

3 We derive here approximations for the change in mean selfing rate during se-
 4 lection and reproduction, neglecting epistasis and genetic associations between loci
 5 affecting fitness. As explained in the main text, we assume that loci affecting the
 6 selfing rate have additive effects:

$$\sigma = \sum_{i=1}^{\ell_{\sigma}} (\sigma_i^M + \sigma_i^P). \quad (\text{C1})$$

7 Neglecting associations between loci affecting the selfing rate, the variance in selfing
 8 rate in the population can be expressed as:

$$V_{\sigma} \approx 2 \sum_i (\tilde{D}_{ii} + D_{i,i}). \quad (\text{C2})$$

9 We first express the change in mean selfing rate in terms of genetic associations, and
 10 then compute expressions for these associations at QLE.

11
 12 **Change in $\bar{\sigma}$ in terms of genetic associations.** The change in $\bar{\sigma}$ during reproduc-
 13 tion, $\Delta_r \bar{\sigma}$, can be obtained as follows. The change in $\bar{\sigma}_i^M$ is given by:

$$\Delta_r \bar{\sigma}_i^M = \bar{\sigma}' E' \left[\frac{\sigma}{\bar{\sigma}'} \frac{\sigma_i^M + \sigma_i^P}{2} \right] + (1 - \bar{\sigma}') E' \left[\frac{1 - \sigma}{1 - \bar{\sigma}'} \frac{\sigma_i^M + \sigma_i^P}{2} \right] - \bar{\sigma}_i^{M'} \quad (\text{C3})$$

14 where E' stands for the average over all individuals after selection, and $\bar{\sigma}'$, $\bar{\sigma}_i^{M'}$ the
 15 averages of σ and σ_i^M over all individuals after selection. The first term of equation C3
 16 corresponds to the contribution from selfed offspring to the average after reproduction
 17 (each parent contributing to the pool of selfed offspring in proportion $\sigma/\bar{\sigma}'$), while the
 18 second term corresponds to the contribution from outcrossed offspring. Using equation

19 2 in the main text, this becomes:

$$\Delta_r \overline{\sigma_i^M} = E' \left[\sigma \frac{\zeta_i^M + \zeta_i^P - \nabla'_i}{2} \right] + E' \left[(1 - \sigma) \frac{\zeta_i^M + \zeta_i^P - \nabla'_i}{2} \right] \quad (\text{C4})$$

20 with $\nabla'_i = \overline{\sigma_i^{M'}} - \overline{\sigma_i^{P'}}$. Using the fact that $E' [\zeta_i^M] = E' [\zeta_i^P] = 0$, equation C4 yields:

$$\Delta_r \overline{\sigma_i^M} = -\frac{\nabla'_i}{2}. \quad (\text{C5})$$

21 Similarly, the change in $\overline{\sigma_i^P}$ is given by:

$$\Delta_r \overline{\sigma_i^P} = \overline{\sigma'} E' \left[\frac{\sigma}{\overline{\sigma'}} \frac{\sigma_i^M + \sigma_i^P}{2} \right] + (1 - \overline{\sigma'}) E' \left[\frac{1 - \kappa \sigma}{1 - \kappa \overline{\sigma'}} \frac{\sigma_i^M + \sigma_i^P}{2} \right] - \overline{\sigma_i^{P'}} \quad (\text{C6})$$

22 where again the first term corresponds to the contribution from selfed offspring while
 23 the second term corresponds to the contribution from outcrossed offspring, each parent
 24 contributing to the pollen pool in proportion $(1 - \kappa \sigma) / (1 - \kappa \overline{\sigma'})$. Equation C6 may
 25 be written as:

$$\Delta_r \overline{\sigma_i^P} = E' \left[\sigma \frac{\zeta_i^M + \zeta_i^P + \nabla'_i}{2} \right] + \frac{1 - \overline{\sigma'}}{1 - \kappa \overline{\sigma'}} E' \left[(1 - \kappa \sigma) \frac{\zeta_i^M + \zeta_i^P + \nabla'_i}{2} \right]. \quad (\text{C7})$$

26 Using equations C1 and equation 2 in the main text, the selfing rate of an individual
 27 is given by:

$$\sigma = \overline{\sigma'} + \sum_i (\zeta_i^M + \zeta_i^P). \quad (\text{C8})$$

28 Replacing σ by equation C8, using equations 4 and 5 in the main text, and neglecting
 29 associations between different loci affecting the selfing rate, equation C7 becomes:

$$\Delta_r \overline{\sigma_i^P} = \frac{1 - \kappa}{1 - \kappa \overline{\sigma'}} \left(\tilde{D}'_{ii} + D'_{i,i} \right) + \frac{\nabla'_i}{2}. \quad (\text{C9})$$

30 From equations C1, C2, C5 and C9, the change in mean selfing rate during reproduc-
 31 tion is:

$$\Delta_r \overline{\sigma} = \frac{1 - \kappa}{1 - \kappa \overline{\sigma'}} \frac{V'_\sigma}{2} \quad (\text{C10})$$

32 where V'_σ is the variance in selfing rate after selection. $\Delta_r \bar{\sigma}$ is positive and represents
33 the effect of the automatic transmission advantage associated with selfing, which van-
34 ishes under complete pollen discounting ($\kappa = 1$). Because we assume that the variance
35 in selfing rate in the population is small, and because the change in $\bar{\sigma}$ during selection
36 should be proportional to V_σ , $\bar{\sigma}'$ can be approximated by $\bar{\sigma}$ in the equation above,
37 yielding:

$$\Delta_r \bar{\sigma} \approx \frac{1 - \kappa}{1 - \kappa \bar{\sigma}} \frac{V'_\sigma}{2}. \quad (\text{C11})$$

38 The change in mean selfing rate during selection is given by:

$$\begin{aligned} \Delta_s \bar{\sigma} &= \text{E} \left[\frac{W}{\bar{W}} \sum_i (\sigma_i^{\text{M}} + \sigma_i^{\text{P}}) \right] - \bar{\sigma} \\ &= \text{E} \left[\frac{W}{\bar{W}} \sum_i (\zeta_i^{\text{M}} + \zeta_i^{\text{P}}) \right]. \end{aligned} \quad (\text{C12})$$

39 Ignoring epistatic interactions between loci affecting fitness, equation 8 in the main
40 text simplifies to:

$$\frac{W}{\bar{W}} = 1 + \sum_j [a_j (\zeta_{j,\emptyset} + \zeta_{\emptyset,j}) + a_{j,j} (\zeta_{j,j} - D_{j,j})]. \quad (\text{C13})$$

41 Using equation C13 and equations 4 and 5 in the main text, this yields:

$$\Delta_s \bar{\sigma} = 2 \sum_{i,j} a_j (\tilde{D}_{ij} + \tilde{D}_{i,j}) + 2 \sum_{i,j} a_{j,j} \tilde{D}_{i,j} \quad (\text{C14})$$

42 where the sums are over all loci i affecting the selfing rate of individuals and all loci
43 j affecting fitness. The first term of equation C14 depends on directional selection
44 at loci affecting fitness (a_j) and on associations between alleles increasing selfing and
45 alleles 1 at fitness loci, either on the same or on different haplotypes (\tilde{D}_{ij} , $\tilde{D}_{i,j}$). As we
46 will see, this term is positive, reflecting the fact that selfing increases the efficiency of
47 selection by increasing homozygosity at fitness loci, so that the favored alleles at these
48 loci tend to be more frequent in genomes coding for higher selfing rates. The second

49 term of equation C14 depends on dominance at fitness loci ($a_{j,j}$) and on associations
50 $\tilde{D}_{i,j}$ between loci affecting selfing and fitness loci. We will see that these associations
51 are positive, reflecting the fact that alleles coding for higher selfing rates tend to be
52 found in more homozygous backgrounds. The coefficients $a_{j,j}$ are negative when dele-
53 terious alleles are partially recessive, in which case the second term of equation C14
54 represents the effect of inbreeding depression disfavoring selfing.

55

56 **Associations generated by selfing.** The genetic associations that appear in equa-
57 tion C14 in turn depend on other associations between pairs of loci. Some of these
58 associations are generated by partial selfing, even in the absence of selection (W con-
59 stant) and in the absence of genetic variation for the selfing rate. In particular, a
60 recursion for the within-locus association $D_{i,i}$ (under neutrality and in the absence of
61 variation for selfing) is given by:

$$D''_{i,i} = \frac{\bar{\sigma}}{2} \left(\tilde{D}_{ii} + D_{i,i} \right). \quad (\text{C15})$$

62 Indeed, the covariance between the effects of the two homologous alleles at locus i
63 in an individual equals zero if this individual has been produced by random mating,
64 while if the individual has been produced by selfing (probability $\bar{\sigma}$) its two alleles are
65 the copies of the same parental allele with probability 1/2, and come from the two
66 parental alleles with probability 1/2. At equilibrium, equation C15 yields:

$$D_{i,i} = F \tilde{D}_{ii}, \quad \text{with} \quad F = \frac{\bar{\sigma}}{2 - \bar{\sigma}}. \quad (\text{C16})$$

67 Similarly, $D_{j,j} = F \tilde{D}_{jj}$ at equilibrium; however, $\tilde{D}_{jj} = p_j q_j$ in the case of biallelic loci
68 (where p_j and q_j are the frequencies of alleles 1 and 0 at locus j), so that:

$$D_{j,j} = F p_j q_j. \quad (\text{C17})$$

69 The association $D_{ij,ij}$ is also generated by selfing, even when loci i and j are neutral.

70 In this case, $D_{ij,ij}$ at the next generation is given by:

$$D''_{ij,ij} = \frac{\bar{\sigma}}{2} \left[(1 - \rho_{ij})^2 \left(\tilde{D}_{iijj} + D_{ij,ij} \right) + 2\rho_{ij} (1 - \rho_{ij}) \left(\tilde{D}_{iij,j} + \tilde{D}_{ijj,i} \right) \right. \\ \left. + \rho_{ij}^2 \left(\tilde{D}_{iijj} + D_{ij,ij} \right) \right] \quad (\text{C18})$$

71 where ρ_{ij} is the recombination rate between loci i and j . From the fact that locus j

72 is biallelic, repeated j indices that appear in the associations above can be eliminated

73 using the relation (e.g., Barton and Turelli, 1991; Kirkpatrick et al., 2002):

$$D_{\mathbb{S}jj} = p_j q_j D_{\mathbb{S}} + (1 - 2p_j) D_{\mathbb{S}j} \quad (\text{C19})$$

74 where \mathbb{S} is any set of loci either on the same or on different haplotypes from the

75 same individual. Furthermore, associations involving a single j index equal zero at

76 equilibrium when locus j is neutral (since alleles 0 and 1 are then equivalent), so that

77 $\tilde{D}_{iijj} = \tilde{D}_{ii} p_j q_j$ and $\tilde{D}_{ijj,i} = D_{i,i} p_j q_j$. Finally, one can show using the same reasoning

78 as above that $\tilde{D}_{ijj,i} = F \tilde{D}_{ii} p_j q_j$ at equilibrium, leading to the following expression for

79 $D_{ij,ij}$ at equilibrium:

$$D_{ij,ij} = \phi_{ij} \tilde{D}_{ii} p_j q_j, \quad \text{with} \quad \phi_{ij} = \frac{\bar{\sigma}}{2 - \bar{\sigma}} \frac{2 - \bar{\sigma} - 2(2 - 3\bar{\sigma}) \rho_{ij} (1 - \rho_{ij})}{2 - \bar{\sigma} [1 - 2\rho_{ij} (1 - \rho_{ij})]} \quad (\text{C20})$$

80 (e.g., Abu Awad and Roze, 2018), ϕ_{ij} representing the joint probability of identity-by-

81 descent at loci i and j .

82

83 **Associations** $\tilde{D}_{ij,j}$. The association $\tilde{D}_{ij,j}$ is generated by the effect of locus i on

84 the selfing rate, even in the absence of selection at locus j . Its equilibrium value can

85 be computed as follows. We have:

$$D''_{ij,j} = \text{E}'' \left[\left(\sigma_i^{\text{M}} - \overline{\sigma_i^{\text{M}''}} \right) \left(X_j^{\text{M}} - p_j^{\text{M}''} \right) \left(X_j^{\text{P}} - p_j^{\text{P}''} \right) \right] \quad (\text{C21})$$

86 where E'' stands for the average among offspring (after reproduction), $\overline{\sigma_i^{M''}}$ is the
87 average of σ_i^M among offspring, and $p_j^{M''}$, $p_j^{P''}$ the frequencies of alleles 1 at locus j in
88 maternally and paternally derived genomes of offspring. In the absence of selection
89 (W constant) we have $\Delta_r \overline{\sigma_i^M} = \overline{\sigma_i^{M''}} - \overline{\sigma_i^M}$, $\Delta_r p_j^M = p_j^{M''} - p_j^M$ and $\Delta_r p_j^P = p_j^{P''} - p_j^P$,
90 and equation C21 can thus be written:

$$D''_{ij,j} = E'' \left[\left(\sigma_i^M - \overline{\sigma_i^M} - \Delta_r \overline{\sigma_i^M} \right) \left(X_j^M - p_j^M - \Delta_r p_j^M \right) \left(X_j^P - p_j^P - \Delta_r p_j^P \right) \right]. \quad (\text{C22})$$

91 Noting that $E'' \left[\sigma_i^M - \overline{\sigma_i^M} \right] = \Delta_r \overline{\sigma_i^M}$, while $E'' \left[X_j^M - p_j^M \right] = \Delta_r p_j^M$ and $E'' \left[X_j^P - p_j^P \right] =$
92 $\Delta_r p_j^P$, expanding equation C22 yields:

$$\begin{aligned} D''_{ij,j} &= D_{ij,j}^{\text{rep}} - \left(\Delta_r \overline{\sigma_i^M} \right) D_{j,j}^{\text{rep}} - \left(\Delta_r p_j^M \right) D_{i,j}^{\text{rep}} - \left(\Delta_r p_j^P \right) D_{i,j,\emptyset}^{\text{rep}} \\ &\quad + 2 \left(\Delta_r \overline{\sigma_i^M} \right) \left(\Delta_r p_j^M \right) \left(\Delta_r p_j^P \right) \end{aligned} \quad (\text{C23})$$

93 where $D_{\text{U,V}}^{\text{rep}}$ stands for associations measured after reproduction, but using “reference
94 values” (the values of $\overline{\sigma_i^M}$, p_j^M and p_j^P that appear in the associations) measured before
95 reproduction: for example, $D_{ij,j}^{\text{rep}} = E'' \left[\left(\sigma_i^M - \overline{\sigma_i^M} \right) \left(X_j^M - p_j^M \right) \left(X_j^P - p_j^P \right) \right]$. Equation
96 C23 can be simplified by noting that $\Delta_r p_j^M = \Delta_r p_j^P = 0$ when allele j is neutral,
97 yielding:

$$D''_{ij,j} = D_{ij,j}^{\text{rep}} - \left(\Delta_r \overline{\sigma_i^M} \right) D_{j,j}^{\text{rep}}. \quad (\text{C24})$$

98 Furthermore, because $\Delta_r \overline{\sigma_i^M}$ is proportional to the amount of genetic variation for
99 selfing (that we suppose small), $D_{j,j}^{\text{rep}}$ can be replaced by its expression in the absence

100 of variation for selfing, given by $\frac{\bar{\sigma}}{2} (p_j q_j + D_{j,j})$ (see above). Finally, $D_{ij,j}^{\text{rep}}$ is given by:

$$\begin{aligned}
D_{ij,j}^{\text{rep}} = & \bar{\sigma} \text{ E} \left[\frac{\sigma}{\bar{\sigma}} \left[\frac{1 - \rho_{ij}}{4} (\zeta_{ijj,\emptyset} + \zeta_{ij,j} + \zeta_{j,ij} - \nabla_i \zeta_{j,j} + \zeta_{\emptyset,ijj} - \nabla_i \zeta_{\emptyset,jj}) \right. \right. \\
& \left. \left. + \frac{\rho_{ij}}{4} (\zeta_{ij,j} + \zeta_{i,jj} + \zeta_{jj,i} - \nabla_i \zeta_{jj,\emptyset} + \zeta_{j,ij} - \nabla_i \zeta_{j,j}) \right] \right] \\
& + (1 - \bar{\sigma}) \text{ E} \left[\frac{1 - \sigma}{1 - \bar{\sigma}} \left[\frac{1 - \rho_{ij}}{2} (\zeta_{ij,\emptyset} + \zeta_{\emptyset,ij} - \nabla_i \zeta_{\emptyset,j}) + \frac{\rho_{ij}}{2} (\zeta_{i,j} + \zeta_{j,i} - \nabla_i \zeta_{j,\emptyset}) \right] \right] \\
& \times \text{ E} \left[\frac{1 - \kappa \sigma}{1 - \kappa \bar{\sigma}} \frac{\zeta_{j,\emptyset} + \zeta_{\emptyset,j}}{2} \right].
\end{aligned} \tag{C25}$$

101 The first term of equation C25 corresponds to the contribution of individuals produced
102 by selfing to $D_{ij,j}^{\text{rep}}$. In this case, the three genes may have been present in different
103 configurations in the parent with different probabilities (corresponding to the different
104 $\zeta_{\text{U,V}}$ terms). The ∇_i terms stem from the fact that $\bar{\sigma}_i^{\text{M}} \neq \bar{\sigma}_i^{\text{P}}$; note that $p_j^{\text{M}} = p_j^{\text{P}}$
105 under the assumption that locus j is neutral, so that ∇_j terms do not arise. The
106 second term of equation C25 corresponds to the contribution of individuals produced
107 by outcrossing: the first average is over all maternal individuals (transmitting genes at
108 loci i and j), while the second average is over all paternal individuals (transmitting one
109 gene at locus j). Using equation C8, neglecting genetic associations between different
110 loci affecting selfing and assuming that the genetic variance in selfing rate is small (so
111 that products between associations involving a single i index may be neglected), one
112 arrives at:

$$\begin{aligned}
D_{ij,j}^{\text{rep}} \approx & \frac{\bar{\sigma}}{2} \left[(1 - \rho_{ij}) \tilde{D}_{ijj} + \rho_{ij} \tilde{D}_{i,jj} + \tilde{D}_{ij,j} - \frac{\nabla_i}{2} (p_j q_j + D_{j,j}) \right] \\
& + \frac{1}{2} \left[(1 - \rho_{ij}) \tilde{D}_{iij} + \rho_{ij} \tilde{D}_{ii,jj} + \tilde{D}_{ijj,i} + \tilde{D}_{iij,j} + D_{ij,ij} \right]
\end{aligned} \tag{C26}$$

113 (note that the second term of equation C25 cancels under the assumption that locus j is
114 neutral). Using equation C19 and the fact that associations with a single j index equal

115 zero when locus j is neutral (in particular, $\tilde{D}_{ij} = \tilde{D}_{i,j} = 0$), we have $\tilde{D}_{ijj} = \tilde{D}_{i,jj} = 0$
 116 while $\tilde{D}_{iij} = \tilde{D}_{ii,j} = \tilde{D}_{ii} p_j q_j$. Using equations C5, C16, C17, C20 and C24, one
 117 finally obtains:

$$D''_{i,j} \approx \frac{\bar{\sigma}}{2} \tilde{D}_{i,j} + \frac{1}{2} [(1+F)^2 + G_{ij}] \tilde{D}_{ii} p_j q_j \quad (\text{C27})$$

118 where $G_{ij} = \phi_{ij} - F^2$ is the identity disequilibrium between loci i and j (e.g., Weir
 119 and Cockerham, 1973).

120 The recursion for $D_{j,ij}$ is obtained using the same general method. We have in
 121 particular:

$$D''_{j,ij} = D_{j,ij}^{\text{rep}} - \left(\Delta_r \bar{\sigma}_i^{\text{P}} \right) D_{j,j}^{\text{rep}}, \quad (\text{C28})$$

122 while:

$$D_{j,ij}^{\text{rep}} \approx \frac{\bar{\sigma}}{2} \left[\tilde{D}_{i,j} + \frac{\nabla_i}{2} (p_j q_j + D_{j,j}) \right] + \frac{1}{2} [(1+F)^2 + G_{ij}] \tilde{D}_{ii} p_j q_j. \quad (\text{C29})$$

123 Equations C9, C28 and C29 yield:

$$D''_{j,ij} \approx \frac{\bar{\sigma}}{2} \tilde{D}_{i,j} + \frac{1}{2} [(1+F)^2 + G_{ij}] \tilde{D}_{ii} p_j q_j \quad (\text{C30})$$

$$- \frac{\bar{\sigma}}{2} \frac{1-\kappa}{1-\kappa\bar{\sigma}} (1+F)^2 \tilde{D}_{ii} p_j q_j.$$

124 Finally, equations C27 and C30 yield the following expression for $\tilde{D}_{i,j}$ at QLE:

$$\tilde{D}_{i,j} \approx \frac{1}{2-\bar{\sigma}} \left[(1+F)^2 \left(1 - \frac{\bar{\sigma}}{2} \frac{1-\kappa}{1-\kappa\bar{\sigma}} \right) + G_{ij} \right] \tilde{D}_{ii} p_j q_j. \quad (\text{C31})$$

125 Equation C31 is equivalent to the result obtained by Epinat and Lenormand (2009)
 126 under strong discounting ($\kappa \approx 1$, their equation A5). Under complete outcrossing, it
 127 simplifies to $\tilde{D}_{i,j} \approx (1/2) \tilde{D}_{ii} p_j q_j$.

128

129 **Associations \tilde{D}_{ij} and $\tilde{D}_{i,j}$.** The associations \tilde{D}_{ij} , $\tilde{D}_{i,j}$ that appear in the first term
 130 of equation C14 are generated by selection acting at locus j and by the effect of locus

131 i on the selfing rate, and can be computed as follows. Using the same reasoning as for
 132 the derivation of equation C24, one obtains:

$$D''_{ij,\emptyset} = D_{ij,\emptyset}^{\text{rep}} - \left(\Delta_r \overline{\sigma_i^M} \right) \left(\Delta_r p_j^M \right) \approx D_{ij,\emptyset}^{\text{rep}}. \quad (\text{C32})$$

133 Indeed, $\Delta_r \overline{\sigma_i^M}$ and $\Delta_r p_j^M$ are both proportional to the variance in selfing rate, and
 134 their product can thus be neglected. $D_{ij,\emptyset}^{\text{rep}}$ is then given by:

$$D_{ij,\emptyset}^{\text{rep}} = (1 - \rho_{ij}) \tilde{D}'_{ij} + \rho_{ij} \tilde{D}'_{i,j}. \quad (\text{C33})$$

135 Similarly,

$$D''_{\emptyset,ij} \approx D_{\emptyset,ij}^{\text{rep}} \quad (\text{C34})$$

136 while $D_{\emptyset,ij}^{\text{rep}}$ is given by:

$$\begin{aligned} D_{\emptyset,ij}^{\text{rep}} = & \bar{\sigma}' \text{E} \left[\frac{\sigma}{\bar{\sigma}'} \left[\frac{1 - \rho_{ij}}{2} (\zeta_{ij,\emptyset} + \nabla'_i \zeta_{j,\emptyset} + \nabla'_j \zeta_{i,\emptyset} + \nabla'_i \nabla'_j + \zeta_{\emptyset,ij}) \right. \right. \\ & \left. \left. + \frac{\rho_{ij}}{2} (\zeta_{i,j} + \nabla'_i \zeta_{\emptyset,j} + \zeta_{j,i} + \nabla'_j \zeta_{\emptyset,i}) \right] \right] \\ & + (1 - \bar{\sigma}') \text{E} \left[\frac{1 - \kappa \sigma}{1 - \kappa \bar{\sigma}'} \left[\frac{1 - \rho_{ij}}{2} (\zeta_{ij,\emptyset} + \nabla'_i \zeta_{j,\emptyset} + \nabla'_j \zeta_{i,\emptyset} + \nabla'_i \nabla'_j + \zeta_{\emptyset,ij}) \right. \right. \\ & \left. \left. + \frac{\rho_{ij}}{2} (\zeta_{i,j} + \nabla'_i \zeta_{\emptyset,j} + \zeta_{j,i} + \nabla'_j \zeta_{\emptyset,i}) \right] \right] \end{aligned} \quad (\text{C35})$$

137 where the first term is the contribution of individuals produced by selfing while the
 138 second term is the contribution of individuals produced by outcrossing, and $\nabla'_j = p_j^{M'} -$
 139 $p_j^{P'}$. The terms in ∇'_i, ∇'_j are proportional to the variance in selfing rate and generate
 140 terms that are proportional to the square of the variance, and are thus neglected. One
 141 obtains:

$$\begin{aligned} D_{\emptyset,ij}^{\text{rep}} = & (1 - \rho_{ij}) \tilde{D}'_{ij} + \rho_{ij} \tilde{D}'_{i,j} \\ & + \frac{1 - \kappa}{1 - \kappa \bar{\sigma}'} \left[(1 - \rho_{ij}) \tilde{D}'_{ii} + \rho_{ij} \tilde{D}'_{ii,j} + \tilde{D}'_{ij,i} \right]. \end{aligned} \quad (\text{C36})$$

142 It is possible to show that \tilde{D}'_{ii} and $\tilde{D}'_{ii,j}$ are negligible compared to $\tilde{D}'_{ij,i}$ (using the
 143 same method as for deriving $\tilde{D}_{ij,i}$ below), and equations C32, C33, C34 and C36 thus
 144 yield:

$$\tilde{D}''_{ij} = (1 - \rho_{ij}) \tilde{D}'_{ij} + \rho_{ij} \tilde{D}'_{i,j} + \frac{1 - \kappa}{2(1 - \kappa \bar{\sigma})} \tilde{D}'_{ij,i}. \quad (\text{C37})$$

145 Similarly, one arrives at:

$$\tilde{D}''_{i,j} = \frac{\bar{\sigma}}{2} (\tilde{D}'_{ij} + \tilde{D}'_{i,j}) + \tilde{D}'_{ij,i}. \quad (\text{C38})$$

146 Using the same reasoning as for deriving equation C32, $D'_{ij,\emptyset}$ is given by:

$$D'_{ij,\emptyset} = D_{ij,\emptyset}^{\text{sel}} - (\Delta_s \bar{\sigma}_i^{\text{M}}) (\Delta_s p_j^{\text{M}}) \quad (\text{C39})$$

147 where $D_{\text{U,V}}^{\text{sel}}$ refers to associations measured after selection, but using as reference values
 148 $\bar{\sigma}_i^{\text{M}}, \bar{\sigma}_i^{\text{P}}, p_j^{\text{M}}$ and p_j^{P} before selection — in particular, $D_{ij,\emptyset}^{\text{sel}} = \text{E}' \left[(\sigma_i^{\text{M}} - \bar{\sigma}_i^{\text{M}}) (X_j^{\text{M}} - p_j^{\text{M}}) \right]$
 149 — and where $\Delta_s \bar{\sigma}_i^{\text{M}}$ and $\Delta_s p_j^{\text{M}}$ are the changes in $\bar{\sigma}_i^{\text{M}}$ and in p_j^{M} due to selection. In
 150 the following, we compute expressions to the first order in the strength of selection
 151 (measured as ϵ , and representing the largest of $a_{\text{U,V}}$ coefficients in absolute value), and
 152 thus neglect the last product of equation C39 (which is of order ϵ^2). We thus have:

$$D'_{ij,\emptyset} \approx D_{ij,\emptyset}^{\text{sel}} = \text{E} \left[\frac{W}{\bar{W}} \zeta_{ij,\emptyset} \right] \quad (\text{C40})$$

153 where the last average is over all individuals before selection. Using equation C13, and
 154 neglecting associations involving different fitness loci, this yields:

$$D'_{ij,\emptyset} \approx D_{ij,\emptyset} + a_j (D_{ijj,\emptyset} + D_{ij,j}) + a_{j,j} (D_{ijj,j} - D_{ij,\emptyset} D_{j,j}). \quad (\text{C41})$$

155 Eliminating repeated j indices from associations, we have $D_{ijj,\emptyset} = (1 - 2p_j) D_{ij,\emptyset}$, and
 156 $D_{ijj,j} = p_j q_j D_{i,j} + (1 - 2p_j) D_{ij,j}$. Because $D_{ij,\emptyset}$ and $D_{i,j}$ are of order ϵ , equation C41
 157 simplifies to:

$$D'_{ij,\emptyset} \approx D_{ij,\emptyset} + [a_j + a_{j,j} (1 - 2p_j)] D_{ij,j} + o(\epsilon). \quad (\text{C42})$$

158 However, using equation C42 (and the equivalent expression for $D'_{i,j}$) yields QLE ex-
159 pressions that diverge when $\rho_{ij}(1 - \bar{\sigma})$ tends to zero (that is, either when ρ_{ij} tends to
160 0 or when $\bar{\sigma}$ tends to 1). Indeed, $D_{ij,\emptyset}$ and $D_{i,j}$ are of order 1 when $\rho_{ij}(1 - \bar{\sigma})$ is of
161 order ϵ , and terms such as $a_j D_{ij,\emptyset}$, $a_{j,j} D_{i,j}$ become of order ϵ , and must be retained in
162 order to prevent divergence at low effective recombination (e.g., Charlesworth, 1990;
163 Roze, 2014; Gervais and Roze, 2017) — note that this implies that allele frequencies at
164 fitness loci are at an equilibrium, in order for the QLE approximation to hold. In that
165 case, and in order to simplify expressions, we will assume that polymorphism stays
166 low at fitness loci and derive expressions to the first order in $p_j q_j$ (“rare alleles approx-
167 imation”). Given that associations involving one or several j indices are proportional
168 to $p_j q_j$, this yields:

$$D'_{ij,\emptyset} \approx [1 + a_j(1 - 2p_j)] D_{ij,\emptyset} + [a_j + a_{j,j}(1 - 2p_j)] D_{i,j} + o(\epsilon). \quad (\text{C43})$$

169 Similarly, one obtains:

$$D'_{i,j} \approx [1 + a_j(1 - 2p_j)] D_{i,j} + [a_j + a_{j,j}(1 - 2p_j)] D_{ij,j} + o(\epsilon) \quad (\text{C44})$$

170 and finally

$$\tilde{D}'_{ij} \approx [1 + a_j(1 - 2p_j)] \tilde{D}_{ij} + [a_j + a_{j,j}(1 - 2p_j)] \tilde{D}_{ij,j} + o(\epsilon) \quad (\text{C45})$$

171

$$\tilde{D}'_{i,j} \approx [1 + a_j(1 - 2p_j)] \tilde{D}_{i,j} + [a_j + a_{j,j}(1 - 2p_j)] \tilde{D}_{ij,j} + o(\epsilon). \quad (\text{C46})$$

172 Equations C37, C38, C45 and C46 finally yield:

$$\tilde{D}_{ij} + \tilde{D}_{i,j} \approx \frac{[a_j + a_{j,j}(1 - 2p_j)](1 + 2\rho_{ij}\bar{\sigma})\tilde{D}_{ij,j} + X\tilde{D}'_{ij,i}}{\rho_{ij}(1 - \bar{\sigma}) - a_j(1 - 2p_j)\left[1 - \frac{\bar{\sigma}}{2} - \rho_{ij}(1 - 2\bar{\sigma})\right]} \quad (\text{C47})$$

173 with:

$$X = \frac{1 - \kappa}{2(1 - \kappa\bar{\sigma})} + 2\rho_{ij} - a_j(1 - 2p_j)(1 - 2\rho_{ij}). \quad (\text{C48})$$

174 Associations \tilde{D}_{ij} and $\tilde{D}_{i,j}$ thus depend on $\tilde{D}_{ij,j}$ (given by equation C31) and on $\tilde{D}'_{ij,i}$,
 175 whose expression at QLE must be computed to the first order in the strength of
 176 selection ϵ . The effect of genetic variance for the selfing rate can be ignored when
 177 computing the effect of reproduction on $\tilde{D}_{ij,i}$, as it would generate second-order terms
 178 in V_σ . This yields:

$$\tilde{D}''_{ij,i} = \tilde{D}_{ij,i}^{\text{rep}} = \frac{\bar{\sigma}}{2} \left[(1 - \rho_{ij}) \tilde{D}'_{ii,j} + \rho_{ij} \tilde{D}'_{ii,j} + \tilde{D}'_{ij,i} \right]. \quad (\text{C49})$$

179 As mentioned above, $\tilde{D}'_{ii,j}$ and $\tilde{D}'_{ii,j}$ are found to equal zero when computed under the
 180 same assumptions as for $\tilde{D}_{ij,i}$, and are thus neglected. We then have:

$$\tilde{D}'_{ij,i} = \tilde{D}_{ij,i}^{\text{sel}} - (\Delta_s p_j) D_{i,i} + o(\epsilon) \quad (\text{C50})$$

181 with (ignoring associations between different loci affecting fitness):

$$\Delta_s p_j = a_j p_j q_j + [a_j + a_{j,j} (1 - 2p_j)] D_{j,j} + o(\epsilon) \quad (\text{C51})$$

182

$$\tilde{D}_{ij,i}^{\text{sel}} = \tilde{D}_{ij,i} + a_j D_{i,i} p_j q_j + [a_j + a_{j,j} (1 - 2p_j)] D_{ij,ij} + o(\epsilon). \quad (\text{C52})$$

183 Therefore:

$$\tilde{D}'_{ij,i} = \tilde{D}_{ij,i} + [a_j + a_{j,j} (1 - 2p_j)] G_{ij} \tilde{D}_{ii} p_j q_j + o(\epsilon) \quad (\text{C53})$$

184 where again $G_{ij} = \phi_{ij} - F^2$ is the identity disequilibrium between loci i and j . Equa-
 185 tions C49 and C53 yield, at QLE:

$$\tilde{D}_{ij,i} = [a_j + a_{j,j} (1 - 2p_j)] F G_{ij} \tilde{D}_{ii} p_j q_j + o(\epsilon) \quad (\text{C54})$$

186 which is equivalent to equation A37 in Roze (2015). Using equation C53, we thus have:

$$\tilde{D}'_{ij,i} = [a_j + a_{j,j} (1 - 2p_j)] (1 + F) G_{ij} \tilde{D}_{ii} p_j q_j + o(\epsilon). \quad (\text{C55})$$

187 Equation C47 shows that associations \tilde{D}_{ij} , $\tilde{D}_{i,j}$ are generated by two different effects.
 188 The first involves the association $\tilde{D}_{ij,j}$, representing the fact that alleles increasing self-
 189 ing at locus i tend to be associated with more homozygous backgrounds at locus j . As
 190 a consequence, selection at locus j is more efficient among individuals carrying alleles
 191 that increase selfing, causing an increased frequency of the best allele at locus j in
 192 these individuals. The second effect involves the association $\tilde{D}'_{ij,i}$ and is generated by
 193 the identity disequilibrium G_{ij} between loci i and j (equation C55). Identity disequi-
 194 librium results from partial selfing, and represents the fact that homozygotes at locus
 195 i tend to be also homozygous at locus j . A consequence of this identity disequilibrium
 196 is that, because selection at locus j is more efficient among homozygotes at this locus,
 197 the frequency of the best allele at locus j tends to be higher among homozygotes at
 198 locus i than among heterozygotes, which is represented by the association $\tilde{D}_{ij,i}$. In
 199 the presence of direct selection at locus i (that is, when $\kappa < 1$), the fact that selection
 200 at locus i is more efficient among homozygotes than among heterozygotes, and the
 201 fact that homozygotes at locus i tend to be associated with the best allele at locus
 202 j generate a positive association between the alleles at both loci that are favored by
 203 direct selection. This corresponds to the previously described result that partial selfing
 204 generates positive associations between selected loci, even in the absence of epistasis
 205 among those loci (Roze and Lenormand, 2005; Kamran-Disfani and Agrawal, 2014).
 206 However, equation C48 indicates that the effect of $\tilde{D}_{ij,i}$ persists even in the absence
 207 of direct selection at locus i ($\kappa = 1$). This effect can be understood as follows. For
 208 simplicity, we will consider the case where only two alleles segregate at locus i , an allele
 209 S coding for more selfing, and an allele O coding for more outcrossing. Furthermore,
 210 we will call A the favored allele at locus j , and a the deleterious allele. $\tilde{D}_{ij,i}$ indicates

211 that A tends to be more frequent in SS and OO individuals, while a tends to be
 212 more frequent in OS individuals. The population thus contains an excess of $SA/S\cdot$,
 213 $OA/O\cdot$, $Sa/O\cdot$ and $Oa/S\cdot$ genotypes, where the dot \cdot stands for any allele at locus
 214 j , and where the slash symbol separates maternal and paternal haplotypes. The as-
 215 sociation between loci present on the same haplotype (\tilde{D}_{ij}) is not affected by whether
 216 individuals have been produced by selfing or outcrossing. However, the association
 217 between genes on different haplotypes ($\tilde{D}_{i,j}$) depends on how offspring are produced:
 218 in particular, selfing from $SA/S\cdot$ individuals maintains an association between S and
 219 A on different haplotypes, while this association is lost when offspring are produced
 220 by outcrossing. Similarly, selfing from $OA/O\cdot$ individuals maintains an association
 221 between O and A on different haplotypes. This effect is stronger in the case of $SA/S\cdot$
 222 individuals, however, due to their higher selfing rate: as a consequence, the net effect
 223 of this process is to generate an association between S and A on different haplotypes
 224 ($\tilde{D}_{i,j}$), which can then be converted into an association between genes on the same
 225 haplotype (\tilde{D}_{ij}) by recombination.

226 Because $\tilde{D}'_{ij,i} = 0$ while $\tilde{D}_{ij,j} = (1/2) \tilde{D}_{ii} p_j q_j$ in the absence of selfing ($\bar{\sigma} = 0$),
 227 the expressions for \tilde{D}_{ij} and $\tilde{D}_{i,j}$ simplify to:

$$\tilde{D}_{ij} \approx \frac{1}{2} \frac{a_j + a_{j,j} (1 - 2p_j)}{\rho_{ij} - a_j (1 - 2p_j) (1 - \rho_{ij})} \tilde{D}_{ii} p_j q_j, \quad \tilde{D}_{i,j} = 0. \quad (\text{C56})$$

228

229 **Variance in selfing rate after selection.** In order to complete the derivations, we
 230 can note that equation C11 involves the variance in selfing rate measured after selection
 231 $V'_\sigma = 2 \left(\tilde{D}'_{ii} + D'_{i,i} \right)$, which must be expressed in terms of V_σ before selection. From
 232 the QLE expressions given above, the term in $a_{j,j} \tilde{D}_{ij,j}$ of equation C14 is proportional

233 to $a_{j,j}$, while the term in $a_j \left(\tilde{D}_{ij} + \tilde{D}_{i,j} \right)$ is proportional to a_j^2 . For consistency, we
 234 thus express V'_σ to the first order in $a_{j,j}$, and to the second order in a_j : this can be
 235 achieved by supposing that a_j is of order η (where η is a small term) while $a_{j,j}$ is of
 236 order η^2 . We have:

$$\tilde{D}'_{ii} = \tilde{D}_{ii}^{\text{sel}} - \frac{1}{2} \left[\left(\Delta_s \overline{\sigma_i^M} \right)^2 + \left(\Delta_s \overline{\sigma_i^P} \right)^2 \right] \quad (\text{C57})$$

237

$$\tilde{D}'_{i,i} = \tilde{D}_{i,i}^{\text{sel}} - \left(\Delta_s \overline{\sigma_i^M} \right) \left(\Delta_s \overline{\sigma_i^P} \right). \quad (\text{C58})$$

238 Because $\Delta_s \overline{\sigma_i^M}$ and $\Delta_s \overline{\sigma_i^P}$ are proportional to the genetic variance in selfing rate, their
 239 products can be neglected, yielding:

$$\tilde{D}'_{ii} \approx \tilde{D}_{ii}^{\text{sel}}, \quad \tilde{D}'_{i,i} \approx \tilde{D}_{i,i}^{\text{sel}}. \quad (\text{C59})$$

240 We then have (from equation C13):

$$\tilde{D}_{ii}^{\text{sel}} = \tilde{D}_{ii} + \sum_j \left[a_j \left(\tilde{D}_{ii,j} + \tilde{D}_{i,i,j} \right) + a_{j,j} \left(\tilde{D}_{ii,j,j} - \tilde{D}_{ii} D_{j,j} \right) \right]. \quad (\text{C60})$$

241 As mentioned above, the associations $\tilde{D}_{ii,j}$ and $\tilde{D}_{i,i,j}$ equal zero at QLE, to the first
 242 order in selection coefficients, while $\tilde{D}_{ii,j,j}$ and $\tilde{D}_{ii} D_{j,j}$ both equal $F \tilde{D}_{ii} p_j q_j$ in the
 243 absence of selection; therefore:

$$\tilde{D}'_{ii} = \tilde{D}_{ii} + o(\eta^2). \quad (\text{C61})$$

244 Then,

$$\tilde{D}'_{i,i} = \tilde{D}_{i,i} + \sum_j \left[2a_j \tilde{D}_{ij,i} + a_{j,j} \left(D_{ij,ij} - D_{i,i} D_{j,j} \right) \right] \quad (\text{C62})$$

245 finally leading to:

$$\tilde{D}'_{ii} + D'_{i,i} \approx \tilde{D}_{ii} + D_{i,i} + \sum_j \left[\left(2a_j^2 F + a_{j,j} \right) G_{ij} p_j q_j \right] \tilde{D}_{ii} \quad (\text{C63})$$

246 and thus:

$$V'_\sigma \approx V_\sigma + 2 \sum_{i,j} [(2a_j^2 F + a_{j,j}) G_{ij} p_j q_j] \tilde{D}_{ii}. \quad (\text{C64})$$

247 Equation C64 stems from the fact that homozygosity at loci coding for the selfing
248 rate is affected by identity disequilibria between these loci and loci affecting fitness;
249 when a_j and $a_{j,j}$ are of the same order of magnitude, the term in a_j^2 in equation C64
250 becomes negligible, and the result becomes equivalent to equation 5 in Roze (2015).
251 Because $F = G_{ij} = 0$ when $\bar{\sigma} = 0$, equation C64 simplifies to $V'_\sigma \approx V_\sigma$ when the mean
252 selfing rate of the population approaches zero.

253

254 **Expressing $\Delta_{\text{purge}}\bar{\sigma}$ in terms of the increase in mean fitness following a**
255 **single generation of selfing.** From the results given above, the contribution of
256 purging to the strength of selection for selfing in a randomly mating population
257 ($\Delta_{\text{purge}}\bar{\sigma} = 2 \sum_{i,j} a_j \tilde{D}_{ij}$) may be expressed in terms of the effect of a single generation
258 of selfing on the mean fitness of offspring. Imagine a parental, non-inbred population
259 from which a pool of selfed offspring and a pool of outcrossed offspring are produced.
260 Inbreeding depression represents the difference in mean fitness between those two pools
261 of offspring. Imagine now that these individuals are let to reproduce in proportion to
262 their fitness, by random mating within each pool. We will call $\overline{W}_{\text{F2,self}}$ and $\overline{W}_{\text{F2,out}}$
263 the mean fitnesses of the offspring of selfed and outcrossed individuals, respectively.
264 Neglecting epistasis, the fitness of an individual relative to the mean fitness of the
265 parental population can be expressed as (from equation A1 in Supplementary File
266 S1):

$$\frac{W}{\overline{W}} = 1 + \sum_j a_j (\zeta_{j,\emptyset} + \zeta_{\emptyset,j}) + \sum_j a_{j,j} \zeta_{j,j} \quad (\text{C65})$$

267 where the allele frequencies p_j within $\zeta_{j,\emptyset}$, $\zeta_{\emptyset,j}$ and $\zeta_{j,j}$ are allele frequencies in the

268 parental population. Averaging $\zeta_{j,\emptyset} + \zeta_{\emptyset,j}$ over the offspring of selfed individuals yields
 269 $2\Delta_{\text{self}}p_j = 2\left(p_j^{\text{F2,self}} - p_j\right)$, where $p_j^{\text{F2,self}}$ is the frequency of allele 1 at locus j among
 270 the offspring of selfed individuals. Given that reproduction among selfed individuals
 271 occurs by random mating, one can show that the average of $\zeta_{i,i}$ among their offspring
 272 is given by $(\Delta_{\text{self}}p_j)^2$. Neglecting terms in $(\Delta_{\text{self}}p_j)^2$, we thus have:

$$\frac{\overline{W}_{\text{F2,self}}}{\overline{W}} \approx 1 + 2 \sum_j a_j \Delta_{\text{self}}p_j \quad (\text{C66})$$

273 and similarly

$$\frac{\overline{W}_{\text{F2,out}}}{\overline{W}} \approx 1 + 2 \sum_j a_j \Delta_{\text{out}}p_j. \quad (\text{C67})$$

274 The gain in mean fitness caused by one generation of purging through selfing can thus
 275 be expressed as:

$$P = \frac{\overline{W}_{\text{F2,self}}}{\overline{W}} - \frac{\overline{W}_{\text{F2,out}}}{\overline{W}} \approx 2 \sum_j a_j (\Delta_{\text{self}}p_j - \Delta_{\text{out}}p_j). \quad (\text{C68})$$

276 The change in p_j due to selection among selfed individuals is given by:

$$\Delta_{\text{self}}p_j = \text{E} \left[\frac{W}{\overline{W}_{\text{self}}} \frac{\zeta_{j,\emptyset} + \zeta_{\emptyset,j}}{2} \right] \quad (\text{C69})$$

277 where the average is over all selfed individuals, and where $\overline{W}_{\text{self}}$ is the mean fitness of
 278 these individuals. From equation C65, we have (given that $D_{j,j}$ among selfed offspring
 279 equals $p_jq_j/2$):

$$\frac{\overline{W}_{\text{self}}}{\overline{W}} = 1 + \frac{1}{2} \sum_j a_{j,j} p_jq_j \quad (\text{C70})$$

280 and thus:

$$\frac{W}{\overline{W}_{\text{self}}} = \frac{W/\overline{W}}{\overline{W}_{\text{self}}/\overline{W}} = 1 + \sum_j a_j (\zeta_{j,\emptyset} + \zeta_{\emptyset,j}) + \sum_j a_{j,j} \left(\zeta_{j,j} - \frac{1}{2}p_jq_j \right) \quad (\text{C71})$$

281 to the first order in $a_{\text{U,V}}$ coefficients. From equations C69 and C71, and neglecting
 282 associations between loci, one obtains:

$$\Delta_{\text{self}}p_j \approx \frac{1}{2} [3a_j + a_{j,j} (1 - 2p_j)] p_jq_j. \quad (\text{C72})$$

283 Similarly, we have:

$$\Delta_{\text{out}} p_j = \mathbb{E} \left[\frac{W}{\overline{W}} \frac{\zeta_{j,\emptyset} + \zeta_{\emptyset,j}}{2} \right] \approx a_j p_j q_j \quad (\text{C73})$$

284 and equation C68 thus leads to:

$$P \approx \sum_j a_j [a_j + a_{j,j} (1 - 2p_j)] p_j q_j. \quad (\text{C74})$$

- 286 Abu Awad, D. and D. Roze. 2018. Effects of partial selfing on the equilibrium ge-
287 netic variance, mutation load, and inbreeding depression under stabilizing selection.
288 *Evolution* 72:751–769.
- 289 Barton, N. H. and M. Turelli. 1991. Natural and sexual selection on many loci. *Genetics*
290 127:229–255.
- 291 Charlesworth, B. 1990. Mutation-selection balance and the evolutionary advantage of
292 sex and recombination. *Genet. Res.* 55:199–221.
- 293 Epinat, G. and T. Lenormand. 2009. The evolution of assortative mating and selfing
294 with in- and outbreeding depression. *Evolution* 63:2047–2060.
- 295 Gervais, C. and D. Roze. 2017. Mutation rate evolution in partially selfing and partially
296 asexual organisms. *Genetics* 207:1561–1575.
- 297 Kamran-Disfani, A. and A. F. Agrawal. 2014. Selfing, adaptation and background
298 selection in finite populations. *J. Evol. Biol.* 27:1360–1371.
- 299 Kirkpatrick, M., T. Johnson, and N. H. Barton. 2002. General models of multilocus
300 evolution. *Genetics* 161:1727–1750.
- 301 Roze, D. 2014. Selection for sex in finite populations. *J. Evol. Biol.* 27:1304–1322.
- 302 ———. 2015. Effects of interference between selected loci on the mutation load, in-
303 breeding depression and heterosis. *Genetics* 201:745–757.

- 304 Roze, D. and T. Lenormand. 2005. Self-fertilization and the evolution of recombination.
305 *Genetics* 170:841–857.
- 306 Weir, B. S. and C. C. Cockerham. 1973. Mixed self and random mating at two loci.
307 *Genet. Res.* 21:247–262.