



**HAL**  
open science

## Establishment of local adaptation in partly self-fertilizing populations

Bogi Trickovic, Sylvain Glémin

► **To cite this version:**

Bogi Trickovic, Sylvain Glémin. Establishment of local adaptation in partly self-fertilizing populations. *Genetics*, 2022, 220 (2), 10.1093/genetics/iyab201 . hal-03467503

**HAL Id: hal-03467503**

**<https://hal.science/hal-03467503>**

Submitted on 9 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 Establishment of local adaptation in partly  
2 self-fertilizing populations

3 Bogi Trickovic<sup>1,\*</sup> and Sylvain Glémin<sup>2,3</sup>

4 <sup>1</sup>Center for Mechanisms of Evolution, Arizona State University, AZ 85281,  
5 Tempe, United States of America

6 <sup>2</sup>University of Rennes, CNRS, ECOBIO [(Ecosystèmes, biodiversité,  
7 évolution)] - UMR 6553, Rennes, France

8 <sup>3</sup>Department of Ecology and Genetics, Evolutionary Biology Centre,  
9 Uppsala University, Uppsala, Sweden

10 \*Corresponding authors: btrickov@asu.edu, sylvain.glemin@univ-rennes1.fr

11  
12 **Running head:** local adaptation with selfing

13  
14 **Keywords:** branching process, local adaptation, migration-selection  
15 balance, protected polymorphism, selfing

16 November 2, 2021

**Abstract**18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

Populations often inhabit multiple ecological patches and thus experience divergent selection, which can lead to local adaptation if migration is not strong enough to swamp locally adapted alleles. Conditions for the establishment of a locally advantageous allele have been studied in randomly mating populations. However, many species reproduce, at least partially, through self-fertilization, and how selfing affects local adaptation remains unclear and debated. Using a two-patch branching process formalism, we obtained a closed-form approximation under weak selection for the probability of establishment of a locally advantageous allele ( $P$ ) for arbitrary selfing rate and dominance level, where selection is allowed to act on viability or fecundity, and migration can occur via seed or pollen dispersal. This solution is compared to diffusion approximation and used to investigate the consequences of a shift in a mating system on  $P$ , and the establishment of protected polymorphism. We find that selfing can either increase or decrease  $P$ , depending on the patterns of dominance in the two patches, and has conflicting effects on local adaptation. Globally, selfing favors local adaptation when locally advantageous alleles are (partially) recessive, when selection between patches is asymmetrical and when migration occurs through pollen rather than seed dispersal. These results establish a rigorous theoretical background to study heterogeneous selection and local adaptation in partially selfing species.

# 1 Introduction

Most advantageous alleles are lost from populations (Haldane, 1927), but those that escape extinction are ultimately destined to fix, providing that the fitness effect of an allele remains positive over space and time. However, fitness effects can vary across species' range, so that an allele that is advantageous in one ecological patch might be deleterious in another. Two mutually exclusive outcomes are then possible: a locally advantageous allele is maintained by divergent selection as polymorphism, or the allele with the best overall performance across the patches becomes fixed. Since ranges of essentially all species are spread across heterogeneous environments, it is of some interest to understand the conditions under which these outcomes are materialized. Although previous works investigated these conditions for randomly mating populations in both a deterministic (Bulmer, 1972, Felsenstein, 1976, Levene, 1953, Maynard Smith, 1970) and a stochastic setting (Tomasini and Peischl, 2018, Yeaman and Otto, 2011), conditions for maintenance of local adaptation in partially self-fertilizing populations remain unexplored. Partial selfing is widespread among eukaryotes and well-described in plants (Igic and Kohn, 2006), animals (Jarne and Auld, 2006), algae (Hanschen *et al.*, 2018), and fungi (Billiard *et al.*, 2012), and considering its effect on spatially heterogeneous selection is especially relevant for at least two reasons. First, partial selfing is more frequently found among sessile or less mobile organisms, for which landscape heterogeneity directly translates into local selective pressures because of the absence of the homogenizing effect of movements across the environment. Second, partial selfing likely influences the conditions for local adaptation by altering the effective population size and effective migration rate, hence the genetic structure of a population, but also by reducing the role of heterozygotes' fitness in the dynamics of allelic frequencies.

In a single population, selfing exerts two opposing effects on the fixation probability of an advantageous allele (Pollak (1987), Pollak and Sabran (1992), Charlesworth (1992), Caballero and Hill (1992)). Firstly, it increases the rate at which a rare advantageous allele spreads through the population. As the selfing rate increases, homozygotes appear more quickly so that the spread of an advantageous allele becomes increasingly decoupled from the fitness of the heterozygote. In other words, selfing increases the effective dominance coefficient: the rate of spread in a partially selfing population with dominance coefficient  $h$ , would be equal to the rate of spread in a randomly mating population with dominance coefficient  $\tilde{h} = F + (1 - F)h$  (where  $F$  is the equilibrium population inbreeding coefficient). Secondly, since alleles making up the progeny are not independently sampled, selfing also reduces the efficacy of selection by reducing the effective population size,  $\tilde{N}$ , by a factor of  $1 + F$  (Gale

71 (1990)). These two opposing effects cancel out when an advantageous allele is codominant  
72 ( $h = 0.5$ ), implying that the fixation probability of a codominant advantageous allele is in-  
73 variant across populations with different selfing rates. If an advantageous allele is partially  
74 recessive ( $h < 0.5$ ), then the effect of increased efficacy of selection through increased ef-  
75 fective dominance outweighs the reduction in  $\tilde{N}$  yielding an increase in the establishment  
76 probability. Conversely, partially dominant advantageous alleles ( $h > 0.5$ ) are less likely to  
77 fix because the reduced efficiency due to lower effective population size trumps increased  
78 efficacy of selection due to increased homozygosity. These predictions are altered for male  
79 and fecundity selection. Under selfing, no selection occurs on the male function, but the  
80 strength of selection on the female function is doubled because two gametes are transmitted  
81 for one selected parent (Damgaard, 2000).

82

83 The evolutionary dynamic becomes further elaborated in structured populations with di-  
84 vergent selection. The simplest scenario involves two patches with an invading allele being  
85 advantageous in one but deleterious in the other patch. The dynamic of local adaptation  
86 then further depends on the effect that selfing has on gene flow. Partial selfing reduces the  
87 effective haploid migration rate due to the reduction in the number of female gametes that  
88 can be sired by immigrating male gametes, for individuals in the selfing population will have  
89 already fertilized themselves to a large extent. This reduction in gene flow has been pro-  
90 posed to prevent the spread of maladaptive alleles from neighboring populations and thus  
91 to promote local adaptation (Linhart and Grant, 1996), but two meta-analyses of reciprocal  
92 transplant experiments reported the absence of correlation between local adaptation and the  
93 mating system (Hereford, 2010, Leimu and Fischer, 2008). However, it is not immediately  
94 clear how the interplay between the mating system, migration, and selection modes affects  
95 the propensity for local adaptation. When local reproduction is panmictic, maintenance of  
96 local polymorphism roughly requires that local selection is stronger than migration. On the  
97 one hand, selfing reduces effective migration. On the other hand, it may either increase or  
98 decrease the efficacy of local selection, depending on the balance between the effect of in-  
99 creasing effective dominance and increasing genetic drift. In addition, in diploids, conditions  
100 where alleles are (partially) dominant when locally beneficial and (partially) recessive when  
101 locally deleterious are especially favorable to the maintenance of polymorphism (Yeaman  
102 and Otto, 2011). By unmasking recessive alleles, selfing is expected to reduce the range of  
103 applicability of those conditions. Overall, we still lack correct theoretical predictions, which  
104 makes it difficult to interpret the empirical results such as in Leimu and Fischer (2008) and  
105 Hereford (2010).

106

107 In the present work, we extend the results of previous theoretical models of local adaptation  
108 to partially self-fertilizing populations. We first derive a closed-form approximation for the  
109 probability that a locally advantageous allele escapes extinction. We then use this result  
110 to examine how previously described effects of selfing affect (1) the probability that allele  
111 escapes extinction, (2) conditions required for the establishment of protected polymorphism,  
112 and (3) the dependence of these properties on the mode of selection and migration.

## 113 2 Methods

114 We work with a hermaphroditic population that is divided into two patches connected by  
115 bidirectional migration. Migration can occur during both the haploid and the diploid phase.  
116 In the following, we will use the example of a plant life cycle, but results can readily be  
117 transposed to other life cycles. Thus, only male migration, pollen dispersal, occurs during  
118 the haploid phase, and diploid migration corresponds to seed dispersal. Fitness is controlled  
119 by a single biallelic locus, with allele  $A$  favored in the first patch and allele  $a$  favored in the  
120 other patch. These patches are referred to as favored and disfavored patch, respectively. The  
121 life cycle of the modeled organism is composed of the following sequential stages: (1) male and  
122 female meiosis leading to the production of male gametophyte (pollen containing sperm  
123 cell) and female gametophyte (embryo sac, containing egg cell), (2) pollen dispersal, (3)  
124 mating (including selfing) and seed formation, (4) seed dispersal, and (5) seed establishment  
125 and development, which yields adults of the next generation. To make the model general,  
126 we allow selection to operate on viability and on male and female fecundity. More precisely,  
127 adults in patch  $i$  with genotype  $k$  ( $AA$ ,  $Aa$ , or  $aa$ ) produce  $W_{i,k}^{\text{♀}}$  female and  $W_{i,k}^{\text{♂}}$  male  
128 gametophytes. We do not consider haploid selection, but this could readily be included  
129 in the model setting. Finally, each offspring with genotype  $k$  survives to maturity with  
130 probability  $V_{i,k}$ . The population selfing rate,  $S$ , is defined as the proportion of individuals in  
131 the population that self-fertilize and is equal across patches. Following previous works (ex:  
132 Hartfield and Glémin, 2016, Hössjer and Tyvand 2016) we assume that the population reaches  
133 equilibrium for inbreeding coefficient  $F$  and genotypic composition on a much shorter time-  
134 scale than the change in allelic frequencies (separation of time scale argument). Within each  
135 patch, we can thus write genotype frequencies directly as a function of allelic frequencies and  
136  $F$ . If selection is not too strong, we can also use the neutral expectation:  $F = S/(2 - S)$ .  
137 This assumption is relaxed in simulations. Mathematical symbols and their meaning are  
138 outlined in Table 1.

## 139 2.1 Deriving the difference equation for allele frequency

140 Using the preceding definitions, one can derive the difference equations for allele frequencies  
 141 in each patch. Let  $X_i$  and  $Y_i$  be the frequencies of genotypes  $AA$  and  $Aa$ , and  $x_i$  be the  
 142 frequency of allele  $A$  in  $i^{th}$  patch. Our system has two degrees of freedom, given that there  
 143 are three genotype frequencies and they must add up to unity. Thus, we keep track of  
 144 frequencies of genotypes  $AA$  and  $Aa$ , while the frequency of  $aa$  genotype is by definition  
 145  $1 - X_i - Y_i$ . Adult genotype frequencies in generation  $t$  in  $i^{th}$  patch are:

$$X_i = x_i^2(t) + Fx_i(t)(1 - x_i(t)) \quad (1a)$$

$$Y_i = 2x_i(t)(1 - x_i(t))(1 - F) \quad (1b)$$

146 Adult genotypes may produce different number of male and female gametophytes. Let  $W_{i,k}^{\sigma}$   
 147 and  $W_{i,k}^{\varphi}$  be the number of male and female gametophytes that an adult with genotype  $k$   
 148 ( $AA$ ,  $Aa$  and  $aa$ ) produces in patch  $i$ . Thus, the frequency of female and male gametophyte  
 149  $A$  in  $i^{th}$  patch are, respectively:

$$x_i^{\varphi} = \left( W_{i,AA}^{\varphi} X_i + W_{i,Aa}^{\varphi} \frac{Y_i}{2} \right) / \bar{W}_i^{\varphi} \quad (2a)$$

$$x_i^{\sigma} = \left( W_{i,AA}^{\sigma} X_i + W_{i,Aa}^{\sigma} \frac{Y_i}{2} \right) / \bar{W}_i^{\sigma} \quad (2b)$$

150 where  $\bar{W}_i^{\varphi} = X_i W_{i,AA}^{\varphi} + Y_i W_{i,Aa}^{\varphi} + (1 - X_i - Y_i) W_{i,aa}^{\varphi}$  and  $\bar{W}_i^{\sigma} = X_i W_{i,AA}^{\sigma} + Y_i W_{i,Aa}^{\sigma} + (1 -$   
 151  $X_i - Y_i) W_{i,aa}^{\sigma}$ . After pollen migration, the frequency of allele  $A$  in pollen in  $i^{th}$  patch is:

$$g_i^{\sigma} = (1 - m_{ij}) x_i^{\sigma} + m_{ij} x_j^{\sigma} \quad (3)$$

152 where  $m_{ij}$  is the fraction of pollen in  $i^{th}$  patch that come from  $j^{th}$  patch. Pollen dispersal is  
 153 followed by mating, which yields offspring genotypes:

$$X'_i = (1 - S) x_i^{\varphi} g_i^{\sigma} + S \left( \frac{W_{i,AA}^{\varphi}}{\bar{W}_i^{\varphi}} X_i + \frac{Y_i}{4} \frac{W_{i,Aa}^{\varphi}}{\bar{W}_i^{\varphi}} \right) \quad (4a)$$

$$Y'_i = (1 - S) \left( x_i^{\varphi} (1 - g_i^{\sigma}) + (1 - x_i^{\varphi}) g_i^{\sigma} \right) + S \frac{W_{i,Aa}^{\varphi}}{\bar{W}_i^{\varphi}} \frac{Y_i}{2} \quad (4b)$$

154 Seed migration changes the genotype frequency to:

$$X''_i = (1 - M_{ij})X'_i + M_{ij}X'_j \quad (5a)$$

$$Y''_i = (1 - M_{ij})Y'_i + M_{ij}Y'_j \quad (5b)$$

155 where  $M_{ij}$  is the fraction of seed in patch  $i$  originating from patch  $j$ . Genotype  $j$  in patch  
 156  $i$  has probability  $V_{i,j}$  to survive to maturity, so the frequency of allele  $A$  in the generation  
 157  $t + 1$  is:

$$x_i(t + 1) = \left( V_{i,AA}X''_i + V_{i,Aa}\frac{Y''_i}{2} \right) / \bar{V}_i \quad (6)$$

158 where  $\bar{V}_i = X''_iV_{i,AA} + Y''_iV_{i,Aa} + (1 - X''_i - Y''_i)V_{i,aa}$ .

159

160 Equation (6) is our difference equation that expresses the frequency of allele  $A$  in the next  
 161 generation in terms of its frequency in the current generation.

## 162 2.2 Analysis of the deterministic model

163 We first analyze the deterministic model that gives the conditions for a protected polymor-  
 164 phism. This extends results of Bulmer (1972) to partial selfing and more general forms of  
 165 selection. Given that investigation of the consequences of an arbitrary selection, regime is  
 166 prohibitively complicated, the proceeding analysis is restricted to a few special cases. We are  
 167 ultimately interested in three selection scenarios: (1) when selection operates on differential  
 168 survival to maturity and when selection operates on (2) female and (3) male fecundity. For  
 169 each of these scenarios, we first consider seed migration only, as in Bulmer (1972) initial  
 170 model (also as in Yeaman and Otto, 2011). In addition to the comparison with previous  
 171 models, it allows analyzing the effect of selfing on selection only, as selfing does not affect  
 172 seed migration. Then, we consider pollen migration thus analyzing the joint effect of selfing  
 173 on both selection and migration. In total, we examine six categories of scenarios.

174

175 We use the standard approach to obtain conditions for stable polymorphism by considering  
 176 the conditions for which both monomorphic equilibria are unstable. Let both patches be fixed  
 177 for allele  $a$  and allele  $A$  acts as an invader. Equivalent results can be obtained by considering  
 178 patches fixed for allele  $A$  and  $a$  acting as the invader. The logic of derivation is the same  
 179 in all cases, wherein we linearize equations (6) around  $\vec{x} = [0, 0]^T$  to obtain the system of  
 180 linear difference equations, where  $\mathbf{J}$  is the associated Jacobian. While  $\mathbf{J}$  captures the invasion



181 dynamics for an arbitrary mode of selection and migration, it is difficult to analyze, so we  
 182 focus on special population genetic cases outlined above. Each of these scenarios will have  
 183 an associated Jacobian  $\mathbf{J}^{\{k\}}$ :

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \mathbf{J}^{\{k\}} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad (J_{ij}^{\{k\}}) = \begin{cases} k = \circ, & M_{ij}\Delta V_{0,i} \\ k = \text{\textcircled{f}}, & M_{ij}\Delta W_{0,i}^{\text{\textcircled{f}}} \\ k = \text{\textcircled{m}}, & M_{ij}\Delta W_{0,j}^{\text{\textcircled{m}}} \end{cases} \quad (7)$$

184 Terms  $\Delta V_{0,i}$ ,  $\Delta W_{0,i}^{\text{\textcircled{f}}}$ , and  $\Delta W_{0,i}^{\text{\textcircled{m}}}$  are the intergenerational change in frequency of  $A$  in  $i^{\text{th}}$   
 185 patch due to selection on a particular fitness component, and  $M_{ij}$  accounts for migration  
 186 between patches, with  $M_{i,i} = 1 - M_{i,j}$ , with  $j \neq i$ . In the case of viability selection, the  
 187 frequency change of  $A$  in the patch 1 reflects either the contribution of alleles  $A$  that stay in  
 188 patch 1 ( $1 - M_{12}$ ) and survives to adulthood ( $\Delta V_{0,1}$ ), or the contribution of alleles  $A$  that  
 189 migrate from patch 2 ( $M_{12}$ ) and then survive to adulthood ( $\Delta V_{0,1}$ ). Similar logic holds for  
 190 female and male fecundity selection but off-diagonal selection terms are exchanged because  
 191 selection occurs prior to migration: note the  $j$  indices in  $\Delta W$  terms instead of  $i$  in the  
 192  $\Delta V$  term in equation (7). Selection terms for the three scenarios are obtained by keeping  
 193 constant the two fitness components not involved in the scenarios in the general recursion  
 194 equations. Thus, under seed migration the three selection terms are:

$$\Delta V_{0,i} = \frac{FV_{i,AA} + (1-F)V_{i,Aa}}{V_{i,aa}} \quad (8a)$$

$$\Delta W_{0,i}^{\text{\textcircled{f}}} = \frac{(1-F)(W_{i,aa}^{\text{\textcircled{f}}} + (1+3F)W_{i,Aa}^{\text{\textcircled{f}}}) + F(1+3F)W_{i,AA}^{\text{\textcircled{f}}}}{2(1+F)W_{i,aa}^{\text{\textcircled{f}}}} \quad (8b)$$

$$\Delta W_{0,i}^{\text{\textcircled{m}}} = \frac{(1-F)((1-F)W_{i,Aa}^{\text{\textcircled{m}}} + FW_{i,AA}^{\text{\textcircled{m}}}) + (1+3F)W_{i,aa}^{\text{\textcircled{m}}}}{2(1+F)W_{i,aa}^{\text{\textcircled{m}}}} \quad (8c)$$

195 These expressions can be made more intuitive by parameterizing fitnesses relative to geno-  
 196 type  $aa$ , where  $s_i^{\circ}$ ,  $s_i^{\text{\textcircled{f}}}$ ,  $s_i^{\text{\textcircled{m}}}$  are the reproductive advantage of  $AA$  homozygote relative to  
 197  $aa$  homozygote when selection acts on viability, female, and male fecundity, respectively.  
 198 Parameters  $\tilde{h}_i^{\circ}$ ,  $\tilde{h}_i^{\text{\textcircled{f}}}$ ,  $\tilde{h}_i^{\text{\textcircled{m}}}$  represent the effective dominance of the heterozygote in  $i^{\text{th}}$  patch  
 199 for the three respective selection modes. These are composite parameters that depend on  $F$   
 200 and the actual dominance and are introduced to capture the fact that selfing increases the  
 201 effective dominance of invading heterozygote due to contribution of mutant homozygotes to  
 202 the invasion process. For the three cases of viability, female, and male fecundity effective

203 dominances are respectively:

$$\tilde{h}_i^\circ = (1 - F)h_i^\circ + F \quad (9a)$$

$$\tilde{h}_i^\ominus = \frac{(1 + 3F)}{2(1 + F)}(F + (1 - F)h_i^\ominus) = \frac{1}{2}(1 + S)(F + (1 - F)h_i^\ominus) \quad (9b)$$

$$\tilde{h}_i^\sigma = \frac{(1 - F)}{2(1 + F)}(F + (1 - F)h_i^\sigma) = \frac{1}{2}(1 - S)(F + (1 - F)h_i^\sigma) \quad (9c)$$

204 These forms capture the fact that selfing increases the effective dominance of invading  
 205 mutants: the invader has dominance  $h_i$  when it invades in a heterozygous form (that is,  
 206  $1 - F$  fraction of time), and dominance of one when it invades in a homozygous form (that  
 207 is,  $F$  fraction of time). Selfing also affects the global selection intensity on male and female  
 208 fecundity. The  $1/2$  factor captures the fact that selection acts only on half of the gametes  
 209 produced by each genotype, the  $1 - S$  factor that selection on male fecundity only operates  
 210 under outcrossing and the  $1 + S$  factor that under selfing an individual contribute two alleles  
 211 through seed production (female fitness). Taken together:

$$\Delta V_{0,i} = 1 + \tilde{h}_i^\circ s_i^\circ \quad (10a)$$

$$\Delta W_{0,i}^\ominus = 1 + \tilde{h}_i^\ominus s_i^\ominus \quad (10b)$$

$$\Delta W_{0,i}^\sigma = 1 + \tilde{h}_i^\sigma s_i^\sigma \quad (10c)$$

212 The fitness of the mutant always take the same form as in a haploid model with an effective  
 213 selective advantage  $\tilde{s} = \tilde{h}s$ .

214

215 The preceding approach can be readily extended to the selective scenarios where migration  
 216 occurs solely via pollen. One can still use the Jacobian in (7), albeit with the modified  
 217 migration rate to account for the facts that (a) pollen is haploid and thus do not contribute  
 218 in the same way as diploid seed to the gene pool of the next generation, and (b) that in all  
 219 three cases migration operates prior to selection rather than after:

$$(J_{ij}^{\{k\}}) = \begin{cases} k = \circ, & \tilde{M}_{ij}^\circ \Delta V_{0,i} \\ k = \ominus, & \tilde{M}_{ij}^\ominus \Delta W_{0,i}^\ominus \\ k = \sigma, & \tilde{M}_{ij}^\sigma \Delta W_{0,i}^\sigma \end{cases} \quad (11)$$

220 Selection terms are still parameterized according to equations (10a–10c), but effective mi-

221 gration is given by:

$$\tilde{M}_{ij}^{\circ} = (1 - S) \frac{m_{ij}}{2} \frac{1 + h_i^{\circ} s_i^{\circ}}{1 + \tilde{h}_i^{\circ} s_i^{\circ}} \quad (12a)$$

$$\tilde{M}_{ij}^{\ominus} = (1 - S) \frac{m_{ij}}{2} \frac{1}{1 + \tilde{h}_i^{\ominus} s_i^{\ominus}} \quad (12b)$$

$$\tilde{M}_{ij}^{\sigma} = (1 - S) \frac{m_{ij}}{2} \frac{1 + \frac{\tilde{h}_j^{\sigma} s_j^{\sigma}}{(1-S)/2}}{1 + (1 - m_{ij})\tilde{h}_i^{\sigma} s_i^{\sigma} + m_{ij}\tilde{h}_j^{\sigma} s_j^{\sigma}} \quad (12c)$$

222 Viability effective migration rate (12a) can be intuitively understood by noting that  $1 - S$   
 223 corresponds to the fraction of pollen that is exchanged between the patches,  $m_{ij}/2$  corrects  
 224 for the fact that haploid phase contributes two-fold less alleles to the gene pool of the  
 225 next generation relative to the diploid progeny dispersal, and the relative fitness ratio  $(1 +$   
 226  $h_i^{\circ} s_i^{\circ}) / (1 + \tilde{h}_i^{\circ} s_i^{\circ})$  accounts for the fold increase in invader frequency in the patch of origin  
 227 due to the inhibitory effect of selfing on gene flow. Similar interpretation holds for (12b)  
 228 and male fecundity selection (12c), except that fitness ratio now corrects for selection acting  
 229 prior to migration. Note that for weak selection, the three effective migration rates reduce to  
 230  $(1 - S) \frac{m_{ij}}{2}$ . The only complication is that when selection acts on the male fitness component,  
 231 one has to modify the effective dominance along with the effective migration rate. More  
 232 precisely,  $\Delta W_{0,i}^{\sigma} = 1 + \omega \tilde{h}_i^{\sigma} s_i^{\sigma}$ , where  $\omega$ :

$$\omega = 1 - m_{ij} \left( 1 - \frac{\tilde{h}_j^{\sigma} s_j^{\sigma}}{\tilde{h}_i^{\sigma} s_i^{\sigma}} \right) \quad (13)$$

233 Overall, a simple parameterization procedure allows the exploration of a wide breadth of  
 234 population genetic scenarios.

### 235 **2.3 Branching process approximation**

236 By analyzing the stability of system 7, one can only examine conditions necessary for the  
 237 invading allele to escape extinction. However, we also want to know the probability of the  
 238 escape. We pursue this by representing the change in the number of invading allele as  
 239 a multi-type branching process. Selection is thus assumed to be stronger than drift (i.e.,  
 240  $|N \tilde{h}_i^{\{k\}} s_i^{\{k\}}| \gg 1$ ). We work with two types, each tracking the number of mutant alleles  
 241 in one of the two patches. The number of the mutant alleles is represented by the vector  
 242  $\vec{N}(t) = [N_1(t), N_2(t)]^T$ , where  $N_1(t)$  and  $N_2(t)$  are the number of mutants in patch 1 and 2 in  
 243 generation  $t$ , respectively. Let  $f_i$  be the probability generating function (PGF) of offspring

244 number of mutant alleles of type  $i$ . The ultimate probability of extinction of type  $i$  ( $1 - P_i$ )  
 245 is given by the smallest positive root of the following system of equations (Haccou *et al.*  
 246 (2005)):

$$1 - P_1 = f_1(1 - P_1, 1 - P_2) \quad (14a)$$

$$1 - P_2 = f_2(1 - P_1, 1 - P_2) \quad (14b)$$

247 As we can assume that the number of offspring that types  $i$  and  $j$  leave are uncorrelated,  
 248 then  $f_k(1 - P_1, 1 - P_2) = f_{k,1}(1 - P_1)f_{k,2}(1 - P_2)$ , with  $k = \{1, 2\}$ . Intuitively, the probability  
 249 that mutant goes extinct conditioning on appearance in patch 1 is the probability that all of  
 250 its offspring left in patch 1 ultimately go extinct ( $f_{1,1}(1 - P_1)$ ), and the probability that all  
 251 offspring left in patch 2 via migration ( $f_{1,2}(1 - P_2)$ ) are ultimately lost. A similar explanation  
 252 holds when mutant originates in patch 2. To solve the system, one needs to obtain the  $f_{i,j}(z)$   
 253 PGFs.

254

255 In the Wright-Fisher model with random mating, the number of mutant offspring is ap-  
 256 proximately Poisson-distributed with mean  $\Delta\mathcal{W}_{0,i}$ , where  $\mathcal{W}$  stands for the different form of  
 257 selection as described above. Partial self-fertilization affects the mutant offspring distribu-  
 258 tion in two ways. First, it increases the mean number of offspring by inflating the effective  
 259 dominance coefficient (see above). Second, selfing increases the variance in offspring number  
 260 by the factor  $1 + F$  relative to the randomly mating population (Caballero and Hill, 1992,  
 261 Pollak, 1987). Therefore, offspring number should follow a distribution with mean  $\Delta\mathcal{W}_{0,i}$   
 262 and variance  $\Delta\mathcal{W}_{0,i}(1 + F)$ . To our knowledge, the full distribution has never been obtained,  
 263 even in single populations, probably because only the two first moments are needed for  
 264 branching process and diffusion approximations. In appendix, we derive the exact probabil-  
 265 ity generating function (PGF) of the distribution in different cases. In a single population  
 266 we have:

$$\begin{aligned}
 f_\phi(z) = (1 - F) & \underbrace{e^{-(1-z) \left( (1-S) \frac{W_{Aa}^\circ + W_{Aa}^\sigma}{2} V_{Aa} + \frac{S}{2} W_{Aa}^\circ V_{Aa} \right)}}_{\text{Transmitted in } Aa} \underbrace{e^{-\frac{1}{2}(1-z^2) \frac{S}{2} W_{Aa}^\circ V_{AA}}}_{\text{Transmitted in } AA \text{ (only selfing)}} \\
 + F & \underbrace{e^{-(1-z) \left( (1-S) \frac{w_{AA}^\circ + w_{AA}^\sigma}{2} V_{Aa} \right)}}_{\text{Transmitted in } Aa \text{ (only outcrossing)}} \underbrace{e^{-\frac{1}{2}(1-z^2) S W_{AA}^\circ V_{AA}}}_{\text{Transmitted in } AA \text{ (only selfing)}} \quad (15)
 \end{aligned}$$

267 The first term in equation (15) corresponds to a  $A$  parental allele coming from an  $Aa$  individ-

268 ual (probability  $1 - F$ ) and the second term to an  $AA$  individual (probability  $F$ ). Offspring  
 269 can be transmitted either in  $AA$  offspring, which occurs only through selfing (because we  
 270 consider a rare mutant) or in  $Aa$  offspring, which can occur either through outcrossing or  
 271 through selfing when the parent is  $Aa$  (so with a probability  $S/2$ ). Note that fecundity selec-  
 272 tion depends on parent genotypes whereas viability selection depends on offspring genotype.  
 273 Using the properties of PGFs we easily retrieve that the mean of the distribution is  $\Delta\mathcal{W}_i$   
 274 obtained in the deterministic analysis and the variance is  $\Delta\mathcal{W}_i(1 + F)$ . For  $F = 0$ , equation  
 275 (15) boils down to the PGF of a Poisson distribution. Hereafter, as for the deterministic  
 276 analysis, we consider the different forms of selection separately by setting constant the other  
 277 fitness components.

278

279 For the two-patch model we need to include migration in the  $f_{i,j}(z)$  PGFs. For seed migra-  
 280 tion, we show that:

$$\begin{aligned}
 f_{i,j}(z) = & (1 - F)e^{- (1-z) \left( (1-S) \frac{W_{Aa,i}^{\circlearrowleft} + W_{Aa,i}^{\circlearrowright}}{2} M_{i,j} V_{Aa,j} + \frac{S}{2} W_{Aa,i}^{\circlearrowleft} M_{i,j} V_{Aa,j} \right)} e^{-\frac{1}{2}(1-z^2) \frac{S}{2} W_{Aa,i}^{\circlearrowleft} M_{i,j} V_{AA,j}} \\
 & + Fe^{- (1-z) \left( (1-S) \frac{W_{AA,i}^{\circlearrowleft} + W_{AA,i}^{\circlearrowright}}{2} M_{i,j} V_{Aa,j} \right)} e^{-\frac{1}{2}(1-z^2) S W_{AA,i}^{\circlearrowleft} M_{i,j} V_{AA,j}}
 \end{aligned} \tag{16}$$

281 recalling that  $M_{i,i} = 1 - M_{i,j}$  with  $j \neq i$ . Note that fecundity selection occurs in patch  $i$ ,  
 282 before seed migration, whereas viability selection occurs in patch  $j$ , after seed migration.  
 283 However, the different orderings of migration and selection yield the same results (see *Math-*  
 284 *ematica* notebook). We also retrieve that the mean of the four distributions are given by the  
 285 Jacobian terms obtained above,  $M_{i,j} \Delta\mathcal{W}_{i,j}$  and the variance is also inflated by  $1 + F$ . For  
 286 pollen migration the form is slightly different because migrants can only contribute to the  
 287 next generation through outcrossing and male gametes (whereas selfed seeds can migrate).  
 288 The form of the PGF is thus different for the resident and migrant contribution. We have:

$$\begin{aligned}
 f_{i,i}(z) = & (1 - F)e^{- (1-z) \left( (1-S) \frac{W_{Aa,i}^{\circlearrowleft} + W_{Aa,i}^{\circlearrowright} (1-m_{i,j})}{2} V_{Aa,i} + \frac{S}{2} W_{Aa,i}^{\circlearrowleft} V_{Aa,i} \right)} e^{-\frac{1}{2}(1-z^2) \frac{S}{2} W_{Aa,i}^{\circlearrowleft} V_{AA,i}} \\
 & + Fe^{- (1-z) \left( (1-S) \frac{W_{AA,i}^{\circlearrowleft} + W_{AA,i}^{\circlearrowright} (1-m_{i,j})}{2} V_{Aa,i} \right)} e^{-\frac{1}{2}(1-z^2) S W_{AA,i}^{\circlearrowleft} V_{AA,i}}
 \end{aligned} \tag{17a}$$

$$f_{i,j}(z) = (1 - F)e^{- (1-z)(1-S) \frac{W_{Aa,i}^{\circlearrowright} m_{i,j}}{2} V_{Aa,j}} + Fe^{- (1-z)(1-S) \frac{W_{AA,i}^{\circlearrowright} m_{i,j}}{2} V_{Aa,j}} \quad \text{for } j \neq i \tag{17b}$$

289 As expected, the mean of the distributions are still given by the Jacobian terms with the

290 appropriate effective migration rates,  $\tilde{M}_{i,j}\Delta\mathcal{W}_{i,j}$  (see above). However, the variances are not  
 291 inflated by  $1 + F$ . For migrant alleles, equation (17b) boils down to a Poisson distribution  
 292 because the contribution is only through outcrossing. In contrast for resident allele, the  
 293 variance is inflated by more than  $1 + F$  (see appendix for details). The reason is that for a  
 294 given selfing rate, only alleles produced by outcrossing can be exported to the other patch,  
 295 which increase the proportion of resident offspring contributed by selfing, hence the variance.  
 296 Note, however, that the difference in PGFs between the two migration modes only affects  
 297 high-order terms of selection and migration so has no effect on branching process approxi-  
 298 mation results, which can all be expressed with the same form using appropriate effective  
 299 parameters as defined in previous section (see *Mathematica* notebook).

300

301 Previous section also showed that  $\Delta\mathcal{W}_{0,i}$  always has the form of  $1 + \tilde{s}_i$ , where  $\tilde{s}_i$  is the effective  
 302 advantage of an invading mutant in patch  $i$ , while  $M_{ij}$  represents the effective migration rate  
 303 between patches  $i$  and  $j$ . Thus, one can derive  $P_i$  for this generic case and then parameterize  
 304 migration and selection to reflect the specific population genetic case. Since we cannot obtain  
 305 the exact solution for extinction probabilities, we approximate to weak selection using the  
 306 approach of Tomasini and Peischl (2018), which is based on the approximation for slightly  
 307 supercritical multitype branching processes (Haccou *et al.*, 2005). All model parameters  
 308 are rescaled by  $\tilde{s}_1$  such that  $M_{ij} := \chi_{ij}\tilde{s}_1$  and  $\tilde{s}_2 := \zeta\tilde{s}_1$ . By linearizing around  $\tilde{s}_1 = 0$   
 309 and neglecting higher-order terms, we obtain a closed-form solution for the probability of  
 310 establishment conditioning on the patch in which locally advantageous allele appeared. All  
 311 analytical solutions are compared to simulated data using the method described in Appendix  
 312 D. Simulations were conducted with the help of GNU Parallel (Tange, 2011).

## 313 2.4 Diffusion process approximation

314 An alternative way of incorporating stochastic transmission of alleles across generations is by  
 315 approximating frequency change to diffusion. This approach has been originally developed  
 316 for the two-patch model in an outcrossing population by Sakamoto and Innan (2019) in the  
 317 case of wholly outcrossing population. Here, we re-capitulate their derivation and extend  
 318 it to the case of partial selfing. The establishment probability satisfies the Kolmogorov  
 319 backward equation (KBE):

$$0 = \sum_{i=1}^2 \mu_i \frac{\partial P(x_1, x_2)}{\partial x_i} + \sum_{i=1}^2 \frac{\sigma_i}{2} \frac{\partial^2 P(x_1, x_2)}{\partial x_i^2} \tag{18}$$

320 where  $\mu_i$  and  $\sigma_i$  are drift and diffusion coefficients for patch  $i$ , and are given by:

$$\mu_i = \tilde{s}_i x_i (1 - x_i) + \tilde{M}_{ij} (x_j - x_i), \quad \sigma_i = \frac{x_i (1 - x_i)}{2N_1 / (1 + F)} \quad (19)$$

321 Note that  $\tilde{s}$  corresponds to effective selective advantage of mutant, and  $\tilde{M}_{ij}$  is the effective  
 322 migration rate; These were derived in section 2.2. Diffusion coefficient is inflated by the  
 323 factor of  $1 + F$  as described previously. Unfortunately, we could not find neither the explicit  
 324 nor approximate solution to this partial differential equation. Following treatment in Barton  
 325 (1987), the problem is simplified by focusing on the first phase of fixation when mutant allele  
 326 is rare. Thus, one can approximate coefficients as:

$$\sigma_i = \frac{x_i}{2N_1 / (1 + F)} + o(x_i^2), \quad \mu_i = \tilde{s}_i x_i + \tilde{M}_{ij} (x_j - x_i) + o(x_i^2) \quad (20)$$

327 which leads to the simplified KBE:

$$\begin{aligned} 0 = & \left( \tilde{s}_1 x_1 + \tilde{M}_{12} (x_2 - x_1) \right) \frac{\partial P(x_1, x_2)}{\partial x_1} + \left( \tilde{s}_2 x_2 + \tilde{M}_{21} (x_1 - x_2) \right) \frac{\partial P(x_1, x_2)}{\partial x_2} + \\ & \frac{x_1}{4N_1 / (1 + F)} \frac{\partial^2 P(x_1, x_2)}{\partial x_1^2} + \frac{x_2}{4N_2 / (1 + F)} \frac{\partial^2 P(x_1, x_2)}{\partial x_2^2} \end{aligned} \quad (21)$$

328 Because the invading allele is rare, the mutants in patch  $i$  go extinct independently of one  
 329 another with probability  $e^{-l_i}$ , which means that the establishment probability takes the form:

$$P(x_1, x_2) = 1 - e^{-2 \left( \frac{N_1}{1+F} x_1 l_1 + \frac{N_2}{1+F} x_2 l_2 \right)} \quad (22)$$

330 Substituting eqn. 22 into 21 results in an algebraic equation which captures the establishment  
 331 probability for an arbitrary number of initial mutants across two patches. We are interested  
 332 in two special cases, when mutant either starts in favored ( $x_1 = 1/(2N_1)$  and  $x_2 = 0$ ) or  
 333 disfavored patch ( $x_1 = 0$  and  $x_2 = 1/(2N_2)$ ). With this parameterization, we obtain the  
 334 system of two non-linear equations:

$$0 = 2\tilde{M}_{21} N_2 l_2 + 2N_1 l_1 \tilde{s}_1 - N_1 l_1 (2\tilde{M}_{12} + l_1) \quad (23)$$

$$0 = 2\tilde{M}_{12} N_1 l_1 + 2N_2 l_2 \tilde{s}_2 - N_2 l_2 (2\tilde{M}_{12} + l_2) \quad (24)$$

335 By solving these for  $l_1$  and  $l_2$ , taking the real root, and substituting into eqn. 22, one retrieves  
 336 a closed-form solution for the establishment probability (see *Mathematica* notebook). With  
 337 all of the simplifying assumptions above, the diffusion approximation works in the same

338 regime when  $N_e s \gg 1$ , as is the case with branching approximation. As will be shown later  
 339 in the text, the comparison of two solutions reveals that diffusion approximation performs  
 340 identically when the allele starts in the favored patch and slightly better when it originates  
 341 in the disfavored one. Because branching process approximation is easier to work with (e.g.,  
 342 series expansion readily yields intuitive special cases), we use this solution throughout our  
 343 analysis and note that the diffusion approach yields marginally better precision.

## 344 3 Results

### 345 3.1 Probability of establishment of a mutant

#### 346 3.1.1 General equations

347 The probability of establishment of a new allele is non-zero if the branching process is super-  
 348 critical, which means that the leading eigenvalue of the mean reproductive matrix must be  
 349 larger than unity (Haccou *et al.*, 2005). For the invasion of allele  $A$  under viability selection,  
 350 it corresponds to:

$$\frac{M_{12}}{\frac{((1-F)V_{1,Aa} + FV_{1,AA}) - V_{1,aa}}{(1-F)V_{1,Aa} + FV_{1,AA}}} + \frac{M_{21}}{\frac{((1-F)V_{2,Aa} + FV_{2,AA}) - V_{2,aa}}{(1-F)V_{2,Aa} + FV_{2,AA}}} < 1 \quad (25)$$

351 The denominator is the relative reproductive advantage of invading mutant against the  
 352 resident allele. The invading population is a combination of mutant heterozygotes (when the  
 353 two gene copies of an individual are not identical by descent,  $1 - F$  fraction of time) and  
 354 homozygotes (when the two gene copies are identical by descent,  $F$  fraction of time). So the  
 355 whole term  $((1 - F)V_{i,Aa} + FV_{i,AA})$  gives the fitness of the invading mutant weighted by the  
 356 contribution of invading genotypes. In an outcrossing population ( $F = 0$ ), condition (25)  
 357 reduces to  $M_{12}/(1 - V_{1,aa}/V_{1,Aa}) + M_{21}/(1 - V_{2,aa}/V_{2,Aa}) < 1$ , which is identical to Bulmer's  
 358 inequality. The same inequality with re-parameterized selection holds for fecundity and male  
 359 sexual selection (see Appendix C). The full equations for the establishment probabilities are  
 360 too long and not very informative and are thus reported in the appendix (see equations  
 361 (A16) and (A17)). Without loss of generality, consider that allele  $A$  is advantageous in  
 362 patch 1 but deleterious in patch 2 ( $s_1 > 0 > s_2$ ); in the case of symmetrical migration  
 363 ( $M_{12} = M_{21} = M/2$ ) and migration prior to selection the general probability of establishment



364 of allele  $A$  in either patch 1 or 2 is given by:

$$P_1 = \max \left[ \frac{1}{1+F} \frac{2\sigma(\sigma + \Delta)(\tilde{h}_1 s_1 + \tilde{h}_2 s_2 + \sigma - M)}{2\Delta(\sigma + \Delta) + M^2 + M(\sigma - \Delta)}, 0 \right] \quad (26)$$

$$P_2 = \max \left[ \frac{1}{1+F} \frac{2M\sigma(\tilde{h}_1 s_1 + \tilde{h}_2 s_2 + \sigma - M)}{2\Delta(\sigma + \Delta) + M^2 + M(\sigma - \Delta)}, 0 \right] \quad (27)$$

365 where  $\Delta = \tilde{h}_1 s_1 - \tilde{h}_2 s_2$  and  $\sigma = \sqrt{\Delta^2 + M^2}$  with  $\tilde{h}$  as defined above for viability, female  
 366 and male fecundity selection. These resemble – but are not equal – to equations (4) and (5)  
 367 in (Tomasini and Peischl, 2018), who obtained an inexact result because of a typo in their  
 368 application of Haccou et al.’s theorem. Initially the typo came from Aeschbacher and Bürger  
 369 (2014) and was independently detected by Pontz and Bürger (2021) (see their Appendix A).  
 370 Interestingly, the correct solutions are more complicated and less elegant than the ones of  
 371 Tomasini and Peischl (2018).

### 372 3.1.2 Limiting cases and comparison with previous results

373 Despite the complexity of the general equations, useful insights can be obtained from simple  
 374 limit conditions. In the limit of no migration ( $M_{12} = M_{21} = 0$ ), we get back to the single-  
 375 patch scenario where selection favors allele  $A$  in patch 1:  $P_1 = 2\frac{\tilde{h}_1 s_1}{1+F}$ . This is equivalent  
 376 to Charlesworth (1992) and Caballero and Hill (1992) for viability selection. For male and  
 377 female fecundity selection, the result is similar to the one intuited in Damgaard (2000), except  
 378 that we also accounted for the reduction in  $P$  due to increased variance of offspring number.  
 379 Conversely, in the limit of full migration we obtain  $P_1 = P_2 = \frac{\tilde{h}_1 s_1 + \tilde{h}_2 s_2}{1+F}$ , which is equivalent  
 380 to Nagylaki (1980), with the appropriate rescaling for dominance and partial selfing. The  
 381 probability of establishment is simply the average over the two patches. Interestingly, the fact  
 382 that the effects of selfing on viability selection cancel out for codominant alleles ( $h_1 = h_2 =$   
 383  $1/2$ ) is no longer true with intermediate migration, even for seed migration that is not affected  
 384 by selfing. In the limit of weak symmetrical migration, we obtain  $P_1 = s_1 - \frac{M}{1+F}$  and  $P_2 =$   
 385  $\frac{M s_1}{(1+F)(s_1 - s_2)}$  and in the limit of strong migration (i.e.  $M \gg s_i$ ):  $P_1 = \frac{s_1 + s_2}{2} + (1+F)\frac{s_1(s_1 - s_2)}{4M}$   
 386 and  $P_2 = \frac{s_1 + s_2}{2} - (1+F)\frac{s_2(s_1 - s_2)}{4M}$ . Thus, unlike in a single population a locally advantageous  
 387 allele that is codominant in both patches is more likely to escape extinction in selfers than  
 388 in outcrossers. The relation between  $F$  and  $P_2$  is more complicated, and the allele can either  
 389 be more likely, for high migration rates, or less likely, for low migration rates, to escape  
 390 extinction. The effects of selfing under codominance only vanish either for no migration or  
 391 full migration.

392 **3.1.3 Comparison with simulations**

393 In general, our approximation for viability shows good correspondence with simulations  
 394 across different selfing rates (Figure 1). This also holds true across different dominance co-  
 395 efficients, and when selection or migration are asymmetrical (Figure 2). Correspondence is  
 396 better in cases where the invading allele initially appears in the favored rather than in the  
 397 disfavored patch. Diffusion approximation gives identical results when allele starts in the  
 398 favored, and slightly better results when in the case of the disfavored patch. The improve-  
 399 ment only occurs when the migration rate is low, and this pattern persists across dominances  
 400 and modes of selection. A possible reason for the slightly worst accuracy of the branching  
 401 process approach in the disadvantageous patch with low migration is that the rescaling of  
 402 the model assumes that selection and migration parameters are of the order of  $\tilde{s}_1$ . For low  
 403 migration, we can have  $m_{21} \ll \tilde{s}_1$ , so that migration could not be sufficient to introduce  
 404 the allele in the favored patch before it goes extinct in the disfavored one.

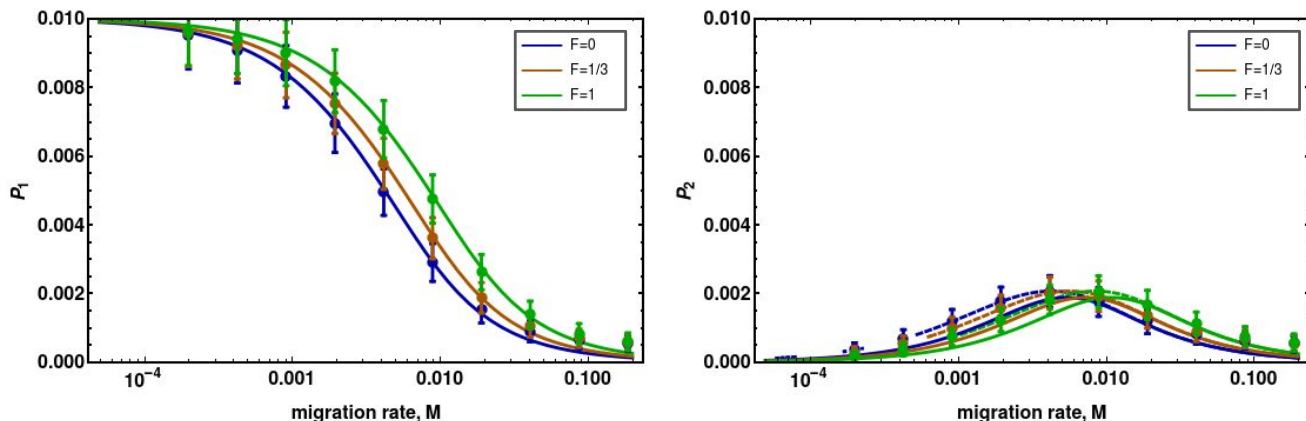


Figure 1: Comparison of analytical solution to simulated data across different selfing rates (corresponding to different equilibrium  $F$ ). Dashed lines are diffusion approximation. Left panel: the probability of establishment when the allele emerges in the favored patch; Right panel: the probability of establishment when the allele emerges in the disfavored patch. Parameters:  $h_1^\circ = h_2^\circ = 1/2$ ,  $s_1^\circ = -s_2^\circ = 0.01$ ,  $M_{12} = M_{21} = M$ .

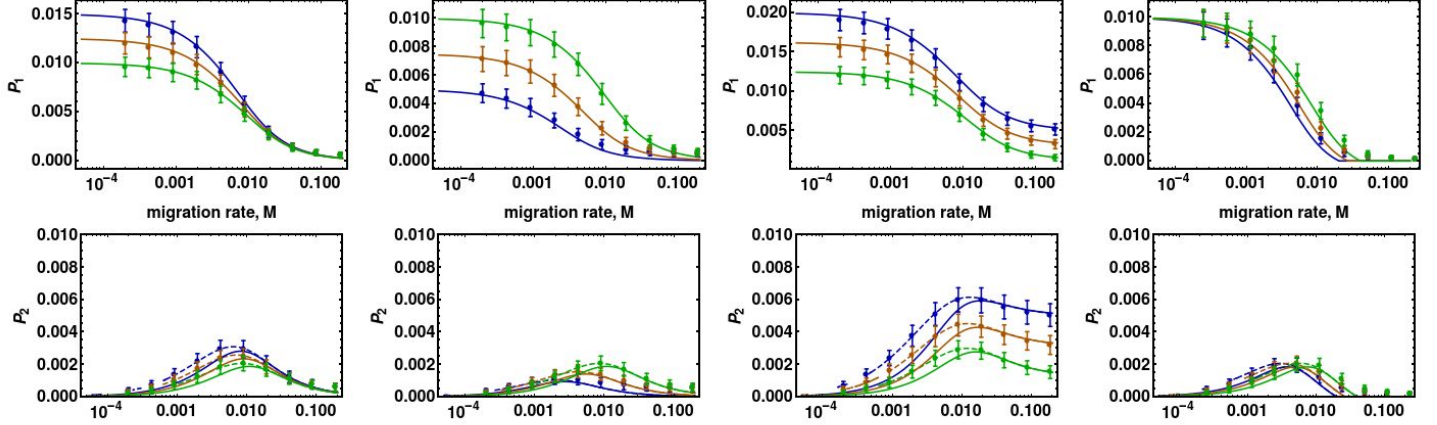


Figure 2: Simulations of different selection regimes. Going from left to right, the examined scenarios are as follows. Dashed lines denote diffusion approximation. First column: dominant case ( $h_1^\circ = h_2^\circ = 3/4$ ,  $s_1^\circ = -s_2^\circ = 0.01$ ,  $M_{12} = M_{21} = M$ ); Second column: recessive case ( $h_1^\circ = h_2^\circ = 1/4$ ,  $s_1^\circ = -s_2^\circ = 0.01$ ,  $M_{12} = M_{21} = M$ ); Third column: Asymmetric selection case ( $s_1^\circ = 0.01$ ,  $s_2^\circ = -0.0125$ ,  $h_1^\circ = h_2^\circ = 1/2$ ,  $M_{12} = M_{21} = M$ ); Fourth column: asymmetric migration ( $s_1^\circ = -s_2^\circ = 0.01$ ,  $h_1^\circ = h_2^\circ = 1/2$ ,  $M_{12} = M$ , and  $M_{21} = 1.25M$ ).

405  $P_1$  is a monotonically decreasing function in  $M$ , because strong migration causes the spread-  
 406 ing of the locally advantageous allele to become swamped by the resident from the disfa-  
 407 vored patch.  $P_2$  is non-monotonic function in  $M$ . Increasing migration initially increases the  
 408 probability that an invading allele escapes from the disfavored patch before it goes extinct.  
 409 However, at large migration rates, the spreading allele escapes the disfavored patch but is  
 410 swamped due to a large influx of deleterious residents from the disfavored patch. For male  
 411 fecundity selection, the analytical solutions for  $P_1$  and  $P_2$  are slightly worse against the sim-  
 412 ulations (Figure A11) than those provided for viability selection. Note, however that as the  
 413 selfing rate tends towards 1, the effective selection coefficient on male function vanishes so  
 414 that the strong selection condition for branching process approximation is not met anymore.  
 415 Slight discrepancies are also observed when selection or migration are asymmetrical (Figures  
 416 A12, and A13).

417

418 We also varied  $F$  continuously to obtain a fine-grained view of the analytic performance.  
 419 Three insights emerge. First, there is a good agreement between simulated data and analy-  
 420 tical solution, although deviations increase as the migration rate increases. Second, the  
 421 branching process poorly captures the dynamics of the invasion when mutants are partially  
 422 recessive. For example, in the limit of complete recessivity ( $h = 0$ ), the mutant does not  
 423 have any reproductive advantage (i.e., equations 10a–10c reduce to unity). We clearly see  
 424 this in Figure 3, where the analytical solution underestimates  $P$  when the allele is partially  
 425 recessive (first column). However, this discrepancy disappears as  $F$  increases, as homozy-

426 goes increasingly contribute to the invading process. Third, diffusion approximation yields  
 427 a better fit when mutant starts in the favored patch, but this improvement generally works  
 428 only when  $M$  is low.

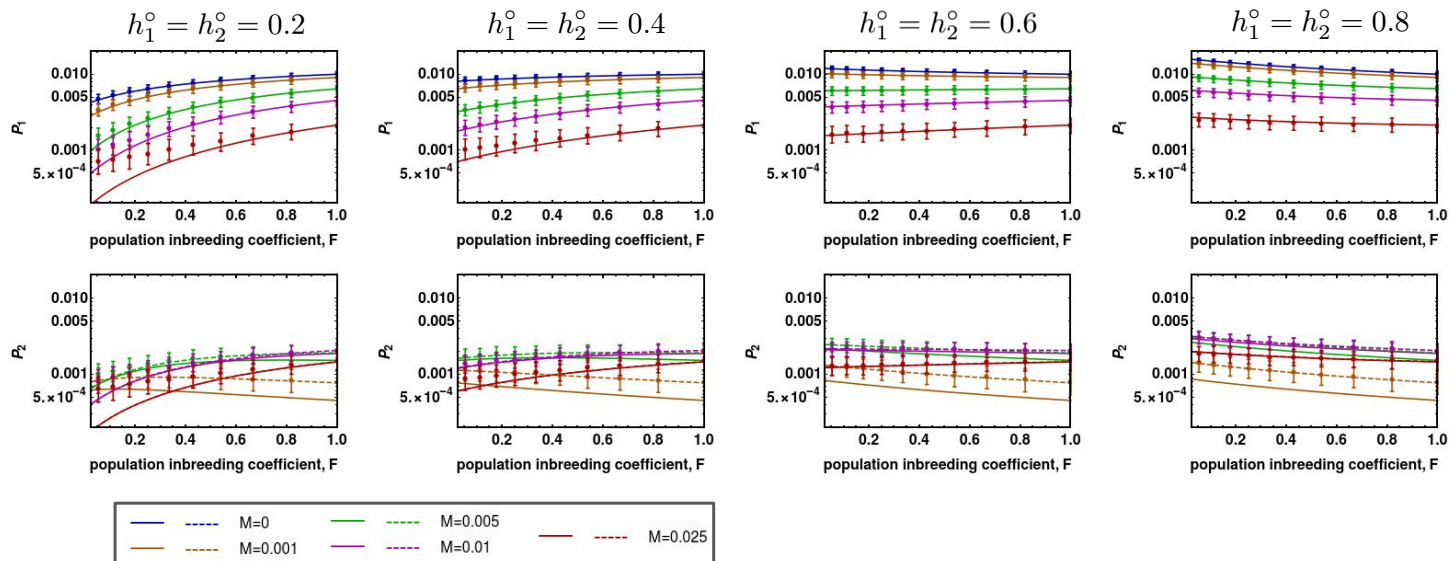


Figure 3: Simulations of different selection regimes. Going from left to right, the examined scenarios are as follows. Dashed lines denote diffusion approximation. Each column denotes the establishment probability for a set of dominance coefficients noted above the panel. Lines correspond to different migration rates (see legend box below). Other parameters:  $s_1^\circ = -s_2^\circ = 0.01$ ,  $M_{12} = M_{21} = M$ . The axes are on logarithmic scale and error bars are excluded.

### 429 3.2 Decomposing the effects of selfing on the establishment prob- 430 ability

431 Selfing impedes the spread of an invading beneficial allele by reducing the effective population  
 432 size and by increasing the effective dominance in the disfavored patch. On the other hand,  
 433 selfing also facilitates the invasion process by increasing the effective dominance in the favored  
 434 patch and reducing the gene flow when dispersal occurs through pollen. In this section, we  
 435 ask: are the impeding effects generally more important than facilitating effects? When  
 436 does the introduction of selfing increase  $P$ ? To this end, we introduce an indicator  $\beta_i(y)$   
 437 which denotes the effect that selfing has on  $P_i$  solely via parameter  $y$ . For example,  $P_i^{\{h_1\}}$   
 438 corresponds to the establishment probability in patch  $i$ , if selfing only exerted its effect by

439 inflating favored dominance. More generally:

$$\beta_i(y) = \left( \frac{dP_i^{\{y\}}}{dF} \right) \Big|_{F=0}, \quad y \in \{\tilde{h}_i, \tilde{N}, \tilde{M}\}, \quad i \in \{1, 2\} \quad (28)$$

440 Consider decomposing the effects of selfing on the establishment of an allele affecting viability.  
 441 For example, the indicator  $\beta_1(\tilde{h}_1^\circ)$  tells us whether a shift to selfing increases  $P_1$ , if selfing  
 442 solely acted through an increase in the effective dominance in the favored patch. This allows  
 443 us to examine the effect of each factor separately and assess when one outweighs the other.  
 444 The indicator is obtained by taking equation A16 and sequentially setting: (1)  $F = 0$  which  
 445 eliminates selfing's effect on  $\tilde{N}$ , then (2) parameterizing  $\tilde{s}_1^\circ := s_1^\circ \tilde{h}_1^\circ$  which introduces the  
 446 effect of selfing on favored dominance, and finally, (3) setting  $\tilde{s}_2^\circ := s_2^\circ h_2^\circ$  which excludes the  
 447 effect of selfing on the effective dominance in the disfavored patch. The outlined method is  
 448 also applicable to other selection and migration scenarios, and the general procedure is given  
 449 in Appendix F.

### 450 3.2.1 Emergence in the favored patch

451 We focus on the establishment conditioning on allele emerging in the favored patch. Consider  
 452 viability selection first (see Figure 4). Given our interest in knowing whether selfing increases  
 453 or decreases the establishment probability, we only focus on the region of parameter space  
 454 where allele can become established in fully outcrossing population (right of the solid line in  
 455 Figure 4). When the invading allele is partially recessive in favored and partially dominant  
 456 in the disfavored patch (white region), the criterion for escaping extinction is not satisfied,  
 457 and a shift to selfing can only have a promoting effect on the establishment in this region of  
 458 parameter space.

459

460 Under seed dispersal, selfing only affects selection via effective dominance and the effective  
 461 population size. One can recognize three regions of selfing effects (Figure 4). Firstly, when  
 462 the invading mutant is partially recessive in the favored and partially dominant in the dis-  
 463 favored patch, selfing increases  $P_1$  via the inflation of  $\tilde{h}_1^\circ$  (blue region). More formally, this  
 464 will be the outcome when  $\beta_k(\tilde{h}_1^\circ) > -(\beta_k(\tilde{h}_2^\circ) + \beta_k(\tilde{N}))$  (dark blue region). This is because  
 465 the mutant is already dominant in the disfavored patch so selfing does not significantly  
 466 increase the rate of purging, but it does increase the rate of spread owing to the mutant  
 467 being partially recessive in the favored patch. Thus, the net effect is increased  $P_1$ . Secondly,  
 468 when the mutant is partially dominant in the favored and partially recessive in the disfa-  
 469 vored patch, selfing decreases  $P_1$  via inflation of  $\tilde{h}_2^\circ$  (darker red region). This happens when

470  $\beta_k(\tilde{h}_1^\circ) < -\beta_k(\tilde{h}_1^\circ)$  (dark red region). The effect is opposite to the first regime: selfing has  
471 a larger effect on the increase in purging from the disfavored patch (since the dominance in  
472 the disfavored patch is low) than on the increase in spreading in the favored patch (because  
473 dominance in the favored patch is already high). The net effect is a decrease in  $P_1$ . The  
474 region of parameter space that lies between these two zones corresponds to cases when the  
475 facilitating effect of selfing (via dominance in the favored patch) roughly cancels out the  
476 impeding effect (via dominance in the disfavored patch) so that the net effect on  $P_1$  is deter-  
477 mined by the reduction in effective population size and the net effect of selfing is to decrease  
478  $P_1$ .

479

480 Under pollen dispersal, we also found the same three regions but with different boundaries  
481 (Figure 4). In addition, selfing also affects pollen dispersal by reducing the effective number of  
482 migrants, so in the intermediate zone the effect of selfing on  $P_1$  depends both on the reduction  
483 in effective population size and effective migration rate, which creates a fourth region – the  
484 region for which all parameter combinations where  $\beta_k(\tilde{h}_1^\circ) + \beta_k(\tilde{M}) > -(\beta_k(\tilde{h}_2^\circ) + \beta_k(\tilde{N}))$   
485 (light blue region). It corresponds to conditions where the reduction in effective migration  
486 rate overwhelms the impeding effects of selfing. Overall, selfing increases the establishment  
487 of new alleles under broader conditions under pollen than under seed dispersal. Note that  
488 selection coefficients are an order of magnitude lower in the right panel of Figure 4, and the  
489 region where selfing has a net-positive effect is comparable (if not larger) to the blue region  
490 in the left panel.

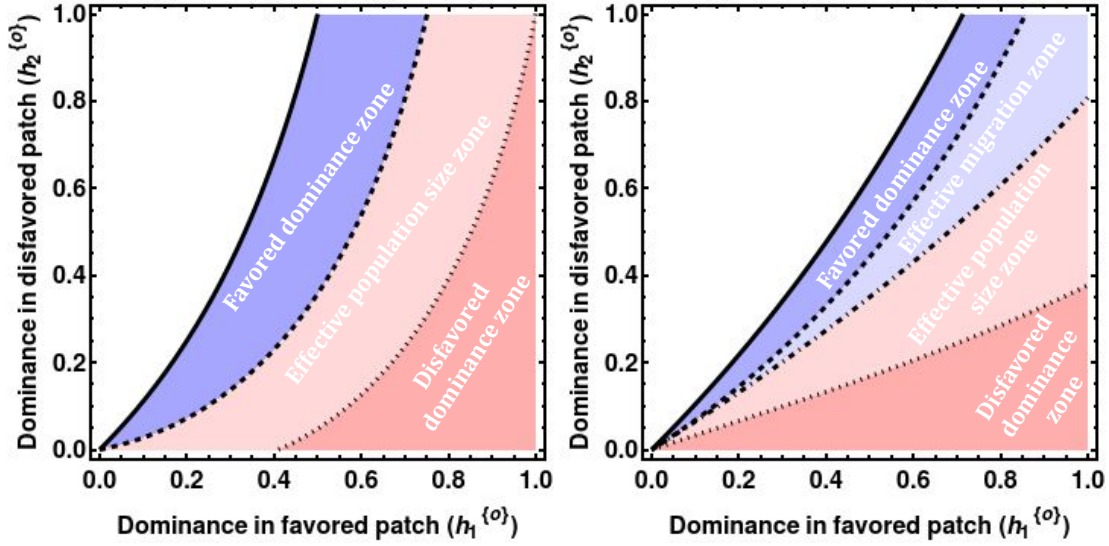


Figure 4: Four zones of selfing's effects on the establishment probability when mutant originates in the favored patch. White region denotes parameter space where the mutant cannot establish under outcrossing. Selection act only on viability, and migration occurs only via pollen (left), or only via seed (right). Blue and red regions correspond to parameter space where selfing facilitates and impedes  $P_1$ , respectively. Lines in the left panel are as follows. Solid black line:  $P_1 = 0$ ; Dashed black line:  $\beta_k(\tilde{h}_1^{\circ}) = -(\beta_k(\tilde{h}_2^{\circ}) + \beta_k(\tilde{N}))$ ; Dotted black line:  $\beta_k(\tilde{h}_1^{\circ}) = -\beta_k(\tilde{h}_2^{\circ})$ . Lines in the right panel are as follows. Solid black line:  $P_1 = 0$ ; Dashed black line:  $\beta_k(\tilde{h}_1^{\circ}) = -\beta_k(\tilde{h}_2^{\circ})$ ; Dashdotted line:  $\beta_k(\tilde{h}_1^{\circ}) + \beta(\tilde{M}) = -(\beta_k(\tilde{h}_2^{\circ}) + \beta_k(\tilde{N}))$ ; Dotted line:  $\beta_k(\tilde{h}_1^{\circ}) + \beta(\tilde{M}) = -\beta_k(\tilde{h}_2^{\circ})$ . Parameters for the left panel:  $s_1^{\circ} = -s_2^{\circ} = 0.01$ ,  $h_1^{\circ} = h_2^{\circ} = 1/2$ ,  $M_{12} = M_{21} = 0.01$ . Parameters for the right panel:  $s_1^{\circ} = -s_2^{\circ} = 0.001$ ,  $h_1^{\circ} = h_2^{\circ} = 1/2$ ,  $M_{12} = M_{21} = 0.005$ .

491 Although qualitatively similar results holds for two other modes of selection (Figure 5), one  
492 notices that the region where the selfing exerts its effect solely via effective dominance is  
493 significantly expanded. This is because an increase in the selfing rate diminishes the male  
494 component's relevance and increases the importance of the female component. The relative  
495 sizes of these regions vary based on parameter values, so more general quantitative statements  
496 are hard to make.

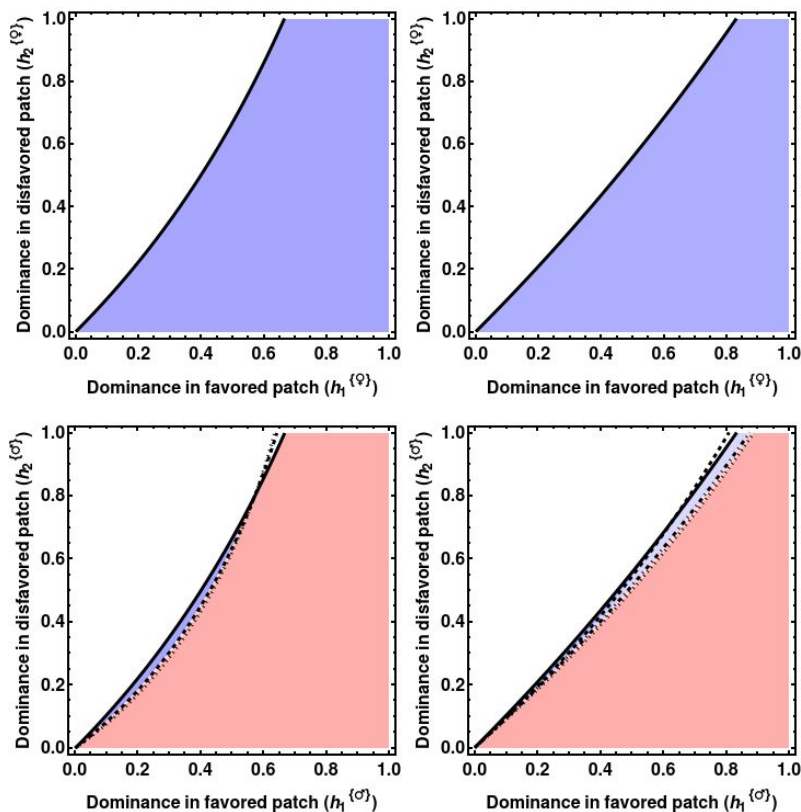


Figure 5: Zones of selfing's effect on  $P_1$  under female fecundity (upper row), and on male fecundity (lower row). Color code as in Figure 4. Migration via seed is depicted in the left column, and via pollen in the right column. Black lines as in the previous figure. Depending on selection and migration scenario, parameters are either:  $h_1^\circ = h_2^\circ = 1/2$  or  $h_1^\sigma = h_2^\sigma = 1/2$ ,  $s_1^\circ = -s_2^\circ = 0.01$  or  $s_1^\sigma = -s_2^\sigma = 0.01$ , and  $M_{12} = M_{21} = 0.01$  or  $m_{12} = m_{21} = 0.01$ .

### 497 3.2.2 Emergence in the disfavored patch

498 Similar conclusions are reached if one considers emergence of mutant in the disfavored patch  
499 (Figure A16). However, under pollen dispersal, selfing reduces the effective migration rate,  
500 which can have either positive or negative effect on  $P_2$  depending on the parameters of the  
501 model. This phenomenon arises due to  $P_2$  being a non-monotonic function of  $m$  (see Figure  
502 2). More specifically, if migration is much stronger than selection, then a mutant can escape



503 to favored patch where it can spread, but the influx of maladapted allele is also high, thus  
 504 having a net impeding effect on invading allele. This means that introduction of selfing will  
 505 reduce the effective migration rate, and thus increase  $P_2$ . Put differently, increasing  $F$  has  
 506 the same kind of effect on  $P_2$  as decreasing  $m$ . However, if the migration rate is already  
 507 low enough, the introduction of selfing further depresses  $m$  and, thus, prevents invader from  
 508 successfully escaping into the favored patch.

### 509 **3.3 Consequence of selfing on the establishment of protected poly-** 510 **morphism**

511 Once the allele  $A$  escapes extinction, it can either fix across both patches or segregate  
 512 at intermediate frequencies for a finite but large number of generations due to divergent  
 513 selection. Because both patches are of finite size, one of the two alleles will ultimately fix  
 514 but this quasi-stationary behavior corresponds to protected polymorphism in deterministic  
 515 models. We wish to delineate the effects of selfing on these two outcomes. Given that an  
 516 invading allele can appear in any of the two patches where the favored patch accounts for  $z$   
 517 fraction of the total population across both patches, the global probability of alleles  $A$  and  
 518  $a$  becoming established is:

$$P^{(A)} = zP_1^{(A)} + (1 - z)P_2^{(A)} \quad (29)$$

$$P^{(a)} = zP_1^{(a)} + (1 - z)P_2^{(a)} \quad (30)$$

519 where  $P_i^{(A)}$  and  $P_i^{(a)}$  are probabilities that  $A$  and  $a$  become established, conditioning on  
 520 initially appearing in patch  $i$ . For the sake of simplicity, we will further assume that patches  
 521 are of equal size ( $z = 1/2$ ), and migration is symmetrical. Probability that allele  $A$  is es-  
 522 tablished is computed by parameterizing equations (26) and (27) where selection coefficients  
 523 were determined from Jacobian associated with boundary  $\vec{x} = [0, 0]^T$ , while the establish-  
 524 ment probability of  $a$  was calculated with elements of Jacobian associated with boundary  
 525  $\vec{x} = [1, 1]^T$ . For example, if selective advantage of invading heterozygote  $Aa$  relative to ho-  
 526 mozygote  $aa$  in patch  $i$  is  $s_i h_i$ , then the advantage of invading  $Aa$  relative to homozygote  
 527  $AA$  is  $-s_i(1 - h_i)/(1 + s_i)$ .

#### 528 **3.3.1 Conditions for protected polymorphism and critical migration rates**

529 Polymorphism is protected if both  $P^{(A)} > 0$  and  $P^{(a)} > 0$ , that is for the conditions under  
 530 which the branching processes for the invasion of  $A$  and for the invasion of  $a$  are supercritical.  
 531 This boils down to the extension of Bulmer's inequality given by equation (25) directly

532 obtained from the Jacobian matrix, which must also be satisfied with genotypes  $AA$  and  
 533  $aa$  shifted (corresponding to the supercriticality for the invasion of allele  $a$ ). From this, we  
 534 can define critical migration rates above which polymorphism is lost and the allele with the  
 535 highest marginal fitness invades as in Yeaman and Otto (2011). Without loss of generality  
 536 we can assume that  $s_1 > -s_2 > 0$ . For seed dispersal and viability selection  $M_c$  is then given  
 537 by:

$$M_c = \begin{cases} 1, & \text{if } \frac{(1+(1-F)h_1s_1)(1+(1-F)h_2s_2)}{(1+s_1)(1+s_2)} > 1 \\ -\frac{(1-(1-F)h_1)(1-(1-F)h_2)s_1s_2}{(1-(1-F)h_1)s_1+(1-(1-F)h_2)s_2+(1-F)(h_1+h_2-2(1-F)h_1h_2)s_1s_2}, & \text{otherwise} \end{cases} \quad (31)$$

538 For  $F = 0$ , (31) reduces to equation (5) in Yeaman and Otto (2011) and for  $F = 1$  to  
 539  $M_c = -\frac{s_1s_2}{s_1+s_2}$ . The first condition corresponds to a form of dominance reversal where the  
 540 allele is dominant when advantageous and recessive when deleterious. Under such a condi-  
 541 tion, polymorphism can be maintained even under full migration.

542

543 Depending on dominance, selfing thus has two opposite effects. On the one hand, by unmask-  
 544 ing recessive alleles, selfing makes selection stronger relative to migration and so enlarges  
 545 the conditions for the maintenance of protected polymorphism. On the other hand, selfing  
 546 makes dominance reversal conditions less likely as selection mainly operates on homozygotes,  
 547 and so restricts the conditions for protected polymorphism. Figure 6 shows the dominance  
 548 conditions under which selfing increases or decreases the critical migration rate. We de-  
 549 termined these conditions in two ways: either when  $M_c$  is higher under full selfing than  
 550 under full outcrossing (i.e.,  $M_c|_{F=0} < M_c|_{F=1}$ ) or whether introduction of selfing in an oth-  
 551 erwise outcrossing population increases  $M_c$ , that is  $\left. \frac{\partial M_c}{\partial F} \right|_{F=0} > 0$ . For symmetrical selection  
 552 ( $s_2 = -s_1$ ), selfing increases migration rate if  $h_2 > h_1$ . Under the opposite condition, the  
 553 critical migration is 1 and local adaptation is always maintained under outcrossing. As se-  
 554 lection becomes more asymmetrical, selfing favors local adaptation for a broader range of  
 555 dominance coefficients. Note that the effect of selfing on the migration rate is not monotonic  
 556 for a narrow range of conditions (light blue region in Figure 6), where little selfing disfavors  
 557 local adaption ( $\left. \frac{\partial M_c}{\partial F} \right|_{F=0} < 0$ ) but high selfing increases it ( $M_c|_{F=0} < M_c|_{F=1}$ ). Finally, it is  
 558 worth noting that when selfing favors local adaptation, the effect is rather weak (blue curves  
 559 on Figure 6). On the contrary, under dominance reversal conditions, above a given threshold,  
 560 selfing dramatically reduces the migration rate from 1 to  $-\frac{s_1s_2}{s_1+s_2}$  under full selfing. Similar  
 561 results are obtained for fecundity selection but with selfing favoring local adaption under  
 562 broader conditions for female fecundity and for narrower conditions for male fecundity (see  
 563 Appendix).

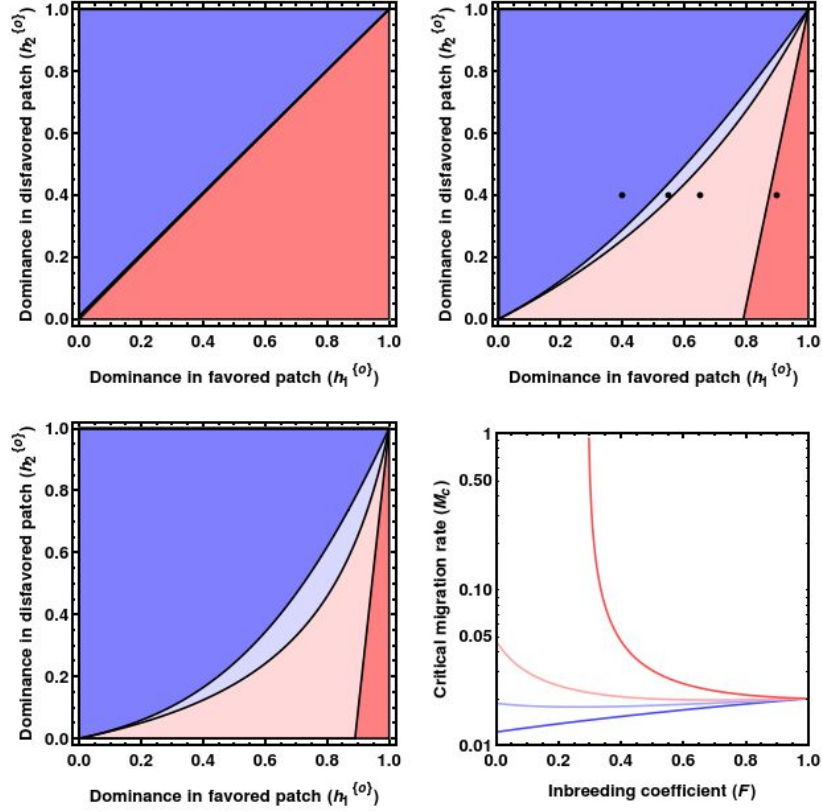


Figure 6: Effect of selfing on the critical migration rate. The three first panels show dominance conditions under which selfing increases (blue regions) or decreases (red regions) the critical migration rate,  $M_c$ . Selection is symmetrical in the first panel ( $s_1 = -s_2 = 0.01$ ), and asymmetrical in the second ( $s_1 = 0.02$  and  $s_2 = -0.01$ ) and third panels ( $s_1 = 0.05$  and  $s_2 = -0.01$ ). In the light blue region,  $M_c$  is not monotonic with  $F$ : introduction of selfing in an outcrossing population decreases  $M_c$  ( $\frac{\partial M_c}{\partial F}|_{F=0} < 0$ ) but above a given selfing rate,  $M_c$  becomes higher than in an outcrossing population and we always have  $M_c|_{F=0} < M_c|_{F=1}$ . In the dark blue region, selfing always increases  $M_c$  ( $M_c|_{F=0} < M_c|_{F=1}$ ). In the dark red region,  $M_c = 1$  in outcrossing populations and  $M_c < 1$  in the light red region. The last panel illustrates how  $M_c$  varies with  $F$  in the four regions, corresponding to the black dots of the second panel. For all curves  $s_1 = 0.02$  and  $s_2 = -0.01$ . Dark blue:  $h_1 = 0.4$  and  $h_2 = 0.4$ ; light blue:  $h_1 = 0.55$  and  $h_2 = 0.4$ ; light red:  $h_1 = 0.65$  and  $h_2 = 0.4$ ; Dark red:  $h_1 = 0.9$  and  $h_2 = 0.4$ . Note that  $M_c$  is in log-scale.

564 The picture is different under pollen dispersal. Under full selfing, effective migration is zero  
565 so selection can act independently in each patch and polymorphism is always maintained. As  
566 a consequence, there is always a threshold selfing rate above which the critical migration rate  
567 is higher than in an outcrossing population. Here, we thus concentrate on the effect of the  
568 introduction of selfing in an outcrossing population, so the conditions for which  $\left. \frac{\partial M_c}{\partial F} \right|_{F=0} > 0$ .  
569 Under symmetrical selection, the conditions are the same as for seed migration. However,  
for asymmetrical selection, selfing increases  $M_c$  for a much broader range of parameters.

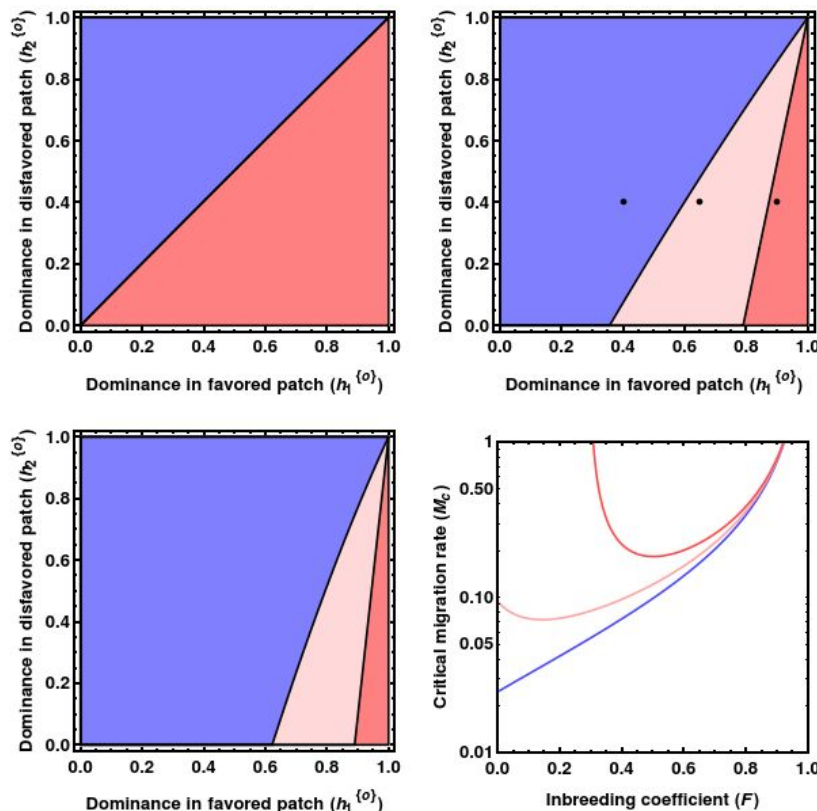


Figure 7: Same legend and same parameters as in Figure 6 but with pollen dispersal instead of seed dispersal.

570

### 571 3.3.2 Establishment and maintenance of protected polymorphism

572 The above analysis showed how selfing affects the critical migration rate, hence the conditions  
573 for local adaptation. However, when local adaptation is possible, selfing also affects the  
574 probability of establishment and maintenance of polymorphism. This can be analyzed by

575 quantifying how selfing affects  $P^{(A)}$  and  $P^{(a)}$ , which is given by:

$$b_k = \left( \frac{dP^{(k)}}{dF} \right) \Big|_{F=0} > 0 \quad \text{where } k = \{A, a\} \quad (32)$$

576 Four outcomes are possible upon the introduction of selfing. First,  $P^{(A)}$  increases, and  $P^{(a)}$   
 577 decreases, meaning that allele  $A$  is more likely to fix across both patches in selfers than  
 578 in outcrossers. Second and conversely,  $P^{(A)}$  decreases, but  $P^{(a)}$  increases, implying that  
 579 selfers are less likely to fix  $A$  compared to an outcrossing population. Third, both  $P^{(A)}$   
 580 and  $P^{(a)}$  increase upon shift to selfing, meaning that selfing increases the probability that  
 581 protected polymorphism is established. Fourth and final, both  $P_A$  and  $P_a$  decrease, making  
 582 the protected polymorphism less likely to become established in the selfing population. Some  
 583 examples are given for different migration rates with symmetrical or asymmetrical selection  
 584 (Figure 8).

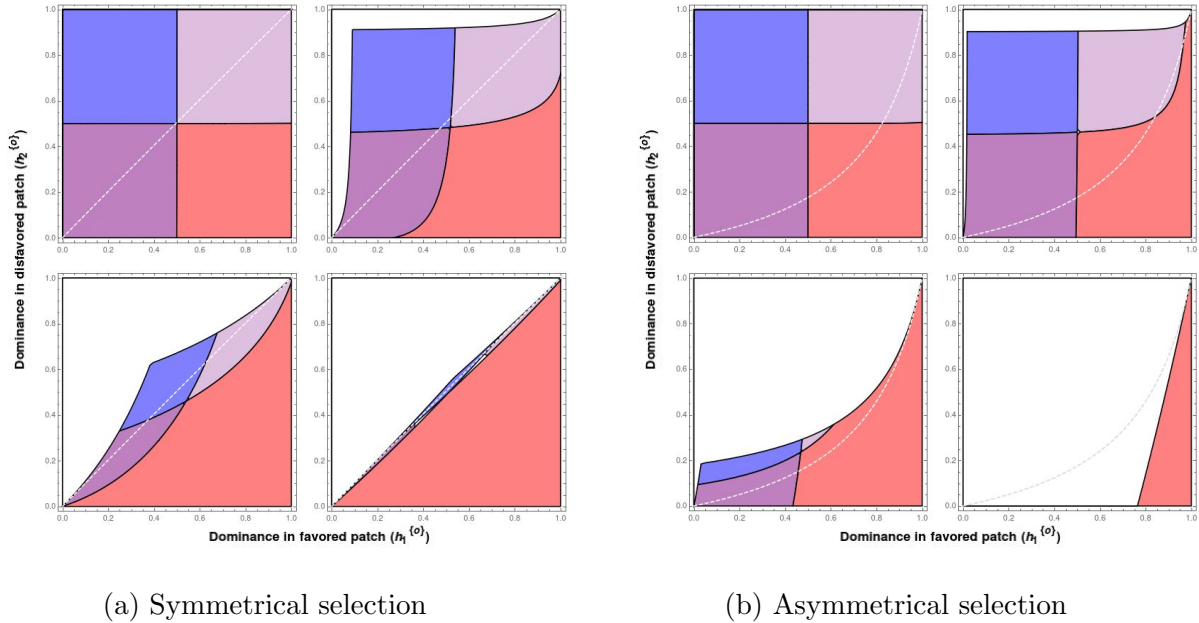


Figure 8: Effect of selfing on the establishment of protected polymorphism under symmetrical (a) and asymmetrical selection (b). Five regions can be distinguished: blue:  $b_A > 0$  and  $b_a > 0$ , red:  $b_A < 0$  and  $b_a < 0$ , dark purple  $b_A > 0$  and  $b_a < 0$ , light purple  $b_A < 0$  and  $b_a > 0$ , and white: polymorphism cannot be maintained under outcrossing. The dotted lines correspond to the limiting conditions for which selfing increases the critical migration rate (above the line) as in Figure 6. (a)  $s_1 = 0.01$  and  $s_2 = -0.01$ ;  $m = 0.0001, 0.001, 0.01, 0.1$  from top left to bottom right. (b)  $s_1 = 0.05$  and  $s_2 = -0.01$ ;  $m =$  as in (a).

585 One can intuitively understand these outcomes by noting that dominance reverses when the  
 586 alternative allele invades. So, if invader  $A$  is dominant, then invader  $a$  will be recessive.

587 The dominant allele  $A$  is likely to escape extinction when invading but unlikely to result  
588 in protected polymorphism because allele  $a$  is recessive and thus likely to go extinct. This  
589 appears clearly in the limit of low migration, where probabilities of emergence tend toward  
590 single population predictions, with four quadrants corresponding to the four dominance  
591 conditions (Figure 8). Importantly, these four regions vary with the amount of migration,  
592 and they do not directly align with the conditions given by the critical migration rates. The  
593 reason is that the sign of  $P^{(A)}$  and  $P^{(a)}$  only depends on the effect of selfing through the  
594 effective dominance, whereas the absolute value also depends on the reduction in effective  
595 size. Here, we only consider the two-fold reduction in  $N_e$  caused by selfing, but the linked  
596 selection can further reduce the effective size of the local population under selfing (Roze,  
597 2016). This will not affect the critical migration rate but can strongly reduce probabilities of  
598 emergence hence the establishment of local adaptation. Following this rationale, Yeaman and  
599 Otto (2011) considered a critical migration rate above which the probability of emergence of  
600 both alleles is higher than  $1/(4N_e)$ . High local drift under selfing can thus strongly reduce  
601 the conditions for the establishment of local adaptation.

## 602 4 Discussion

603 We investigated the effect of selfing on the establishment of an allele in a population sub-  
604 divided into two patches, where only one locus determines the fitness of an organism. By  
605 representing the spread of an invading allele as a branching process, we obtained a closed-  
606 form analytical solution for the probability of establishment of the locally advantageous  
607 allele for an arbitrary population selfing rate. This extends the work of Tomasini and Peis-  
608 chl, 2018 to include diploidy, dominance, and partial selfing. We also corrected for a typo in  
609 the derivation of their equations but which surprisingly lead to less elegant results. Below,  
610 we discuss the implication for adaptation in partially selfing species.

611  
612 A well-known result is that selfing favors the establishment (and fixation) of recessive alleles  
613 but disfavors the establishment of dominant ones, with no effect under exact codominance  
614 ( $h = 1/2$ ) (Caballero and Hill, 1992, Charlesworth, 1992). The pattern is more complex in  
615 a subdivided population with heterogeneous selection. First, the codominant allele has a  
616 higher probability of becoming established in the selfing population than in the outcrossing  
617 population if it emerges in the favored patch and generally smaller probability if it emerges  
618 in the disfavored patch. Therefore, the probability of establishment of a codominant allele is  
619 not independent of the mating system, as was the case for a single partly selfing population  
620 (Charlesworth (1992)). Second, assuming that migration and selection are symmetrical, self-

621 ing will increase the establishment probability when the invading mutant is recessive in the  
622 favored patch (resulting in a maximal increase in  $\tilde{h}_1$ ) and dominant in the disfavored patch  
623 (yielding a minimal increase in  $\tilde{h}_2$ ). If the two dominances are roughly equal, the opposed  
624 effects of selfing on effective dominances cancel each other out, and the net effect is set by the  
625 interplay between the positive effect of reduced effective migration and the negative effect of  
626 the reduction in  $N_e$ .

627

628 Once the invading allele escapes extinction, it can either fix in both populations, thus one  
629 population is fixed for a maladaptive allele, or segregate at intermediate frequencies, thus  
630 contributing to protected polymorphism, and both populations can be considered as locally  
631 adapted. It is usually thought that selfers are better locally adapted because of the reduced  
632 gene flow between patches with different selection demands (Linhart and Grant (1996)).  
633 However, selfing affects in a complex way three key parameters determining local adaptation  
634 - selection, drift, and gene flow. By separately considering seed and pollen migration, we  
635 were able to distinguish the straightforward effect of reducing gene flow from the more subtle  
636 effects of altering genotypic frequencies and drift induced by selfing.

637

638 Under seed dispersal only, selfing does not affect gene flow but still alters the propensity  
639 for local adaptation. The conditions under which selfing favors local adaptation strongly  
640 depend on the dominance of alleles in the two patches. Under outcrossing, the most fa-  
641 vorable condition for local adaptation is under dominance reversal, that is, when the allele  
642 is dominant when locally beneficial and recessive when locally deleterious. If so, protected  
643 polymorphism can be maintained even under full migration ( $M_c = 1$ ), which corresponds to  
644 conditions for polymorphism in Levene (1953)'s model. Under this condition, selfing desta-  
645 bilizes polymorphism, which cannot be maintained under full dispersal: above a given selfing  
646 rate, the critical migration rate strongly decreases (Figure 6, and see also Glémin, 2021, for  
647 the Levene's model with selfing). Conversely, under most other conditions, especially when  
648 selection is asymmetrical, selfing tends to increase the critical migration rate, hence favors  
649 local adaptation. What dominance patterns across heterogeneous habitat are the most fre-  
650 quent in natural populations is unknown. However, dominance reversal is maybe not as  
651 unlikely as it may seem because it can naturally arise in simple fitness landscape models  
652 (Connallon and Chenoweth, 2019). If we now consider pollen dispersal, selfing enhances  
653 local adaptation by reducing gene flow. Under dominance reversal conditions, this leads to  
654 non-monotonic behaviors where local adaptation can be the most easily maintained, either  
655 under outcrossing or under high selfing (Figure 7). Characterizing the pollination and seed  
656 dispersal modes and their quantitative impacts on gene flow appears thus crucial to make

657 proper predictions on the effect of selfing. Finally, if we also consider the effect of linked se-  
658 lection that reduces local effective size beyond the two-fold level in selfing populations (Roze,  
659 2016), selfing reduces the possibility of local adaptation, which requires stronger selection  
660 as exemplified by Hodgins and Yeaman (2019) who simulated local adaptation in selfing  
661 populations with and without background selection. Overall, the complex and contradic-  
662 tory effects of selfing on local adaptation may explain why no general empirical pattern has  
663 emerged so far (Hereford, 2010, Leimu and Fischer, 2008).

664

665 Despite the large uncertainties about the global effect of selfing on local adaptation, the  
666 analysis of the model yields some predictions about the genetic architecture of local adap-  
667 tation in selfing versus outcrossing species. In outcrossing species, the global adaptation  
668 from de novo mutations is predicted to be biased towards dominant mutations, the so-called  
669 Hadane’s sieve (Haldane, 1927, Ronfort and Glémin, 2013), and local adaptation to benefi-  
670 cial dominant/deleterious recessive mutations (Yeaman and Otto, 2011). On the contrary,  
671 no sieve related to dominance is expected under high selfing. As for other forms of selection,  
672 we also predict that local adaptation should be more prominent on female traits because  
673 selfing reinforces selection on female fecundity components at the cost of male fecundity  
674 components. A similar conclusion was reached by Olito *et al.* (2018) from a model including  
675 male/female antagonism and heterogeneous habitat with full dispersal. Finally, beyond the  
676 current work, simulations showed that local adaptation genes tend to aggregate into clusters  
677 in outcrossers (Yeaman and Whitlock, 2011) whereas the genetic architecture tends to be  
678 more diffuse in selfers (Hodgins and Yeaman, 2019, Le Thierry d’Ennequin *et al.*, 1999).  
679 Thus, a natural extension of the model would be to consider local adaptation at two loci to  
680 dissect the additional effect of selfing on genetic linkage. However, using the same formalism  
681 would require following a multi-type branching process in a higher dimension (at least six),  
682 which remains highly challenging.

## 683 5 Data availability

Supplemental files are available at [https://github.com/BogiTrick/local\\_adaptation\\_single\\_locus](https://github.com/BogiTrick/local_adaptation_single_locus) and at figshare (<https://doi.org/10.25386/genetics.16664650>). These entries contain: *Mathematica* notebook required for reproduction of all reported theoretical results and figures; C++ source code of the custom simulator used for generation of simulated data for comparison with analytics; Shell scripts for the bulk run of the simulator; and R scripts used to process the raw simulated data for plotting in the above-mentioned notebook.



## 691 **6 Acknowledgements**

692 We are thankful to Guillaume Martin and two anonymous reviewers for helpful comments,  
693 especially for suggesting to formally derive the offspring distribution and to alternatively  
694 solve our main problem by diffusion approximation. We also thank Matthew Hartfield for  
695 the critical reading of the manuscript.

## 696 **7 Funding**

697 B. T. was supported by Erasmus Mundus Joint Masters Degree Scholarship awarded by  
698 Erasmus+ programme. S.G. was supported by the Agence National de la Recherche (ANR  
699 SEAD-ANR-13-ADAP-0011).

## 700 **8 Competing interests**

701 Authors declare no competing interests.

## 702 **References**

- 703 Aeschbacher, S. and R. Bürger, 2014 The effect of linkage on establishment and survival of  
704 locally beneficial mutations. *Genetics* **197**: 317–336.
- 705 Barton, N., 1987 The probability of establishment of an advantageous mutant in a subdivided  
706 population. *Genet. Res., Camb.* **50(1)**: 35–40.
- 707 Billiard, S., M. Lopez-Villavicencio, M. E. Hood, and T. Giraud, 2012 Sex, outcrossing and  
708 mating types: unsolved questions in fungi and beyond. *Journal of Evolutionary Biology*  
709 **25**: 1020–38.
- 710 Bulmer, M., 1972 Multiple niche polymorphism. *The American Naturalist* **106**: 254–257.
- 711 Caballero, A. and W. G. Hill, 1992 Effects of partial inbreeding on fixation rates of mutant  
712 genes. *Genetics* **131**: 493–507.
- 713 Charlesworth, B., 1992 Evolutionary rates in partially self-fertilizing species. *The American*  
714 *Naturalist* **140**: 126–148.
- 715 Connallon, T. and S. F. Chenoweth, 2019 Dominance reversals and the maintenance of  
716 genetic variation for fitness. *PLoS Biol* **17**: e3000118.

- 717 Damgaard, C., 2000 Fixation of advantageous alleles in partially self-fertilizing populations:  
718 the effect of different selection modes. *Genetics* **154**(2): 813–821.
- 719 Felsenstein, J., 1976 The theoretical population genetics of variable selection and migration.  
720 *Annual Review of Genetics* **10**: 253–280.
- 721 Gale, J. S., 1990 *Theoretical population genetics*. Springer Science & Business Media.
- 722 Glémin, S., 2021 Balancing selection in self-fertilizing populations. *Evolution* **75**(5): 1011–  
723 1029.
- 724 Haccou, P., P. Jagers, and V. A. Vatutin, 2005 *Branching processes: variation, growth, and*  
725 *extinction of populations*. Cambridge University Press.
- 726 Haldane, J., 1927 The mathematical theory of natural and artificial selection. *Proc. Camb.*  
727 *Philos. Soc.* **23**: 838—844.
- 728 Hanschen, E. R., M. D. Herron, J. J. Wiens, H. Nozaki, and R. E. Michod, 2018 Repeated  
729 evolution and reversibility of self-fertilization in the volvocine green algae. *Evolution* **72**:  
730 386–398.
- 731 Hereford, J., 2010 Does selfing or outcrossing promote local adaptation? *American Journal*  
732 *of Botany* **97**: 298–302.
- 733 Hodgins, K. A. and S. Yeaman, 2019 Mating system impacts the genetic architecture of  
734 adaptation to heterogeneous environments. *New Phytol.* **224**: 1201–1214.
- 735 Igic, B. and J. R. Kohn, 2006 The distribution of plant mating systems: study bias against  
736 obligately outcrossing species. *Evolution Int J Org Evolution* **60**: 1098–103.
- 737 Jarne, P. and J. R. Auld, 2006 Animals mix it up too: the distribution of self-fertilization  
738 among hermaphroditic animals. *Evolution* **60**: 1816–24.
- 739 Le Thierry d’Ennequin, M., B. Toupance, T. Robert, B. Godelle, and P. H. Gouyon, 1999  
740 Plant domestication: a model for studying the selection of linkage. *Journal of Evolutionary*  
741 *Biology* **12**: 1138–1147.
- 742 Leimu, R. and M. Fischer, 2008 A meta-analysis of local adaptation in plants. *PLOS ONE*  
743 **3**: 1–8.
- 744 Levene, H., 1953 Genetic equilibrium when more than one ecological niche is available. *The*  
745 *American Naturalist* **87**: 331–333.

- 746 Linhart, Y. B. and M. Grant, 1996 Evolutionary significance of local differentiation in plants.  
747 *Annu. Rev. Ecol. Syst.* **27**: 237–277.
- 748 Maynard Smith, J., 1970 Genetic polymorphism in a varied environment. *The American*  
749 *Naturalist* **104**: 487–490.
- 750 Nagylaki, T., 1980 The strong-migration limit in geographically structured populations.  
751 *Journal of Mathematical Biology* **9**: 101—114.
- 752 Olito, C., J. K. Abbott, and J. C. Y., 2018 The interaction between sex-specific selection and  
753 local adaptation in species without separate sexes. *Phil. Trans. R. Soc. B* **373**: 20170426.
- 754 Pollak, E., 1987 On the theory of partially inbreeding finite populations. i. partial selfing.  
755 *Genetics* **117**: 353–360.
- 756 Pollak, E. and M. Sabran, 1992 On the theory of partially inbreeding finite populations. iii.  
757 fixation probabilities under partial selfing when heterozygotes are intermediate in viability.  
758 *Genetics* **131**: 979–985.
- 759 Pontz, M. and R. Bürger, 2021 The effects of epistasis and linkage on the invasion of locally  
760 beneficial mutations and the evolution of genomic islands. *bioRxiv* .
- 761 Ronfort, J. and S. Glémin, 2013 Mating system, Haldane’s sieve, and the domestication  
762 process. *Evolution* **67**: 1518–26.
- 763 Roze, D., 2016 Background selection in partially selfing populations. *Genetics* **203**: 937–57.
- 764 Sakamoto, T. and H. Innan, 2019 The evolutionary dynamics of a genetic barrier to gene  
765 flow: From the establishment to the emergence of a peak of divergence. *Genetics* **212(4)**:  
766 1383–1398.
- 767 Tange, O., 2011 GNU Parallel – The Command-Line Power Tool. *The USENIX Magazine*  
768 **36**: 42–47.
- 769 Tomasini, M. and S. Peischl, 2018 Establishment of locally adapted mutations under diver-  
770 gent selection. *Genetics* **209**: 885–895.
- 771 Yeaman, S. and S. P. Otto, 2011 Establishment and maintenance of adaptive genetic diver-  
772 gence under migration, selection, and drift. *Evolution* **65**: 2123–2129.
- 773 Yeaman, S. and M. C. Whitlock, 2011 The genetic architecture of adaptation under  
774 migration-selection balance. *Evolution* **65**: 1897–911.

Symbol	Meaning
$x_i(t)$	Frequency of allele A in patch $i$ at the start of the life cycle (generation $t$ )
$x_i^{\ominus}$	Frequency of allele A in patch $i$ after female fecundity selection
$x_i^{\omin�}$	Frequency of allele A in patch $i$ after male fecundity selection
$g_i^{\omin�}$	Frequency of allele A in patch $i$ after pollen migration
$X_i$	Frequency of genotype $AA$ at the start of the life cycle
$Y_i$	Frequency of genotype $Aa$ at the start of the life cycle
$X'_i$	Frequency of genotype $AA$ after mating
$Y'_i$	Frequency of genotype $Aa$ after mating
$X''_i$	Frequency of genotype $AA$ after seed migration
$Y''_i$	Frequency of genotype $Aa$ after seed migration
$N_i^*$	Population size in patch $i$
$N_i$	The number of invading alleles in patch $i$
$\tilde{N}$	The effective population size
$M_{ij}$	The fraction of seed (or the probability of seed) in patch $i$ originating from patch $j$
$m_{ij}$	The fraction of pollen (or the probability of pollen) in patch $i$ originating from patch $j$
$\Delta W_{0,i}$	The intergenerational $A$ frequency change in patch $i$ , when $A$ rare (generic selection)
$\Delta V_{0,i}$	The intergenerational $A$ frequency change in patch $i$ , when $A$ rare (only viability selection)
$\Delta W_{0,i}^{\ominus}$	The intergenerational $A$ frequency change in patch $i$ , when $A$ rare (only female fecundity selection)
$\Delta W_{0,i}^{\omin�}$	The intergenerational $A$ frequency change in patch $i$ , when $A$ rare (only male fecundity selection)
$V_{i,k}$	The progeny number left by genotype $k$ ( $k \in \{AA, Aa, aa\}$ ) that survives to maturity in patch $i$
$W_{i,k}^{\ominus}$	Number of female gametophytes produced by genotype $k$ ( $k \in \{AA, Aa, aa\}$ ) in patch $i$
$W_{i,k}^{\omin�}$	Number of male gametophytes produced by genotype $k$ ( $k \in \{AA, Aa, aa\}$ ) in patch $i$
$\bar{V}_i$	Mean population viability fitness in patch $i$
$\bar{W}_i^{\ominus}$	Mean population fecundity fitness in patch $i$
$\bar{W}_i^{\omin�}$	Mean population male sexual fitness in patch $i$
$F$	Equilibrium population inbreeding coefficient. Set to $S/(2 - S)$
$S$	Fraction of the population that reproduces by selfing (i.e., the population selfing rate)
$s_i^{\{k\}}$	The relative selective advantage of invading $A$ , with selection via $k^{th}$ ( $k \in \{V, F, M\}$ ) fitness component
$h_i^{\{k\}}$	The dominance coefficient of allele $A$ , when selection acts only via $k^{th}$ fitness component
$\tilde{M}_{ij}$	The effective fraction of seed (or the probability of seed) in patch $i$ that originates from patch $j$
$\tilde{h}_i^{\{k\}}$	The effective dominance coefficient of allele $A$ , when selection acts only via $k^{th}$ fitness component
$\omega$	Correction for the effective dominance coefficient for male selection and pollen dispersal
$\tilde{s}_i^{\{k\}}$	The effective selective advantage of invading allele $A$ with selection via $k^{th}$ fitness component

**Table 1:** List of symbols used in the text. As a general rule, parameters with tilde signify an effective parameter. Upper-case symbols are reserved for processes affecting the diploid phase of a life cycle, whereas the lower-case represents the process of the haploid phase.

# 775 Appendices

## 776 A Deriving the PGF of the distribution of allele off- 777 spring number

### 778 A.1 Single population

779 We first derive the full probability generating function (PGF) of the number of mutant alleles  
780 under partial selfing in a single population. When the  $A$  mutant is rare, it can be found either  
781 in a  $Aa$  individual with probability  $1 - F$  and in a  $AA$  individual with probability  $F$ . The  
782 PGF can thus be written as  $(1 - F)f_{Aa}(z) + Ff_{AA}(z)$  where  $f_{Aa}$  and  $f_{AA}$  are the PGF for the  
783 parent allele being in each genotype, respectively. We can then decompose the total number  
784 of alleles,  $G_i$ , contributed by one individual of genotype  $i$  as the sum of alleles transmitted  
785 through outcrossed ovule,  $N_{i,o}$ , through exported pollen,  $N_{i,p}$ , and through selfed ovule,  $N_{i,s}$ .  
786 Because the mutant is rare, it can only be transmitted in heterozygote after outcrossing, so  
787 in a single copy. After selfing, two copies are always transmitted if the parent is  $AA$  and  
788 either two or one copies are transmitted if the parent is  $Aa$ . In this last case we note  $N_{i,s}^{AA}$   
789 and  $N_{i,s}^{Aa}$  are the number of homozygote and heterozygote seeds produced under selfing. The  
790 total number of offspring alleles transmitted at the next generation is thus:

$$G_{AA} = N_{AA,o} + N_{AA,p} + 2N_{AA,s} \quad (\text{A1a})$$

$$G_{Aa} = N_{Aa,o} + N_{Aa,p} + 2N_{Aa,s}^{AA} + N_{Aa,s}^{Aa} \quad (\text{A1b})$$

791 We assume that the number of exported pollen and the different numbers of seeds follow  
792 independent Poisson distributions. The different means depends on the fecundity of the  
793 parent, the viability of the seed produced and the proportion of each category:

$$\begin{aligned} \mathbb{E}[N_{Aa,o}] &= \frac{1}{2}(1 - S)W_{Aa}^{\circlearrowleft} V_{Aa} & \mathbb{E}[N_{AA,o}] &= \frac{1}{2}(1 - S)W_{AA}^{\circlearrowleft} V_{Aa} \\ \mathbb{E}[N_{Aa,p}] &= \frac{1}{2}(1 - S)W_{Aa}^{\circlearrowright} V_{Aa} & \mathbb{E}[N_{AA,p}] &= \frac{1}{2}(1 - S)W_{AA}^{\circlearrowright} V_{Aa} \\ \mathbb{E}[N_{Aa,s}^{AA}] &= \frac{1}{2} \frac{S}{2} W_{Aa}^{\circlearrowleft} V_{AA} & \mathbb{E}[N_{AA,s}] &= \frac{1}{2} S W_{AA}^{\circlearrowleft} V_{AA} \\ \mathbb{E}[N_{Aa,s}^{Aa}] &= \frac{1}{2} \frac{S}{2} W_{Aa}^{\circlearrowright} V_{AA} & & \end{aligned} \quad (\text{A2})$$

794 The factor  $1/2$  corresponds to the fact that the focal  $A$  is chosen with probability  $1/2$ . Put  
795 another way, on average each individual transmits two alleles so the contribution for a single

796 allele must be halved. Finally, we use the property that the PGF of  $2X$  is  $f(z^2)$  where  $f(z)$   
797 is the PGF of  $X$  and that the PGF of the sum of independent variables is the product of the  
798 PGFs. Noting  $g(\lambda, z) = e^{-(1-z)\lambda}$ , the PGF of a Poisson distribution with mean  $\lambda$ , we have:

$$\begin{aligned} f_\phi(z) &= (1 - F)g(\mathbb{E}[N_{Aa,o}], z)g(\mathbb{E}[N_{Aa,p}], z)g(\mathbb{E}[N_{Aa,s}^{Aa}], z)g(\mathbb{E}[N_{Aa,s}^{AA}], z^2) \\ &\quad + Fg(\mathbb{E}[N_{AA,o}], z)g(\mathbb{E}[N_{AA,p}], z)g(\mathbb{E}[N_{AA,s}], z^2) \end{aligned} \quad (\text{A3})$$

799 which yields equation (15) of the main text.

800 From the properties of PGFs we can easily obtain the moments of the distribution for the  
801 different forms of selection (see *Mathematica* notebook):

$$\begin{aligned} \mathbb{E}_\phi &= f'_\phi(1) \\ &= 1 + \tilde{h}s \end{aligned} \quad (\text{A4a})$$

$$\begin{aligned} \mathbb{V}_\phi &= f''_\phi(1) + f'_\phi(1) - f'_\phi(1)^2 \\ &= (1 + F)(1 + \tilde{h}s) + o(s) \end{aligned} \quad (\text{A4b})$$

802 So we retrieved the well known result that selfing increases variance in allele offspring number  
803 by  $1 + F$ . Interestingly, the full distribution presents a peculiar non-monotonic behavior for  
804 high selfing, with an excess (resp. a deficit) in even (resp. odd) numbers. Note also that the  
805 full distributions are not exactly the same under the different modes of selection.

## 806 A.2 Two-patch model

807 We need to derive the distribution of the number of mutant alleles issued from one patch  
808 and staying in the same patch, with PGF  $f_{i,i}(z)$ , and the distribution of those establish-  
809 ing in the other patch, with PGF  $f_{i,j}(z)$ . Assuming that the distribution of resident and  
810 migrant alleles are independent, the total number of alleles produced by patch  $i$  has PGF:  
811  $f_i(z) = f_{i,i}(z)f_{i,j}(z)$ . The  $f_{i,j}(z)$  can be obtained using the same equations as for a single  
812 population by paying attention to the order of events to correctly set indices: male fecundity  
813 selection, pollen migration, female fecundity selection, reproduction, seed migration and vi-  
814 ability selection. Under seed migration, fecundity selection occurs in patch  $i$  but viability

815 selection in patch  $j$ . The means given in equation (A2) become:

$$\begin{aligned}
\mathbb{E}[N_{Aa,o,ij}] &= \frac{1}{2}(1-S)W_{Aa,i}^{\ominus}M_{i,j}V_{Aa,j} & \mathbb{E}[N_{AA,o,ij}] &= \frac{1}{2}(1-S)W_{AA,i}^{\ominus}M_{i,j}V_{Aa,j} \\
\mathbb{E}[N_{Aa,p,ij}] &= \frac{1}{2}(1-S)W_{Aa,i}^{\ominus}M_{i,j}V_{Aa,j} & \mathbb{E}[N_{AA,p,ij}] &= \frac{1}{2}(1-S)W_{AA,i}^{\ominus}M_{i,j}V_{Aa,j} \\
\mathbb{E}[N_{Aa,s,ij}^{AA}] &= \frac{1}{2}\frac{S}{2}W_{Aa,i}^{\ominus}M_{i,j}V_{AA,j} & \mathbb{E}[N_{AA,s,ij}] &= \frac{1}{2}SW_{AA,i}^{\ominus}M_{i,j}V_{AA,j} \\
\mathbb{E}[N_{Aa,s,ij}^{Aa}] &= \frac{1}{2}\frac{S}{2}W_{Aa,i}^{\ominus}M_{i,j}V_{Aa,j} & & 
\end{aligned} \tag{A5}$$

816 with  $M_{i,j} = 1 - M_{i,i}$ . Plugging (A5) into (A3) yields equation (16) in the main text. From  
817 the corresponding PGFs we then retrieve the same mean as obtained by the deterministic  
818 analysis (see below), and variances inflated by  $1 + F$  (see *Mathematica* notebook):

$$\mathbb{E}_{\phi,ij} = M_{i,j}(1 + \tilde{h}_i s_i) \quad \text{for fecundity selection} \tag{A6a}$$

$$= M_{i,j}(1 + \tilde{h}_j s_j) \quad \text{for viability selection} \tag{A6b}$$

$$\mathbb{V}_{\phi,ij} = (1 + F)\mathbb{E}_{\phi,ij} + o(s_i, s_j) \tag{A6c}$$

819 Note that in subsequent analyses the order of migration and selection terms yields the same  
820 results. Under pollen migration, the PGF for resident and migrant contribution have different  
821 forms because an allele can contribute offspring to the other patch only through outcrossing  
822 and through the male pathway. For offspring contributing to the resident patch we have:

$$\begin{aligned}
\mathbb{E}[N_{Aa,o,ii}] &= \frac{1}{2}(1-S)W_{Aa,i}^{\ominus}V_{Aa,i} & \mathbb{E}[N_{AA,o,ii}] &= \frac{1}{2}(1-S)W_{AA,i}^{\ominus}V_{Aa,i} \\
\mathbb{E}[N_{Aa,p,ii}] &= \frac{1}{2}(1-S)W_{Aa,i}^{\ominus}(1 - m_{i,j})V_{Aa,i} & \mathbb{E}[N_{AA,p,ii}] &= \frac{1}{2}(1-S)W_{AA,i}^{\ominus}(1 - m_{i,i})V_{Aa,i} \\
\mathbb{E}[N_{Aa,s,ii}^{AA}] &= \frac{1}{2}\frac{S}{2}W_{Aa,i}^{\ominus}V_{AA,i} & \mathbb{E}[N_{AA,s,ii}] &= \frac{1}{2}SW_{AA,i}^{\ominus}V_{AA,i} \\
\mathbb{E}[N_{Aa,s,ii}^{Aa}] &= \frac{1}{2}\frac{S}{2}W_{Aa,i}^{\ominus}V_{Aa,i} & & 
\end{aligned} \tag{A7}$$

823 whereas for offspring contributing to the other patch:

$$\begin{aligned}
\mathbb{E}[N_{Aa,o,ii}] &= 0 & \mathbb{E}[N_{AA,o,ii}] &= 0 \\
\mathbb{E}[N_{Aa,p,ii}] &= \frac{1}{2}(1-S)W_{Aa,i}^{\ominus}m_{i,j}V_{Aa,j} & \mathbb{E}[N_{AA,p,ii}] &= \frac{1}{2}(1-S)W_{AA,i}^{\ominus}m_{i,i}V_{Aa,j} \\
\mathbb{E}[N_{Aa,s,ii}^{AA}] &= 0 & \mathbb{E}[N_{AA,s,ii}] &= 0 \\
\mathbb{E}[N_{Aa,s,ii}^{Aa}] &= 0 & & 
\end{aligned} \tag{A8}$$

824 Plugging (A7) and (A8) into (A3) yields equations (17a) and (17b) in the main text. The  
 825 mean and variance for resident offspring are:

$$\mathbb{E}_{\phi,ii} = (1 - \tilde{M}_{i,j})(1 + \tilde{h}_i s_i) \quad \text{for viability and female fecundity selection} \quad (\text{A9a})$$

$$= (1 - \tilde{M}_{i,j})(1 + \omega \tilde{h}_i s_i) \quad \text{for male fecundity selection} \quad (\text{A9b})$$

$$\mathbb{V}_{\phi,ii} = \mathbb{E}_{\phi,ij} \left( 1 + \frac{F}{1 - \tilde{M}_{ij}} \right) + o(s_i, s_j) \quad (\text{A9c})$$

826 and for migrant offspring:

$$\mathbb{E}_{\phi,ij} = \tilde{M}_{i,j}(1 + \tilde{h}_j s_j) \quad \text{for viability fecundity selection} \quad (\text{A10a})$$

$$= 0 \quad \text{for female fecundity selection} \quad (\text{A10b})$$

$$= \tilde{M}_{i,j}(1 + \omega \tilde{h}_i s_i) \quad \text{for male fecundity selection} \quad (\text{A10c})$$

$$\mathbb{V}_{\phi,ij} = \mathbb{E}_{\phi,ij} + o(s_i, s_j) \quad (\text{A10d})$$

827 where  $\tilde{M}_{i,j} \approx m_{i,j}(1-S)/2$  are effective migration rates as defined in the main text (equations  
 828 12a to 12c) and  $\omega$  is a correcting factor defined in equation (13). Compared to seed migration,  
 829 the variance is not uniformly increased by  $1 + F$ . As migrant offspring can only be produced  
 830 through outcrossing, the distribution is simply Poisson and the variance equal to the mean.  
 831 On the contrary, because the proportion of outcrossed offspring contributing to the resident  
 832 patch is reduced due to pollen migration, the variance is inflated by more than  $1 + F$ .  
 833 However, the difference between the two modes of migration does not affect the following  
 834 approximations for weak selection.

## 835 **B The establishment probability approximated to weak** 836 **selection**

837 Following the procedure outlined in Haccou *et al.* (2005) (Section 5.6.2) and previously  
 838 adapted for a similar problem by Tomasini and Peischl (2018), we seek to approximate  $P$  by  
 839 working with slightly supercritical process. Let  $\rho$  be the leading eigenvalue and  $\mathbf{u} = [u_1, u_2]^T$   
 840 and  $\mathbf{v} = [v_1, v_2]^T$  normed left and right eigenvectors of of the mean reproductive matrix  $\mathbf{M}$ .



841 Eigenvectors are normalized such that:

$$u_i > 0, \quad v_i > 0, \quad i \in \{1, 2\}$$

$$\sum_i u_i v_i = \sum_i u_i = 1$$

842 We choose a parameter  $\epsilon$  in the model, such that the process is slightly supercritical when  $\epsilon$   
843 is small. All other parameters in the model are rescaled by  $\epsilon$ . More formally:

$$\rho(\epsilon) \rightarrow 1 \quad \text{as} \quad \epsilon \rightarrow 0$$

844 Eigenvalues and eigenvectors of  $\mathbf{M}$  are dependent on  $\epsilon$ . Probability of establishment starting  
845 from a single copy of allele in patch  $i$  is given by:

$$P_i = 2(\rho(\epsilon) - 1) \frac{v_i(\epsilon)}{B(\epsilon)} + o(\epsilon) \quad (\text{A11})$$

$$B(\epsilon) = \sum_i u_i(\epsilon) \text{Var} \left( \phi_{ij} \sum_j v_j(\epsilon) \right) + (\rho(\epsilon)(\rho(\epsilon) - 1)) \sum_j u_j(\epsilon) v_j(\epsilon)^2 \quad (\text{A12})$$

846 As  $\epsilon \rightarrow 0$ ,  $\rho$  approaches unity and the second term in B (equation (A12)) can be neglected.  
847 As per equation 5.85 in Haccou *et al.* (2005), we take that  $B(\epsilon) \rightarrow B(0)$  and  $v_i(\epsilon) \rightarrow v_i(0)$ :

$$P_k \approx 2(\rho(\epsilon) - 1) \frac{v_k(0)}{B(0)} \quad (\text{A13})$$

$$B(0) \approx \sum_k u_k(0) \text{Var} \left( \phi_{kj} \sum_j v_j(0) \right) \quad (\text{A14})$$

848 Recall that  $\tilde{s}_i$  is the advantage of an allele in patch  $i$  accounting for the effect of self-  
849 fertilization. Assuming weak selective advantage,  $\tilde{s}_1$  is taken as  $\epsilon$ . All other parameters are  
850 expressed in terms of  $\tilde{s}_1$ :

$$\tilde{M}_{ij} = \tilde{s}_1 \chi_{ij}, \quad \tilde{s}_2 = \zeta \tilde{s}_1$$

851 The leading eigenvalue of  $\mathbf{M}$  can be written as  $\rho = 1 + \tilde{c} \tilde{s}_1 + o(\tilde{s}_1^2)$ . Dropping the higher-order  
852 terms in  $\tilde{s}_1$ , we retrieve the expression for  $\tilde{c}$ :

$$\tilde{c} = \frac{1}{2} \left( \sqrt{(\zeta + \chi_{12} - 1)^2 + 2\chi_{21}(-\zeta + \chi_{12} + 1) + \chi_{21}^2 + \zeta - \chi_{12} - \chi_{21} + 1} \right)$$

853 By taking only the zeroth term of  $v_i$  and  $B$  in Taylor expansion about  $\tilde{s}_1 = 0$ , we obtain the

854 approximation of the probability of establishment conditioning on a single mutant appearing  
 855 in patch  $i$ :

$$P_k \approx 2\tilde{c} \frac{v_k(0)}{B(0)} \tilde{s}_1 \quad (\text{A15})$$

856 Transforming back to original variables  $\tilde{M}_{ij}$  and  $\tilde{s}_i$  and letting

$$= \sqrt{2\tilde{M}_{12} \left( \tilde{M}_{21} - \tilde{s}_1 + \tilde{s}_2 \right) + \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 \right)^2 + \tilde{M}_{12}^2}$$

857 we have:

$$P_1 = \frac{\left( -\tilde{M}_{12} \left( -2\tilde{M}_{21} + 2\tilde{s}_1 - 2\tilde{s}_2 + \psi \right) + \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 \right) \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 + \psi \right) + \tilde{M}_{12}^2 \right) \left( -\tilde{M}_{12} + \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 + \psi \right) \left( -\tilde{M}_{12} - \tilde{M}_{21} + \tilde{s}_1 + \tilde{s}_2 + \psi \right)}{2(F+1) \left( \tilde{M}_{12}^2 \left( 2\tilde{M}_{21} + 3\tilde{s}_1 - 3\tilde{s}_2 + \psi \right) - \tilde{M}_{12} \left( \tilde{M}_{21} \left( 3\tilde{s}_1 - 3\tilde{s}_2 + \psi \right) + \left( \tilde{s}_1 - \tilde{s}_2 \right) \left( 3\tilde{s}_1 - 3\tilde{s}_2 + 2\psi \right) \right) + \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 \right)^2 \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 + \psi \right) - \tilde{M}_{12}^3 \right)} \quad (\text{A16})$$

$$P_2 = \frac{\tilde{M}_{12} \left( -\tilde{M}_{12} \left( -2\tilde{M}_{21} + 2\tilde{s}_1 - 2\tilde{s}_2 + \psi \right) + \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 \right) \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 + \psi \right) + \tilde{M}_{12}^2 \right) \left( -\tilde{M}_{12} - \tilde{M}_{21} + \tilde{s}_1 + \tilde{s}_2 + \psi \right)}{(F+1) \left( \tilde{M}_{12}^2 \left( 2\tilde{M}_{21} + 3\tilde{s}_1 - 3\tilde{s}_2 + \psi \right) - \tilde{M}_{12} \left( \tilde{M}_{21} \left( 3\tilde{s}_1 - 3\tilde{s}_2 + \psi \right) + \left( \tilde{s}_1 - \tilde{s}_2 \right) \left( 3\tilde{s}_1 - 3\tilde{s}_2 + 2\psi \right) \right) + \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 \right)^2 \left( \tilde{M}_{21} + \tilde{s}_1 - \tilde{s}_2 + \psi \right) - \tilde{M}_{12}^3 \right)} \quad (\text{A17})$$

858

859 When migration is symmetrical ( $\tilde{M}_{12} = \tilde{M}_{21} = \tilde{M}$ ),  $\psi$  reduces to  $\sqrt{4\tilde{M}^2 + (\tilde{s}_1 - \tilde{s}_2)}$ , pre-  
 860 viously defined as the scaled measure of the heterogeneity in selection and migration. Set-  
 861 ting  $F = 0$ , the equations (A16) and (A17) behave similarly to those derived previously  
 862 (Tomasini and Peischl (2018)). Comparing the analytical solution against simulations, one  
 863 can see that the solution reported here fits simulations slightly better than the Tomasini-  
 864 Peischl result (Figure A9). The discrepancy is probably caused by the latter's use of B term  
 865 from Aeschbacher and Bürger (2014) (see their equation S22), which neglects to square ele-  
 866 ments of the eigenvectors after factoring them out of the variance. We, on the other hand,  
 867 computed B directly from Haccou *et al.* (2005), which does not suffer from this error. Cu-  
 868 riously, the Tomasini-Peischl approximation is much more elegant as the denominator can  
 869 be interpreted as the measure of the heterogeneity of migration and selection. Relative to  
 870 the result reported in Sakamoto and Innan (2019), our solution had identical performance in  
 871 the favored patch, and worse performance in the disfavored patch (pink dashed lines). This  
 872 discrepancy occurs only when the migration rate is low.

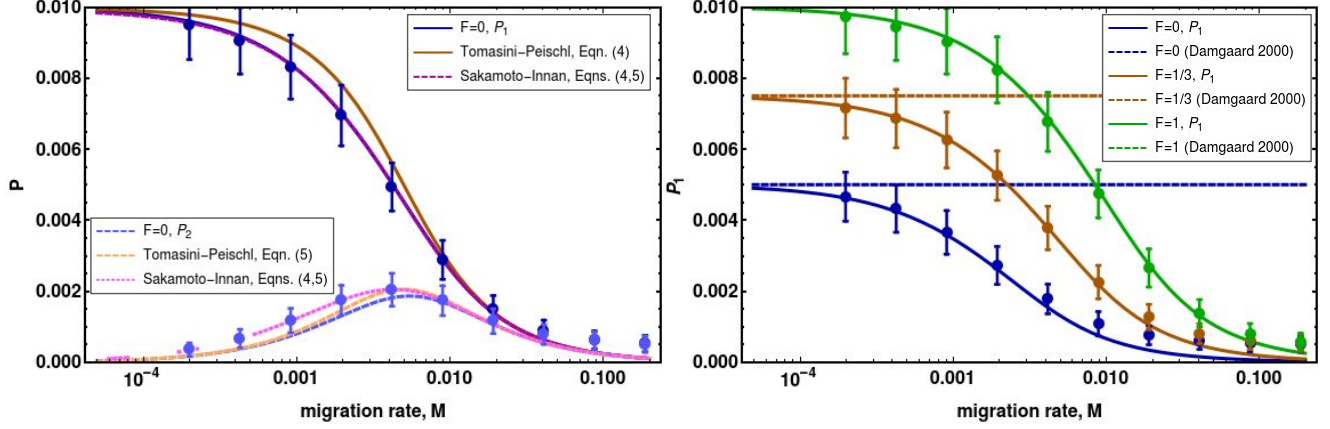


Figure A9: Comparison of analytical solution to simulated data and previous results. Comparison to approximation Tomasini and Peischl (2018) (left), and comparison to the single-patch heuristic under fecundity selection Damgaard (2000) (right). All equations parameterized with  $h_1^+ = h_2^+ = h_1^0 = -h_2^0 = 1/2$ ,  $s_1^+ = -s_2^+ = s_1^0 = -s_2^0 = 0.01$ ,  $M_{12} = M_{21} = M$ .

## 873 C The criterion for escaping extinction

### 874 C.1 Seed dispersal

875 If the selective disadvantage in the disfavored patch is too large or migration is too strong,  
 876 the spreading locally advantageous allele can be swamped by its deleterious counterpart.  
 877 The range of parameters that are necessary but not sufficient for a successful invasion are  
 878 obtained by linearizing the system about  $\vec{x} = [0, 0]^T$  and investigating the conditions required  
 879 for this equilibrium to be locally unstable. If migration occurs prior to selection – which is  
 880 the case when seed disperses and selection acts on viability, then Jacobian  $\mathbf{J}$  is:

$$\mathbf{J} = \begin{bmatrix} (1 - M_{12})\Delta\mathcal{W}_{0,1} & M_{12}\Delta\mathcal{W}_{0,1} \\ M_{21}\Delta\mathcal{W}_{0,2} & (1 - M_{21})\Delta\mathcal{W}_{0,2} \end{bmatrix} \quad (\text{A18})$$

$$\Delta\mathcal{W}_{0,i} = (F\mathcal{W}_{i,AA} + (1 - F)\mathcal{W}_{i,Aa})/\mathcal{W}_{i,aa} \quad (\text{A19})$$

881 Symbol  $\mathcal{W}$  is a place-holder for any of the three selection modes ( $V$ ,  $W^\varphi$ , or  $W^\sigma$ ). If  
 882 selection occurs prior to migration – which happens when seed disperses and selection acts  
 883 on sexual components– then Jacobian is:

$$\mathbf{J} = \begin{bmatrix} (1 - M_{12})\Delta\mathcal{W}_{0,1} & M_{12}\Delta\mathcal{W}_{0,2} \\ M_{21}\Delta\mathcal{W}_{0,1} & (1 - M_{21})\Delta\mathcal{W}_{0,2} \end{bmatrix} \quad (\text{A20})$$

884 However, the eigenvalues of these two matrices are identical, so we here use (A18). The  
 885 equilibrium is locally unstable whenever the leading eigenvalue of  $\mathbf{J}$  is greater than unity:

$$\rho = \frac{1}{2} \left( (1 - M_{12})\Delta\mathcal{W}_{0,1} + (1 - M_{21})\Delta\mathcal{W}_{0,2} + \sqrt{4(M_{12} + M_{21} - 1)\Delta\mathcal{W}_{0,1}\Delta\mathcal{W}_{0,2} + \left((M_{21} - 1)\Delta\mathcal{W}_{0,1} + (M_{21} - 1)\Delta\mathcal{W}_{0,2}\right)^2} \right) \quad (\text{A21})$$

886 Term  $\Delta\mathcal{W}_{0,i}$  can be thought of as the rate of spread of  $A$  allele in  $i^{\text{th}}$  patch. Then  $\Delta\mathcal{W}_{0,1} > 1$   
 887 (because  $A$  is advantageous in the first patch), and  $0 < \Delta\mathcal{W}_{0,2} < 1$  (because  $A$  is deleterious  
 888 in the second patch). It is important to note that these inequalities hold regardless of the  
 889 mode of selection. More formally:

$$\Delta V_{0,1} > 1, \quad 1 > \Delta V_{0,2} > 0 \quad (\text{A22})$$

$$\Delta W_{0,1}^{\circ} > 1, \quad 1 > W_{0,2}^{\circ} > 0 \quad (\text{A23})$$

$$\Delta W_{0,1}^{\sigma} > 1, \quad 1 > W_{0,2}^{\sigma} > 0 \quad (\text{A24})$$

890 Substituting (A19) in (A21), and rearranging (see *Mathematica* notebook for details), we  
 891 find that  $\rho > 1$  when:

$$\frac{M_{12}}{\frac{((1-F)\mathcal{W}_{1,Aa} + F\mathcal{W}_{1,AA}) - \mathcal{W}_{1,aa}}{(1-F)\mathcal{W}_{1,Aa} + F\mathcal{W}_{1,AA}}} + \frac{M_{21}}{\frac{((1-F)\mathcal{W}_{2,Aa} + F\mathcal{W}_{2,AA}) - \mathcal{W}_{2,aa}}{(1-F)\mathcal{W}_{2,Aa} + F\mathcal{W}_{2,AA}}} < 1 \quad (\text{A25})$$

892 Terms in denominators of inequality A25 represent the relative fitness of the invading allele  
 893  $A$ . Letting  $F = 0$  and parameterizing such that  $\mathcal{W}_{i,AA} = 1 - s_1^{\circ}$ ,  $\mathcal{W}_{i,Aa} = 1$ , and  $\mathcal{W}_{i,aa} = 1 - t_1^{\circ}$ ,  
 894 we retrieve Bulmer's inequality  $M_{12}/t_1^{\circ} + M_{21}/t_2^{\circ}$ . Parameterizing according to our fitness  
 895 scheme ( $\mathcal{W}_{i,AA} = 1 + s_i^{\circ}$ ,  $\mathcal{W}_{i,Aa} = 1 + s_i^{\circ}\tilde{h}_i^{\circ}$ , and  $\mathcal{W}_{i,aa} = 1$ ) yields:

$$\frac{(1 + s_1^{\circ}\tilde{h}_1^{\circ})M_{12}}{s_1^{\circ}\tilde{h}_1^{\circ}} + \frac{(1 + s_2^{\circ}\tilde{h}_2^{\circ})M_{21}}{s_2^{\circ}\tilde{h}_2^{\circ}} \quad (\text{A26})$$

896 Conversely to the previous case,  $0 < \Delta\mathcal{W}_{1,1} < 1$  (as invading allele  $a$  is deleterious in the  
 897 first patch) and  $\Delta\mathcal{W}_{1,2} > 1$  (given that  $a$  is beneficial in the second patch). Once the allele  $A$   
 898 has escaped extinction, it can either fix in both patches or be maintained for finite number  
 899 of generations by divergent selection. Thus, the criterion for protected polymorphism is that  
 900 both allele  $A$  and  $a$  can escape extinction. By linearizing the system of replicator equations

901 about  $\vec{x} = [1, 1]^T$  we obtain Jacobian:

$$\mathbf{J} = \begin{bmatrix} (1 - M_{12})\Delta\mathcal{W}_{1,1} & M_{12}\Delta\mathcal{W}_{1,1} \\ M_{21}\Delta\mathcal{W}_{1,2} & (1 - M_{21})\Delta\mathcal{W}_{1,2} \end{bmatrix} \quad (\text{A27})$$

$$\Delta\mathcal{W}_{1,i} = (F\mathcal{W}_{i,aa} + (1 - F)\mathcal{W}_{i,Aa})/\mathcal{W}_{i,AA} \quad (\text{A28})$$

902 Therefore, allele  $a$  is allowed to invade whenever:

$$\frac{M_{12}}{\frac{((1-F)\mathcal{W}_{1,Aa} + F\mathcal{W}_{1,aa}) - \mathcal{W}_{1,AA}}{(1-F)\mathcal{W}_{1,Aa} + F\mathcal{W}_{1,aa}}} + \frac{M_{21}}{\frac{((1-F)\mathcal{W}_{2,Aa} + F\mathcal{W}_{2,aa}) - \mathcal{W}_{2,AA}}{(1-F)\mathcal{W}_{2,Aa} + F\mathcal{W}_{2,aa}}} < 1 \quad (\text{A29})$$

903 which reduces to  $M_{12}/s_1^\circ + M_{21}/s_2^\circ$  with fitness notation of Bulmer.

## 904 C.2 Pollen dispersal

905 The analysis is more complicated for the three selection scenarios when pollen disperses  
 906 because one has to re-parameterize migration rate in addition to dominance coefficients. We  
 907 use Jacobian of the form:

$$\mathbf{J} = \begin{bmatrix} (1 - \tilde{M}_{12})\Delta\mathcal{W}_{0,1} & \tilde{M}_{12}\Delta\mathcal{W}_{0,1} \\ \tilde{M}_{21}\Delta\mathcal{W}_{0,2} & (1 - \tilde{M}_{21})\Delta\mathcal{W}_{0,2} \end{bmatrix} \quad (\text{A30})$$

908 The leading eigenvalue of (A30) is:

$$\rho = \frac{1}{2} \left( (1 - \tilde{M}_{12})\Delta\mathcal{W}_{0,1} + (1 - \tilde{M}_{21})\Delta\mathcal{W}_{0,2} + \sqrt{4(\tilde{M}_{12} + \tilde{M}_{21} - 1)\Delta\mathcal{W}_{0,1}\Delta\mathcal{W}_{0,2} + \left( (\tilde{M}_{21} - 1)\Delta\mathcal{W}_{0,1} + (\tilde{M}_{21} - 1)\Delta\mathcal{W}_{0,2} \right)^2} \right), \quad (\text{A31})$$

909 where  $\Delta\mathcal{W}$  and  $\tilde{M}$  are parameterized as outlined in Section 2.2. Given that  $1/2 \geq \tilde{M}_{ij} \geq 0$   
 910 for all  $i$  and  $j$ , and

$$\Delta V_{0,1} > 1, \quad 1 > \Delta V_{0,2} > 0 \quad (\text{A32})$$

$$\Delta W_{0,1}^\circ > 1, \quad 1 > W_{0,2}^\circ > 0, \quad (\text{A33})$$

911 the same inequalities hold as in the case of seed dispersal. However, under male fecundity

912 selection migration rates have to be:

$$m_{ij} < \frac{W_{j,aa}^{\sigma} \left( W_{i,aa}^{\sigma} - (1-F)W_{i,Aa}^{\sigma} - FW_{i,AA}^{\sigma} \right)}{W_{i,aa}^{\sigma} \left( (1-F)W_{j,Aa}^{\sigma} + FW_{j,AA}^{\sigma} \right) - W_{j,aa}^{\sigma} \left( (1-F)W_{i,Aa}^{\sigma} + FW_{i,AA}^{\sigma} \right)}, \quad (\text{A34})$$

913 for both  $m_{12}$  and  $m_{21}$ , or:

$$m_{ij} > \frac{W_{j,aa}^{\sigma} \left( W_{i,aa}^{\sigma} - (1-F)W_{i,Aa}^{\sigma} - FW_{i,AA}^{\sigma} \right)}{W_{i,aa}^{\sigma} \left( (1-F)W_{j,Aa}^{\sigma} + FW_{j,AA}^{\sigma} \right) - W_{j,aa}^{\sigma} \left( (1-F)W_{i,Aa}^{\sigma} + FW_{i,AA}^{\sigma} \right)}, \quad (\text{A35})$$

914 for both  $m_{12}$  and  $m_{21}$ . The conditions (A34) and (A35) are identical to saying that  $\omega_i > 0$   
 915 for all  $i$ , or  $\omega_i < 0$  for all  $i$ . A possible intuitive explanation for these conditions is as follows.  
 916 If migration from favored to disfavored patch is too high, the mutant alleles are transferred  
 917 to a disfavored patch where they are purged, thus causing the mutant to go extinct. If,  
 918 on the other hand, migration from disfavored to favored patch is high, then the spreading  
 919 mutant is swamped by the influx of deleterious residents from the opposite patch.

## 920 D Simulation method

921 Selection and migration are assumed to alter genotype frequencies deterministically. Genetic  
 922 drift is implemented by randomly drawing genotypes from multinomial distribution right  
 923 after reproduction. We relax the assumption that the population has to reach equilibrium in  
 924 the inbreeding coefficient  $F$ . Each simulation run terminates in a successful or failed invasion  
 925 and is composed of the following four steps:

- 926 1. Inject a single heterozygote containing allele  $A$  in a population that is fixed for  $a$  allele.  
 927 When interested in  $P_i$ , the mutant is injected in  $i^{\text{th}}$  patch.
- 928 2. Update genotype frequencies by applying equations (2a)–(4b); This emulates selection  
 929 on sexual components, pollen dispersal, and reproduction (including selfing).
- 930 3. Sample the genotypes from multinomial distribution to determine the genotype fre-  
 931 quencies after the reproduction step:  $\text{MN}([X_i, Y_i, Z_i], N_i^*)$ , where  $N_i^*$  is the size of  $i^{\text{th}}$   
 932 patch.
- 933 4. Update genotype frequencies due to seed dispersal by implementing (5a) and (5b).
- 934 5. Compute the number of each genotype after viability selection as  $X_i^* := X''_i V_{i,AA}$ ,  
 935  $Y_i^* := Y'''_i V_{i,Aa}$ , and  $Z_i^* := 1 - X_i^* - Y_i^*$ .

936 6. Is the number of invading mutant alleles larger than 1000? If not, convert genotype  
937 numbers to genotype frequencies by dividing drawn genotype numbers with the total  
938 patch size and begin the new generation by going to step 2. If yes, count it as a  
939 successful invasion and terminate the simulation run.

940 If the mutant allele does not invade in 10,000 generations, we count the simulation run as  
941 a failed invasion. The establishment probability is obtained by running 10,000 simulations  
942 and computing the fraction of runs that ended in the successful invasion. All error bars  
943 in the figures correspond to the standard deviation of this metric. The chosen number for  
944 invasion threshold is well over  $1/s_1$ , given that we use  $s_1 = 0.01$ . Both patches contain 10,000  
945 individuals, and populations are always in  $|Ns| \gg 1$  regime.

## 946 **E Simulations under various selection and migration** 947 **modes**

948 Our approximation gives a good fit for simulated data under female and male fecundity,  
949 although the fit is not excellent as in the case of viability selection (Figures A10–A13). In  
950 the case of male fecundity selection under complete selfing ( $S = 1$ ), the male component  
951 does not contribute to the fitness, and the allele is expected to behave neutrally. One can see  
952 that this occurs because the probability that the invading allele reaches the threshold that  
953 we use to determine whether the invasion is successful is inversely proportional to the size  
954 of the threshold,  $1/N_{\text{thres}}$ , (grey dashed line in figures below). The figures below compare  
955 the analytics to simulated data, with upper row denoting the case when allele originates in  
956 the favored patch, and bottom row depicts the same dynamic but when allele appears in  
957 the disfavored patch. From left to right, columns show codominant, partially dominant, and  
958 partially recessive case, respectively.

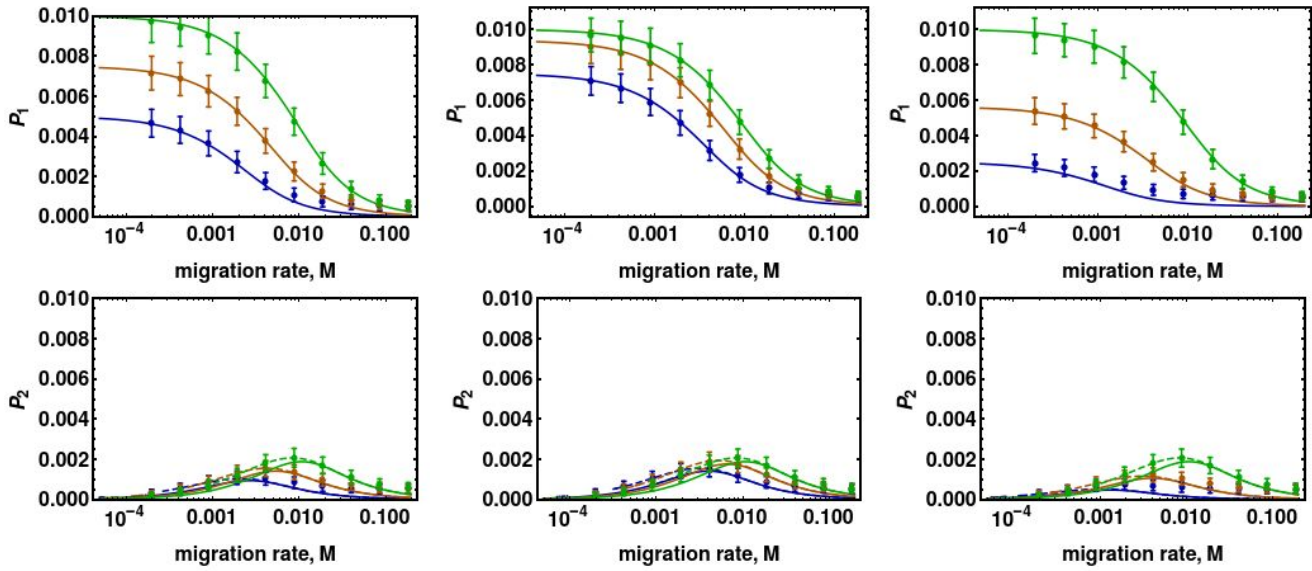


Figure A10: Comparison of analytical solution to simulations when selection acts on the female fecundity and only seed disperses. Left column: codominant case ( $h_1^{\circ} = h_2^{\circ} = 1/2$ ); Middle column: dominant case ( $h_1^{\circ} = h_2^{\circ} = 3/4$ ); Right column: recessive case ( $h_1^{\circ} = h_2^{\circ} = 1/4$ ). Upper and lower panels depict the establishment probability conditioning on allele emerging in favored and disfavored patch, respectively. Parameters:  $s_1^{\circ} = -s_2^{\circ} = 0.01$ ,  $M_{12} = M_{21} = M$ .



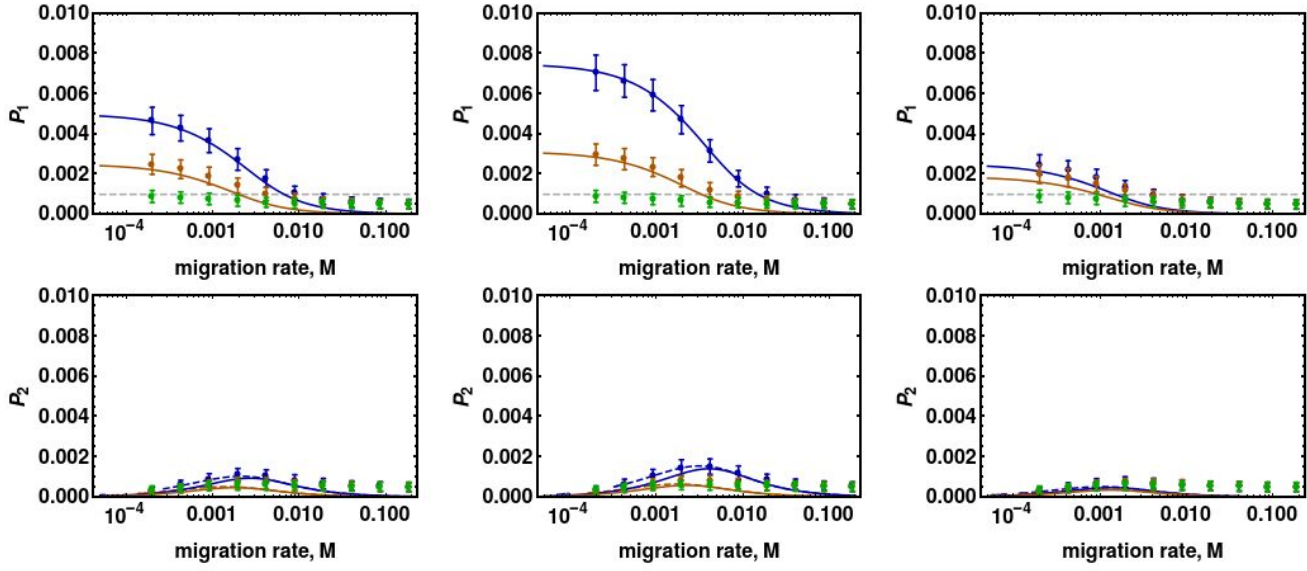


Figure A11: Comparison of analytical solution to simulations when selection acts on male fecundity and only seed disperses. Left column: codominant case ( $h_1^{\sigma} = h_2^{\sigma} = 1/2$ ); Middle column: dominant case ( $h_1^{\sigma} = h_2^{\sigma} = 3/4$ ); Right column: recessive case ( $h_1^{\sigma} = h_2^{\sigma} = 1/4$ ). Upper and lower panels depict the establishment probability conditioning on allele emerging in favored and disfavored patch, respectively. Parameters:  $s_1^{\sigma} = -s_2^{\sigma} = 0.01$ ,  $M_{12} = M_{21} = M$ .

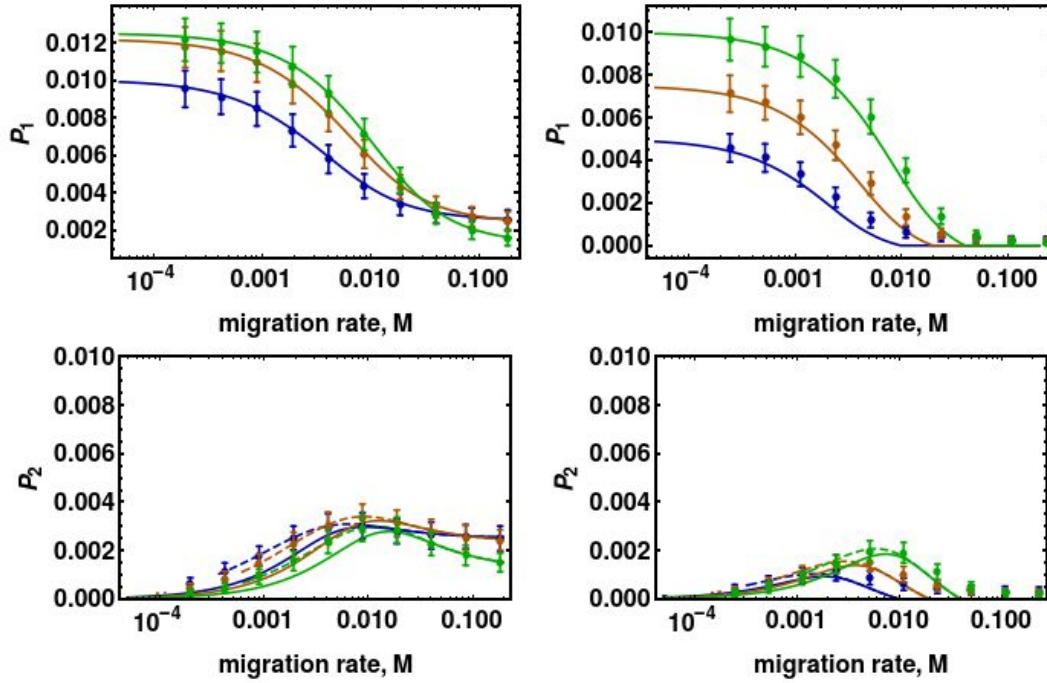


Figure A12: Comparison of analytical solution to simulations when selection acts on the female fecundity, only seed disperses, and selection or migration are asymmetrical. Left column: asymmetric selection ( $s_1^\circ = 0.01$ ,  $s_2^\circ = -0.02$ ,  $h_1^\circ = h_2^\circ = 1/2$ ); Right column: asymmetric migration ( $s_1^\circ = -s_2^\circ = 0.01$ ,  $h_1^\circ = h_1^\circ = 1/2$ ,  $M_{12} = M$ , and  $M_{21} = 1.25M$ ).

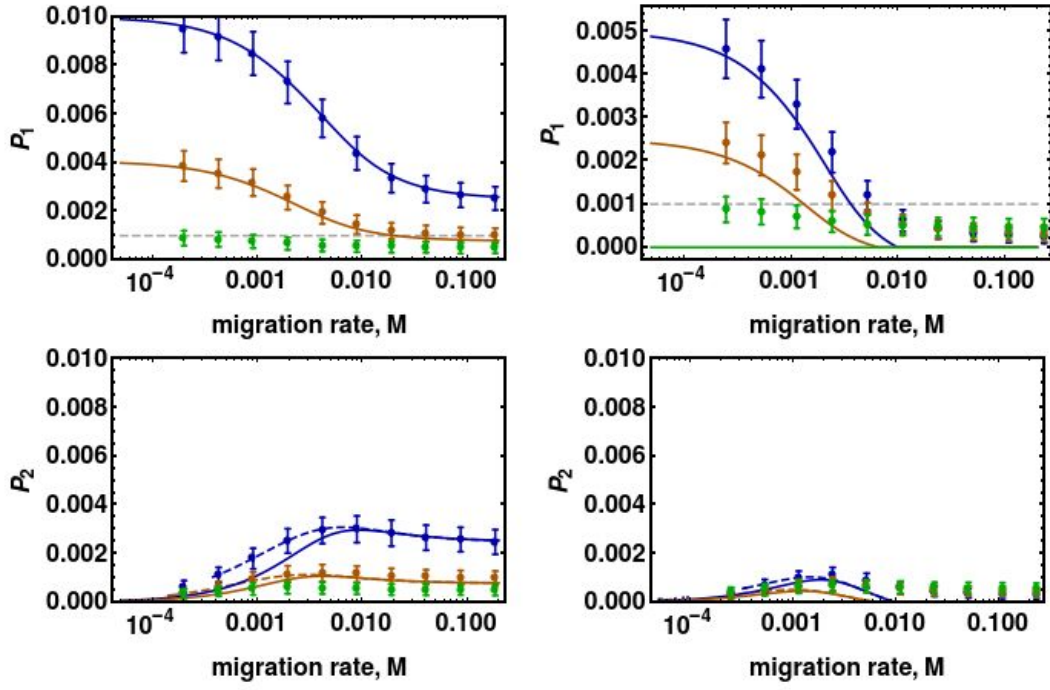


Figure A13: Comparison of analytical solution to simulations when selection acts on the male fecundity, only seed migrates, and selection or migration are asymmetrical. Left column: asymmetric selection ( $s_1^\sigma = 0.01$ ,  $s_2^\sigma = -0.02$ ,  $h_1^\sigma = h_2^\sigma = 1/2$ ); Right column: asymmetric migration ( $s_1^\sigma = -s_2^\sigma = 0.01$ ,  $h_1^\sigma = h_2^\sigma = 1/2$ ,  $M_{12} = M$ , and  $M_{21} = 1.25M$ ).

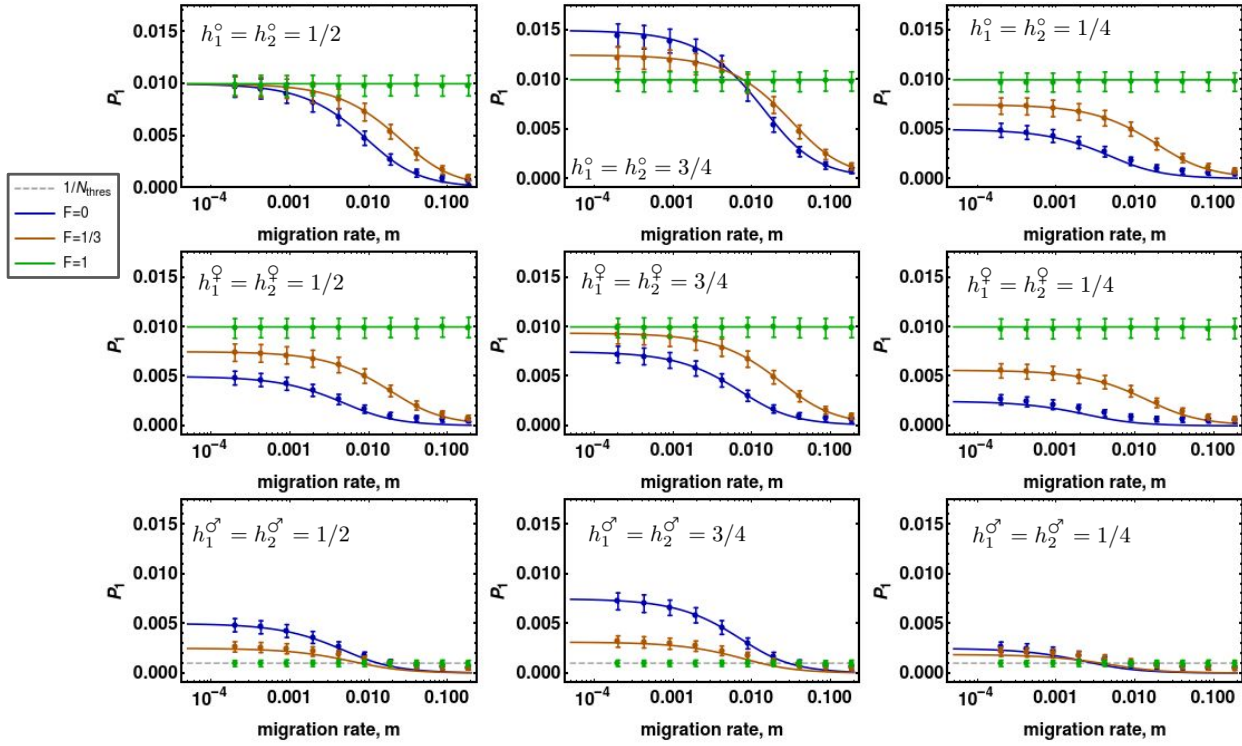


Figure A14: Comparison of analytical solution to simulations when only pollen disperses. Allele starts in the favored patch. Top row: viability selection; Middle row: female fecundity selection; Bottom row: Male fecundity selection. Dominances reported in panels. Other parameters: corresponding selection coefficients are always  $s_1 = -s_2 = 0.01$ , and  $m_{12} = m_{21} = m$ .

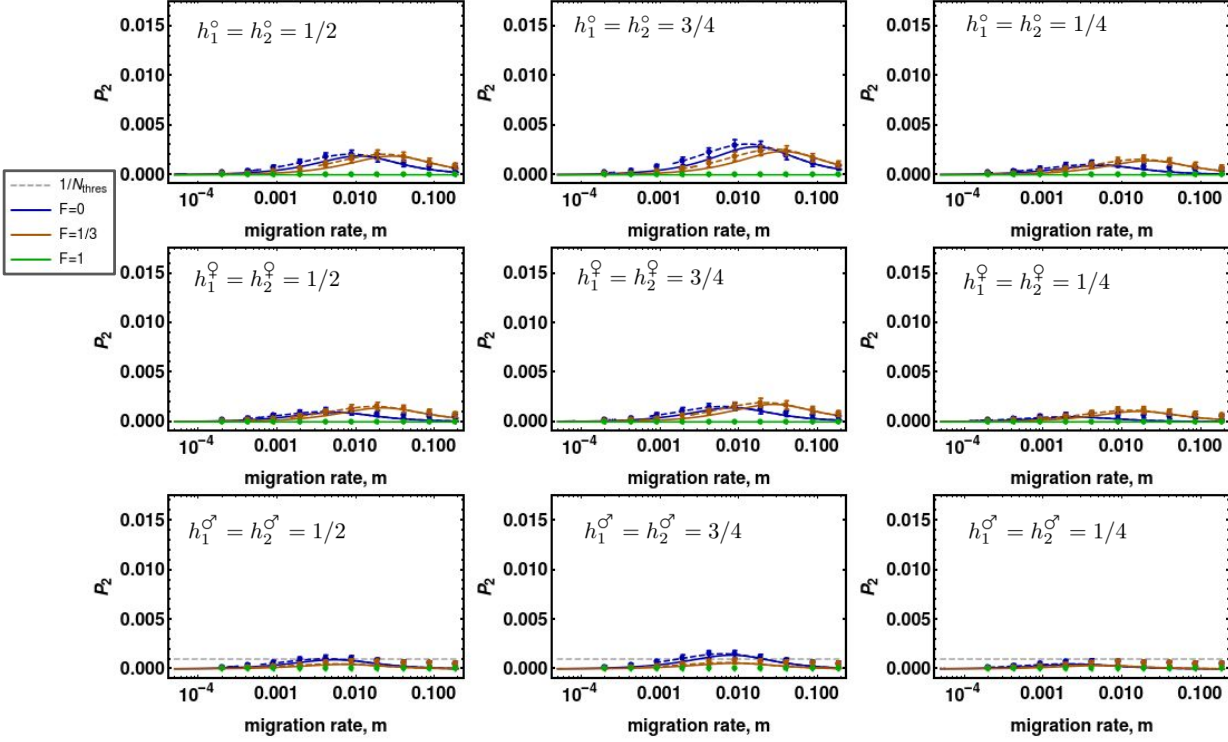


Figure A15: Same as above, but allele starts in the disfavored patch.

## F Procedure for computing $\beta$ indicators

959

960 Suppose one wants to derive  $\beta$  indicator when selection operates on  $j^{\text{th}}$  fitness component  
 961 (where  $j \in \{V, F, M\}$ ) in  $k^{\text{th}}$  patch (where  $k \in \{1, 2\}$ ). Selfing affects the establishment  
 962 probability through three factors: the effective population size ( $\beta_k(\tilde{N})$ ), effective favored  
 963 and disfavored dominance ( $\beta_k(\tilde{h}_1)$  and  $\beta_k(\tilde{h}_2)$ ), and the effective migration rate ( $\beta_k(\tilde{M})$ ).  
 964 Starting from equations (A16) and (A17), expressions for  $\beta$  are obtained using the following  
 965 steps in a sequential manner:

- 966 •  $\beta_k(\tilde{N})$ : Parameterize migration according to eqns. 12a–12c, and then set  $F = 0$  to  
 967 exclude the effect via  $\tilde{M}$ . Let  $\tilde{s}_i^{\{j\}} = s_i^{\{j\}} \tilde{h}_i^{\{j\}}$  for  $i \in \{1, 2\}$ . Next, set  $F = 0$  to exclude  
 968 the effect on dominances. Take a derivative in  $F$  and evaluate at  $F = 0$ .
- 969 •  $\beta_k(\tilde{h}_1)$ : Parameterize migration according to eqns. 12a–12c, and then set  $F = 0$  to  
 970 exclude the effect via  $\tilde{M}$ . Set  $F = 0$  to exclude the effect via  $\tilde{N}$ . Let  $\tilde{s}_2^{\{j\}} = s_2^{\{j\}} \tilde{h}_2^{\{j\}}$   
 971 and then set  $F = 0$  to exclude the effect via  $\tilde{h}_2$ . Let  $\tilde{s}_1^{\{j\}} = s_1^{\{j\}} \tilde{h}_1^{\{j\}}$ . Take a derivative  
 972 in  $F$  and evaluate at  $F = 0$ .

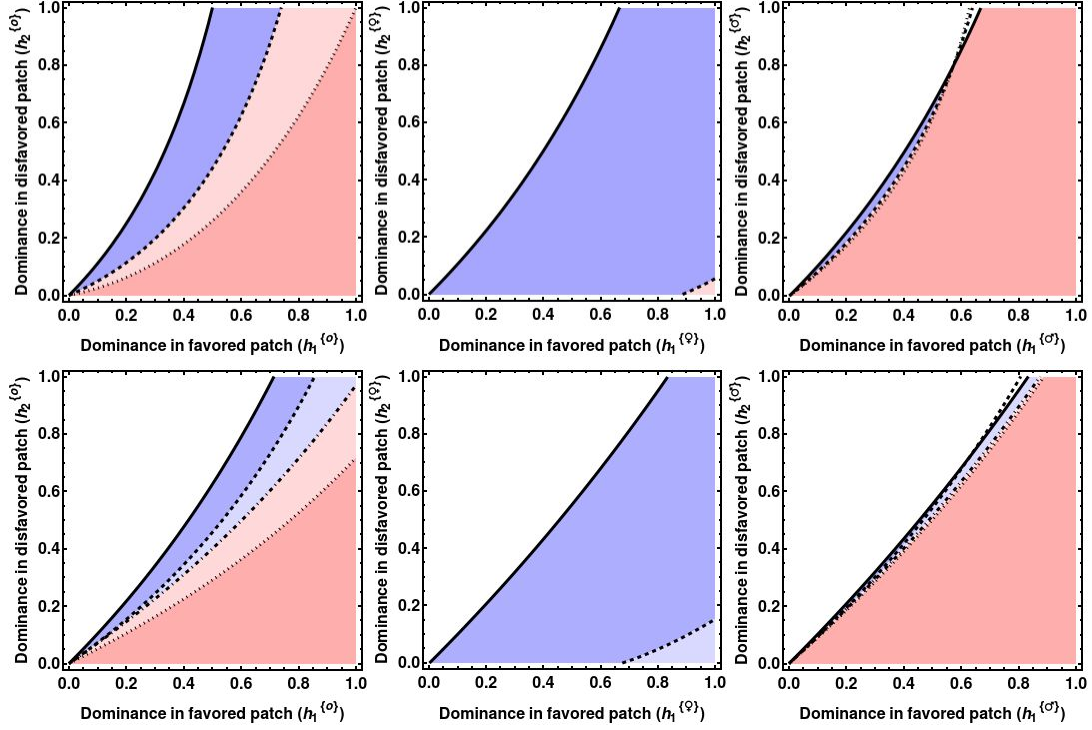


Figure A16: Zones of selfing's effect on the probability of establishment across viability (left panel), female fecundity (middle panel), and male fecundity (right panel), conditioning on allele appearing in the disfavored patch. Only seed migrates. Parameters are identical to those used in Figure 4. Upper row: scenarios with seed migration; Bottom row: scenarios with pollen migration.

- 973 •  $\beta_k(\tilde{h}_2)$ : Parameterize migration according to eqns. 12a–12c, and then set  $F = 0$  to  
 974 exclude the effect via  $\tilde{M}$ . Set  $F = 0$  to exclude the effect via  $\tilde{N}$ . Let  $\tilde{s}_1^{\{j\}} = s_1^{\{j\}}\tilde{h}_1^{\{j\}}$   
 975 and then set  $F = 0$  to exclude the effect via  $\tilde{h}_1$ . Let  $\tilde{s}_2^{\{j\}} = s_2^{\{j\}}\tilde{h}_2^{\{j\}}$ . Take a derivative  
 976 in  $F$  and evaluate at  $F = 0$ .
- 977 •  $\beta_k(\tilde{M})$ : Set  $F = 0$  to exclude the effect via  $\tilde{N}$ . Let  $\tilde{s}_i^{\{j\}} = s_i^{\{j\}}\tilde{h}_i^{\{j\}}$  for  $i \in \{1, 2\}$ . Then  
 978 set  $F = 0$  to exclude the effect via  $\tilde{h}_1$ , and  $\tilde{h}_2$ . Next, parameterize migration according  
 979 to eqns. 12a–12c. Take a derivative in  $F$  and evaluate at  $F = 0$ .

980 Parameters  $\tilde{h}_i^\circ$ ,  $\tilde{h}_i^\circ$ , and  $\tilde{h}_i^\sigma$  are given by equations (9a)–(9c). Whenever a  $\beta_i(x)$  is greater  
 981 than zero, a shift to selfing increases the probability that an allele becomes established,  
 982 conditioning on starting in patch  $i$ . Evaluating the derivative in point other than  $F = 0$  will  
 983 change the results quantitatively, but not qualitatively.

984 **G The consequences of shift to selfing on local adap-**  
 985 **tation**

986 Where fecundity selection is examined, the establishment probability equations were param-  
 987 eterized with  $s_i^{\ominus}$ , and  $h_i^{\ominus}$ , while selection on male fitness component was done by paramete-  
 988 rizing with  $s_i^{\ominus}$ , and  $h_i^{\ominus}$ . Other than that, parameters were quantitatively identical to the  
 989 viability selection case. Migration rates were also kept constant across different scenarios,  
 990 and only the migration type has changed.

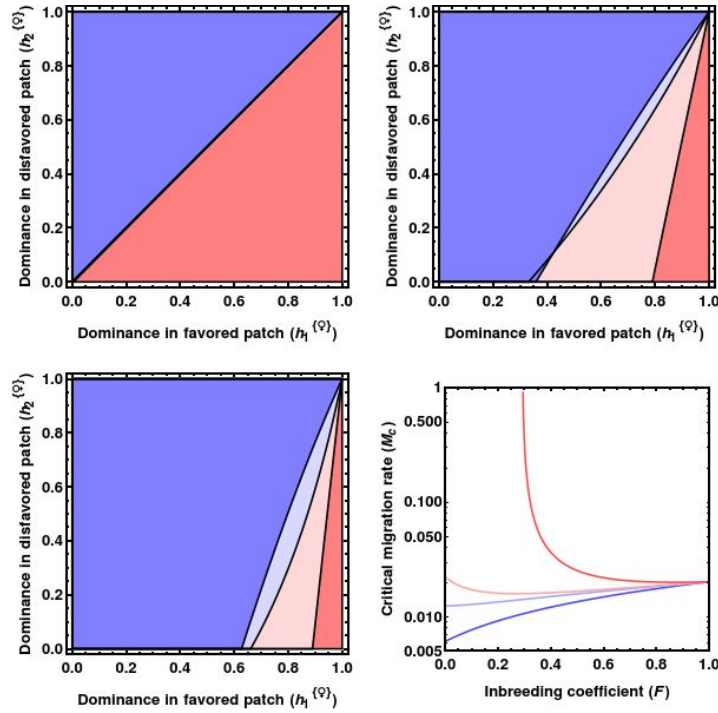


Figure A17: The consequences of a shift to selfing on establishment of local adaptation under female fecundity selection and seed dispersal. Color-coding and parameters as in the main text and parameters as in the main text.

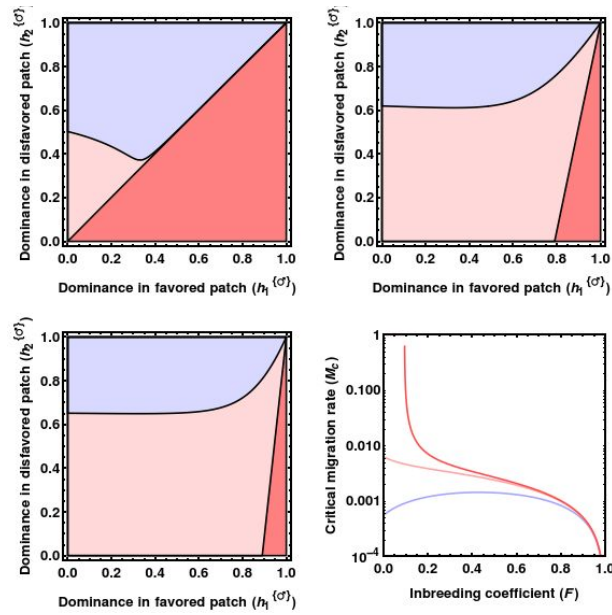


Figure A18: The consequences of a shift to selfing on establishment of local adaptation under male fecundity selection and seed dispersal. Color-coding and parameters as in the main text and parameters as in the main text.

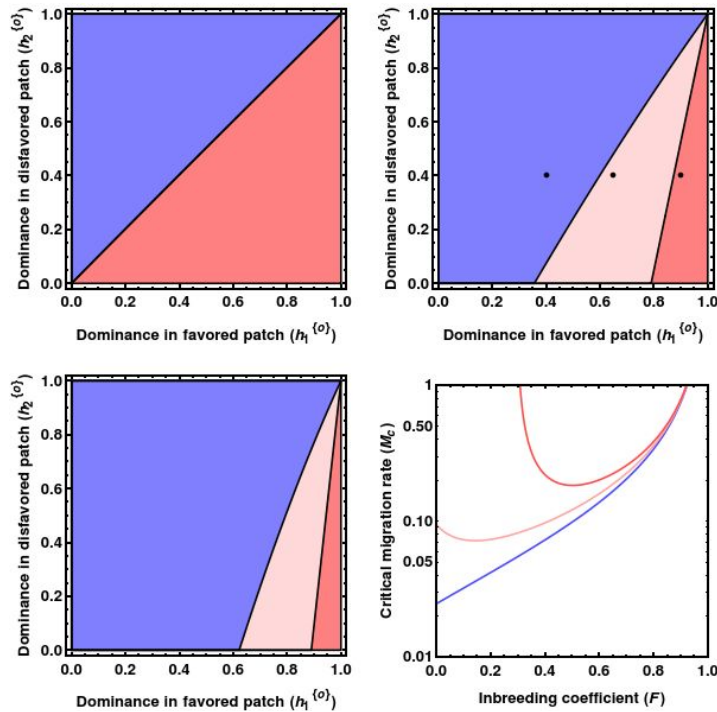


Figure A19: The consequences of a shift to selfing on establishment of local adaptation under viability selection and pollen dispersal.