

# Supplemental Materials: Natural selection and the advantage of recombination [1] and companion publications [2, 47].

## S1. COVARIANCE, THE MEASURE OF RECOMBINANT ADVANTAGE

### Classical discrete-time definitions

Using standard notation and definitions, we let  $w_i$  denote the mean number of offspring that the  $i^{\text{th}}$  individual contributes to the next generation, and  $\bar{w} = \frac{1}{N} \sum_{i=1}^N w_i$ , the grand mean number of offspring taken across all individuals in the population. The relative fitness of the  $i^{\text{th}}$  individual is simply:

$$\frac{w_i}{\bar{w}} \quad (\text{S1})$$

We emphasize here that this definition of relative fitness is in *discrete time*, where the discrete time interval is a generation. While it makes biological sense and is the most commonly-used definition of relative fitness, technically speaking it implicitly makes the questionable assumption of synchronous reproduction, which can introduce error.

We now suppose that each individual has two genes, having genic fitnesses  $x$  and  $y$ . In the absence of epistasis, the relative fitness of the  $i^{\text{th}}$  individual is:

$$\frac{x_i y_i}{\bar{x} \bar{y}} \quad (\text{S2})$$

If a recombinant is produced that carries the  $x$  allele from the  $i^{\text{th}}$  individual and the  $y$  allele from the  $j^{\text{th}}$  individual, the relative fitness of the recombinant will be:

$$\frac{x_i y_j}{\bar{x} \bar{y}} \quad (\text{S3})$$

If a recombinant is produced that carries the  $x$  allele from the  $i^{\text{th}}$  individual and the  $y$  allele from a randomly chosen individual in the population, then on average, the relative fitness of the recombinant will be:

$$\frac{x_i \frac{1}{N} \sum_{j=1}^N y_j}{\bar{x} \bar{y}} \quad (\text{S4})$$

Following the same logic, if a recombinant is produced from two randomly-chosen parents, then on average, the relative fitness of the recombinant will be:

$$\bar{w}_r = \frac{\frac{1}{N} \sum_{i=1}^N x_i \frac{1}{N} \sum_{j=1}^N y_j}{\bar{x} \bar{y}} \xrightarrow{N \rightarrow \infty} \frac{\mathbb{E}[x] \mathbb{E}[y]}{\mathbb{E}[xy]} \quad (\text{S5})$$

The standard discrete-time definition of selective advantage is mean relative fitness after one generation minus one. By this definition, the mean selective advantage of recombinants is:

$$\bar{s}_r = \bar{w}_r - 1 = \frac{\mathbb{E}[x] \mathbb{E}[y]}{\mathbb{E}[xy]} - 1 = \frac{-\text{Cov}(x, y)}{\bar{w}} \quad (\text{S6})$$

where  $\bar{w} = \mathbb{E}[xy]$ , the mean fitness of the population.

### Continuous time

Elsewhere, we have shown that time-integrated recombinant selective advantage evolves as:

$$\int_0^t s_r(u) du = C_0(t, 0) + C_0(0, t) - C_0(t, t)$$

where  $C_0(\varphi, \theta)$  is the *cgf* of the initial fitness distribution.

Average recombinant advantage over the course of its first generation of growth is therefore:

$$\begin{aligned}\bar{s}_r &= \int_0^1 s_r(u) du = \mathcal{C}_0(1, 0) + \mathcal{C}_0(0, 1) - \mathcal{C}_0(1, 1) \\ &= \ln \frac{\mathbb{E}[e^X] \mathbb{E}[e^Y]}{\mathbb{E}[e^{X+Y}]} \\ &\approx -\sigma_{XY}\end{aligned}$$

Figure S1 shows the approximation in the last step to be extremely accurate. We believe the reason this approximation is so accurate is that the Jensen gap for numerator and denominator cancel each other out:

Jensen's inequality in two dimensions gives  $\ln \mathbb{E}[e^X] \mathbb{E}[e^Y] = \mathbb{E}[X] \mathbb{E}[Y] + JG1$  and  $\ln \mathbb{E}[e^{X+Y}] = \mathbb{E}[XY] + JG2$ , where  $JG1$  and  $JG2$  are Jensen gaps one and two. If  $JG1 \approx JG2$  then  $\ln \mathbb{E}[e^X] \mathbb{E}[e^Y] - \ln \mathbb{E}[e^{X+Y}] \approx \mathbb{E}[X] \mathbb{E}[Y] - \mathbb{E}[XY] = -\sigma_{XY}$ . Figure S1 reveals that  $JG1$  and  $JG2$  are indeed extremely close and thus effectively cancel each other out.

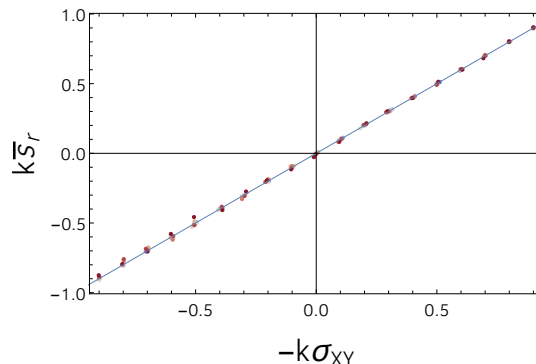


FIG. S1. Mean selective advantage of recombinants over their first generation of growth is extremely well approximated by minus the covariance ( $-\sigma_{XY}$ ). Dots plot values based on 10,000 points drawn from a bivariate normal distribution with zero means and standard deviations varying between zero and one. Quantities were standardized to fit on the same plot by multiplying by  $k = (\sigma_X \sigma_Y)^{-1}$ . The thin line plots the  $y = x$  line.

## S2. EVOLUTIONARY MODELS

### One-locus model with selection and mutation

Let  $u_t(x)$  denote probability density in fitness  $x$  at time  $t$  (i.e.,  $\int_x u_t(x) = 1$  for all  $t$ ) for an evolving population. Dropping the subscript  $t$ , we have that, under selection and mutation,  $u$  evolves as:

$$\partial_t u(x) = (x - \bar{x})u(x) + U \int_{\gamma} u(x - \gamma)g(\gamma) - Uu(x)$$

where  $U$  is genomic mutation rate, and  $g(\gamma)$  is probability density for fitness changes incurred by mutation, i.e.,  $g(\gamma)$  is the “distribution of fitness effects” of newly-arising mutations, or DFE.

Let  $M(\varphi)$  denote the moment-generating function (*mgf*) for  $u(x)$ , i.e.,  $M(\varphi) = \mathbb{E}_u[e^{\varphi X}]$  and let  $G(\varphi)$  denote the *mgf* for the DFE, i.e.,  $G(\varphi) = \mathbb{E}[e^{\varphi x}]$ . The transformed equation is:

$$\partial_t M(\varphi) = \partial_{\varphi} M(\varphi) - \partial_{\varphi} M(0)M(\varphi) + UM(\varphi)(G(\varphi) - 1).$$

Over the time interval in question (assumed to be short on evolutionary time scales), we will suppose the DFE is invariant such that  $G(\varphi)$  is a constant function. We define cumulant-generating function (*cgf*)  $\mathcal{C}(\varphi) = \ln M(\varphi)$ ; noting that  $\partial_{\varphi} \mathcal{C}(\varphi) = (\partial_{\varphi} M(\varphi))/M(\varphi)$ , and  $\partial_t \mathcal{C}(\varphi) = (\partial_t M(\varphi))/M(\varphi)$  we find that the *cgf* evolves as:

$$\partial_t \mathcal{C}(\varphi) = \partial_{\varphi} \mathcal{C}(\varphi) - \partial_{\varphi} \mathcal{C}(0) + U(G(\varphi) - 1).$$

This equation is a variant of the transport equation and, when boundary condition  $\mathcal{C}(0) = 0 \forall t$  is applied, it has solution:

$$\mathcal{C}_t(\varphi) = \mathcal{C}_0(\varphi + t) - \mathcal{C}_0(t) + U \int_0^t (G(\varphi + \gamma) - G(\gamma))d\gamma \quad (\text{S7})$$

where the subscripts are now necessary again:  $\mathcal{C}_t(\varphi)$  is the *cgf* of the fitness distribution  $u_t(x)$  at time  $t$ . We note that the evolution of a population can thus be projected into the future based only on the present fitness distribution (i.e., at  $t = 0$ ).

### Two-locus model with selection and mutation

We now suppose that there are two “genes” that determine fitness, such that total fitness is determined by their sum. Letting fitness contributions of the two genes be denoted by  $x$  and  $y$ , respectively, the total fitness is then simply  $x + y$ . The extension of the previous one-dimensional *pde* is therefore immediate:

Let  $u_t(x, y)$  denote probability density in fitness contributions  $x$  and  $y$  at time  $t$  for an evolving population. Dropping the subscripts again, under selection and mutation,  $u$  evolves as:

$$\begin{aligned} \partial_t u(x, y) &= (x + y - \bar{x} - \bar{y})u(x, y) \\ &+ U \int_{\gamma, \phi} u(x - \gamma, y - \phi)g(\gamma, \phi) - Uu(x, y) \end{aligned}$$

where again  $U$  is genomic mutation rate, and  $g(\gamma, \phi)$  is again the “distribution of fitness effects” of newly-arising mutations, or DFE, only now it is a bivariate distribution.

Let  $M(\varphi, \theta)$  denote the *mgf* for  $u(x, y)$ , i.e.,  $M(\varphi, \theta) = \mathbb{E}_u[e^{\varphi X + \theta Y}]$  and let  $G(\varphi, \theta)$  denote the *mgf* for the DFE, i.e.,  $G(\varphi, \theta) = \mathbb{E}_g[e^{\varphi X + \theta Y}]$ . Again, over the time interval in question (assumed to be short on evolutionary time scales), we will suppose the DFE is invariant such that  $G(\varphi, \theta)$  is a constant function. As before, defining *cgf*  $\mathcal{C}(\varphi, \theta) = \ln M(\varphi, \theta)$ , we have:

$$\begin{aligned} \partial_t \mathcal{C}(\varphi, \theta) &= \partial_\varphi \mathcal{C}(\varphi, \theta) + \partial_\theta \mathcal{C}(\varphi, \theta) - \partial_\varphi \mathcal{C}(0, 0) - \partial_\theta \mathcal{C}(0, 0) \\ &+ U(G(\varphi, \theta) - 1). \end{aligned} \quad (\text{S8})$$

This equation is a two-dimensional variant of the transport equation and has more possible solution forms than the one-dimensional case, namely, solutions can be of the form:  $\mathcal{C}(t + \varphi, \theta - \varphi)$ ,  $\mathcal{C}(t + \theta, \varphi - \theta)$ , or  $\mathcal{C}(t + \varphi, t + \theta)$ . The consistent solution is the last of these. When boundary condition  $\mathcal{C}(0, 0) = 0 \forall t$  is applied, it has solution:

$$\begin{aligned} \mathcal{C}_t(\varphi, \theta) &= \mathcal{C}_0(\varphi + t, \theta + t) - \mathcal{C}_0(t, t) \\ &+ U \int_0^t (G(\varphi + \gamma, \theta + \gamma) - G(\gamma, \gamma))d\gamma \end{aligned} \quad (\text{S9})$$

where the subscripts have again become necessary. We again note that the evolution of a population can thus be projected into the future based only on the present fitness distribution (i.e., at  $t = 0$ ).

### Two-locus model with selection, mutation and drift

For our first few steps, we follow the logic outlined in Ewens’ book [S48], section 4.10. We use the same notation as in that reference, with the exception of his use of the variable  $x$  which we replace with  $q$ , because we will later use  $X$  as we have before, to denote the fitness contribution of one of the two genes. We define the expectation of some arbitrary function  $g(\mathbf{q})$ , to be:

$$\mathbb{E}[g(\mathbf{q})] = \int g(\mathbf{q})f(\mathbf{q}; \mathbf{p}, t)d\mathbf{q}$$

where  $f(\mathbf{q}; \mathbf{p}, t)$  is a probability density of a diffusion process in  $\mathbf{q}$  with initial frequencies  $\mathbf{p}$ .

Ewens [S48] (p.154, eq.4.83) gives the rate of change of the expectation of  $g(\mathbf{q})$ :

$$\frac{\partial}{\partial t} \mathbb{E}[g(\mathbf{q})] = \mathbb{E} \left[ \sum a_i(\mathbf{q}) \frac{\partial g(\mathbf{q})}{\partial q_i} + \frac{1}{2} \sum b_i(\mathbf{q}) \frac{\partial^2 g(\mathbf{q})}{\partial q_i^2} + \sum \sum c_{ij}(\mathbf{q}) \frac{\partial^2 g(\mathbf{q})}{\partial q_i \partial q_j} \right] \quad (\text{S10})$$

We now define the function  $g(\mathbf{q})$  to be:

$$g(\mathbf{q}) = g(q_1, q_2, \dots, q_n) = \text{Log} \left( \sum_{i=1}^n q_i e^{\varphi X_i + \theta Y_i} \right) = \tilde{\mathcal{C}}(\varphi, \theta) \quad (\text{S11})$$

where, adhering to our previous notation, the  $X_i$  and  $Y_i$  are fitness contributions of the two genes in question, as before, and  $q_i$  denotes the frequency of individuals with total fitness  $Z_i = X_i + Y_i$ .

The first term on the right-hand side of (S10) is the selection term, with:

$$a_i(\mathbf{q}) dt + \mathcal{O}(\delta t) = \mathbb{E}[\delta q_i] = q_i(Z_i - \mathbb{E}[Z])$$

and the second two terms are drift terms, with:

$$b_i(\mathbf{q}) dt + \mathcal{O}(\delta t) = \text{Var}[\delta q_i] = q_i(1 - q_i)/n$$

and

$$c_{ij}(\mathbf{q}) dt + \mathcal{O}(\delta t) = \text{Cov}[\delta q_i, \delta q_j] = -q_i q_j / n$$

Plugging these definitions into Eq. (S10), we have:

First term (selection):

$$\begin{aligned} \sum_i a_i(\mathbf{q}) \frac{\partial g(\mathbf{q})}{\partial q_i} &= \frac{\sum_i q_i (X_i + Y_i) e^{\varphi X_i + \theta Y_i}}{\sum_i q_i e^{\varphi X_i + \theta Y_i}} - \sum_i q_i (X_i + Y_i) \\ &= \tilde{\mathcal{C}}_0^{(1,0)}(\varphi, \theta) + \tilde{\mathcal{C}}_0^{(0,1)}(\varphi, \theta) - \tilde{\mathcal{C}}_0^{(1,0)}(0, 0) - \tilde{\mathcal{C}}_0^{(0,1)}(0, 0) \end{aligned}$$

as before, where  $\tilde{\mathcal{C}}_0(\theta, \phi) = \tilde{\mathcal{C}}(\theta, \phi, 0)$  is the cumulant-generating function associated with random variables  $X$  and  $Y$  at time zero, as defined above in (S11).

Second term (drift 1):

$$\frac{1}{2} \sum_i b_i(\mathbf{q}) \frac{\partial^2 g(\mathbf{q})}{\partial q_i^2} = -\frac{1}{2} \frac{\frac{1}{n} \sum_i q_i (1 - q_i) e^{2\varphi X_i + 2\theta Y_i}}{(\sum_i q_i e^{\varphi X_i + \theta Y_i})^2}$$

Third term (drift 2):

$$\begin{aligned} &\sum_i \sum_{j>i} c_{ij}(\mathbf{q}) \frac{\partial^2 g(\mathbf{q})}{\partial q_i \partial q_j} \\ &= \frac{-\frac{1}{n} \sum_i \sum_{j>i} q_i q_j e^{\varphi(X_i + X_j) + \theta(Y_i + Y_j)}}{(\sum_i q_i e^{\varphi X_i + \theta Y_i})^2} \end{aligned}$$

The two drift terms have a common denominator so the numerators can simply be added. When added, it can be rearranged so that the sum of the two drift terms is:

$$\begin{aligned} &\frac{\frac{1}{2} (\sum q_i e^{\varphi X_i + \theta Y_i})^2 - \frac{1}{2} \sum q_i e^{2\varphi X_i + 2\theta Y_i}}{(\sum q_i e^{\varphi X_i + \theta Y_i})^2} \\ &= \frac{1}{2n} \frac{\tilde{M}(\varphi, \theta)^2 - \tilde{M}(2\varphi, 2\theta)}{\tilde{M}(\varphi, \theta)^2}, \end{aligned}$$

where  $\tilde{M}(\varphi, \theta) = \sum_i q_i e^{\varphi X_i + \theta Y_i}$ , the moment-generating function associated with random variables  $X$  and  $Y$ . And of course we have that  $\tilde{M}(\varphi, \theta) = e^{\tilde{\mathcal{C}}(\varphi, \theta)}$ , so that the drift term may be rewritten as:

$$\frac{1}{2n} \left( 1 - e^{\tilde{\mathcal{C}}(2\varphi, 2\theta) - 2\tilde{\mathcal{C}}(\varphi, \theta)} \right)$$

The CGF equation now becomes:

$$\frac{\partial}{\partial t} \mathbb{E}[g(\mathbf{q})] = \frac{\partial}{\partial t} \mathbb{E}[\tilde{\mathcal{C}}(\varphi, \theta)] = \mathbb{E} \left[ \tilde{\mathcal{C}}^{(1,0)}(\varphi, \theta) + \tilde{\mathcal{C}}^{(0,1)}(\varphi, \theta) - \tilde{\mathcal{C}}^{(1,0)}(0, 0) - \tilde{\mathcal{C}}^{(0,1)}(0, 0) + \frac{1}{2n} \left( 1 - e^{\tilde{\mathcal{C}}(2\varphi, 2\theta) - 2\tilde{\mathcal{C}}(\varphi, \theta)} \right) \right] \quad (\text{S12})$$

Dropping the tildes and the expectations to reduce clutter, we can now write the full evolutionary model that incorporates selection, mutation, recombination (main text) and drift:

$$\frac{\partial}{\partial t} \mathcal{C}(\varphi, \theta) = \mathcal{C}^{(1,0)}(\varphi, \theta) + \mathcal{C}^{(0,1)}(\varphi, \theta) - \mathcal{C}^{(1,0)}(0, 0) - \mathcal{C}^{(0,1)}(0, 0) + R(e^{\mathcal{C}(\varphi, 0) + \mathcal{C}(0, \theta) - \mathcal{C}(\varphi, \theta)} - 1) + \frac{1}{2n} \left( 1 - e^{\mathcal{C}(2\varphi, 2\theta) - 2\mathcal{C}(\varphi, \theta)} \right) \quad (\text{S13})$$

where  $R$  is recombination rate.

We suspect this equation can no longer be solved because of the non-local arguments in the drift term (the  $2\varphi$  and  $2\theta$ ), but we can immediately see how drift affects the covariance by taking derivatives with respect to  $\varphi$  and  $\theta$  and setting these equal to zero:

$$\partial_t \sigma_{XY}(t) = \kappa_{1,2}(t) + \kappa_{2,1}(t) - \frac{1}{n} \sigma_{XY}(t)$$

where  $\kappa_{i,j}(t)$  is the  $(i, j)^{th}$  joint cumulant of  $X$  and  $Y$  at time  $t$ , and we recall  $\sigma_{XY}(t) = \kappa_{1,1}(t)$ . This shows that drift will tend to weakly push covariance towards zero from either side.

Now writing out the full equation for covariance dynamics, with selection, mutation, drift and recombination, we have:

$$\partial_t \sigma_{XY}(t) = \kappa_{1,2}(t) + \kappa_{2,1}(t) + U m_{1,1} - \left( R + \frac{1}{n} \right) \sigma_{XY}(t)$$

where  $m_{1,1} = \mathbb{E}(\Delta X \Delta Y)$  and  $\Delta X$  and  $\Delta Y$  are changes in fitness due to mutation; i.e.,  $m_{1,1}$  is the first joint moment of the DFE; and  $R$  is recombination rate.

In Ewen's book [S48],  $n$  is interpreted as the number of alleles plus one:  $n = K - 1$ , where  $K$  is the number of alleles. And if you start with an asexual population in which all individuals have different fitnesses, then in effect you have  $K = N$  alleles, where  $N$  is population size, so that  $n = N - 1 \approx N$ .

The foregoing developments allow us to study the effects of drift in isolation:

$$\partial_t \sigma_{XY}(t) = -\frac{1}{n} \sigma_{XY}(t) ,$$

giving rise to the prediction that under drift only,

$$\sigma_{XY}(t) = \sigma_{XY}(0) e^{-t/n} . \quad (\text{S14})$$

We compare these predictions with simulations in Fig S2. Simulations were individual-based and stochastic; they started with a population that was heterogeneous in with  $\sigma_{XY}(0) = -0.025$ , and proceeded with no selection and no mutation. Results of the simulations support the interpretation of  $n$  as  $N$ .

### S3. ONE GENERATION OF SELECTIVE SORTING

We recall that  $\sigma_{XY} = \kappa_{1,1}$ ,  $\sigma_X^2 = \kappa_{2,0}$ , and  $\sigma_Y^2 = \kappa_{0,2}$ . We are interested in the changes in these quantities, especially in the change in  $\kappa_{1,1}$  over the course of a single generation. These results are expressed most compactly in terms of cumulants, so we use the cumulant notation, recalling that  $\kappa_{i,j}$  is the  $(i, j)^{th}$  cumulant in  $X$  and  $Y$ .

Change in covariance over time interval  $(t, t + \delta t)$  can be written as follows:

$$\kappa_{1,1}(t + \delta t) = \kappa_{1,1}(t) + \delta t \kappa'_{1,1}(t) + \frac{1}{2} \delta t^2 \kappa''_{1,1}(t) + \mathcal{O}(\delta t^3) + \dots$$

We define total change in covariance between  $t = 0$  and time  $t = \delta t$  to be:

$$\begin{aligned} \Delta \kappa_{1,1} &= \kappa_{1,1}(\delta t) - \kappa_{1,1}(0) \\ &= \delta t \kappa'_{1,1}(0) + \frac{1}{2} \delta t^2 \kappa''_{1,1}(0) + \mathcal{O}(\delta t^3) + \dots \end{aligned} \quad (\text{S15})$$

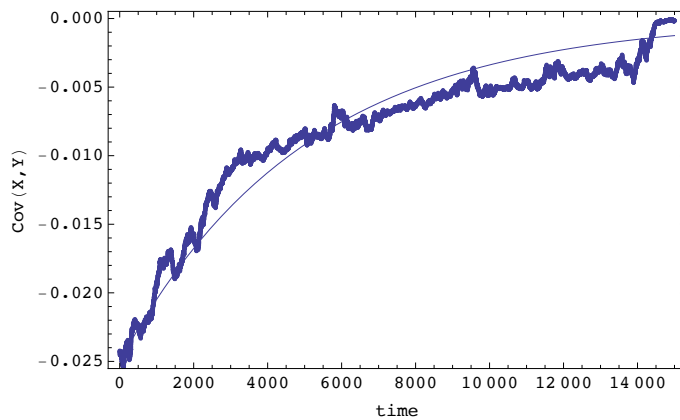


FIG. S2. Covariance dynamics under drift only. Solid curve plots the predicted dynamics given by Eq (S14) and assuming  $n = N$ ; points plot the averages of 50 simulations. The initial state of the population was a set of 5000  $(X, Y)$  pairs ( $N = 5000$ ) drawn at random from a bivariate normal distribution with  $\sigma_{XY}(0) = -0.025$ .

We will see that:

$$\kappa'_{1,1}(0) = NG_1 + D \quad (\text{S16})$$

$$\kappa''_{1,1}(0) = NG_2 + G \quad (\text{S17})$$

where  $G$ ,  $NG$  and  $D$  refer to “gaussian”, “non-gaussian” and “drift” components, respectively. We note that Eq (S16) has no gaussian component; likewise, (S17) has no drift component. Gaussian component are functions of cumulants  $\kappa_{i,j}$  where  $i + j \leq 2$ ; non-gaussian components are functions of cumulants  $\kappa_{i,j}$  where  $i + j > 2$ .

To de-clutter the notation here, in what follows we let  $\kappa_{i,j} = \kappa_{i,j}(0)$ . The components come directly from Eq (S12); they are:

$$\begin{aligned} D &= -k_N \kappa_{1,1} \\ G &= -2k_N (4\kappa_{1,1}^2 + 3\kappa_{1,1}(\kappa_{2,0} + \kappa_{0,2}) + 2\kappa_{2,0}\kappa_{0,2}) \\ NG_1 &= \kappa_{1,2} + \kappa_{2,1} \\ NG_2 &= \kappa_{1,3} + 2\kappa_{2,2} + \kappa_{1,3} \end{aligned}$$

where  $k_N = \frac{N-1}{N^2}$ . Letting  $\delta t = 1$ , i.e., over the course of a single generation, we have:

$$\Delta\kappa_{1,1} = D + NG_1 + \frac{1}{2}G + \frac{1}{2}NG_2,$$

which very accurately predicts the change in covariance when compared to stochastic simulations.

We note the explicit dependence of  $D$  and  $G$  on population size  $N$ , which suggests these components can be significant in small populations but become negligible in large populations. In large populations, it is the non-gaussian components that determine the change in covariance. We also note that  $D$  is a linear function in  $\kappa_{1,1}$  with slope  $-k_N$  and passing through the origin.

Gaussian component,  $G$ , is a concave parabolic function in  $\kappa_{1,1}$  with maximum value:

$$\hat{G} = -k_N \left( 4\kappa_{2,0}\kappa_{0,2} - \frac{9}{8}(\kappa_{2,0} + \kappa_{0,2}) \right)$$

occurring at:

$$\hat{\kappa}_{1,1} = -\frac{3}{8}(\kappa_{2,0} + \kappa_{0,2})$$

If we suppose the variances in  $X$  and  $Y$  are equal (i.e.,  $\sigma_X^2 = \kappa_{2,0} = \kappa_{0,2} = \sigma_Y^2$ ), the foregoing maybe rewritten as maximum:  $\hat{G} = \frac{1}{2}k_N\sigma_X^4$  occurring at:  $\hat{\kappa}_{1,1} = \hat{\sigma}_{XY} = -\frac{3}{4}\sigma_X^2$ . In this case of equal variances, the maximum should thus occur 3/4 of the way between 0 and the minimum value for covariance (in this case, the minimum value is

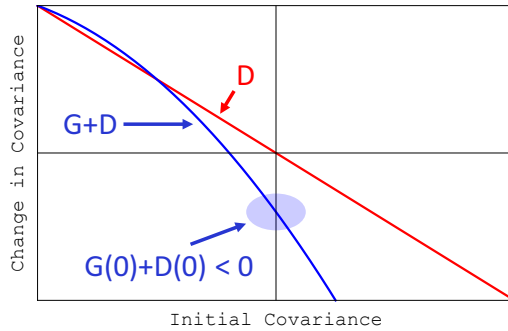


FIG. S3. Schematic plot of change in covariance due to drift ( $D$ , in red) and drift plus the gaussian component of fitness ( $G+D$ , in blue) as a function of initial covariance. Three key features are: 1) Drift alone (red) is linear and symmetric passing through the origin and with an average value of zero; put differently, drift by itself has no net effect on the covariance. 2) Drift plus the gaussian component of fitness is unconditionally negative for independent  $X$  and  $Y$  (i.e., for initial covariance of zero); because the drift component is zero in this case, it is the gaussian component of fitness by itself that creates negative covariance. 3) On average, the gaussian component of fitness is negative; put differently, there is a net tendency to decrease covariance.

$-\sigma_X\sigma_Y = -\sigma_X^2$ ). This can be seen in our simulations and in the schematic plot in Fig S3 by the fact that the  $G+D$  curve rises slightly above the  $D$ -only curve toward the negative end of the  $\kappa_{1,1}$  domain and is otherwise concave, becoming increasingly smaller than  $D$  as  $\kappa_{1,1}$  increases.

Of special interest is the fact that the intercept of  $G$  is unconditionally negative; i.e., when  $\kappa_{1,1} = 0$ , we have:

$$G(0) = -4k_N\kappa_{2,0}\kappa_{0,2} = -4k_N\sigma_X^2\sigma_Y^2$$

Add this to the fact that the intercept of  $D$  is zero, and we have that when  $\kappa_{1,1} = 0$ :

$$G(0) + D(0) = -4k_N\sigma_X^2\sigma_Y^2 \leq 0 \quad (\text{S18})$$

Non-gaussian component  $NG_1 = \kappa_{1,2} + \kappa_{2,1}$  may be rewritten as:

$$NG_1 = \frac{1}{3}(\kappa_3(X+Y) - \kappa_3(X) - \kappa_3(Y))$$

i.e., it is the discrepancy between the skewness of the sum and the sum of the skewnesses of  $X$  and  $Y$ .

## RECOMBINANT ADVANTAGE WITH EPISTASIS

**COROLLARY 8.** For any real number  $\xi$ , let us consider a fitness function of the form  $\phi_\xi(x, y) = ax + by + \xi g(x, y)$ , where  $a, b > 0$  and  $g$  is a function independent of  $\xi$ . Let  $Z(\xi) = \phi_\xi(X_2, Y_2) - \phi_\xi(X_1, Y_1)$ . Assume that for some  $\varepsilon > 0$ ,

$$\mathbb{E} \left[ \sup_{|\xi| < \varepsilon} \frac{|\Delta X \Delta Y|}{|Z(\xi)|} \right] < \infty, \quad (\text{S19})$$

and that  $\mathbb{P}(\Delta X \Delta Y = 0) < 1$ . Then, there is  $\varepsilon_0 \in (0, \varepsilon)$ , such that for all  $\xi \in (-\varepsilon_0, \varepsilon_0)$ , we have

$$\mathbb{E} \left[ \frac{\Delta X \Delta Y}{|Z(\xi)|} \right] < 0.$$

*Proof.* Condition (S19) implies that the function  $h : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$  defined via

$$h(\xi) = \mathbb{E} \left[ \frac{\Delta X \Delta Y}{|Z(\xi)|} \right]$$

is continuous. Moreover, since  $\mathbb{P}(\Delta X \Delta Y = 0) < 1$ , proceeding as in the proof of Proposition ??, we obtain that  $h(0) < 0$ . Hence, by continuity of  $h$ , we infer that there is  $\varepsilon_0 \in (0, \varepsilon)$  such that  $h$  is negative in  $(-\varepsilon_0, \varepsilon_0)$ , which concludes the proof.  $\square$

Let us now focus our attention on the fitness function  $\phi_\xi(X, Y) = aX + bY + \xi XY$  with  $a, b > 0$  and  $\xi \in \mathbb{R}$ . As before, let  $Z(\xi) = \phi_\xi(X_2, Y_2) - \phi_\xi(X_1, Y_1) = (a + \xi Y_1)\Delta X + (b + \xi X_2)\Delta Y$ . The case where the random variables  $(|\Delta X \Delta Y|/|Z(\xi)|)_{\xi \in (-\varepsilon, \varepsilon)}$  are uniformly integrable (i.e. condition (S19) is satisfied) is covered already by Corollary 8. As a counterpart, the next result considers the case where the expectation of  $|\Delta X \Delta Y|/|Z(\xi)|$  is infinite, and provides a simple condition to assure that the expectation of  $\Delta X \Delta Y/|Z(\xi)|$  is negative (in fact, equal to  $-\infty$ ).

**COROLLARY 9.** *Assume that the distribution of  $(X_i, Y_i)$  has finite support, i.e. there is  $K > 0$  such that  $\mathbb{P}(X_i \in [-K, K], Y_i \in [-K, K]) = 1$  and that  $|\xi| < (a \wedge b)/K$ , where  $a \wedge b$  denotes the minimum between  $a$  and  $b$ . If we have*

$$\mathbb{E} \left[ \frac{|\Delta X \Delta Y|}{|Z(\xi)|} \right] = \infty, \quad (\text{S20})$$

then

$$\mathbb{E} \left[ \frac{\Delta X \Delta Y}{|Z(\xi)|} \right] = -\infty.$$

*Proof:* Note first that, if  $|\xi| < (a \wedge b)/K$ , then  $\mathbb{P}(a + \xi Y_1 \geq a - |\xi K|, b + \xi X_2 \geq b - |\xi K|) = 1$ , and hence

$$\begin{aligned} \mathbb{E} \left[ \frac{\Delta X \Delta Y \mathbb{1}_{\Delta X \Delta Y > 0}}{|Z(\xi)|} \right] &= E \left[ \frac{\Delta X \Delta Y \mathbb{1}_{\Delta X \Delta Y > 0}}{|(a + \xi Y_1)\Delta X + (b + \xi X_2)\Delta Y|} \right] \\ &= E \left[ \frac{\Delta X \Delta Y \mathbb{1}_{\Delta X \Delta Y > 0}}{|(a + \xi Y_1)\Delta X| + |(b + \xi X_2)\Delta Y|} \right] \\ &\leq \frac{2}{(a \wedge b) - |\xi K|} E \left[ \frac{|\Delta X \Delta Y|}{|\Delta X| + |\Delta Y|} \right] \\ &\leq \frac{1}{(a \wedge b) - |\xi K|} E \left[ \sqrt{|\Delta X \Delta Y|} \right] \leq \frac{K}{a \wedge b} < \infty. \end{aligned}$$

Therefore, condition (S20) implies that

$$\mathbb{E} \left[ \frac{|\Delta X \Delta Y| \mathbb{1}_{\Delta X \Delta Y < 0}}{|Z(\xi)|} \right] = \infty,$$

and thus,

$$\mathbb{E} \left[ \frac{\Delta X \Delta Y}{|Z(\xi)|} \right] = \mathbb{E} \left[ \frac{\Delta X \Delta Y \mathbb{1}_{\Delta X \Delta Y > 0}}{|Z(\xi)|} \right] - \mathbb{E} \left[ \frac{|\Delta X \Delta Y| \mathbb{1}_{\Delta X \Delta Y < 0}}{|Z(\xi)|} \right] = -\infty,$$

achieving the proof. □

## ASYMPTOTIC MODIFIER FREQUENCY WITH EPISTASIS

Here, we assess the sensitivity to epistasis of our theoretical predictions for asymptotic modifier frequency. To this end, we performed stochastic simulations with epistasis and compared the final modifier frequency with the asymptotic predictions given by Eqs (30) and (31). Results of these simulations are plotted in Fig S4. This figure shows that asymptotic modifier frequency is little-affected by epistasis.

## COVARIANCE DYNAMICS IN A METAPOPULATION: SIMULATIONS

Detailed metapopulation simulations in which natural selection was modeled explicitly in a finite population are plotted in Fig ???. Initially, each deme of the metapopulation is assigned  $n$  distinct genotypes. Natural selection then acts on this variation for 100 generations (the first bout of selection), at which time gaussian noise (mutation) was injected to initialize the next bout of selection, and the process was repeated to mimic several bouts of selection. Figure S5 reveals that within-deme (within-population) covariance becomes negative in the first bout of selection,



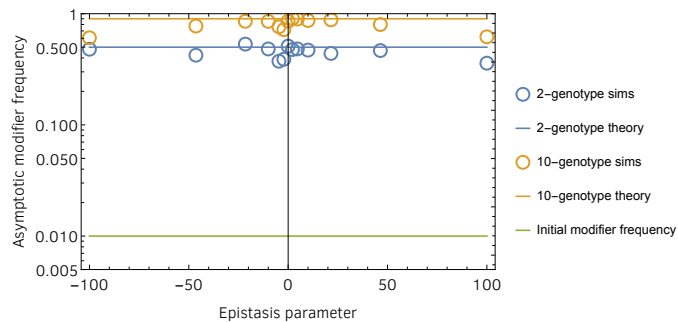


FIG. S4. Effect of epistasis on asymptotic modifier frequency. Open circles plot means of 500 stochastic simulations. Blue and orange lines plot theoretical predictions given by Eq (30) for 2 and 10 genotypes, respectively. Simulations had population size 5000.

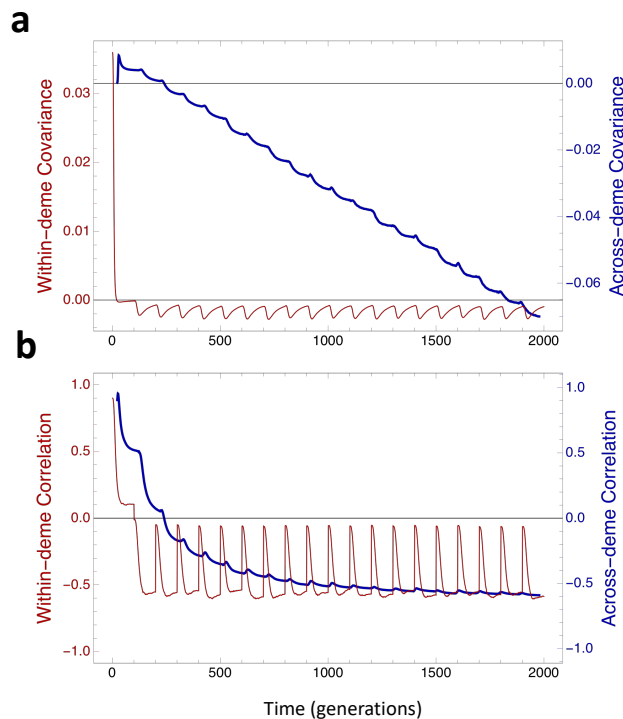


FIG. S5. Covariance and correlation dynamics in a metapopulation. Simulated metapopulations of size  $N = 500$  begin with all individuals being assigned unique genic fitness pairs,  $(X, Y)$ , drawn at random from a common bivariate normal distribution with correlation coefficient 0.9, means  $-0.1$  and variances 0.2. Every 100 generations, uncorrelated gaussian noise was injected as follows:  $X' = X + Q$  and  $Y' = Y + Q$ , where  $Q \sim \mathcal{N}(-.1, .1)$ . Plotted is **a**) mean covariance and **b**) mean correlation, where means were computed from 2000 simulations.

despite starting with a correlation coefficient close to +1, as the theory in our companion paper [S47] predicts. Across-deme covariance also starts close to +1; it is reduced in the first bout of selection but does not become negative until the third bout of selection.

- 
- [S1] P. J. Gerrish, B. Galeota-Sprung, F. Cordero, P. Sniegowski, A. Colato, N. Hengartner, V. Vejalla, J. Chevallier, and B. Ycart, Natural selection and the advantage of recombination, *Phys. Rev. Lett.* **In Review** (2021).  
[S2] P. J. Gerrish, B. Galeota-Sprung, P. Sniegowski, J. Chevallier, and B. Ycart, Natural selection promotes the evolution of recombination 1: among selected genotypes, *Physical Review E* **In Review** (2021).

- [S3] S. P. Otto and T. Lenormand, enResolving the paradox of sex and recombination, *Nat. Rev. Genet.* **3**, 252 (2002).
- [S4] N. H. Barton and B. Charlesworth, enWhy sex and recombination?, *Science* **281**, 1986 (1998).
- [S5] S. P. Otto, enSelective interference and the evolution of sex, *J. Hered.* **112**, 9 (2021).
- [S6] S. p. Otto, The evolutionary enigma of sex, *Am. Nat.* **174**, S1 (2009).
- [S7] J. A. G. M. de Visser and S. F. Elena, enThe evolution of sex: empirical insights into the roles of epistasis and drift, *Nat. Rev. Genet.* **8**, 139 (2007).
- [S8] M. Hartfield and P. D. Keightley, enCurrent hypotheses for the evolution of sex and recombination, *Integr. Zool.* **7**, 192 (2012).
- [S9] S. C. Lee, M. Ni, W. Li, C. Shertz, and J. Heitman, enThe evolution of sex: a perspective from the fungal kingdom, *Microbiol. Mol. Biol. Rev.* **74**, 298 (2010).
- [S10] R. A. Fisher, *The genetical theory of natural selection* (Oxford Clarendon Press, 1930) p. 302.
- [S11] H. J. Muller, Some genetic aspects of sex, *Am. Nat.* **66**, 118 (1932).
- [S12] D. Roze and N. H. Barton, enThe Hill-Robertson effect and the evolution of recombination, *Genetics* **173**, 1793 (2006).
- [S13] W. G. Hill and A. Robertson, enThe effect of linkage on limits to artificial selection, *Genet. Res.* **8**, 269 (1966).
- [S14] N. H. Barton, enLinkage and the limits to natural selection, *Genetics* **140**, 821 (1995).
- [S15] N. H. Barton, enA general model for the evolution of recombination, *Genet. Res.* **65**, 123 (1995).
- [S16] N. H. Barton, enGenetic linkage and natural selection, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 2559 (2010).
- [S17] S. P. Otto and N. H. Barton, enThe evolution of recombination: removing the limits to natural selection, *Genetics* **147**, 879 (1997).
- [S18] J. F. Crow and M. Kimura, Evolution in sexual and asexual populations, *Am. Nat.* **99**, 439 (1965).
- [S19] Y. Kim and H. A. Orr, enAdaptation in sexuals vs. asexuals: clonal interference and the Fisher-Muller model, *Genetics* **171**, 1377 (2005).
- [S20] J. M. Smith and J. Maynard-Smith, *The evolution of sex*, Vol. 4 (Cambridge University Press Cambridge, 1978).
- [S21] J. Maynard Smith, The evolution of sex, .
- [S22] B. Charlesworth, enMutation-selection balance and the evolutionary advantage of sex and recombination, *Genet. Res.* **89**, 451 (1990).
- [S23] S. P. Otto and M. W. Feldman, enDeleterious mutations, variable epistatic interactions, and the evolution of recombination, *Theor. Popul. Biol.* **51**, 134 (1997).
- [S24] A. Blachford and A. F. Agrawal, enAssortative mating for fitness and the evolution of recombination, *Evolution* **60**, 1337 (2006).
- [S25] N. H. Barton and S. P. Otto, enEvolution of recombination due to random drift, *Genetics* **169**, 2353 (2005).
- [S26] S. P. Otto and N. H. Barton, enSelection for recombination in small populations, *Evolution* **55**, 1921 (2001).
- [S27] P. D. Keightley and S. P. Otto, enInterference among deleterious mutations favours sex and recombination in finite populations, *Nature* **443**, 89 (2006).
- [S28] S. P. Otto, enThe evolutionary enigma of sex, *Am. Nat.* **174 Suppl 1**, S1 (2009).
- [S29] J. Felsenstein, enThe evolutionary advantage of recombination, *Genetics* **78**, 737 (1974).
- [S30] M. Slatkin, Linkage disequilibrium—understanding the evolutionary past and mapping the medical future, *Nat. Rev. Genet.* **9**, 477 (2008).
- [S31] G. Martin and L. Roques, enThe nonstationary dynamics of fitness distributions: Asexual model with epistasis and standing variation, *Genetics* **204**, 1541 (2016).
- [S32] M.-E. Gil, F. Hamel, G. Martin, and L. Roques, Mathematical properties of a class of integro-differential models from population genetics, *SIAM J. Appl. Math.* **77**, 1536 (2017).
- [S33] Y. Anciaux, A. Lambert, O. Ronce, L. Roques, and G. Martin, enPopulation persistence under high mutation rate: From evolutionary rescue to lethal mutagenesis, *Evolution* **73**, 1517 (2019).
- [S34] P. J. Gerrish and P. D. Sniegowski, enReal time forecasting of near-future evolution, *J. R. Soc. Interface* **9**, 2268 (2012).
- [S35] R. Bürger, enMoments, cumulants, and polygenic dynamics, *J. Math. Biol.* **30**, 199 (1991).
- [S36] M. Smerlak and A. Youssef, enLimiting fitness distributions in evolutionary dynamics, *J. Theor. Biol.* **416**, 68 (2017).
- [S37] S. P. Otto and T. Lenormand, enResolving the paradox of sex and recombination, *Nat. Rev. Genet.* **3**, 252 (2002).
- [S38] B. Charlesworth and N. H. Barton, Recombination load associated with selection for increased recombination (1996).
- [S39] S. P. Otto, enUnravelling gene interactions, *Nature* **390**, 343 (1997).
- [S40] W. R. Rice, enExperimental tests of the adaptive significance of sexual recombination, *Nat. Rev. Genet.* **3**, 241 (2002).
- [S41] E. Baake and M. Baake, Haldane linearisation done right: Solving the nonlinear recombination equation the easy way, *Discrete & Continuous Dynamical Systems - A* **36**, 6645 (2016).
- [S42] D. Mchale and G. A. Ringwood, enHaldane linearisation of baric algebras, *J. Lond. Math. Soc.* **s2-28**, 17 (1983).
- [S43] M. Baake and E. Baake, An exactly solved model for mutation, recombination and selection, *Canad. J. Math.* **55**, 3 (2003).
- [S44] E. Baake, enMutation and recombination with tight linkage, *J. Math. Biol.* **42**, 455 (2001).
- [S45] M. Esser, S. Probst, and E. Baake, enPartitioning, duality, and linkage disequilibria in the moran model with recombination, *J. Math. Biol.* **73**, 161 (2016).
- [S46] E. Baake, Deterministic and stochastic aspects of single-crossover recombination, in *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010)* (Published by Hindustan Book Agency (HBA), India. WSPC Distribute for All Markets Except in India, 2011) pp. 3037–3053.
- [S47] P. J. Gerrish, F. Cordero, B. Galeota-Sprung, A. Colato, V. Vejalla, and P. Sniegowski, Natural selection promotes the evolution of recombination 2: during the selective process, *Physical Review E* **In Review** (2021).

[S48] W. J. Ewens, *Mathematical Population Genetics: I. Theoretical Introduction* (Springer, New York, NY, 2004).