SUPPLEMENTARY INFORMATION

Supplementary Information 1: Stability of three-way associations with multilocus underdominance/assortment

We will assume that fitness acts in the same way at each locus, and so depends only on the number of heterozygous loci. However, linkage relations will in general differ amongst loci: indeed, they must differ if the loci are linked and there is no interference between crossovers. In this model, we have $s_1 = s_2 = s_3$ and $s_{12} = s_{23} = s_{13}$. Close to LE, and assuming symmetrical allele frequencies, the recursions for linkage disequilibria are special cases of Eq. 15:

$$\begin{aligned} D_{12}^{*} &= D_{12}\left((1 - r_{1,2})\left(1 + \frac{s_{12}}{16}\right) + \frac{s_1}{2} r_{1,2}\right) \\ D_{123}^{*} &= D_{123}\left((1 - r_{123})\left(1 + \frac{s_{123}}{64}\right) + r_{123}\left(\frac{s_1}{4} + \frac{s_{12}}{16}\right)\right) \end{aligned} \tag{35}$$

and similarly for the other D_U. With no interference, the chance of a crossover between any of the three loci is $r_{123} = 1 - (1 - r_{1,2})(1 - r_{2,3})$, assuming that the loci are in order 1,2,3.

The conditions for instability are

$$\begin{aligned} r_{1,2} &< \frac{s_{12}}{16 - 8 s_1 + s_{12}} \text{ for } D_{12} \text{ unstable} \\ r_{123} &< \frac{s_{123}}{64 - 16 s_1 - 4 s_{12} + s_{123}} \text{ for } D_{123} \text{ unstable} \end{aligned} \tag{36}$$

We can get a concrete understanding of the model by relating the selection coefficients to the fitnesses W_i of individuals with i heterozygous loci:

$$\begin{aligned} W_0 &= 1 + \frac{3}{4} s_1 + \frac{3}{16} s_{1,2} + \frac{1}{64} s_{1,2,3} > 0 \\ W_1 &= 1 + \frac{1}{4} s_1 - \frac{1}{16} s_{1,2} - \frac{1}{64} s_{1,2,3} > 0 \\ W_2 &= 1 - \frac{1}{4} s_1 - \frac{1}{16} s_{1,2} + \frac{1}{64} s_{1,2,3} > 0 \\ W_3 &= 1 - \frac{3}{4} s_1 + \frac{3}{16} s_{1,2} - \frac{1}{64} s_{1,2,3} > 0 \end{aligned} \tag{37}$$

and with $\overline{W} = \frac{1}{8}(W_0 + 3W_1 + 3W_2 + W_3) = 1$ always.

giving:

$$\begin{aligned} s_1 &= 4 - W_1 - 2W_2 - W_3 \\ s_{12} &= 8(2 - W_1 - W_2) \\ s_{123} &= 16(4 - 3W_1 - W_3) \end{aligned} \tag{38}$$

We can now rewrite the thresholds as:

$$r_{1,2} < \frac{(2 - W_1 - W_2)}{(W_2 + W_3)} \text{ for } D_{12} \text{ unstable}$$
$$r_{123} < \frac{4 - 3W_1 - W_3}{4 W_2} \text{ for } D_{123} \text{ unstable} \qquad (39)$$

We can find parameter combinations that give all four possible stability regimes. We focus on the extreme case where complete reproductive isolation can be achieved ($W_3 = 0$), and then vary the remaining parameters W_1, W_2; numerical examples are given in Fig. 2. If the pairwise $D's$ are unstable, but the third-order association is stable, then we expect the final outcome to be fixation of two complementary genotypes (000 and 111, 001 and 110, etc.), since in these states all the pairwise $D's$ are maximal and the third-order association is zero (Fig. 3a). When both second and third-order associations are unstable, D_{123} may increase initially, but then decrease as pairwise $D's$ increase. The most interesting case is when D_{123} is unstable, but the pairwise D's are stable (Fig. 3c). Then, four complementary genotypes (e.g. {0,0,1},{0,1,0},{1,0,0},{1,1,1}) predominate, in equal proportions, and any combination of these leads to precisely two loci being heterozygous. Note, however, that the stability at linkage equilibrium does not necessarily tell us the final outcome (e.g., Fig. 3b).

Supplementary Information 2: Analysis of Felsenstein's (1981) model, assuming $p_b = p_c = \frac{1}{2}$

We can write the net contribution of each genotype to the next generation as the average over the two niches. The full expression is complex, but simplifies when we assume equal allele frequencies at loci b, c:

$$W = \left(1 + \frac{\alpha}{p_a\, q_a}\, \zeta_a\, \zeta_a^*\right)\, (1 + \gamma\, (\zeta_b\, \zeta_c - D_{bc}))\, (1 + \gamma\, (\zeta_b^*\, \zeta_c^* - D_{bc})) \qquad (40)$$

where $\gamma = \frac{\epsilon}{\bar{w}_I}$, $\bar{w}_I = \bar{w}_{II} = 1 + s + \epsilon \left(\frac{1}{4} + D_{bc}\right)$, $p_b = p_c = \frac{1}{2}$.

$\bar{w}_I = \bar{w}_{II}$ denote the average fitness within each niche. The parameter ϵ measures the deviation from additivity within each niche, and would equal s^2 with multiplicative fitness. The dynamics depend on the scaled coefficient of epistasis, $\gamma = \frac{\epsilon}{\bar{w}_I}$. Thus, the following analysis applies to any two-locus model with net epistasis of strength γ, provided that allele frequencies are symmetric ($p_b = p_c = \frac{1}{2}$): Felsenstein's (1981) two-niche model is one of many ways of maintaining a stable polymorphism despite disruptive selection.

Rather than working with the full selection coefficients that give the total effect of selection on haploid and diploid stages (Eq. 1), it is easier to separate the effects of viability selection on haploids, and of assortment amongst haploids (or, underdominance of diploids). Haploid viability is written in terms of a set of selection coefficients a_X. Using the haploid equivalent of Eq. 2, $D_U^* = D_U + \sum_X a_X\, (D_{UX} - D_U\, D_X)$. In this case, the only selection coefficient for haploid viability is $a_{bc} = \gamma$, and so we find the pairwise associations after viability selection:

$$\begin{aligned} D_{ab}' &= D_{ab} + \gamma\, (h_b\, D_{ac} - D_{bc}\, D_{ab}) \\ D_{ac}' &= D_{ac} + \gamma\, (h_c\, D_{ab} - D_{bc}\, D_{ac}) \\ D_{bc}' &= D_{bc} + \gamma\, (h_b\, h_c - D_{bc}^2) \end{aligned} \qquad (41)$$

where we have used the reduction formula, Eq. 3, and assumed $p_b = p_c = \frac{1}{2}$. The product of allele frequencies is denoted by $h_x = p_x\, q_x = \frac{1}{4}$; it is convenient still to express the formulae in terms of h_b, h_c, so as to keep track of the sources of the various terms.

From Eqs. 2, 4, the effect of assortment and recombination is simply:

$$\begin{aligned} D_{ab}'' &= (1-\rho_b)\, D_{ab}' \\ D_{ac}'' &= (1-\rho_c)\, D_{ac}' \\ D_{bc}'' &= (1-r_{b,c})\, D_{bc}' + r_{b,c}\, \frac{D_{ab}'\, D_{ac}'}{h_a} \end{aligned} \tag{42}$$

where $\rho_x = r_{a,x}\,(1-\alpha)$ measures the effective rate of recombination between the assorting locus and locus $x \in \{b, c\}$, which is reduced by assortment or underdominance at locus a. Note that these equations apply for arbitrary $D_{ab},\ D_{ac}$: we have not made any linear approximation, and yet the recursions for $D_{ab},\ D_{ac}$ are linear. Combining Eqs. 41, 42, these recursions can be rewritten in matrix form:

$$\begin{pmatrix} D_{ab}'' \\ D_{ac}'' \end{pmatrix} = \begin{pmatrix} (1-\rho_b)\,(1-\gamma\, D_{bc}) & (1-\rho_b)\,\gamma h_b \\ (1-\rho_c)\,\gamma h_c & (1-\rho_c)\,(1-\gamma\, D_{bc}) \end{pmatrix} \cdot \begin{pmatrix} D_{ab} \\ D_{ac} \end{pmatrix} \tag{43}$$

Crucially, this pair of linear equations applies *in general*, not just for small D_{ab}- although D_{bc} will change as a function of the $D_{ab},\ D_{ac}$. We can therefore identify two kinds of equilibria. There is always a symmetrical equilibrium with $D_{ab} = D_{ac} = 0$. This will be stable if the leading eigenvalue of the matrix in Eq. 43, λ, has magnitude less than 1. If it is unstable, then associations with locus a will grow, and as they do so, the association between the selected loci, D_{bc}, will also grow. We see from Eqs. 41 that as D_{bc} increases, the rates of growth of $D_{ab},\ D_{ac}$ decrease. At some point, the leading eigenvalue falls to precisely λ=1, and we have an alternative asymmetric equilibrium in which both pairwise associations are either positive, or negative.

The leading eigenvalue can be written explicitly as the solution to a quadratic equation, and takes a simple form when loci $b,\ c$ are equivalent (i.e., when they are unlinked, or when a is midway between $b,\ c$):

$$\lambda = (1-\rho)\,(1+\gamma\,(h - D_{bc})) \tag{44}$$

The association D_{bc} is given by setting $D_{ab},\ D_{ac}$ to zero in Eqs. 41, 42:

$$D_{bc} = D_{bc}'' = (1-r_{b,c})\,(D_{bc} + \gamma\,(h_b\, h_c - D_{bc}^2)) \tag{45}$$

This has the solution:

$$\begin{aligned} D_{bc} &= \frac{r}{2\,(1-r)\,\gamma}\left(\sqrt{1+\frac{\gamma^2\,(1-r)^2}{4\,r^2}} - 1\right) = \\ &\frac{1}{4}\left(\sqrt{(1-r) + 4\,\frac{r^2}{\epsilon^2}\left(1+s+\frac{\epsilon}{4}\right)^2} - 4\,\frac{r}{\epsilon}\left(1+s+\frac{\epsilon}{4}\right)\right) \end{aligned} \tag{46}$$

where we have dropped the subscript from $r_{b,c}$ for compactness. It may be convenient to measure the strength of epistasis by either ϵ or $\gamma = \epsilon/(1 + s + \epsilon(\frac{1}{4} + D_{bc}))$, and so we give two forms for this solution.

This equilibrium is unstable to growth of $D_{ab},\ D_{ac}$ if λ>1. For $\rho_b = \rho_c = \rho$, this condition can be rewritten as:

$$\frac{\rho}{1-\rho} < \gamma\left(\frac{1}{4} - D_{bc}\right) \tag{47}$$

for instability. This condition compares the strength of D_{bc} (which depends on γ and r_{bc}) with the effective rates of recombination between a and $b,\ c$. By substituting for $D_{b,c}$ from Eq. 46, the condition for instability is given explicitly as:

$$\frac{\rho\ (r_{b,c} - \rho)}{(1-\rho)\ (r_{b,c}\ (1+\rho) - 2\ \rho)} < \frac{\gamma}{4} \tag{48}$$

Figure 4 shows that for the two incompatibilities to become coupled, there must be sufficiently strong assortment, and sufficiently strong epistasis, relative to the rates of recombination.

Supplementary Information 3: Stability of Felsenstein's (1981) three-locus model

Allele frequency at the assorting locus, p_a can be arbitrary if there is no association with other loci, but once associations do build up, p_a will tend towards $\frac{1}{2}$. First, suppose that $D_{ab},\ D_{ac} = 0$. Then, D_{abc} can be non-zero, and p_a will change if D_{abc} is non-zero. Perturbations to D_{abc} always decay to zero, but as they decay, they lead to a shift in p_a. So, we can justify assuming that D_{abc}=0, but can keep p_a arbitrary. Now, suppose that $D_{ab},\ D_{ac}$ are non-zero. With arbitrary p_a, D_{abc} then becomes non-zero, p_a is selected to converge to $\frac{1}{2}$, and thereafter D_{abc} converges to 0. However, the transient stage becomes complicated, because the allele frequencies $p_b,\ p_c$ temporarily deviate from $\frac{1}{2}$. Thus, it is consistent to assume that all allele frequencies are at $\frac{1}{2}$. This stability analysis, detailed below, justifies setting all allele frequencies to $\frac{1}{2}$, on the grounds that while p_a is in principle arbitrary, perturbations will send it towards $p_a = \frac{1}{2}$.

The change in allele frequencies is obtained from Eq. 2:

$$\begin{aligned}
\triangle p_a &= D_{abc}\,(\tilde{a}_{bc} + \tilde{a}_{abc}\,h_a) \;-\; a_{abc,abc}\,(p_a - q_a)\,D_{abc}^2 \\
\triangle p_b &= D_{abc}\,(a_{a,abc}\,D_{ab} + a_{abc,abc}\,h_b\,D_{ac}) \\
\triangle p_c &= D_{abc}\,(a_{a,abc}\,D_{ac} + a_{abc,abc}\,h_c\,D_{ab})
\end{aligned} \qquad (49)$$

Here, $\tilde{a}_{bc} = a_{bc,\emptyset} + a_{bc,bc}\,D_{bc}$, $\tilde{a}_{abc} = a_{abc,a} + a_{abc,abc}\,D_{bc}$. The last term in the first equation arises because ($a_{abc,abc}\,D_{aabc}\,D_{abc} = a_{abc,abc}(h_a\,D_{bc} - (p_a - q_a)\,D_{abc})\,D_{abc}$). We denote the product of allele frequencies by $h_a = p_a\,q_a$ etc. Since we assume $p_b = p_c = \frac{1}{2}$ throughout, $h_b = h_c = \frac{1}{4}$. However, it is convenient to retain the h_b, h_c since this makes the origin of the terms clearer.

Because assortative mating is defined to be cost-free, allele frequencies at locus a only change through associations with the selected loci. Because the latter are at a stable equilibrium with $p_b = p_c = \frac{1}{2}$, these loci experience no directional selection (i.e., $a_{b,\emptyset} = a_{c,\emptyset} \ldots = 0$). Therefore, locus a only experiences indirect selection via the three way association D_{abc}. This selection acts in two ways. First, if allele a^1 is associated with the fitter *combinations* of alleles $b^1\,c^1$ and $b^0\,c^0$ (i.e. if $\tilde{a}_{bc} > 0$ and $D_{abc} > 0$), then allele a^1 will increase ($\Delta p_a = D_{abc}\,\tilde{a}_{bc} > 0$). The second term in the equation for Δp_a is due to the third-order selection coefficient $\tilde{a}_{abc} = a_{abc,a} + a_{abc,abc}\,D_{bc}$, which arises through the multiplication of fitness components in Eq. 40, and has the same direction as the leading term This coefficient involves an association between an allele a^X in one individual, and alleles $a^X\,b^1\,c^1$, $a^X\,b^0\,c^0$ in the other, represented by $a_{a,abc}$. Thus, if allele a^1 (say) is associated with fitter combinations at loci $b,\ c$ ($D_{abc} > 0$), and if allele a^1 in another individual tends to mate with those fitter combinations ($a_{a,abc}$), then allele a^1 tends to increase. The changes in allele frequencies at the selected loci also arise through this

second effect (i.e., through terms such as $D_{abc}\, a_{a,abc}\, D_{ab}$). To summarise: alleles at the assortment locus increase if they are associated with fit combinations of alleles at the selected loci; and alleles at all three loci increase via preferences of alleles in one haploid individual for fit combinations in another individual.

We see that allele frequencies at loci $b,\ c$ will stay fixed at $\frac{1}{2}$ if either D_{abc} is zero, or if D_{ab} and D_{ac} are zero. In the following, we examine two kinds of deviation from the symmetric equilibrium: changes in $p_a,\ D_{abc}$, with $D_{ab} = D_{ac} = 0$, or conversely, changes in pairwise associations, with all allele frequencies and D_{abc} symmetrical. In both cases, we maintain $p_b = p_c = \frac{1}{2}$, and $D_{bc} > 0$ because of epistatic selection on loci $b,\ c$.

First, assume that $D_{ab}, D_{ac} = 0$. The recursion for p_a is given in Eqs. 49. For D_{bc}:

$$D'_{bc} = (1 - r_{b,c}) \\ (D_{bc} + \tilde{a}_{bc}\, (h_b\, h_c - D_{bc}^2) + (a_{a,abc} - a_{abc,abc}\, D_{bc})\, D_{abc}^2) \qquad (50)$$

The pairwise association between the selected loci is built up by epistasis between them ($\tilde{a}_{bc}$), and broken down by recombination $(1 - r_{b,c})$. In the presence of a three-way association of either sign, it is also generated by the higher-order coefficient $a_{a,abc}$.

For the three-way association:

$$\begin{aligned} D'_{abc} = {} & r_{abc,\emptyset} \\ & (D_{abc} - \tilde{a}_{bc}\, D_{abc}\, D_{bc} + a_{a,abc}\, (h_a\, D_{bc} - (p_a - q_a)\, D_{abc}) \\ & D_{abc} + a_{abc,abc}\, (h_a\, h_b\, h_c\, D_{abc} - D_{abc}^3)) + \\ & r_{a,bc}\, (a_{bc,\emptyset}\, D_{abc}\, D_{bc} + a_{bc,bc}\, D_{abc}\, h_b\, h_c + a_{a,a}\, h_a\, D_{abc} + \\ & a_{abc,a}\, (h_a\, D_{bc} - (p_a - q_a)\, D_{abc})\, D_{abc}) - \Delta p_a\, D^*_{bc} \end{aligned} \qquad (51)$$

This can be rearranged to give:

$$\begin{aligned} D'_{abc} = {} & r_{abc,\emptyset}\, (D_{abc} - \tilde{a}_{bc}\, D_{abc}\, D_{bc} + \\ & a_{abc,abc}\, (h_a\, h_b\, h_c\, D_{abc} - D_{abc}^3)) + \\ & r_{a,bc}\, (a_{bc,\emptyset}\, D_{abc}\, D_{bc} + a_{bc,bc}\, D_{abc}\, h_b\, h_c + \\ & a_{a,a}\, h_a\, D_{abc}) - \\ & (1 - r_{b,c})\, D_{abc}\, (\tilde{a}_{abc}\, (p_a - q_a)\, D_{abc} + \tilde{a}_{bc}\, D_{bc} + \\ & a_{abc,abc}\, h_a\, D_{bc}^2) - \Delta p_a\, (1 - r_{b,c}) \\ & (\tilde{a}_{bc}\, (h_b\, h_c - D_{bc}^2) + (a_{a,abc} - a_{abc,abc}\, D_{bc})\, D_{abc}^2) \end{aligned} \qquad (52)$$

In terms of the parameters $\alpha,\ s,\ \gamma$:

$$\bar{W}\, \Delta p_a = (1 + \alpha)\, \gamma\, D_{abc} - \gamma^2\, \frac{\alpha}{p_a\, q_a}\, {D_{abc}}^2\, (p_a - q_a)$$

$$\bar{W}\, D'_{bc} =$$

$$(1 - r_{b,c}) \left(\frac{\gamma}{16} + \left(D_{bc} + \frac{\alpha}{p_a\, q_a}\, \gamma\, D_{abc}^2 \right) (1 - \gamma\, D_{bc}) \right)$$

$$\bar{W}\, D'_{abc} = (1 - r_{b,c})\, D_{abc}\, \frac{(1 + s)\, (1 + s + \frac{\epsilon}{2})}{(1 + s + \epsilon\, (\frac{1}{4} + D_{bc}))^2}$$

$$\left(A + \frac{D_{abc}}{\bar{W}}\, \frac{\alpha}{p_1\, q_1}\, \gamma\, ((q_1 - p_1) - \gamma\, (1 + \alpha)\, D_{abc}) \right)$$

where A=1 - r_{ab} (1-α), $\overline{W} = 1 + \frac{\alpha}{p_a\, q_a}\, \gamma^2\, D_{abc}^2$.

The equilibrium at D_{abc}=0 is stable if :

$$(1 - r_{b,c})\, \frac{(1 + s)\, (1 + s + \frac{\epsilon}{2})}{(1 + s + \epsilon\, (\frac{1}{4} + D_{bc}))^2}\, (1 - r_{ab}\, (1 - \alpha)) < 1 \qquad (54)$$

Now, because $0 < D_{bc} < \frac{1}{4}$, the middle term in Eq. 54 is necessarily less than one, and so D_{abc}=0 is always stable. If a transient D_{abc} is introduced, it will decay to zero, but while it decays, the frequency of the assortment alleles will shift (Fig. XXX). In the following, therefore, we will assume $D_{abc} = 0$ and p_a arbitrary. Then, the equilibrium D_{bc} is, from Eq. 50:

$$D_{bc} = \frac{r}{2\, (1 - r)\, \gamma} \left(\sqrt{1 + \frac{\gamma^2\, (1 - r)^2}{4\, r^2}} - 1 \right) = \frac{1}{4} \left(\sqrt{(1 - r) + 4\, \frac{r^2}{\epsilon^2} \left(1 + s + \frac{\epsilon}{4}\right)^2} - 4\, \frac{r}{\epsilon} \left(1 + s + \frac{\epsilon}{4}\right) \right) \qquad (55)$$

(Note that it may be convenient to measure the strength of epistasis by either ϵ or $\gamma = \epsilon/(1 + s + \epsilon(\frac{1}{4} + D_{bc}))$).

We can now examine the stability of the pairwise associations, assuming that $D_{abc} = 0$, $p_b = p_c = \frac{1}{2}$. The association between the selected loci b, c is given by

$$D'_{bc} = (1 - r_{b,c}) \left(D_{bc} + \gamma \left(\frac{1}{16} - D_{bc}^2 \right) \right) + \alpha\, r_{b,c} \left((D_{ab}^2 + D_{ac}^2)\, \gamma\, (1 - D_{bc}\, \gamma) + 4\, D_{ab}\, D_{ac} \left((1 - \gamma\, D_{bc})^2 + \frac{\gamma^2}{16} \right) \right) \qquad (56)$$

This simplifies to Eq. 50, as expected, if D_{ab}, $D_{ac} = 0$ (first term in Eq. 56). The second term in Eq. 56 is positive for $D_{ab} > 0$, $D_{ac} > 0$, showing that pairwise associations with an assorting locus increase the strength of linkage disequilibrium D_{bc}.

With arbitrary p_a, the three-way association becomes non-zero, and so we need to consider the full system. The change in allele frequencies is:

$$\begin{aligned}
\bar{w}\,\Delta p_a &= \gamma\,(1+\alpha)\,D_{abc} - (p_a - q_a)\,\frac{\alpha}{p_a\,q_a}\,\gamma^2\,D_{abc}^2 \\
\bar{w}\,\Delta p_b &= \frac{\alpha}{p_a\,q_a}\,\gamma\,D_{abc}\left(D_{ab}\,(1-\gamma D_{bc}) + \frac{\gamma}{4}\,D_{ac}\right) \\
\bar{w}\,\Delta p_c &= \frac{\alpha}{p_a\,q_a}\,\gamma\,D_{abc}\left(D_{ac}\,(1-\gamma D_{bc}) + \frac{\gamma}{4}\,D_{ab}\right)
\end{aligned} \tag{57}$$

We see that three-way associations of either sign impose a stabilising force on p_a, through the term $(p_a - q_a)\,\frac{\alpha}{p_a\,q_a}\,\gamma^2\,D_{abc}^2$. Numerical work suggests that D_{abc} always tends to zero when p_a is $\frac{1}{2}$, and that together, p_a tends to 1/2 and D_{abc} tends to zero. This drastically simplifies the dynamics. One can show that near the symmetric equilibrium, D_{abc} is stable, and so can at most exert a transient force on p_a. Moreover, whenever D_{abc} is nonzero, p_a will be pushed towards $\frac{1}{2}$.

Supplementary Information 4: Coupling between multiple pairs of interacting loci

By analogy with Eqs. 11, 40, we define fitness as an arbitrary function of the number of pairs that are in coupling (00 or 11) or repulsion (01 or 10). We assume random mating, so that selection acts independently on the haploid genomes. Thus, relative fitness is the product of two similar terms for relative viability of haploids, which is written in terms of the general selection coefficients a_U:

$$\frac{W}{\bar{W}} = \frac{V}{\bar{V}} \frac{V^*}{\bar{V}} \text{ where } \frac{V}{\bar{V}} = 1 + \sum_{U \subseteq \Omega} a_U \, (\zeta_U - D_U) \tag{58}$$

Ω denotes the set of pairs of interacting loci, $\{i, \; j, \; \ldots\}$, and i denotes a *pair* of loci, so that $\zeta_i = \zeta_{i_1} \, \zeta_{i_2}$ and $\zeta_U = \prod_{i \in U} \zeta_i$. We assume that viability depends only on the set of pairs that is in coupling or repulsion, and so the only coefficients are those which involve *pairs* of interacting loci. Thus, $a_i \equiv a_{i_1 \, i_2}$, $a_{ij} \equiv a_{i_1 \, i_2 \, j_1 \, j_2}$ will in general be non-zero, but coefficients such as $a_{i_1 \, j_1}$ or $a_{i_1 \, i_2 \, j_1}$ will not contribute.

In the special case where the effects of different incompatibilities multiply, we can write viability as:

$$V = \prod_{i \in \Omega} (1 + \gamma_i \, (\zeta_i - D_i)) \tag{59}$$

If there were no associations between different incompatibilities (i.e., if $D_U = \prod_{i \in \Omega} D_i$, so that state of each pair is statistically independent), then the mean viability would be $\overline{V} = 1$. However, in general we have:

$$\bar{V} = 1 + \sum_{U \subseteq \Omega} \left(\prod_{i \in U} \gamma_i \right) \left(D_U - \left(\prod_{i \in \Omega} D_i \right) \right) \prod_{i \in \Omega \setminus U} (1 - \gamma_i \, D_i) \tag{60}$$

where $D_U - \prod_{i \in \Omega} D_i$ measures the association amongst incompatibilities. The mean viability at linkage equilibrium is just $\overline{V}_{\mathrm{LE}} = \prod_{i \in \Omega} (1 - \gamma_i \, D_i)$. By identifying the coefficients of ζ_U in Eqs. 58, 59, we see that in this special case where the effects of incompatibilities are multiplicative, we have:

$$a_U = \phi \left(\prod_{i \in U} \frac{\gamma_i}{1 - \gamma_i \, D_i} \right) \tag{61}$$

where $\phi = \bar{V}_{\mathrm{LE}} \, / \, \bar{V}$ is the reduction in viability which would be caused if the population suddenly moved to linkage equilibrium; $| U |$ is the number of pairs in the set U. Equation 61 implies a special relationship amongst the a_U:

$$a_U = \phi^{1-|U|} \left(\prod_{i \in U} a_i \right) \tag{62}$$

In general, however, multiple incompatibilities may interact. For example, the effects of increasing numbers of incompatibilities may accumulate faster or slower than multiplicatively. However, allowing these kinds of interaction between incompatibilities does not amount to assuming a completely general model of epistasis. Our key assumption is that the effect of a pair of loci $i = \{i_1, i_2\}$ on fitness depends only on $\zeta_i \equiv \zeta_{i_1} \zeta_{i_2}$, and not on ζ_{i_1}, ζ_{i_2} separately. Combined with the assumption of equal allele frequencies, this greatly simplifies the recursions. (Indeed, if allele frequencies are symmetrical, and at equilibrium, then there can be no directional selection, and hence no terms ζ_{i_1}).

The change in allele frequency is:

$$\triangle p_{i_1} = \sum_U a_U D_{i_1 U} \tag{63}$$

We can separate the selection coefficients a_U into those that include the pair i, and those that do not. Then, $D_{i_1 i_1 i_2 U}$ simplifies to $h_{i_1} D_{i_2 U}$ when $p_{i_1} = p_{i_2} = \frac{1}{2} \; \forall \, i$. On this assumption of symmetrical allele frequencies, $h_{i_1} = \frac{1}{4}$, but it is convenient to leave it as h_{i_1}.

$$\triangle p_{i_1} = \sum_{U \subseteq \Omega \setminus i} a_U D_{i_1 U} + \sum_{U \subseteq \Omega \setminus i} a_{i_1 U} h_{i_1} D_{i_2 U} \tag{64}$$

Therefore, allele frequencies will be at equilibrium if $D_{i_1 U} = D_{i_2 U} = 0 \; \forall \, U \subseteq \Omega$.

We denote associations after viability selection by D'_U, and after selection and recombination, by D''_U. The association within pairs changes as:

$$\begin{aligned} D'_i &= D_i + \sum_{U \subseteq \Omega} a_U \left(D_{iU} - D_i D_U \right) \\ D''_i &= (1 - r_i) D'_i \end{aligned} \tag{65}$$

Separating the cases where U does or does not contain i:

$$D'_i = D_i + \sum_{U \subseteq \Omega \setminus i} a_{iU} \left(h_i D_U - D_i D_{iU} \right) + \sum_{U \subseteq \Omega \setminus i} a_U \left(D_{iU} - D_i D_U \right) \tag{66}$$

In general, this has no closed-form solution, because it depends on higher-order associations, D_{iU}; the recursions for the D_{iU} depend on both higher and lower-order associations, and so no general explicit solution is possible. In the special case where there is no association between incompatibilities ($D_U = \prod_{i \in U} D_i$), the second term in Eq. 66 vanishes, and the equilibrium for D_i depends only on selection coefficients that involve the

pair i. With the further assumption that the effects of incompatibilities are multiplicative, Eq. 66 gives essentially the same solution as the two-locus model (Eq. 46).

We are primarily concerned with whether pairwise associations between loci ($D_{i_1\, j_1}$ etc,) in different pairs will grow, thus coupling together different incompatibilities. Before considering these, however, we must establish whether different incompatibilities will become associated with each other. We will see that even when effects on viability are multiplicative across pairs, associations between incompatibilities may still be generated by recombination, complicating the analysis.

Consider the four-way association, $D_{ij} = D_{i_1\, j_1\, i_2\, j_2}$. We assume that only associations involving *pairs* of interacting loci are non-zero. Then, the association between pairs i and j after recombination is a sum over meioses where all four genes stay together (at rate $r_{ij,\emptyset}$), and meioses where one pair is inherited from one parent, and the other from the other parent (at rate $r_{i,j}$):

$$D''_{ij} = r_{ij,\emptyset}\, D'_{ij} + r_{i,j}\, D'_i\, D'_j \tag{67}$$

where the D'_U are the associations after viability selection. Writing the deviation from random association between pairs as $\theta_{ij} = D_{ij} - D_i\, D_j$ we have:

$$\theta''_{ij} = r_{ij,\emptyset}\, \theta'_{ij} + \chi\, D'_i\, D'_j \tag{68}$$

where $\chi = r_{ij,\emptyset} + r_{i,j} - (1 - r_i)\,(1 - r_j)$, and $r_i \equiv r_{i_1,i_2}$ is the recombination rare between the genes in pair i. We see that there is a contribution to the four-way association proportional to χ, which is just the covariance between recombination events involving pair i, and pair j. If the pairs do not overlap on the genetic map, and there is no interference between crossovers, then χ=0, and recombination does not generate any association between the pairs. However, if the interacting pairs overlap, then χ>0, because a single recombination event can break up both pairs; this generates a four-way association, θ_{ij}, even when there is no epistasis between different pairs of loci.

To find the effect of viability selection on the four-way association D_{ij}, we proceed as in Eq. 66, for D_i, by separating selection coefficients according to whether they include loci i and j:

$$\begin{aligned} D'_{ij} = D_{ij} &+ \sum_{U \subseteq \Omega \setminus ij} a_{ijU}\, (h_i\, h_j\, D_U - D_{ij}\, D_{ijU}) + \\ &\sum_{U \subseteq \Omega \setminus ij} a_{iU}\, (h_i\, D_{jU} - D_{ij}\, D_{iU}) + \\ &\sum_{U \subseteq \Omega \setminus ij} a_{jU}\, (h_j\, D_{iU} - D_{ij}\, D_{jU}) + \sum_{U \subseteq \Omega \setminus \mathrm{ij}} a_U\, (D_{ijU} - D_{ij}\, D_U) \end{aligned} \tag{69}$$

Now, suppose that there are no associations between incompatibilities, *except* for the focal pair {*i*, *j*}. Writing $D_{ij} = D_i D_j + \theta_{ij}$:

$$\begin{aligned} D'_{ij} = & \\ & D_{ij} + \sum_{U \subseteq \Omega \setminus ij} a_{ijU} D_U \left(h_i h_j - D_i^2 D_j^2 - 2\,\theta_{ij} D_i D_j - \theta_{ij}^2\right) + \\ & \sum_{U \subseteq \Omega \setminus ij} a_{iU} D_U \left(h_i D_j - D_i^2 D_j - \theta_{ij} D_i\right) + \\ & \sum_{U \subseteq \Omega \setminus ij} a_{jU} D_U \left(h_j D_i - D_i D_j^2 - \theta_{ij} D_j\right) \end{aligned} \tag{70}$$

We define the marginal selection on pair *i* and on the pair of pairs *i j* as:

$$\tilde{a}_i \equiv \sum_{U \subseteq \Omega \setminus i} a_{iU} D_U, \quad \tilde{a}_{ij} \equiv \sum_{U \subseteq \Omega \setminus i,j} a_{ijU} D_U \tag{71}$$

In the special case of multiplicative effects, these marginal coefficients simplify to $\tilde{a}_U = \left(\prod_{i \in U} \gamma_i\right) / \bar{V}$.

After some rearrangement, and using the relation $\tilde{a}_i = \tilde{a}_{ij} D_j + \sum_{U \subseteq \Omega \setminus ij} a_{iU} D_U$, Eq. 70 leads to:

$$\begin{aligned} \theta'_{ij} = \theta_{ij} - \Delta D_i \Delta D_j + \tilde{a}_{ij} & \left(\left(h_i - D_i^2\right)\left(h_j - D_j^2\right) - \theta_{ij}^2\right) \\ & - 2\,\theta_{ij}\left(\tilde{a}_i D_i + \tilde{a}_j D_j - 2\,\tilde{a}_{ij} D_i D_j\right) \end{aligned} \tag{72}$$

The change in the pairwise associations due to selection ($\Delta D_i = D_i^* - D_i$ etc.), assuming that only pairs *i*, *j* are associated via θ_{ij}, can be found in a similar way from Eq. 66:

$$\Delta D_i = \tilde{a}_i \left(h_i - D_i^2\right) + \left(\tilde{a}_j - 2\,\tilde{a}_{ij} D_i\right) \theta_{ij} \tag{73}$$

We see that when $\theta_{ij} = 0$, $\theta^*_{ij} = \left(\tilde{a}_{ij} - \tilde{a}_i \tilde{a}_j\right)\left(h_i - D_i^2\right)\left(h_j - D_j^2\right)$; thus, four-way associations are not generated by multiplicative selection. Indeed, under the multiplicative model, Eq. 72 simplifies to:

$$\theta'_{ij} = \frac{\theta_{ij}}{\bar{V}^2}\left(\left(1 - \gamma_i D_i\right)^2 - \gamma_i^2 h_i\right)\left(\left(1 - \gamma_j D_j\right)^2 - \gamma_j^2 h_j\right) \tag{74}$$

which is always smaller than 1.

Thus, incompatibilities will not be associated with each other ($D_{ij} = 0$) if there are multiplicative effects, if pairs of loci do not overlap on the genetic map, and if there is no interference between crossovers (χ=0). However, even with no epistasis between incompatibilities, incompatibilities will be associated if pairs of loci overlap (χ>0). The equilibrium association, θ_{ij}, is given by Eqs. 68, 74; from Eq. 74, its effect is to increase the strength of each pairwise association D_i, D_j.

Selection changes associations between loci in different pairs $i,\ j$:

$$D'_{i_1\ j_1} = D_{i_1\ j_1} + \sum_{U\subseteq\Omega} a_U\ (D_{i_1\ j_1\ U} - D_{i_1\ j_1}\ D_U) \tag{75}$$

Dividing the sum according to whether the set U includes $i,\ j$:

$$\begin{aligned} D'_{i_1\ j_1} = D_{i_1\ j_1} &+ \sum_{U\subseteq\Omega\setminus\{i,j\}} \tilde{a}_U\ (D_{i_1\ j_1\ U} - D_{i_1\ j_1}\ D_U) + \\ &\sum_{U\subseteq\Omega\setminus\{i,j\}} a_{i\,U}\ (h_{i_1}\ D_{i_2\ j_1\ U} - D_{i_1\ j_1}\ D_{i\,U}) + \\ &\sum_{U\subseteq\Omega\setminus\{i,j\}} a_{j\,U}\ (h_{j_1}\ D_{i_1\ j_2\ U} - D_{i_1\ j_1}\ D_{j\,U}) + \\ &\sum_{U\subseteq\Omega\setminus\{i,j\}} a_{i\,j\,U}\ (h_{i_1}\ h_{j_1}\ D_{i_2\ j_2\ U} - D_{i_1\ j_1}\ D_{i\,j\,U}) \end{aligned} \tag{76}$$

This does not have a closed form solution. However, it does include only associations of the form $D_{i_a\ j_b\ U}$, where $U \subseteq \Omega\setminus\{i,\ j\}$; the same is true for the more general recursion for the $D_{i_a\ j_b\ U}$. We can, therefore, examine the growth of just this class of associations.

We now assume that there are no associations between incompatibilities, except between pairs $i,\ j$. As above, the recursions now simplify to depend only on the four loci $\{i,\ j\}$, and on the marginal selection coefficients $\tilde{a}_i,\ \tilde{a}_j,\ \tilde{a}_{ij}$. We have not made any explicit assumption about interactions between incompatibilities. However, such interactions would generate associations such as D_{ik}, and would also enter into the marginal coefficients $\tilde{a}_U$. Therefore, the following analysis will not apply if any of the background pairs of loci interact with loci $i,\ j$, or if any of them overlap on the genetic map. Proceeding as before:

$$\begin{aligned} D'_{i_1\ j_1} = D_{i_1\ j_1} &+ \left(\tilde{a}_i - \tilde{a}_{ij}\ D_j\right)\ (h_{i_1}\ D_{i_2\ j_1} - D_{i_1\ j_1}\ D_i) + \\ &\left(\tilde{a}_j - \tilde{a}_{ij}\ D_i\right)\ (h_{j_1}\ D_{i_1\ j_2} - D_{i_1\ j_1}\ D_j) + \\ &\tilde{a}_{ij}\ (h_{i_1}\ h_{j_1}\ D_{i_2\ j_2} - D_{i_1\ j_1}\ (D_i\ D_j + \Theta_{ij})) \end{aligned} \tag{77}$$

This, together with the corresponding equations for the other three cross-locus associations, defines a 4×4 matrix whose eigenvalues determine whether loci in different pairs will become associated with each other.

With multiplicative fitnesses, the matrix simplifies to:

$$\frac{1}{\bar{V}}\begin{pmatrix} \phi_i\,\phi_j & \gamma_i\,h\phi_j & \gamma_j\,h\phi_i & \gamma_i\,\gamma_j\,h^2 \\ \gamma_i\,h\phi_j & \phi_i\,\phi_j & \gamma_i\,\gamma_j\,h^2 & \gamma_j\,h\phi_i \\ \gamma_j\,h\phi_i & \gamma_i\,\gamma_j\,h^2 & \phi_i\,\phi_j & \gamma_i\,h\phi_j \\ \gamma_i\,\gamma_j\,h^2 & \gamma_j\,h\phi_i & \gamma_i\,h\phi_j & \phi_i\,\phi_j \end{pmatrix}$$

where $h = \frac{1}{4}$ and $\phi_i = 1 - \gamma_i D_i$ is the reduction in viability at LE due to locus i. Remarkably, this matrix does not depend explicitly on any association between the pairs of loci, θ_{ij}, except through the mean viability, $\overline{V} = 1 + \gamma_i\,\gamma_j\,\theta_{ij}$, and indirectly, through the strength of linkage disequilibria $D_i,\ D_j$. It has leading eigenvalue $\lambda = \frac{1}{\overline{V}}\,(1 + \gamma_i(h - D_i))\,(1 + \gamma_j(h - D_j))$ (assuming $\gamma_i,\ \gamma_j > 0$). Since $h = \frac{1}{4} > D_i,\ D_j$, we see that under selection alone, pairs of loci will tend to become associated with each other; this tendency is countered by recombination, which reduces $D_{i_x\,j_y}$ by a factor $(1 - r_{i_x,j_y})$. With no linkage, all the r_{i_x,i_y} are $\frac{1}{2}$, leading to Eq. 18.

The stability of this system cannot be determined explicitly for general recombination rates. We deal with three special cases: no linkage; very tight linkage; and selection as strong as possible. With maximal selection, where recombinants {01, 10} at each pair are completely inviable, the least fit genotype cannot have negative fitness, and so $1 - \gamma_i(\frac{1}{4} + D_i) \geq 0$. Substituting for the equilibrium from Eq. 46, we find that $\gamma_i \leq \frac{4}{2 - r_i}$, which is $\frac{8}{3}$ with no linkage. At this maximum value of γ_i, we have that $\hat{D}_i = \frac{(1 - r_i)}{4}$; this corresponds simply to the linkage disequilibrium after selection, but before recombination, taking its maximum value of $\frac{1}{4}$.

With complete selection, θ_{ij} after selection is necessarily zero. However, θ_{ij} is generated by recombination, and so equals $\chi D_i\,D_j$ after recombination. Since mean viability depends on θ_{ij} ($\overline{V} = 1 + \gamma_i\ \gamma_j\,\theta_{ij}$), the rate of growth of pairwise associations is reduced if pairs of interacting loci overlap, so that χ>0. Complete selection on two pairs of interacting loci is equivalent to underdominance at two loci, since recombinants within pairs do not survive. There is a chance r_i that pair i will be broken up, and so in effect, there is selection against heterozygotes ($\frac{00}{11}$) of r_i, and similarly for pair j. If the pairs do not overlap on the genetic map (χ=0), then recombination events within each pair are independent. Otherwise, however, recombination events r_i and r_j are correlated, which is equivalent to an interaction in the fitness effects of the two pairs: hence, the dependence of the leading eigenvalue (Eq. 17) on χ.

Supplementary Information 5: Coupling between an assortment locus and multiple selected loci

We measure the net association between locus a and the set of selected loci Ω by:

$$C_{\underline{1}} = \sum_{U \subseteq \Omega} 2^{|U|-n} D_{aU} \tag{79}$$

This measure is equal to the association between a and the genotype $\underline{1} = \{1, 1 \ldots\}$. To see this, recall that, D_{aU} is defined as $E[\zeta_a \zeta_U]$, where $\zeta_a = X_a - p_a$, $\zeta_U = \prod_{i \in U} \zeta_i$, and $\zeta_i = X_i - \frac{1}{2}$; the X_a, X_i indicate allelic state, and take values 0, 1. We can rewrite $C_{\underline{1}}$ as:

$$\begin{aligned} C_{\underline{1}} &= \sum_{U \subseteq \Omega} E\left[2^{|U|-n} \zeta_a \zeta_U\right] = E\Big[\zeta_a \sum_{U \subseteq \Omega} 2^{|U|-n} \zeta_U\Big] \\ &= E\Big[\zeta_a \sum_{U \subseteq \Omega} \prod_{i \in \Omega \setminus U} \left(\frac{1}{2}\right) \prod_{i \in U} \left(X_i - \frac{1}{2}\right)\Big] \\ &= E\Big[\zeta_a \left(\left(\prod_{i \in U} X_i\right) - \left(\frac{1}{2}\right)^n\right)\Big] = E\Big[\zeta_a \left(\prod_{i \in U} X_i\right)\Big] \end{aligned} \tag{80}$$

where we have used the relations $E[\zeta_a] = 0$ and $\sum_{U \subseteq \Omega} (\prod_{i \in U} a_i)(\prod_{j \in (\Omega \setminus U)} b_i) = \prod_{i \in \Omega} (a_i + b_i) - \prod_{i \in \Omega} b_i$; the last term is subtracted because the sum over U does not include the empty set. This shows that $C_{\underline{1}}$ is just the excess frequency of allele $X_a = 1$ within genotype $\underline{1}$ (indicated by $(\prod_{i \in U} X_i) = 1$), multiplied by the frequency of that background.

The change in each D_{aU} due to selection is given by the haploid version of Eq. 2. Summing over $U \subseteq \Omega$:

$$\begin{aligned} C'_{\underline{1}} &= C_{\underline{1}} + \sum_{U \subseteq \Omega} \sum_{V \subseteq \Omega} 2^{|U|-n} a_V \left(D_{aUV} - D_{aU} D_V\right) \\ &= C_{\underline{1}} + \sum_{U \subseteq \Omega} \sum_{V \subseteq \Omega} 2^{|U|-n} a_V D_{aUV} - C_{\underline{1}} \left(\sum_{V \subseteq \Omega} a_V D_V\right) \end{aligned} \tag{81}$$

Reversing the order of sums, we can separate the sum over U into the sum over all subsets X of the selected set V, and the sum over all subsets that do not overlap with V. Let $V = YZ$:

$$C'_{\underline{1}} = C_{\underline{1}} + \sum_{V \subseteq \Omega} a_V \sum_{\left(X \subseteq \Omega \setminus V\right)} \sum_{Y \subseteq V} 2^{|XY|-n} D_{aXYYZ} - C_{\underline{1}} \left(\sum_{V \subseteq \Omega} a_V D_V\right) \tag{82}$$

Now, linkage disequilibria with repeated indices D_{aXYYZ} reduce to $2^{-|YY|}\, D_{aXZ}$ (Eq. 3). Hence:

$$
\begin{aligned}
C_1' &= \\
&C_1 + \sum_{V\subseteq\Omega} a_V\, 2^{-|V|} \sum_{\substack{Y\subseteq V \\ X\subseteq\Omega\setminus V}} \left(2^{|XZ|-n}\, D_{aXZ}\right) - C_1 \left(\sum_{V\subseteq\Omega} a_V\, D_V\right) \\
&= C_1 + \sum_{V\subseteq\Omega} a_V\, 2^{-|V|} \sum_{U\subseteq\Omega} \left(2^{|U|-n}\, D_{aU}\right) - C_1 \left(\sum_{V\subseteq\Omega} a_V\, D_V\right) \\
&= C_1 \left(1 + \sum_{V\subseteq\Omega} a_V\, \left(2^{-|V|} - D_V\right)\right)
\end{aligned}
\tag{83}
$$

which gives Eq. 26. Note that from the definition of fitness (Eq. 1), the increase is by a factor equal to the relative fitness of genotype $\underline{1}$, $\overline{W}_{\underline{1}} / \overline{W}$. (This derivation has been given for the genotype $\underline{1}$, but applies for any genotype: the D_{aU} are weighted by $\prod_{i\in U} (2\, X_i - 1)$, where $(2\, X_i - 1) = \pm 1$).

From Eq. 21, assortment at locus a is represented by the coefficient $a_{a,a} = \frac{\alpha}{h_a}$, where $h_a = p_a\, q_a$. Substituting into Eq. 2, we see that this does not alter genotype frequencies within haploids (D_{aU}), but does generate associations between haploids, $D''_{aS,T} = D^*_{aS}\, D^*_T + \alpha D^*_S\, D^*_{aT}$. (The D'_X denote associations after viability selection). Substituting into Eq. 4, the change due to assortment followed by recombination is:

$$
D''_{aU} = \sum_{ST=U} r_{aS,T}\, \left(D'_{aS}\, D'_T - D'_{aU}\right) + \alpha \sum_{ST=U} r_{aS,T}\, D'_{aT}\, D'_S \tag{84}
$$

Now, we add $\sum_{ST=U} r_{aS,T}\, D'_{aT}\, D'_S$ to the first term, and subtract it from the second. Exchanging S, T in the first term, and noting that $r_{aS,T} + r_{aT,S} = r_{S,T}$ gives Eq. 27.

First, we show that the fitness of the fittest genotypes, relative to the population mean, depends on the mean recombination rate; it is this which determines the increase of $C_{\underline{1}}$ under viability selection (Eq. 26). With tight linkage, D_U is close to its maximum, $2^{-|U|}$ for U even; by symmetry, D_U=0 for $|U|$ odd. The change due to recombination is given by Eq. 4, as $\sum_{ST=U} r_{s,T}\, D_S^*\, D_T^*$. We are concerned with even-numbered sets of loci, which can be broken up into either two even-sized sets, or two odd-sized sets. The former make no contribution, because $D_S^*\, D_T^* \sim D_U^*$; for the latter, $D_S^* = D_T^* = 0$, and so the net change due to recombination is $-\tilde{R}_U\, 2^{-|U|}$, where $\tilde{R}_U$ is defined as $\sum_{ST=U} r_{S,T}$, with the sum is taken over odd sets S, T. This must be balanced against the increase in D_U due to selection. Let $D_U = 2^{-|U|}\,(1-\theta_U)$, where θ_U is small and of the same order as the recombination rates. Substituting into the haploid version of Eq. 2, setting equal to the change due to recombination, and keeping leading terms:

$$\tilde{R}_U = \sum_{V\in\Omega} a_V\, 2^{-|V|}\ (\theta_U + \theta_V - \theta_{UV})$$
$$= \left(\sum_{V\in\Omega} a_V\, 2^{-|V|}\ \theta_V\right) + \left(\sum_{V\in\Omega} a_V\, 2^{-|V|}\right)\theta_U - \left(\sum_{V\in\Omega} a_V\, 2^{-|V|}\ \theta_{UV}\right) \qquad (85)$$

Equation 85 shows that deviations of linkage disequilibria from their maxima are proportional to the rate of recombination events, $\tilde{R}_U$, that split even-sized sets U into two odd-sized sets. Moreover, the first term in Eq. 85 is equal to the recombination load of Eq. 26, which determines the net rate of increase of associations under viability selection. To find this term, we show that when we sum over all even-sized sets, U, the last two terms in Eq. 85 cancel. Separate $U = XY$ into a component $Y \subseteq V$ and a component $X \subseteq \Omega \setminus V$. Then, $\theta_{UV} = \theta_{XYYZ}$ includes duplicate indices Y; applying Eq. 3, θ_{UV} reduces to θ_{XZ}. Therefore, summing Eq. 85 over all the 2^{n-1} even-sized subsets of Ω:

$$\sum_{|U|\ \text{even}} \tilde{R}_U = 2^{n-1} \sum_{V\in\Omega} a_V\, 2^{-|V|}\ \theta_V \qquad (86)$$

This can be simplified further. Each of the $\tilde{\mathrm{R}}_{\mathrm{U}}$ is a sum over all partitions of U into odd-sized subsets of U. This can be rewritten as a sum over all the partitions of the complete set of n selected loci. (For example, with 4 loci $\tilde{R}_{12} = r_{13,24} + r_{14,23} + r_{1,234} + r_{134,2}$). This sum includes every distinct recombination 2^{n-2} times, and so we have $\sum_{|\mathrm{U}|\ \mathrm{even}} \tilde{\mathrm{R}}_{\mathrm{U}} = 2^{n-2}\, R_\Omega$. Thus, viability selection increases $C_{\underline{1}}$ by a factor $(1+\sum_{\mathrm{V}\in\Omega} a_{\mathrm{V}}\, 2^{-|\mathrm{V}|}\ \theta_{\mathrm{V}}) = (1 + \frac{1}{2} R_\Omega)$, which is independent of selection.

When linkage is tight, the population is dominated by the two fittest genotypes, which we label as $\underline{0}$, $\underline{1}$. Because the D_{aU} are defined as the covariance between alleles at locus a and

sets of selected loci, U, they must necessarily be close to each other in the limit: when only those two fittest genotypes are present, $D_{aU} = 2^{1-|U|}\, C_\perp \;\forall\; U$. We therefore assume that D_{aU} differs from this limiting value by a small amount, of order recombination rate. Equation 27 shows that recombination has two effects on the D_{aU}. The first term is a sum over $r_{S,T}(D^*_{aS}\, D^*_T - D^*_{aU})$, which cancels when $D^*_{aU} \sim 2^{1-|U|}\, C^*_\perp,\; D^*_T \sim 2^{-|T|}$. The second term gives the approximation:

$$\begin{aligned} C_1'' &= C_1' - (1-\alpha) \sum_{U\subseteq\Omega} 2^{|U|-n} \sum_{ST=U} r_{aS,T}\, 2^{1-|S|}\, C_1'\, 2^{-|T|} \\ &= C_1' \left(1 - (1-\alpha)\, \tilde{r}_a\right) \\ \tilde{r}_a &= \frac{\sum_{U\subseteq\Omega} \sum_{ST=U} r_{aS,T}}{2^{n-1}} \end{aligned} \tag{87}$$

where $\tilde{r}_a$ is the unweighted average of $r_{aS,T}$, summed over all odd-sized sets U, with S even and T odd, and including $S = \{\}$. (The denominator arises because there are 2^{n-1} odd-sized subsets of a set of n selected loci).

It remains to rewrite $\tilde{r}_a$ in a simpler form, in the same way that we related the sum of the $\tilde{R}_U$ over all sets $U \subseteq \Omega$ to the total recombination rate, R_Ω. Each of the recombination rates involving a set $U \subseteq \Omega$ is a sum over possible partitions of the full set, Ω. Therefore, the sum $\sum_{U\subseteq\Omega} \sum_{ST=U} r_{aS,T}$ can be rewritten as a sum of the $r_{aX,Y}$, where $XY = \Omega$. Now, each $r_{aX,Y}$ contributes to all $r_{aS,T}$ where S is an even-sized subset of X, and T is an odd-sized subset of Y. A set of size $|X|$ has $2^{|X|-1}$ odd subsets and a set Y has $2^{|Y|-1}$ even-sized subsets, for |Y|>0. Therefore, $r_{aX,Y}$ contributes to $(2^{|X|-1})\,(2^{|Y|-1}) = 2^{n-2}$ terms $r_{aS,T}$, except for $r_{a,\Omega}$, which contributes to the 2^{n-1} odd-sized subsets of Ω. Therefore, $\tilde{r}_a = \frac{1}{2}(R_\Omega + r_{a,\Omega})$. Combining the successive effects of selection, assortment and recombination yields Eq. 30.

Supplementary Information 6: Invasion of an assortment allele

Figure 7. The invasion of an assortment allele. The frequency of the assorting allele (upper grey curve) is increased through associations with pairs of selected loci (D_{ajk}, lower light dotted curve), which in turn are generated as a result of pairwise associations between the assortment and the selected loci (D_{aj}, dark dotted curve). The pairwise associations amongst selected loci (D_{jk}) are shown by the heavy curve; these rapidly approach equilibrium, and then increase slightly as the assortment allele becomes more common. Note that initially, p_a, D_{aj} and D_{ajk} all increase exponentially at the same rate, which appears linear on this log scale. All loci are unlinked, so that all associations of the same kind are equal Selection is as in Fig. 7, with four unlinked selected loci, and selection at its maximum value of $s = 0.2$.

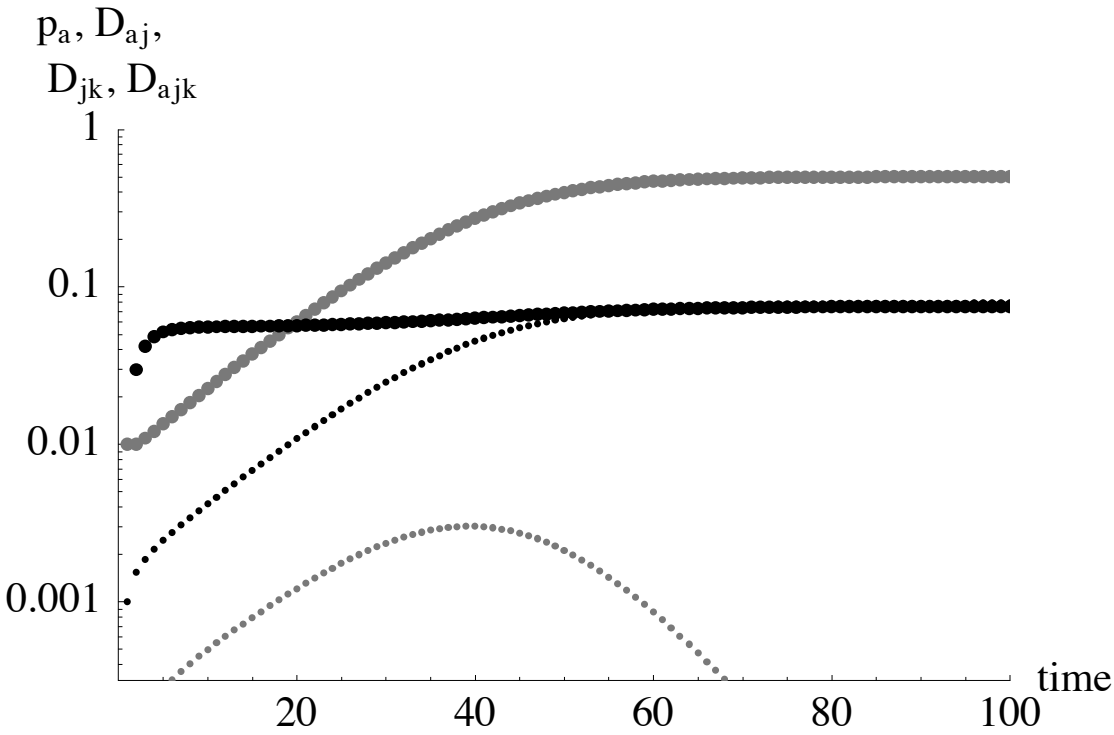

Figure 7

Supplementary Information 7: Generating Figs. 1-6, 8

Supplementary Information 8: Generating Fig. 7

Supplementary Information 9: Mathematica definitions