribution with mean U (the genomic mutation rate). A mutation is represented by a change to the value of one of the γn^2 nonzero entries in $\mathbf{R}$ chosen at random; the mutated value is drawn randomly from a standard normal distribution.

Migration: Every generation, a number of individuals from each deme is randomly chosen (without replacement) to migrate to either neighboring deme, with no bias for fitness or reproductive mode. The number of migrants from one deme to one of its neighbors is Poisson distributed with parameter $mN_d/2$, where m is the migration rate. The numbers of migrants to and from individual demes are not guaranteed to be equal. Thus, following migration the population may transiently show variation in deme sizes. Selection and reproduction restore the size of every deme to N_d.

Reproductive mode

The reproductive mode of an individual is determined by its genotype at a modifier locus, *rec*, unlinked to the genes involved in the gene network. There are two alleles at the modifier locus: rec^- and rec^+. If a population is fixed for the rec^- allele, every individual reproduces asexually, and if it is fixed for the rec^+ allele, every individual reproduces sexually.

The sexual and asexual subpopulations are reproductively isolated from each other. Sexual organisms do not experience a frequency-dependent cost of finding mates. One individual is chosen for every reproductive event with probability proportional to its fitness. If it carries the rec^- allele, it reproduces asexually. If it carries the rec^+ allele, a second individual carrying an rec^+ allele is chosen with probability proportional to its fitness, and the two individuals reproduce sexually and produce one recombinant offspring. The second individual may be the same as the first individual because sampling is done with replacement; the resulting reproductive event is equivalent to asexual reproduction.

Invasion analyses

Populations were evolved for 10^4 generations under asexual (or sexual) reproduction to allow sufficient time for the population to approach mutation-recombination-selection-drift-migration equilibrium. We then mutated the allele at the modifier locus *rec* ($rec^- \leftrightarrow rec^+$, see "Reproductive mode" above) in a single randomly chosen individual, causing it to switch to the opposite reproductive mode, from asexual to sexual or vice-versa. We measured the fixation probability of the novel modifier allele, u, relative to that of a neutral mutation ($u^* = 1/N$) in replicate invasion trials. During the invasion trials, individuals carrying the sexual modifier allele rec^+ experience a fixed cost of sex, C (see "Fitness" above). For invasion trials in which the novel modifier allele fixed, we recorded the generation in which it fixed as T_{fix}.

Population metrics

Mean fitness: The mean fitness, $\overline{W}$, of all individuals present in the population at a given time (see Equation 2). The mean fitness of sexual and asexual individuals is denoted by $\overline{W}_{\text{sex}}$ and $\overline{W}_{\text{asex}}$, respectively.

Rate of adaptation: Defined as

$$\Delta W / \Delta t = \frac{\ln W_2 - \ln W_1}{\ln t_2 - \ln t_1} \ ,$$

where W_i is mean fitness at generation t_i. We used $t_1 = 562$ and $t_2 = 1778$ generations in our calculations because variance among treatments in $\Delta W / \Delta t$ was maximized over this time interval.

Deleterious mutation rate: Defined as $U_d = U(p_l + p_d)$, where U is the genomic mutation rate and p_l and p_d are the proportion of lethal and non-lethal deleterious muta-

tions, respectively. U is constant throughout the course of a simulation but p_l and p_d can evolve. We estimated the quantity $p_l + p_d$ for a genotype by generating 100 copies of the genotype, each carrying one independently generated mutation, and evaluating the proportion of them that have lower fitness than the original genotype.

Epistasis: We define multiplicative epistasis between two nonlethal mutations, i and j, as $\varepsilon = W_{ij}/W - W_i W_j/W^2$, where W is the fitness of the unmutated genotype, W_i and W_j are the fitnesses of the single mutants, and W_{ij} is the fitness of the double mutant. We calculated mean epistasis coefficients, ε, across a random sample of 100 pairs of mutations, introduced individually and in combination into each viable individual in the population.

Mean fitness of recombinant offspring: Defined as $\overline{W}_R$, the mean fitness of offspring produced by recombination (without mutation) between individuals of the same deme, W_r, relative to the average fitness of its parents $\overline{W}_p$. We averaged $W_r/\overline{W}_p$ across N independently chosen pairs of individuals, where each parent was chosen with probability proportional to its fitness (i.e., in the same way the population reproduced in the evolutionary simulations).

Mean fitness of migrant offspring: Defined as $\overline{W}_M$, the mean fitness of offspring produced by recombination (without mutation) between individuals of neighboring demes, W_m, relative to the average fitness of its parents $\overline{W}_p$. We averaged $W_m/\overline{W}_p$ across N independent offspring, generated by recombination between N_d pairs of parents randomly chosen from each of the D pairs of neighboring demes. Both parents were chosen from among the individuals in their respective demes with probability proportional their fitness.

Population differentiation: Measured as $Q_{ST} = V_b/(V_w + V_b)$, where V_w and V_b are the variance within and among demes at a neutral locus, respectively. Each individual

carries 20 neutral loci, which were neither linked to each other nor to any of the gene network loci. The founding genotype had a value of 0 at all 20 neutral loci. Every generation, each neutral locus in each individual acquired a random number of mutations drawn from a Poisson distribution with mean 1.0 (the neutral locus mutation rate). Each mutation changed the value of the neutral locus by adding a random value drawn from the standard normal distribution. V_w and V_b were calculated as variance components in a random-effects analysis of variance. We averaged Q_{ST} across the 20 neutral loci.

Maximum cost of sex: Defined as origin-C_{max} and maintenance-C_{max} to describe the maximum sustainable cost of sex during the origin and maintenance of sex, respectively. Origin-C_{max} is the maximum value of C for which $u_{\text{sex}} > u^*$ in sexual invasion trials. Maintenance-C_{max} is the maximum value of C for which $u_{\text{asex}} < u^*$ in asexual invasion trials (see "Invasion analyses"). For each set of equilibrium populations evolved at a particular m and N_d, we measured the modifier fixation probability, u, at each of at least 4 values of C and estimated C_{max} from the resulting data using logistic regression (Figures S2 and S3).

Equilibrium values: Defined as values of population metrics at generation 10^4 and indicated by a hat symbol (e.g., $\widehat{W}_{\text{sex}}$).

Parameter values

By default, we used the same parameters as in Whitlock et al. (2016): networks of $n = 100$ genes with connectivity density $c = 0.05$, a genomic mutation rate of $U = 1$ per generation, and moderate stabilizing selection ($\sigma = 0.2$). Using these parameters, mutations have a broad range of potential fitness effects, including beneficial, neutral, slightly deleterious and lethal (Figure S3). Random genotypes have a deleterious mutation rate of $U_d \approx 0.5$ (Figure S2, Whitlock et al. 2016). This value is comparable to those estimated

for fruitflies ($U_d = 0.6$, Haag-Liautard et al. 2008) and rodents ($U_d = 0.5$, Gaffney and Keightley 2006).

We simulate populations of $N = 10^3$ individuals. We chose this value because it is the largest population size in which sex is approximately neutral in the short term in panmictic populations (see Fig. 4 of Whitlock et al. 2016).

Statistical Analysis

All statistics were conducted using the R statistical package, version 3.4.1 (Ihaka and Gentleman 1996). Logistic regression was conducted using the function *glm* from the *stats* package. Linear models that investigated the contributions of various population metrics to the long- and short-term advantages of sex were conducted using the function *lm* from the *stats* package.

All data were generated from 50 pairs of sexual and asexual populations. Each pair was founded by an independent, randomly generated genotype. Combined metrics from sexual and asexual populations, like $\widehat{W}_{\text{sex}}/\widehat{W}_{\text{asex}}$, were calculated separately for each pair, and the resulting 50 composite values were used for plotting and statistical tests.

Data Availability

Programs used to run all simulations were written in Python 2.7 and are available at `https://bitbucket.org/aobwhitlock/whitlock-et-al-2018`. Supplemental figures are available in the Supporting Information.

Results

Population structure has a complex effect on the long-term advantage of sex

Previously, it was shown that panmictic (i.e., unstructured) sexual populations of $N = 10^3$ individuals experiencing a genomic mutation rate of $U = 1$ and moderate stabilizing selection ($\sigma = 0.2$) evolved a mean fitness at equilibrium $5.7\% \pm 0.8\%$ (mean and 95% confidence interval) higher than that of otherwise identical asexual populations (Whitlock et al. 2016). Here we investigate the extent to which population structure increases the long-term advantage of sex.

We manipulated population structure in two ways. First, we subdivided the total population into $D = 10$ to 100 demes of identical size such that $N_d = 10^3/D$ (note that D and N_d were not varied independently). Demes were arranged around a circle and migration occurred via a one-dimensional stepping-stone model (Kimura 1952). Second, we varied the migration rate (m) between neighboring demes from 2×10^{-4} to 0.2. There was no migration (i.e., $m = 0$) between non-neighboring demes. Structured populations consisting entirely of either sexual or asexual individuals were allowed to evolve for 10^4 generations, by which time most populations had reached an approximate mutation-selection-drift-migration fitness equilibrium (Figure S1). The least structured populations (i.e., those subdivided into $D = 10$ demes of $N_d = 100$ individuals each, experiencing the highest migration rate, $m = 0.2$) evolved approximately like panmictic populations, achieving similar equilibrium mean fitness and $Q_{ST} \approx 0$ (Figure 1). Decreasing either the deme size (N_d) or the migration rate (m) caused the level of genetic differentiation among demes, Q_{ST}, to increase, confirming that populations became more structured (Figure 1).

Relative to panmictic populations, small increases in population structure, imposed by decreasing either N_d or m, caused the magnitude of the long-term advantage of sex ($\widehat{W}_{\text{sex}}/\widehat{W}_{\text{asex}}$) to increase (Figure 1). This pattern was true provided the population con-

sisted of relatively large demes ($N_d \gtrsim 50$). At $N_d = 50$ and $m = 2 \times 10^{-4}$, the average advantage of sex achieved a maximum that was 5.8-fold larger than it was in a panmictic population of the same total size.

Indeed, for each $m \leq 2 \times 10^{-2}$, there was an intermediate deme size, N_d^*, at which the advantage of sex was maximized. Further subdivision into demes that were smaller than N_d^* eroded the advantage of sex. Surprisingly, the effect of small deme size was most pronounced at the intermediate migration rate of $m = 2 \times 10^{-3}$, generating a strong *disadvantage* of sex in the smallest deme sizes ($N_d \leq 33$) at this migration rate.

The variance we observed in the long-term advantage of sex was almost entirely explained by the relative rate of adaptation in sexual compared to asexual populations ($\Delta W_{\text{sex}}/\Delta t - \Delta W_{\text{asex}}/\Delta t$; $R^2 = 0.95, F_{1,34} = 669.3, p < 10^{-15}$; Figure 1). We note that $\Delta W/\Delta t$ is a composite metric that increases with adaptation rate and decreases with the accumulation of genetic load. Although $\Delta W/\Delta t$, by definition, measures the rate of adaptation, in our experiments differences in $\Delta W/\Delta t$ may better reflect differences in the rate of accumulation of load. $\Delta W_{\text{sex}}/\Delta t - \Delta W_{\text{asex}}/\Delta t$ took on its highest values for combinations of m and N_d in which migration effectively countered the accumulation of drift load in sexual ($\Delta W_{\text{sex}}/\Delta t > 0$) but not asexual ($\Delta W_{\text{asex}}/\Delta t < 0$) populations. In these conditions, asexual populations experienced Muller's Ratchet. Conditions in which drift load accumulated in both sexuals and asexuals, but did so more rapidly in sexuals ($\Delta W_{\text{sex}}/\Delta t < \Delta W_{\text{asex}}/\Delta t < 0$), were entirely unexpected because Muller's Ratchet is not expected to operate in sexual populations. We next investigated the cause of the accelerated accumulation of drift load in some of the highly structured sexual populations.

Dobzhansky-Muller incompatibilities accelerate the accumulation of drift load in structured sexual populations

We hypothesized that evolving properties of the genetic architecture made substantial contributions to rate at which drift load accumulated, thus we measured the mutation,

recombination, and migration loads of equilibrium structured populations. Both the mutation and recombination loads evolved in ways that contributed to the long-term advantage of sex (Figure 2), as was observed previously in panmictic populations (Whitlock et al. 2016). The lower equilibrium $\widehat{U}_d$ of sexual, compared to asexual, populations generated a fitness advantage for sexual populations of 6–9% ($1.06 < \exp(-\widehat{U}_{d,\text{sex}})/\exp(-\widehat{U}_{d,\text{asex}}) < 1.09$). Recombination load at equilibrium generated a fitness disadvantage for sexual populations of 5–8% ($0.92 < \hat{W}_R < 0.95$). However, the differences among treatments in mutation and recombination loads (Figure 2) cannot explain the differences among treatments in the overall advantage to sex (Figure 1; linear regression of $\widehat{W}_{\text{sex}}/\widehat{W}_{\text{asex}}$ against $\exp(-\widehat{U}_{d,\text{sex}})/\exp(-\widehat{U}_{d,\text{asex}})$: $R^2 = 0.0937$, $F_{1,34} = 3.514$, $p = 0.0695$; linear regression against $\widehat{W}_R$: $R^2 = 0.0153$, $F_{1,34} = 0.5275$, $p = 0.4726$).

A third property of the genetic architecture, the relative fitness of sexual migrant offspring, $\widehat{W}_M$, also evolved in our simulations (Figure 2). At equilibrium, $\widehat{W}_M$ depended predictably on both N_d and m. For large N_d and m, the fitness of migrant offspring was indistinguishable from that of resident offspring (i.e., $\overline{W}_M \approx \overline{W}_R$), as expected when there is relatively little differentiation between demes (low $\widehat{Q}_{ST}$). As N_d and m declined, so did $\widehat{W}_M$, until when $N_d \lesssim 50$ and $m \leq 2 \times 10^{-3}$, demes were strongly differentiated (high $\widehat{Q}_{ST}$) and $\widehat{W}_M \approx 0$. These reductions in $\widehat{W}_M$ must have resulted from the accumulation of alleles in some demes that show strong negative epistatic interactions with alleles present in other demes—i.e., Dobzhansky-Muller incompatilities. This extreme variation in $\widehat{W}_M$ had only a negligible direct effect on population mean fitness because $\widehat{W}_M$ was low only when migration rates were also low ($m \lesssim 2 \times 10^{-3}$). Instead, the main consequence of a low $\widehat{W}_M$ was to reduce the effective gene flow between demes, which in turn intensified the accumulation of drift load in sexual populations.

In highly structured populations with moderate migration rates (e.g., $N_d < 50$ and $m = 2 \times 10^{-3}$; hot pink line in Figure 2), the accumulation of Dobzhansky-Muller incompatibilities ($W_M \approx 0$) explains the faster fitness declines in sexual than in asexual

populations ($\Delta W_{\text{sex}}/\Delta t < \Delta W_{\text{asex}}/\Delta t$; hot pink line in Figure 1). In this parameter range, the moderate migration experienced only by asexual populations more effectively countered Muller's Ratchet than the recombination experienced only by sexual populations. Outside this parameter range, the effective migration rate was similar in asexual and sexual populations, either because Dobzhansky-Muller incompatibilities didn't accumulate in sexual populations ($m > 10^{-3}$) or because the migration rate was also near zero in asexual populations ($m = 10^{-4}$).

The long-term advantage of sex is a good predictor of the evolutionary success of sex

Previously, it was shown that sexual modifier mutations were approximately neutral in panmictic asexual populations of $N = 10^3$; asexual modifier mutations were also approximately neutral in panmictic sexual populations of the same size. Thus, despite the existence of a long-term advantage of sex, sexual reproduction neither evolved nor was maintained in panmictic populations of $N = 10^3$ individuals (Whitlock et al. 2016). The addition of population structure is expected to promote the maintenance of costly sex — perhaps even costly sex— by increasing the time (T_{fix}) that asexual modifier mutations spend in transit to fixation. In turn, increasing T_{fix} is expected to increase the accumulation of deleterious mutations in the invading asexual lineage, depressing its mean fitness and probability of fixation. This effect, of population structure on T_{fix}, is not expected to influence the *origin* of costly sex. Here we investigate the extent to which the predicted effect of population structure on T_{fix} combines with the complex effect of population structure on equilibrium mean fitness (Figure 1) to promote both the origin and maintenance of costly sex.

As expected, population structure promoted the maintenance of costly sex (Figures 3B and 3D). For each combination of m and N_d we used invasion assays to estimate the maximum cost of sex (maintenance-C_{max}) under which the fixation probability of an

asexual modifier mutation was lower than the neutral expectation ($u < u^*$, Figure S2). maintenance-$C_{\max} > 1$ for almost all the parameter combinations we examined and rose as high as 1.36 for the most favorable parameter combination. Maintenance-$C_{\max}$ significantly increased with both the mean T_{fix} among successful invaders ($F_{1,13} = 142.5$, $p < 10^{-7}$) and the equilibrium advantage to sex ($\widehat{W}_{\text{sex}}/\widehat{W}_{\text{asex}}$, $F_{1,13} = 343.7$, $p < 10^{-10}$). Indeed, T_{fix} and the equilibrium advantage of sex predicted maintenance-$C_{\max}$ almost perfectly (multiple regression model containing only these two predictors, $R^2 = 0.98$).

Population structure also promoted the origin of costly sex (Figures 3A and 3C). To investigate the origin of sex we used invasion assays to estimate the maximum cost of sex (origin-$C_{\max}$) under which the fixation probability of a sexual modifier mutation was higher than the neutral expectation ($u > u^*$, Figure S3). Similar to maintenance-$C_{\max}$, origin-$C_{\max} > 1$ for almost all the parameter combinations we examined and increased significantly with the equilibrium advantage to sex ($F_{1,11} = 12.71$, $p = 0.0044$). However, origin-$C_{\max}$ rose only as high as 1.1 for the most favorable parameter combination and showed no significant relationship with T_{fix} ($F_{1,11} = 3.62$, $p = 0.0835$). In addition, origin-$C_{\max}$ was only somewhat predictable ($R^2 = 0.60$) from the quantities $\widehat{W}_{\text{sex}}/\widehat{W}_{\text{asex}}$ and T_{fix}.

We note that these results are for modifiers of sex that cause free recombination between rows of the genotype matrix, $\mathbf{R}$. We also examined the behavior of modifiers of recombination that cause only small changes in recombination rate, using a model that allows recurrent mutation at the modifier locus. Modifiers of recombination behaved similarly to modifiers of sex. Using the parameter combination that was most favorable to the origin of sex, recombination rate increased over time when $C = 1.1$, but not when $C = 1.2$ (Figure 4). This similarity between the two types of modifiers was not observed in panmictic populations, where modifiers of recombination were more likely to evolve than modifiers of sex (Whitlock et al. 2016).

Discussion

Our work builds on prior observations that Hill-Robertson interference generates an advantage of sex that, in structured populations, can be large enough to explain the evolutionary maintenance of *costly* sex (Peck et al. 1999; Salathé et al. 2006; Martin and Lenormand 2006; Hartfield et al. 2012). We employed an artificial gene network model that explicitly includes genetic interactions, resulting in two limits on the evolution of costly sex that were not present in the prior work—recombination load and migration load. First, population structure improved the ability of the short-term advantage of sex to counter recombination load. As a result, population structure facilitated the *origin* of sex, in addition to its maintenance. Second, highly structured populations experienced migration load in the form of Dobzhansky-Muller incompatibilities, accelerating the accumulation of drift load in the most-structured sexual populations and reducing the long-term advantage of sex. This effect was responsible for the observation that the maximum advantage of sex was realized in moderately structured populations.

Our first observation, that structure promoted the origin of sex, was not predictable from prior investigations of the evolution of sex because it resulted from the complex genetic architecture in the gene network model. Sexual lineages experience a short-term fitness decline, in the form of recombination load, before they evolve robustness that ameliorates the accumulating recombination load and produces a long-term fitness advantage in the form of lower U_d. (Figures S1, 1, and 2). In panmictic populations, successful sexual invaders must arise by chance in a high fitness genetic background and quickly hitchhike with that genetic background to a frequency that prevents their loss during the short-term fitness decline (Whitlock et al. 2016). Population structure facilitated the origin of sex by slowing the loss of such sexual invaders during the short-term fitness decline. In other words, population structure reduced the limitation imposed by recombination load by increasing the time over which recombinational robustness can evolve (T_{fix}) in invading sexual lineages.

Our second observation, that the long-term advantage of sex was maximized at intermediate structure, was also not predictable from prior work conducted using simpler genetic architectures. In the gene network model, differences in the rate at which drift load accumulated in sexual and asexual populations depended, in turn, on the effective rate of migration between demes, $m_e \approx m\,\overline{W}_M$. Increasing structure promoted the accumulation of Dobzhansky-Muller incompatibilities in sexual populations (Dobzhansky 1937; Muller 1942; Palmer and Feldman 2009; Bank et al. 2012), which in turn decreased the effective migration rate and resulted in substantial increases in migration load. In the most structured sexual populations, the mean fitness of hybrids between demes dropped to $\overline{W}_M \approx 0$ and the effective migration rate dropped to $m_e \approx 0$. Asexual populations did not experience migration load (i.e., $\overline{W}_M = 1$ and $m_e = m$) because asexual migrant alleles did not change genetic backgrounds and were equally well-adapted to local and distant environments. Thus, asexual populations benefited from migration in some high-structure conditions (e.g., $m = 10^{-3}$ and $N_d < 33$) where sexual populations did not. In these conditions, drift load actually accumulated faster in sexual than in asexual populations. Strong Dobzhansky-Muller genetic incompatilities are observed in other similar models (Palmer and Feldman 2009), and both within (Corbett-Detig et al. 2013) and among (Presgraves 2010; Maheshwari and Barbash 2011) biological species.

Earlier work has shown that migration load could generate a cost of sex in structured populations experiencing spatially *heterogeneous* selection (García-Ramos and Kirkpatrick 1997; Kirkpatrick and Barton 1997; Hendry et al. 2001; Alleaume-Benharira et al. 2006; Bolnick and Nosil 2007). Our results extend the potential role of migration load as a cost of sex to structured populations under spatially *homogeneous* selection. Our data also revealed a general mechanism for the cost of sex caused by migration load. Migration load, whether generated by the accumulation of Dobzhansky-Muller incompatibilities or by local maladaptation, reduces the effective migration rate, which intensifies the accumulation of drift load in sexual populations.

In sum, our results demonstrate that complex genetic architecture interacts with population structure to both promote and constrain the origin and maintenance of costly sex. The magnitude of both the long- and short-term advantages of sex are likely to be affected by many factors not considered here, including deviations from random mating other than those arising from population structure, ploidy, and environmental change (Shields 1982; Agrawal 2001; Siller 2001; Blachford and Agrawal 2006; Kirkpatrick and Jenkins 1989; Kondrashov and Crow 1991; Agrawal and Chasnov 2001; Otto 2003; Charlesworth 1993; Barton 1995; Otto and Nuismer 2004; Carja et al. 2014; Nowak et al. 2014). We posit that these factors may also interact with a complex genetic architecture to influence the evolution of sex, making them promising avenues for future inquiry.

Acknowledgements

We thank Sonia Singhal for her comments on the manuscript. The University of North Carolina at Chapel Hill and the Research Computing group provided computational resources and support. This work was funded by grant DEB-1355084 awarded to C.L.B. and grant DEB-1354952 awarded to R.B.R.A. by the National Science Foundation and by a Dissertation Completion Fellowship awarded to A.O.B.W. by the University of North Carolina at Chapel Hill. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

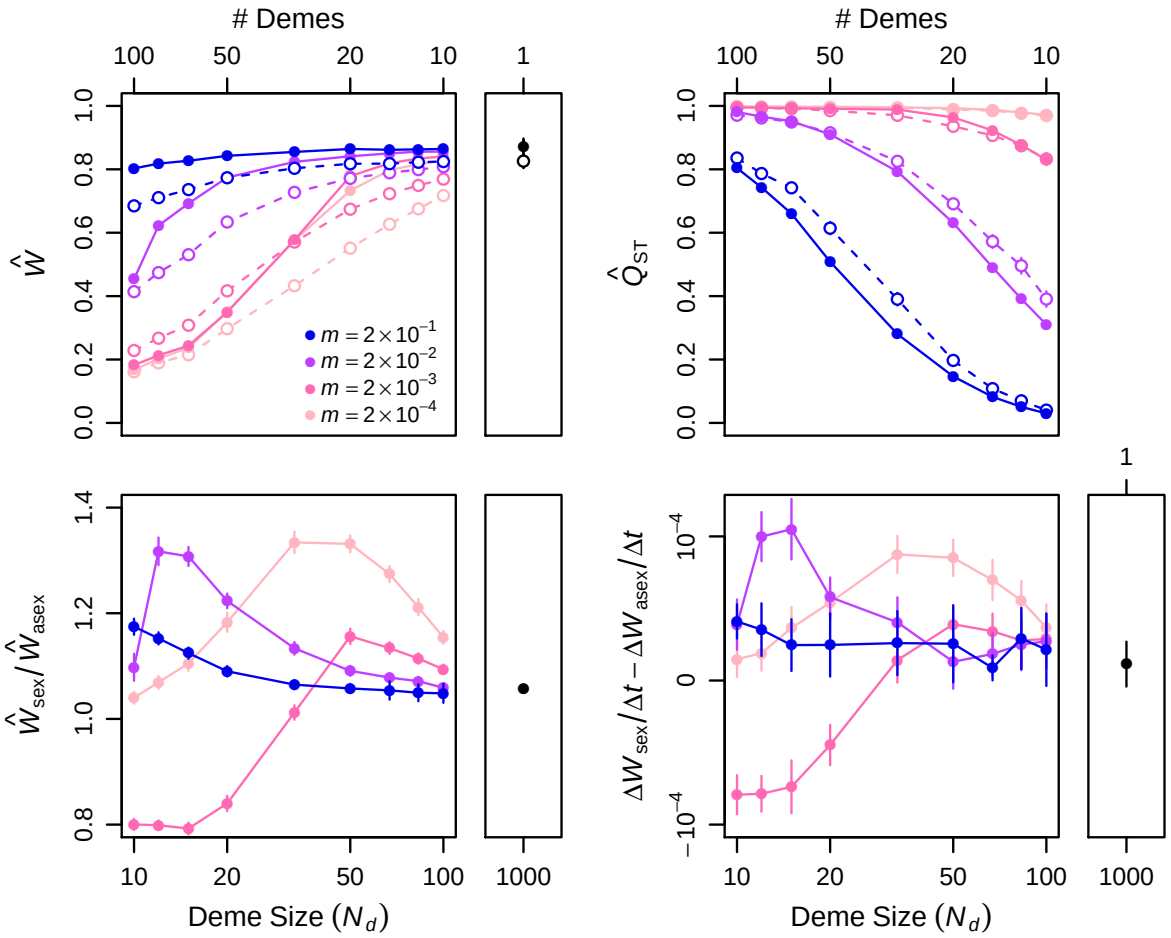

Figure 1: Migration rate and population subdivision interact in a complex way to determine the long-term advantage of sex. Mean fitness ($\hat{W}$), the fitness advantage of sexuals ($\hat{W}_{\text{sex}}/\hat{W}_{\text{asex}}$), and population differentiation ($\widehat{Q}_{ST}$) were calculated at generation 10^4, after populations of all migration rates and deme sizes had achieved approximate fitness equilibria. The rate of adaptation relative ($\Delta W_{\text{sex}}/\Delta t - \Delta W_{\text{asex}}/\Delta t$) was calculated at generation 10^3. In the top panels, sexual and asexual populations are represented by closed and open circles, respectively. Large panels show the relationships between these equilibrium metrics and the deme number and size of subdivided populations. Small panels show the equilibrium metrics calculated from panmictic asexual and sexual populations of the same total size, $N = 10^3$ (obtained as in Whitlock et al. 2016). Values are means and 95% confidence intervals based on 50 replicate populations initiated from different randomly chosen founders.

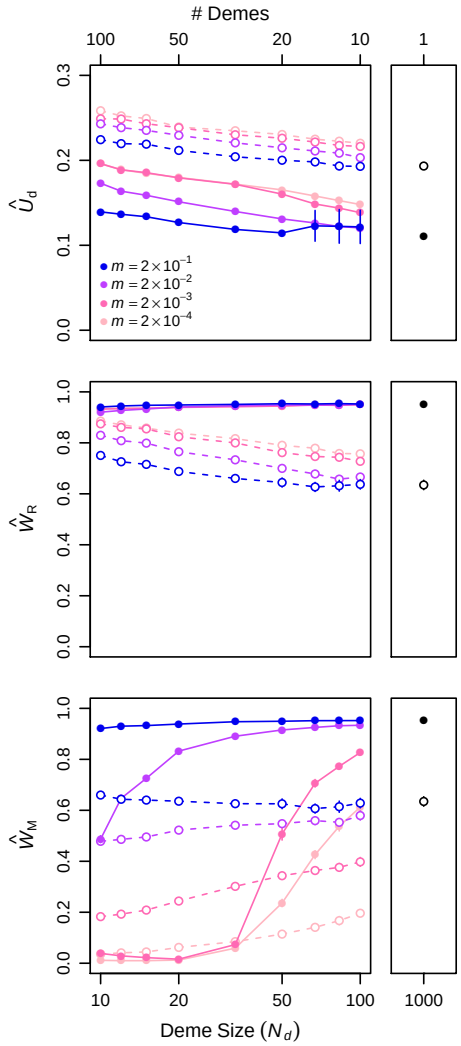

Figure 2: Migration rate and deme size affect three aspects of the equilibrium genetic architecture. Deleterious mutation rate ($\widehat{U}_d$), the mean fitness of recombinant offspring of parents from the same deme ($\widehat{W}_R$), and the mean fitness of recombinant offspring of parents from different demes ($\widehat{W}_M$) at generation 10^4, after populations of all migration rates and deme sizes achieved approximate fitness equilibria. As in Figure 1, sexual and asexual populations are represented by closed and open circles, respectively, and values are means and 95% confidence intervals based on 50 replicate populations initiated from different randomly chosen founders. Large panels show the relationships between these equilibrium metrics and the deme number and size of subdivided populations. Small panels show the equilibrium metrics calculated from panmictic asexual and sexual populations of size $N = 10^3$ (obtained as in Whitlock et al. 2016). Note that measures of $\widehat{W}_R$ and $\widehat{W}_M$ in asexual populations are provided for comparison only. During the evolution experiments, offspring produced by asexual reproduction experience $W_R = 1$ and $W_M = 1$, by definition.

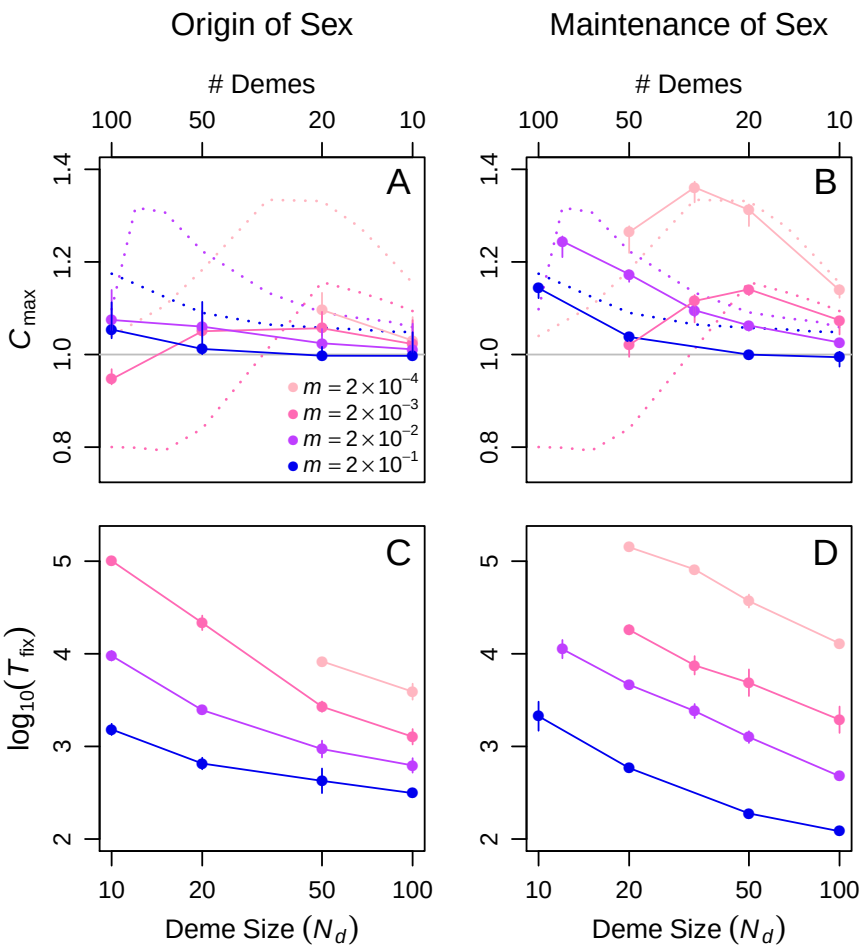

Figure 3: Population structure increases the transit time (T_{fix}) of invading sexual and asexual mutants, facilitating both the origin and maintenance of costly sex. (A) Points connected by solid lines show the maximum cost of sex (C_{max}) at which sexual mutants successfully invaded equilibrium asexual populations in each deme size and migration rate treatment. (B) Points connected by solid lines show the maximum cost of sex at which equilibrium sexual populations successfully resisted invasion by asexual mutants. Dotted lines in both (A) and (B) show the long-term advantage of sex, $\widehat{W}_{\text{sex}}/\widehat{W}_{\text{asex}}$, replotted from Figure 1, for comparison. (C and D) The average fixation time (T_{fix}) of sexual (C) and asexual (D) mutants that successfully invaded equilibrium asexual and sexual populations, respectively. Values are means and 95% confidence intervals based on all of the successful invasions in each condition. Missing points in the $m = 2 \times 10^{-3}$ and $m = 2 \times 10^{-4}$ series correspond to parameter combinations for which invasion trials were not run because they were not computationally feasible (i.e., T_{fix} was too long).

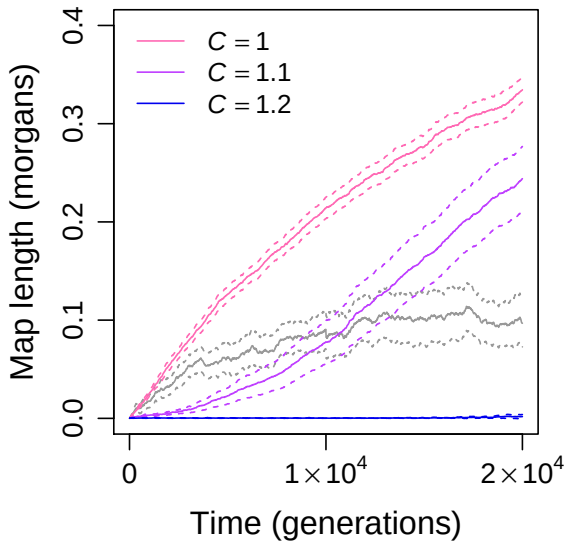

Figure 4: Evolution of the recombination rate under recurrent mutation at the recombination modifier locus. Color lines show the change in mean and 95% confidence interval, respectively, of the genome map length (i.e., mean crossover probability) over time in structured populations with $N_d = 50$ and $m = 2 \times 10^{-4}$ and costs of sex ranging from none ($C = 1$, pink) to substantial (purple and blue). Panmictic populations of the same total size ($N = 10^3$) experiencing no cost of sex ($C = 1$) are shown in gray for comparison. Data are from 50 replicates initiated with the equilibrium structured asexual populations shown in Figures 1 and S1 (pink, purple, and blue) or the equilibrium panmictic asexual populations of size $N = 10^3$ (gray, as in Whitlock et al. 2016).

Literature Cited

Agrawal, A. F., 2001. Sexual selection and the maintenance of sexual reproduction. Nature 411:692–695.

Agrawal, A. F. and J. R. Chasnov, 2001. Recessive mutations and the maintenance of sex in structured populations. Genetics 158:913–917.

Alleaume-Benharira, M., I. R. Pen, and O. Ronce, 2006. Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. J. Evol. Biol. 19:203–215.

Azevedo, R. B. R., R. Lohaus, S. Srinivasan, K. K. Dang, and C. L. Burch, 2006. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. Nature 440:87–90.

Bank, C., R. Bürger, and J. Hermisson, 2012. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. Genetics 191:845–863.

Barton, N. H., 1995. A general model for the evolution of recombination. Genet. Res. 65:123–145.

Barton, N. H. and S. P. Otto, 2005. Evolution of recombination due to random dirift. Genetics 169:2353–2370.

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

Blachford, A. and A. F. Agrawal, 2006. Assortative mating for fitness and the evolution of recombination. Evolution 60:1337–1343.

Bolnick, D. I. and P. Nosil, 2007. Natural selection in populations subject to a migration load. Evolution 61:2229–2243.

Burt, A., 2000. Perspective: Sex, recombination, and the efficacy of selection—Was Weismann right? Evolution 54:337–351.

Carja, O., U. Liberman, and M. W. Feldman, 2014. Evolution in changing environments:

Modifiers of mutation, recombination, and migration. Proc. Natl. Acad. Sci. U. S. A. 111:17935–17940.

Charlesworth, B., 1993. Directional selection and the evolution of sex and recombination. Genet. Res. 61:205–224.

Comeron, J. M., A. Williford, and R. M. Kliman, 2008. The Hill-Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. Heredity 100:19–31.

Corbett-Detig, R. B., J. Zhou, A. G. Clark, D. L. Hartl, and J. F. Ayroles, 2013. Genetic incompatibilities are widespread within species. Nature 504:135–7.

Dobzhansky, T., 1937. Genetics and the origin of species. Columbia Univ. Press, New York.

Felsenstein, J., 1974. The evolutionary advantage of recombination. Genetics 78:737–756.

Gaffney, D. J. and P. D. Keightley, 2006. Genomic selective constraints in murid noncoding dna. PLoS Genetics 2:e204.

García-Ramos, G. and M. Kirkpatrick, 1997. Genetic models of adaptation and gene flow in peripheral populations. Evolution 51:21–28.

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 1:6–15.

Gordo, I. and P. R. A. Campos, 2008. Sex and deleterious mutations. Genetics 179:621–626.

Haag-Liautard, C., N. Coffey, D. Houle, M. Lynch, B. Charlesworth, and P. D. Keightley, 2008. Direct estimation of the mitochondrial dna mutation rate in *Drosophila melanogaster*. PLoS Biol. 6:e204.

Haigh, J., 1978. The accumulation of deleterious genes in a population—Muller's ratchet. Theor. Popul. Biol. 14:251–267.

Hartfield, M., S. P. Otto, and P. D. Keightley, 2010. The role of advantageous mutations in enhancing the evolution of a recombination modifier. Genetics 184:1153 –1164.

———, 2012. The maintenance of obligate sex in finite, structured populations subject to

recurrent beneficial and deleterious mutation. Evolution 66:3658–3669.

Hendry, A. P., T. Day, and E. B. Taylor, 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. Evolution 55:459–466.

Hill, W. G. and A. Robertson, 1966. The effect of linkage on limits to artificial selection. Genet. Res. 8:269–294.

Ihaka, R. and R. Gentleman, 1996. R: a language for data analysis and graphics. J. Comp. Graph. Stat. 5:299–314.

Iles, M. M., K. Walters, and C. Cannings, 2003. Recombination can evolve in large finite populations given selection on sufficient loci. Genetics 165:2249–2258.

Keightley, P. D. and S. P. Otto, 2006. Interference among deleterious mutations favours sex and recombination in finite populations. Nature 443:89–92.

Kimura, M., 1952. "stepping-stone" model of population. Natl. Inst. Genet. Japan 3:62–63.

Kirkpatrick, M. and N. H. Barton, 1997. Evolution of a species' range. Am. Nat. 150:1–23.

Kirkpatrick, M. and C. D. Jenkins, 1989. Genetic segregation and the maintenance of sexual reproduction. Nature 339:300–301.

Kondrashov, A. S. and J. F. Crow, 1991. Haploidy or diploidy: which is better? Nature 351:314–315.

Lehtonen, J., M. D. Jennions, and H. Kokko, 2012. The many costs of sex. Tr. Ecol. Evol. 27:172–178.

Lohaus, R., C. L. Burch, and R. B. R. Azevedo, 2010. Genetic architecture and the evolution of sex. J. Hered. 101:S142–157.

Maheshwari, S. and D. A. Barbash, 2011. The genetics of hybrid incompatibilities. Annu. Rev. Genet. 45:331–355.

Martin, G. and T. Lenormand, 2006. The fitness effect of mutations across environments: a survey in light of fitness landscape models. Evolution 60:2413–2427.

Martin, G., S. P. Otto, and T. Lenormand, 2006. Selection for recombination in structured

populations. Genetics 172:593–609.

Maynard Smith, J., 1978. The Evolution of Sex. Cambridge University Press, Cambridge.

Muller, H. J., 1942. Isolating mechanisms, evolution and temperature. Biol. Symp. 6:71–125.

———, 1964. The relation of recombination to mutational advance. Mutat. Res. 1:2–9.

Nowak, S., J. Neidhart, I. G. Szendro, and J. Krug, 2014. Multidimensional Epistasis and the Transitory Advantage of Sex. PLoS Comput. Biol. 10:e1003836.

Otto, S. P., 2003. The advantages of segregation and the evolution of sex. Genetics 164:1099–1118.

Otto, S. P. and N. H. Barton, 2001. Selection for recombination in small populations. Evolution 55:1921–1931.

Otto, S. P. and T. Lenormand, 2002. Resolving the paradox of sex and recombination. Nat. Rev. Genet. 3:252–261.

Otto, S. P. and S. L. Nuismer, 2004. Species interactions and the evolution of sex. Science 304:1018–1020.

Palmer, M. E. and M. W. Feldman, 2009. Dynamics of hybrid incompatibility in gene networks in a constant environment. Evolution 63:418–431.

Peck, J. R., J. Yearsley, and G. Barreau, 1999. The maintenance of sexual reproduction in a structured population. Proc. R. Soc. B 266:1857–1863.

Presgraves, D. C., 2010. The molecular evolutionary basis of species formation. Nat. Rev. Genet. 11:175–180.

Salathé, M., R. Salathé, P. Schmid-Hempel, and S. Bonhoeffer, 2006. Mutation accumulation in space and the maintenance of sexual reproduction. Ecol. Lett. 9:941–946.

Shields, W. M., 1982. Philopatry, inbreeding, and the evolution of sex. State University of New York Press, Albany.

Siller, S., 2001. Sexual selection and the maintenance of sex. Nature 411:689–692.

Vrijenhoek, R. C., 1998. Animal Clones and Diversity. BioScience 48:617–628.

Wagner, A., 1994. Evolution of gene networks by gene duplications: A mathematical model and its implications on genome organization. Proc. Natl. Acad. Sci. U. S. A. 91:4387–4391.

———, 1996. Does evolutionary plasticity evolve? Evolution 50:1008–1023.

Whitlock, A. O. B., K. M. Peck, R. B. R. Azevedo, and C. L. Burch, 2016. An evolving genetic architecture interacts with Hill–Robertson interference to determine the benefit of sex. Genetics 203:923–936.