

Divergence and Evolution of Assortative Mating in a Polygenic Trait Model of Speciation with Gene Flow

Himani Sachdeva and Nicholas H. Barton

Institute of Science and Technology Austria (IST Austria), Am Campus
1, Klosterneuburg A-3400, Austria

Abstract

Assortative mating is an important driver of speciation in populations with gene flow and is predicted to evolve under certain conditions in few-locus models. However, the evolution of assortment is less understood for mating based on quantitative traits, which are often characterized by high genetic variability and extensive linkage disequilibrium between trait loci. We explore this scenario for a two-deme model with migration, by considering a single polygenic trait subject to divergent viability selection across demes, as well as assortative mating and sexual selection within demes, and investigate how trait divergence is shaped by various evolutionary forces. Our analysis reveals the existence of sharp thresholds of assortment strength, at which divergence increases dramatically. We also study the evolution of assortment via invasion of modifiers of mate discrimination and show that the evolutionarily stable assortment strength has an intermediate value under a range of migration-selection parameters, even in diverged populations, due to subtle effects which depend sensitively on the extent of phenotypic variation within these populations. The evolutionary dynamics of the polygenic trait is studied using the hypergeometric and infinitesimal models. We further investigate the sensitivity of

23 our results to the assumptions of the hypergeometric model, using individual-based
24 simulations.

25 The extent of gene flow between diverging populations and its implications for specia-
26 tion have been extensively investigated in both empirical and theoretical studies (Dobzhan-
27 sky 1940; Felsenstein 1981; Neimiller et al. 1997; Servedio and Kirkpatrick 1997; Wang
28 et al. 1997; Kirkpatrick 2000; Servedio 2000; van Doorn et al. 2000; Servedio and
29 Noor 2003; Coyne and Orr 2004; Gavrilets 2004; Hey 2006; Bolnick and Fitzpatrick
30 2007; Servedio 2011). A typical scenario involves divergence of two populations in al-
31 lopatry followed by secondary contact; selection against the resultant, possibly inviable,
32 hybrids can spur the reinforcement of prezygotic mating preferences, leading to complete
33 reproductive isolation and the emergence of ‘good species’ (Dobzhansky 1940; Serve-
34 dio and Noor 2003; Coyne and Orr 2004). Alternatively, divergence may occur in the
35 face of continuous gene flow even without an initial allopatric phase, if populations face
36 divergent selection due to ecological heterogeneity, competition and/or selective mating
37 (Bolnick and Fitzpatrick 2007). Prezygotic isolation is in this case both a driver of initial
38 divergence, and may itself be selected for or against as divergence increases.

39 In this paper, we focus on the latter case of parapatric populations with ongoing migra-
40 tion and ask: How do the competing forces of divergent selection and homogenizing gene
41 flow shape the evolution of prezygotic isolation and resultant divergence levels between
42 populations, when the traits under natural and sexual selection are highly polygenic?

43 In considering this question, it is useful to distinguish between two scenarios. Rein-
44 forcement may occur via the divergence of an assortment trait between populations, on
45 a timescale for which the strength of assortment is assumed fixed. Here, strength of
46 assortment refers to female ‘choosiness’ or the the extent to which females discriminate
47 between similar and dissimilar males. Alternatively, if there is heritable variation in the
48 degree of mate discrimination, as observed in several natural populations (Butlin 1993;

49 Ortiz-Barrientos and Noor 2005), then the assortment strength may itself evolve due to
50 direct or indirect selection. Reinforcement, in this case, involves *co-evolution* of the degree
51 of discrimination with the assortment traits on which it acts. These two views of rein-
52 forcement are complementary—the ‘established assortment’ model is useful in surveying
53 possible outcomes, while the ‘evolving assortment’ model elucidates whether outcomes
54 are evolutionary attractors, i.e., if long-term evolution over a higher-dimensional trait
55 space is likely to converge to them.

56 Both established and evolving assortment scenarios have been studied extensively using
57 Adaptive Dynamics (AD) and in models with one or two (or very few) loci under
58 disruptive selection (Felsenstein 1981; Kelly and Noor 1996; Servedio 2000; Matessi et
59 al. 2001; Pennings et al. 2008; Ripa 2009; Servedio 2011; Rettelbach et al. 2013).
60 These models are often chosen for their analytical tractability and can disentangle the
61 roles of different evolutionary processes during divergence in a simple setting. However,
62 many putative assortment traits, such as body size, beak length or display traits such
63 as colorful plumage may be highly polygenic and differ from oligogenic traits in two key
64 ways.

65 First, polygenic traits harbour substantial genetic variation in spite of being under sta-
66 bilizing selection (Barton and Keightley 2002) and typically exhibit a continuous range
67 of phenotypes. Frequency-dependent selection acts on this continuous variation very dif-
68 ferently from how it acts on few-locus traits, with divergence in the latter case sometimes
69 arising only as an artefact of the small number of accessible phenotypes (Polechová and
70 Barton 2005). Moreover, as we argue in this paper, even within diverged populations,
71 the extent of genetic variation and the frequency of atypical, ‘outlier’ phenotypes can
72 qualitatively impact whether these populations evolve complete reproductive isolation.

73 A second distinctive feature of polygenic traits is that the response of their distribu-
74 tion to selection is driven primarily by changes in associations or linkage disequilibria

75 (LD) between loci (Bulmer 1980). Assortative mating, especially, generates strong LD
76 (Felsenstein 1981; Barton and de Cara 2009), which inflates trait variance. For highly
77 polygenic traits, this LD-driven increase in variance can be stronger and more rapid than
78 loss of variation due to sexual selection (Kirkpatrick and Nuismer 2004; Bürger and
79 Schneider 2006). In fact, we find that the buildup of LD across loci triggers synergistic
80 changes in allele frequencies, resulting in a steep increase in divergence between the two
81 populations beyond a threshold assortment strength, akin to a speciation event.

82 To clarify shared features as well as key differences between assortment based on highly
83 polygenic traits versus assortment based on similarity at a few loci, we consider a well-
84 studied ‘magic trait’ scenario of speciation (Gavrilets 2004; Servedio et al 2011), wherein
85 two initially identical populations with ongoing gene flow express a continuous trait that
86 is subject to divergent viability selection across demes and is also the basis of assorta-
87 tive mating by female choice, which generates sexual selection on males. We focus on
88 the following questions: First, are there general conditions on the strengths of selec-
89 tion, migration and assortment for which polygenic traits diverge? In particular, what
90 role does assortative mating play in generating divergence? Second, when assortment
91 strength can evolve through the invasion of modifiers of mate discrimination, is there
92 an evolutionarily stable (ES) level of assortment? Is it sufficient to produce divergence
93 and eventually complete reproductive isolation? Third, under what conditions is genetic
94 variation maintained within demes, and how does the extent of variation affect modifier
95 evolution?

96 In considering the last question, it is important to distinguish between genetic (and
97 phenotypic) variation that exists in populations before the onset of divergent selection
98 (for instance due to migration or frequency-dependent or fluctuating selection in the past)
99 and variation that is dynamically maintained during divergence by de novo mutation or
100 gene flow between demes. Here, we take the initial phenotypic variance as a parameter,
101 but assume that it is not so low as to inhibit a response to selection. In principle,

102 initial variation may be depleted by stabilizing selection within each deme, but for highly
103 polygenic traits, this occurs more slowly than the changes in trait mean involved in
104 divergence. Further, even for traits determined by a moderate number of loci, variation
105 may be replenished by the influx of genetically dissimilar individuals. In order to clarify
106 when stable variation is maintained simply by gene flow, we mostly study the model with
107 no mutation. The effects of mutation are discussed in Appendix S6.

108 It is also useful to contextualize the model within the main themes emerging from
109 other theoretical models of prezygotic isolation and divergence (Kirkpatrick and Ravigné
110 2002). A crucial determinant of prezygotic isolation is whether the isolating mechanism
111 has a one-allele or two-allele basis (Felsenstein 1981). One-allele mechanisms, e.g., those
112 involving the spread of modifiers that increase habitat preference or mate discrimination,
113 are not hindered by migration or recombination, as they require substitution of the same
114 allele in both populations. By contrast, two-allele mechanisms, e.g., those involving
115 divergence of sexual display or assortment traits, require different alleles to be maintained
116 in incipient species (possibly at multiple loci), and are unlikely to evolve with high gene
117 flow. However, many plausible scenarios of reinforcement have both one and two-allele
118 components (Kirkpatrick and Ravigné 2002; Servedio and Noor 2003). For instance,
119 in our model, while isolation increases through substitution of modifiers that increase
120 mate discrimination (one-allele mechanism), it also requires the emergence of distinct
121 phenotypes in each deme (which has a two-allele basis, in that different alleles must be
122 maintained in the two populations at *several* trait loci). Moreover, migration influences
123 these two components in opposite directions, with high migration promoting hybridization
124 and making it harder for distinct phenotypes to emerge, while also increasing the tendency
125 of populations to evolve stronger discrimination that reduces hybridization. We explore
126 the dual role of migration in assortment evolution subsequently.

127 While populations can split purely due to strong assortment under certain special
128 conditions (Kondrashov and Shpak 1998), divergence is greatly facilitated if the as-

129 assortment trait is a ‘magic trait’ under direct disruptive selection (Servedio et al 2011),
130 which leads to a strong coupling between postzygotic and prezygotic isolation. Here,
131 we consider a magic assortment trait subject to both viability and sexual selection, and
132 demonstrate how these two components of direct selection can drive mate discrimina-
133 tion antagonistically—an effect also observed in few-locus models (Kelly and Noor 1996;
134 Matessi et al. 2001; Kirkpatrick and Nuismer 2004)

135 Even with direct selection on the assortment trait, selection on modifiers of assortment
136 may be indirect, e.g., when it is due to linkage disequilibrium between the modifier locus
137 and assortment trait loci. Direct selection on modifiers arises if choice is ‘costly’, for
138 instance, if more discriminating females suffer reduced mating opportunities. Such costs,
139 unless very minor, inhibit the evolution of mate discrimination (Kopp and Hermisson
140 2008; Otto et al. 2008) and are excluded from the model.

141 **The model**

142 Our model assumes two demes, in which haploid organisms undergo viability selec-
143 tion towards distinct optima in each deme, followed by migration between demes, and
144 finally assortative mating within each deme, driven by female preference for phenotypi-
145 cally similar males. We consider a scenario especially favourable to speciation, namely,
146 where the phenotype X , expressed in both sexes, simultaneously affects viability and
147 assortment. The phenotypic distributions in the two demes are assumed to be initially
148 identical and Gaussian (with mean X_0 and variance V_0), but evolve differently over time
149 due to divergent selection.

150 Generations are assumed to be discrete and non-overlapping. In each generation, pop-
151 ulations first undergo stabilizing selection with fitness $W_i(X)=\exp(-(X-\mu_i)^2/2V_{s_i})$ for
152 deme i . Density is regulated independently in each deme (‘soft selection’). Selection is
153 strongly divergent across the two demes when the difference between the optima μ_1 and

154 μ_2 is large, and the strength of stabilizing selection ($\sim 1/V_{s_1}, 1/V_{s_2}$) high. Selection is
155 followed by migration, where a fraction m_{ij} of the population of deme i is replaced by
156 migrants from deme j .

157 After migration, each population mates assortatively, followed by free recombination.
158 Assortment is based on female choice, with females preferentially mating with males
159 within a phenotypic range α . We consider a Gaussian preference function: the probability
160 of mating, given an encounter between a female with phenotype Y and a male with
161 phenotype Z is proportional to $\exp[-(Y - Z)^2/2\alpha^2]$, where α is the preference range of
162 the female, and $1/\alpha^2$ represents the degree of mate discrimination. If all females have
163 the same preference range, then $1/\alpha^2$ also represents the strength of assortative mating
164 in the population. Each female, irrespective of phenotype, has the same probability of
165 mating, but males with common phenotypes have a mating advantage, resulting in sexual
166 selection on males. Population census is carried out after mating and recombination to
167 obtain the distribution $P(X)$ of phenotypes.

168 The preference range α is determined by a single unlinked locus which is present in
169 both sexes, and can be inherited from either parent. However, strength of preference
170 is expressed only by females. In the established assortment case, a single allele with
171 preference range α is fixed at this locus, while in the evolving assortment scenario, alleles
172 with preference ranges $\alpha_1, \alpha_2, \dots$ segregate.

173 The magic trait X is determined additively by L autosomal, haploid, unlinked, diallelic
174 loci. Allelic effects at each locus are assumed to be identical, and chosen to ensure that
175 the typical segregation variance, i.e., the phenotypic variance of offspring of parents near
176 the center of the phenotypic range is 1 in the large L limit (Appendix S1); then, the
177 phenotypic range extends from $-\sqrt{2L}$ to $\sqrt{2L}$. Further, all phenotypic ‘distances’ such
178 as μ , $\sqrt{V_s}$, $\sqrt{V_0}$ and α are specified in units of this segregation variance, which can be
179 easily measured in populations.

180 We study the evolutionary dynamics of the trait X using the hypergeometric model
181 (HM) which assumes that all genotypes corresponding to any phenotype are represented
182 equally in the population (‘symmetry assumption’, details in Appendix S1), resulting in
183 all trait loci being polymorphic with allele frequencies equal to $1/2$ for $X=0$ (Kondrashov
184 1984; Barton 1992; Doebeli 1996). In the large L limit, our results are also recovered
185 from the infinitesimal model (Bulmer 1980), which makes the less restrictive assumption
186 of a constant, non-zero segregation variance (which nevertheless requires high genetic
187 polymorphism). The infinitesimal model encompasses a wider class of genetic architec-
188 tures and also describes traits determined by many loci of unequal effects with weak,
189 random epistasis (Barton et al. 2016).

190 A limitation of this approach is that small deviations from the symmetry assumption
191 can be amplified by stabilizing selection, resulting in much less genetic variation than
192 predicted by the HM (Barton and Shpak 2000). To test the stability of HM predictions,
193 we perform individual-based simulations of finite populations, for which deviations from
194 symmetry assumptions arise simply due to drift. To suppress other confounding effects of
195 genetic drift (for instance, in small populations, where it may swamp weak, indirect selec-
196 tion on modifiers), we simulate very large populations ($N \sim 10^6$), for which changes due to
197 the intrinsic instability of the HM are expected to dominate over stochastic fluctuations
198 due to drift.

199 Individual-based simulations are initialized by assigning to each individual in the two
200 demes a phenotype X drawn from a Gaussian distribution with mean X_0 and variance
201 V_0 , and a uniformly distributed genotype corresponding to this phenotype. Selection is
202 implemented by drawing N individuals in each deme (with replacement) from the previous
203 generation with probability proportional to their fitness. Subsequently, a fraction m_{12}
204 of individuals in deme 1 is replaced by randomly chosen individuals from deme 2 and
205 vice versa. Assortative mating is implemented by choosing a female at random, and then
206 iteratively drawing males (allowing for mating with probability equal to the Gaussian

207 preference function) until mating is successful. The offspring genotype has loci drawn
208 with equal probability from either parent. This procedure is repeated until N offspring
209 are created. Simulations are run for $2\text{-}5 \times 10^4$ generations to test for deviations from the
210 HM, which emerge slowly.

211 Most of our analysis deals with a highly symmetric scenario with reciprocal migration
212 between demes ($m_{12}=m_{21}=m$), and equally strong selection ($V_{s1}=V_{s2}=V_s$) on the two
213 populations towards optima with $\mu_1=-\mu_2=\mu$, which are symmetric about $X=0$. We
214 explore both the established and evolving assortment scenarios for the symmetric model
215 in detail, and then briefly comment on versions with asymmetric gene flow (see Appendix
216 S7).

217 **Established assortment**

218 With established assortment, every female has the same preference range α . Starting
219 with identical Gaussian phenotypic distributions with mean $X_0=0$ and low initial variance
220 ($\sqrt{V_0} \ll \mu$), the two populations evolve under divergent selection, migration and assortative
221 mating (see Appendix S1), until a long-term steady state is reached (see Appendix S2
222 for a discussion of alternative initial conditions). We first characterize how divergence
223 between populations in this long-term state depends on various evolutionary parameters
224 under HM assumptions.

225 In general, strongly divergent selection (low V_s , high μ) facilitates local adaptation
226 while strong migration (high m) tends to create homogenized, generalist populations.
227 In an *intermediate selection* regime, where these two opposing forces are comparable in
228 magnitude, assortment plays a crucial role. An increase in assortment strength, i.e., a
229 decrease in α , has two somewhat contrary effects—first, it reduces the mating success of
230 rare, outlier males by creating positive frequency-dependent sexual selection (especially
231 at intermediate α); second, it leads to a stronger correlation between male and female

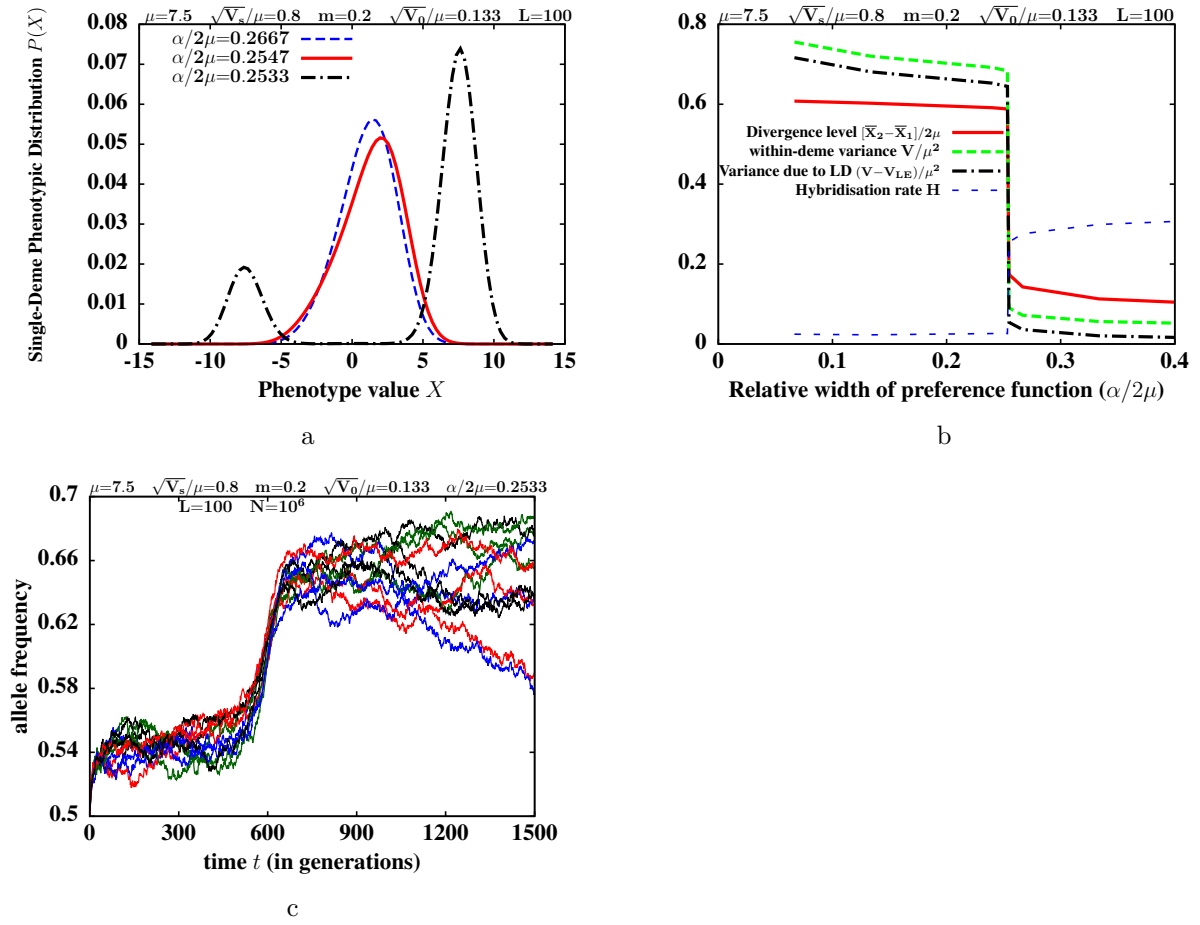


Figure 1: (a) Long-term phenotypic distributions $P(X)$ vs. X for a single deme with optimum at $+\mu$, for various values of α (HM predictions). The unimodal distribution becomes broader as α decreases and switches to a bimodal distribution at a threshold preference range $\alpha_c/2\mu$, here approximately 0.254. The bimodal distribution has clearly differentiated peaks at $+\mu$ and $-\mu$ corresponding respectively to diverged resident and migrant phenotypes, and may also have a small number of hybrids at $X=0$ (not visible on this plot), depending on the value of α_c . (b) Various statistics of the population (as predicted by the HM) vs. relative width ($\alpha/2\mu$) of preference function. Statistics include (i) trait divergence measured as the relative difference between the trait means (\bar{X}_1 and \bar{X}_2) of the two demes (ii) variance of the phenotypic distribution in a deme (iii) variance due to LD (iv) hybridization rate. All statistics exhibit a sharp change at $\alpha=\alpha_c$. (c) Allele frequency vs. time for randomly chosen loci in individual-based simulations for a population near the divergence threshold ($\alpha \sim \alpha_c$). Allele frequencies increase synergistically over a short period of time, resulting in a sharp increase in divergence and the emergence of bimodal $P(X)$.

232 phenotypes within each mating pair—thus when outlier males do mate, they mate pri-
233 marily with outlier females and produce outlier offspring. This increased correlation
234 builds up LD between loci, *increasing the variance* of the unimodal trait distribution as
235 well as its skew towards the optimum (fig. 1a), in spite of sexual selection against outlier
236 phenotypes. This, in turn, leads to a stronger response to natural selection; in fact at a
237 critical preference range α_c , the LD becomes high enough for selection to drive *synergistic*
238 changes in allele frequencies across all loci (fig. 1c), causing the mean phenotype to shift
239 towards the selection optimum of the deme. Thus populations with $\alpha < \alpha_c$ have clearly
240 distinguishable resident and migrant phenotypes in each deme (represented by the two
241 peaks of the bimodal distribution in fig. 1a), along with a small fraction of hybrids.

242 The qualitative change in the phenotypic distribution at $\alpha = \alpha_c$ manifests itself as a
243 sharp increase in the divergence, defined as the difference between mean trait values in
244 the two demes, as well as in the trait variance within each deme (fig. 1b, also fig. 2a). The
245 divergence in figs. 1b and 2a is scaled by 2μ , which is the maximum divergence possible
246 (with zero migration), while the variance is scaled by μ^2 . The increase in variance at α_c
247 is due to an increase in the disequilibrium component of the variance (see fig. 1b), which
248 is just the pairwise LD summed over all pairs of loci. Elevated genome-wide LD and
249 variance simply reflect the coexistence of diverged residents and migrants within each
250 deme. Note also the corresponding drop in the rate of hybridization H at α_c , where H
251 is defined as the fraction of individuals in a deme produced by resident-migrant pairings
252 in each generation ($H = 2m(1 - m)$ in a randomly mating population).

253 In the following, we refer to populations with bimodal phenotypic distributions within
254 demes (having $\alpha < \alpha_c$) as diverged, and populations with unimodal distributions ($\alpha > \alpha_c$) as
255 hybridized (even though trait means may differ slightly in the latter state). Some degree
256 of gene flow can persist between diverged populations, as evinced by the small number
257 of hybrids in populations with $\alpha < \alpha_c$. Thus divergence, as defined here, represents an
258 important step in the speciation process, but does not necessarily complete reproductive

259 isolation. We refer to α_c , the preference range at which phenotypic distributions become
260 bimodal, as the preference threshold for divergence, and typically specify it as the ratio
261 $\alpha_c/2\mu$, thus indicating how narrow preferences must be relative to the phenotypic distance
262 between selection optima, for divergence to occur. Alternatively, we can specify the
263 threshold assortment strength $(\alpha_c/2\mu)^{-2}$.

264 Note that α_c decreases as selection becomes weaker or migration stronger (fig. 2c),
265 pointing towards the high levels of assortment required for divergence under these condi-
266 tions. We also develop a Gaussian approximation for the infinitesimal model (Appendix
267 S4) and find that this predicts the divergence threshold very well, especially for moderate
268 selection (fig. 2c).

269 In addition to selection and migration parameters, α_c also depends on the initial phe-
270 notypic variance V_0 of the populations (fig. 2d). With no selection or assortment, V_0
271 rapidly equilibrates to twice the segregation variance, but diversifying selection, assorta-
272 tive mating and/or a history of secondary contact can cause it to differ significantly
273 from this value; we thus treat V_0 as a parameter which encapsulates population history.
274 First consider populations with mean phenotype $X \sim 0$ and $\sqrt{V_0}/\mu \ll 1$. Such populations
275 are unlikely to include the optimal phenotype at $t=0$, and divergence, if it occurs, is
276 preceded by a gradual buildup of variance via strong assortative mating. On the other
277 hand, for $\sqrt{V_0}/\mu \gg 1$, optimal phenotypes are present in the populations at the outset,
278 irrespective of assortment level (as would be natural for populations that have just come
279 into secondary contact after an allopatric phase). This results in an initial divergence,
280 which is eroded subsequently due to continued migration and hybridization if assorta-
281 tive mating is weak. Crucially, stronger assortment is required to create divergence in
282 unimodal populations than to prevent hybridization among diverged populations, as also
283 noted in Kondrashov and Shpak (1998). Thus for assortment strengths which lie within
284 the rising part of the α_c vs. $\sqrt{V_0}/\mu$ curve (fig. 2d), divergence is possible if V_0 is high
285 and optimal phenotypes already present (i.e., for V_0 to the right of the α_c vs. $\sqrt{V_0}/\mu$

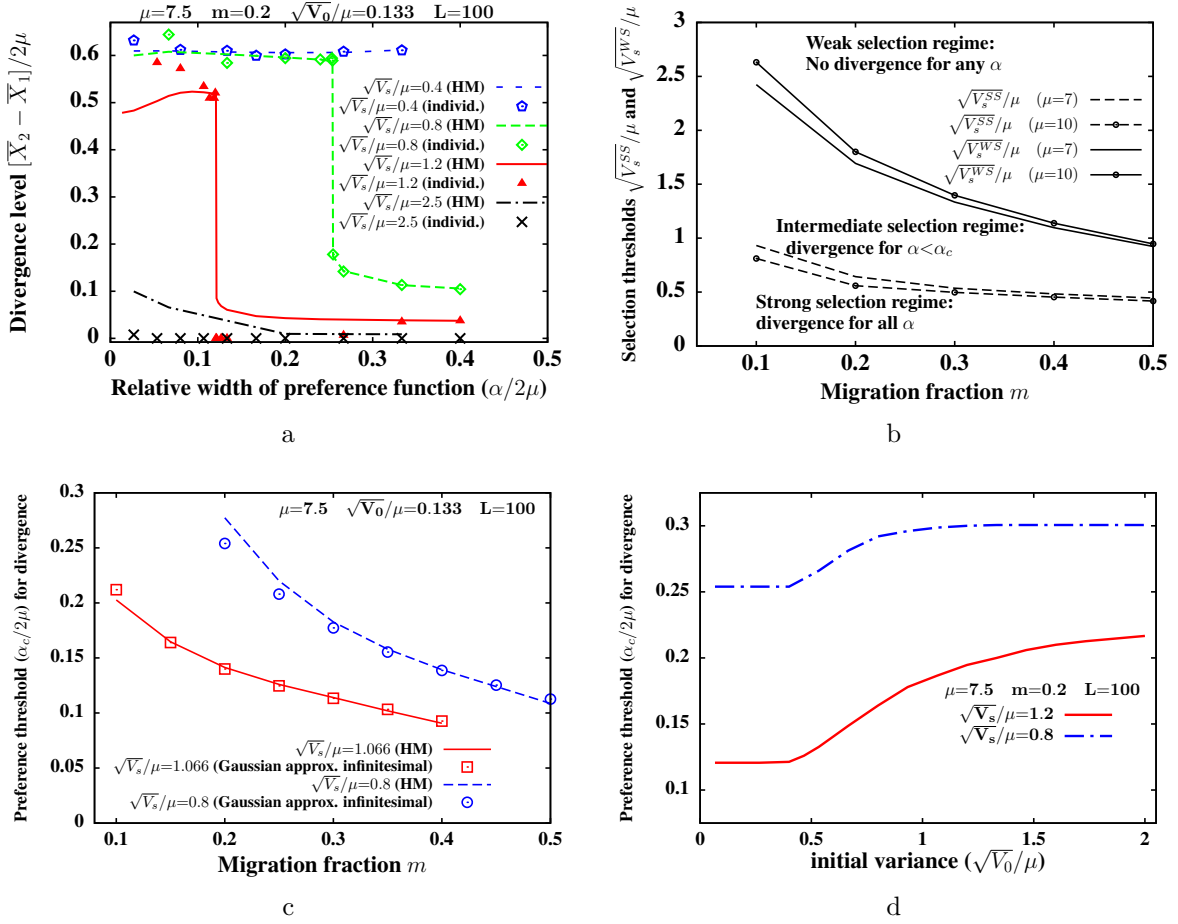


Figure 2: (a) Divergence measured as the difference between the trait means of the two demes vs. relative width $(\alpha/2\mu)$ of preference function, for various V_s . Lines represent HM predictions and points are from individual-based simulations with $N=10^5$. For intermediate selection ($\sqrt{V_s}/\mu=0.8, 1.2$), trait divergence undergoes a sharp drop at a threshold α_c , signifying a switch from a bimodal to a unimodal $P(X)$. For strong selection ($\sqrt{V_s}/\mu=0.4$) and weak selection ($\sqrt{V_s}/\mu=2.5$), no sharp transitions occur and populations remain diverged and hybridized respectively at all values of α . HM correctly predicts the qualitative state (diverged vs. hybridized) and α_c in individual-based simulations, but not the exact divergence levels. (b) Selection-migration parameters corresponding to the three selection regimes (HM predictions): Populations are always bimodal (or trimodal) for strong selection (for V_s below the V_s^{SS} threshold) and always unimodal for weak selection (for $V_s > V_s^{WS}$), irrespective of α . For $V_s^{SS} < V_s < V_s^{WS}$, phenotypic distributions shift from unimodal to bimodal at $\alpha=\alpha_c$. (c) Preference threshold $\alpha_c/2\mu$ vs. migration rate m : Predictions for α_c from a Gaussian approximation of the infinitesimal model (lines) are in good agreement with HM predictions (points). (d) Preference threshold $\alpha_c/2\mu$ vs. the initial phenotypic variance V_0 of populations (HM predictions): Stronger assortment is required for divergence (lower α_c) when initial variance is low.

286 curve), but cannot be induced in low- V_0 populations (V_0 to the left of the curve), leading
 287 to a bistability, wherein the initial variance determines whether divergence is possible in
 288 the long run (see also Appendix S4). This bistability has important consequences for
 289 modifier evolution, as discussed later.

290 In contrast to the intermediate selection regime, divergence levels change very little
 291 with assortment strength when selection is much stronger or weaker than migration (fig.
 292 2a). In the *strong selection regime*, i.e., for V_s smaller than the threshold V_s^{SS} in fig.
 293 2b, the distribution in each deme is bimodal (or trimodal if m is large and hybrids
 294 numerous enough), and divergence correspondingly high even with random mating ($\alpha \rightarrow$
 295 ∞). At the other extreme, in the *weak selection regime*, i.e., for V_s larger than the weak
 296 selection threshold V_s^{WS} in fig. 2b, populations remain hybridized and exhibit a unimodal
 297 distribution even for maximum assortment ($\alpha \rightarrow 0$).

298 The existence of a weak selection threshold V_s^{WS} is explained by considering the lifetime
 299 fitness of different phenotypes in the $\alpha \rightarrow 0$ limit. Neglecting the effect of segregation
 300 (due to mating between genotypically distinct but phenotypically identical individuals),
 301 the relative change in frequency of phenotype X in deme 1 in one generation is $\sim (1 -$
 302 $m)W_1(X)/\overline{W}_1 + mW_2(X)/\overline{W}_2$, where \overline{W}_1 and \overline{W}_2 denote the mean fitness in the two
 303 demes. For $m=0.5$ and $\overline{W}_1 \sim \overline{W}_2$ (as expected for a hybridized population), this relative
 304 change is maximum for the generalist phenotype $X=0$ when selection is weak ($\sqrt{V_s} > \mu$),
 305 but is (locally) minimum at $X=0$ and maximum near the optima $\pm\mu$ for stronger selection
 306 ($\sqrt{V_s} < \mu$). Thus, weak stabilizing selection within demes effectively generates stabilizing
 307 selection on the whole population around the generalist phenotype. The threshold V_s^{WS}
 308 for $m < 0.5$ can be derived similarly, see also van Doorn et al. (2000). Even in the weak
 309 selection regime, the HM does predict a modest assortment-driven increase in divergence
 310 ($\sqrt{V_s}/\mu = 2.5$ curve in fig. 2a) due to a shift of the distributions towards extreme or edge
 311 phenotypes. However, finite populations cannot sustain these shifts and congeal into a
 312 few genotypes near $X=0$, ultimately resulting in zero divergence (illustrated in detail in

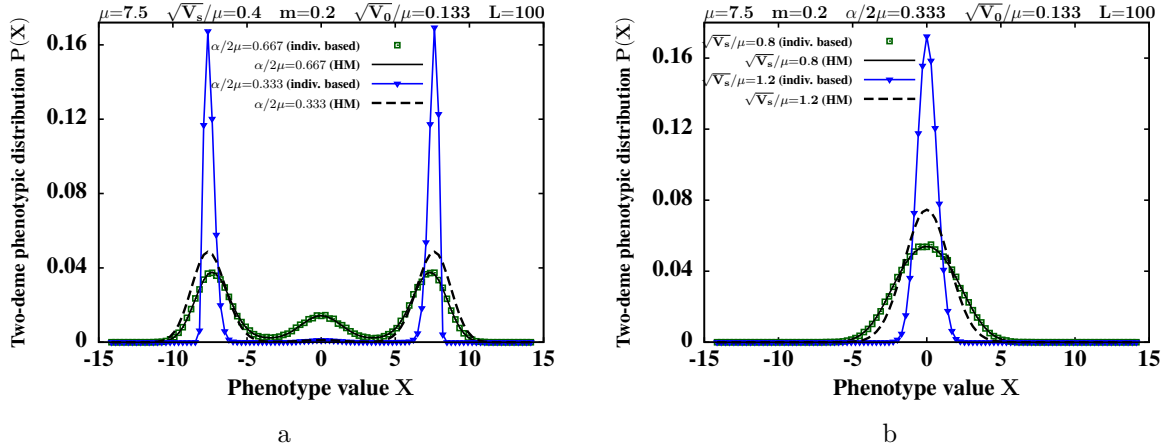


Figure 3: Long-term phenotypic distribution $P(X)$ of the whole population comprising both demes from individual-based simulations (for $N=10^5$) vs. HM predictions, $t=50000$ generations after onset of divergent selection. (a) HM accurately predicts $P(X)$ for diverged populations in individual-based simulations if there is significant hybridization between them (close match between squares and the solid line for trimodal distribution with numerous hybrids), but not if hybridization is low (triangles vs. dashed line). (b) HM predictions are accurate for hybridized populations when they are subject to strong divergent selection (squares), but not if selection on populations is weak (triangles vs. dashed line). The HM predicts the bimodal/trimodal (fig. 2(a)) or unimodal (fig. 2(b)) character of the phenotypic distributions accurately.

313 Appendix S5 for very weak selection).

314 Thus, a scan across parameter space reveals three qualitatively different scenarios: the
 315 strong selection scenario with purely ecological divergence, the intermediate selection
 316 scenario with assortment-dependent divergence and the weak selection scenario with no
 317 divergence. Figure 2b depicts typical (V_s, m) combinations corresponding to the three
 318 scenarios, by plotting the strong selection V_s^{SS} and weak selection V_s^{WS} thresholds against
 319 migration rate m .

320 **Stability of hypergeometric model (HM) predictions.** While the HM correctly
 321 predicts the qualitative state (diverged vs. hybridized) of finite populations, it may not
 322 always predict the exact divergence level (fig. 2a). This discrepancy, when it appears,
 323 is due to the fact that the high genetic polymorphism, implicit in the HM, is disfavored
 324 by stabilizing selection (Barton and Shpak 2000). With migration, however, each deme

325 can also replenish genetic variation by mating between phenotypically distinct residents
326 and migrants, raising the possibility that the HM is more robust in a two-deme than in
327 a one-deme scenario.

328 To test this, we compare the long-term phenotypic distributions $P(X)$ of large pop-
329 ulations ($N=10^5$ - 10^6) from individual-based simulations with the corresponding HM
330 predictions (fig. 3). This reveals certain general trends. First, the HM correctly pre-
331 dicts $P(X)$ for finite, initially polymorphic populations, whether diverged (bimodal) or
332 hybridized (unimodal) over short timescales (results not shown). Second, for diverged
333 populations, long-term stability of the HM prediction for $P(X)$ depends on the extent
334 of gene flow: populations with minimal gene flow between migrant and resident pheno-
335 types (as indicated by rare hybrids) eventually lose phenotypic variation around the two
336 optima due to stabilizing selection, while diverged populations with significant hybridiza-
337 tion maintain variation and have distributions that are accurately predicted by the HM
338 (fig. 3a). Thus in the diverged state, strong assortment (low α), reduced migration (low
339 m) or strong selection (low V_s) tend to destabilize the HM prediction.

340 Finally, for hybridized populations, the unimodal distribution $P(X)$ agrees with the
341 HM for intermediate selection, but collapses into a single-phenotype distribution if V_s
342 is large (fig. 3b). As discussed above, this is due to the fact that weak stabilizing
343 selection within demes generates net stabilizing selection about $X \sim 0$. In fact, an increase
344 in assortment in the weak-selection regime can actually destabilize $P(X)$ by depleting
345 polymorphic loci, an effect which precludes the buildup of LD with assortment. Loss of
346 variation in the weak selection limit is also observed in single diploid-locus and oligogenic
347 models (van Doorn et al. 2000; Kirkpatrick and Nuismer 2004; Bürger and Schneider
348 2006; Pennings et al. 2008; Rettelbach et al. 2013).

349 The instabilities discussed above merely cause loss of variation around the peak(s)
350 of the phenotypic distributions, while typically maintaining their bimodal or unimodal

351 character (fig. 3, see Appendix S5 for some exceptions). Thus, the HM accurately predicts
352 the qualitative state of populations, as well as the threshold α_c for divergence (fig. 2a),
353 as long as populations have a sufficient number of polymorphic loci in the initial state.

354 **Evolving assortment**

355 We now consider the evolving assortment scenario, where a modifier (associated with
356 preference range α_{mod}) is segregating at a small frequency p_0 in an ancestral popula-
357 tion (preference range α_{anc}). Our goal is to understand how modifier frequency changes
358 in response to selective forces, identify conditions for invasion of assortment modifiers,
359 and use this to determine the evolutionarily stable (ES) assortment strength for various
360 migration-selection scenarios.

361 To study modifier evolution, we make two simplifying assumptions: that there is
362 no initial asymmetry between demes (modifier frequency in each deme is p_0 at $t=0$),
363 and that there is no initial disequilibrium between the modifier locus and the trait loci
364 ($P(X, \alpha_{mod})=p_0P(X)$ at $t=0$). We first examine HM predictions in detail, and then
365 compare these with individual-based simulations.

366 **Assortment ESS and its dependence on the evolutionary history of the ances-**
367 **tral population.** Consider an ancestral population with selection-migration param-
368 eters ($\mu=7.5$, $\sqrt{V_s}=0.8\mu$, $m=0.2$) for which divergence occurs at a threshold α_c that
369 depends on V_0 , the initial phenotypic variance of the population (*before* it equilibrated
370 at α_{anc}). For various values of α_{anc} , this population may equilibrate to qualitatively dif-
371 ferent (unimodal vs. bimodal) states depending on V_0 (see fig. 2d). Thus the fate of
372 an assortment modifier introduced in this population depends not only on the preference
373 ranges (α_{anc} and α_{mod}) of the ancestral and modifier alleles, but also on the evolution-
374 ary history of the ancestral population. To explore this dependence, we contrast how
375 modifiers evolve in ancestral populations with a history of low variance ($\sqrt{V_0}/\mu \ll 1$) with

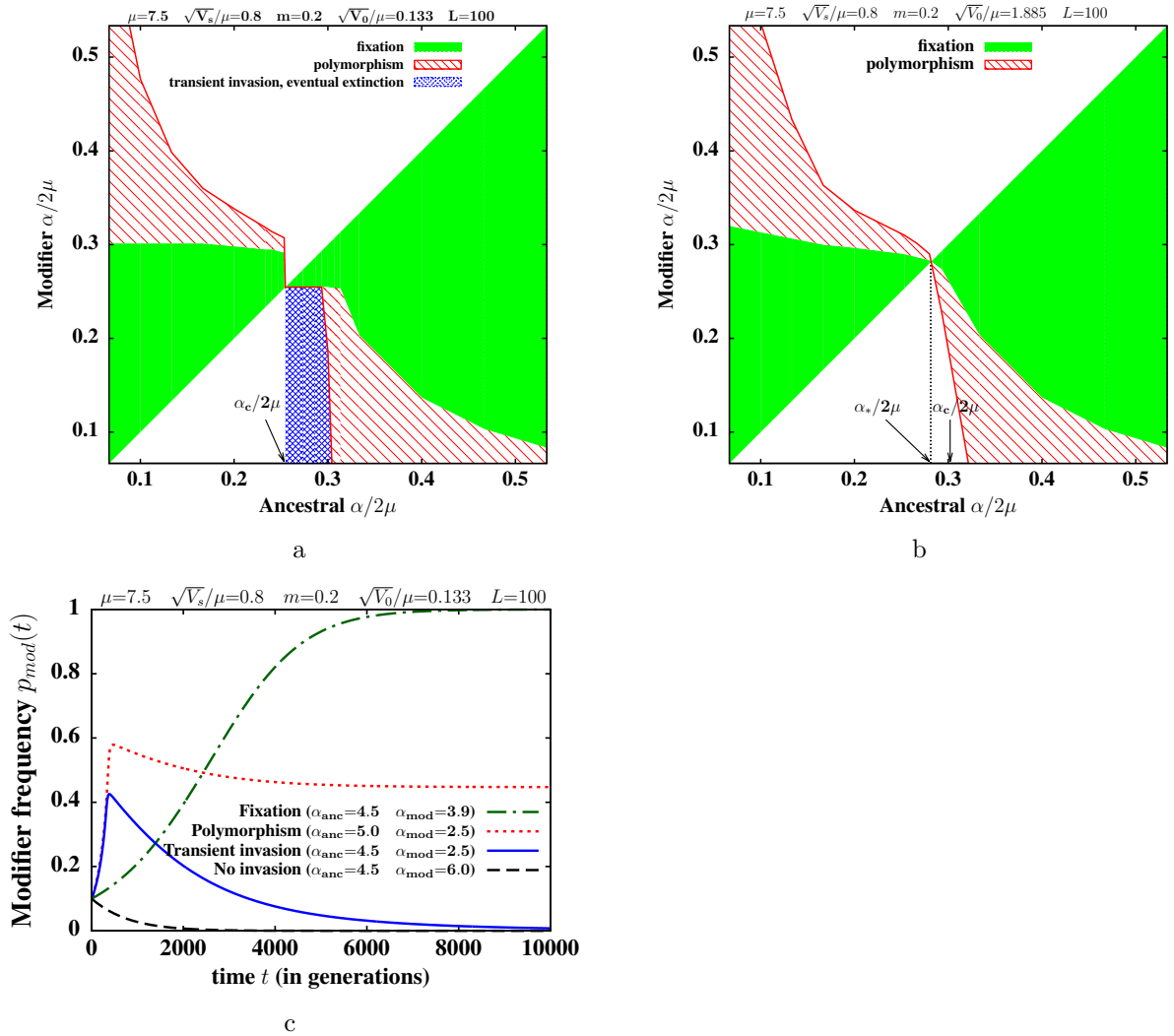


Figure 4: (a)-(b) Pairwise invasibility plots (as predicted by HM) for modifier with preference range α_{mod} introduced in an ancestral population in a long-term stationary state at some preference range α_{anc} , where the long-term state depends on the initial (pre-equilibration) variance V_0 of the ancestral population. PIPs are shown for the case where V_0 is (a) low ($\sqrt{V_0} \ll \mu$) and (b) high ($\sqrt{V_0} > \mu$). A modifier may invade and go to fixation (filled region of both PIPs) or invade but remain in a polymorphic equilibrium with the ancestral allele (diagonally shaded regions), or transiently invade an ancestral population with (historically) low V_0 , induce divergence, and then decline in frequency due to negative selection on the modifier in the diverged state (chequer shaded region in fig. 4a). The unusual topology of the PIP in fig. 4a is due to the fact that diverged populations ($\alpha_{anc} < \alpha_c$) can be invaded by any modifier with α_{mod} sufficient to maintain divergence. (c) Modifier frequency as a function of time t for four different parameter combinations (α_{anc} , α_{mod}) corresponding to the four scenarios: invasion and fixation, invasion and stable polymorphism at the preference locus, transient invasion and subsequent decline of modifier, and no invasion of modifier (HM predictions).

376 modifier evolution in ancestral populations which had high initial variance ($\sqrt{V_0}/\mu > 1$),
377 by plotting pairwise invasibility plots (PIPs) for $(\alpha_{anc}, \alpha_{mod})$ combinations in both cases
378 (figs. 4a and 4b). Since the PIPs assume a specific evolutionary history of the ancestral
379 population, they only illustrate the outcome of a single mutation event and cannot be
380 used to infer the outcome of a sequence of mutations at the modifier locus.

381 Consider first the PIP for the ancestral population with high V_0 before equilibration
382 (fig. 4b). This plot reveals the existence of an *intermediate* evolutionarily stable (ES)
383 preference range α_* , somewhat lower than α_c . Ancestral populations with $\alpha_{anc} > \alpha_c$ are
384 invaded by modifiers that reduce the preference range (filled region to the right of α_* in
385 fig. 4b), thus driving populations towards the divergence threshold. However, diverged
386 ancestral populations with $\alpha_{anc} \ll \alpha_c$, are invaded by modifiers that *increase* α (filled
387 region to the left of α_*) and introduce some hybridization between resident and migrant
388 phenotypes in each deme. Positive selection on such modifiers is due to their preferential
389 association with relatively fit hybrids at the expense of highly unfit migrant phenotypes
390 (details below).

391 The PIP also reveals various combinations of very high assortment and very low assort-
392 ment alleles which form protected polymorphisms (diagonally shaded region in fig. 4b; the
393 phase boundary of this region can be also derived from a simple geometric construction,
394 see Geritz et al. (1998)). High-assortment alleles ($\alpha_{mod} \ll \alpha_*$) have a fitness advantage
395 in weakly assortative, unimodal populations ($\alpha_{anc} \gg \alpha_*$), as they tend to associate with
396 phenotypic outliers close to the selection optimum. This causes the high-assortment allele
397 to invade and the populations to diverge. In the bimodal state, low-assortment alleles as-
398 sociate with more hybrids and fewer phenotypes near the migrant optimum, giving them
399 a selective advantage, which persists as long as they are too rare to collapse divergence.
400 Thus both high and low assortment alleles have a fitness advantage at low frequencies,
401 resulting in a polymorphic equilibrium between them, which allows populations to remain
402 diverged while maintaining significant hybridization. Such polymorphisms are, however,

403 ultimately unstable— alleles with α value intermediate to these may either supplant both
404 of the original alleles or form a new dimorphism with the low-assortment or the high-
405 assortment allele, eventually bringing the population close to α_* (Geritz et al. 1998).

406 We next examine the PIP (fig. 4a) for an ancestral population with a history of low vari-
407 ance V_0 (before equilibration) and a correspondingly stringent divergence threshold (low
408 α_c). Consider a situation with α_{anc} slightly higher than α_c , for which the ancestral pop-
409 ulations are hybridized, but could have diverged had the initial variance been sufficiently
410 high. As before, small-effect modifiers that reduce preference range ($\alpha_c < \alpha_{mod} < \alpha_{anc}$),
411 successfully invade and fix. However, when modifiers that reduce α below the divergence
412 threshold ($\alpha_{mod} \leq \alpha_c < \alpha_{anc}$) are introduced, very different dynamics ensue—the modifier
413 initially shows a sharp increase in frequency, causing the populations to diverge, but once
414 divergence sets in, there is a resurgence of the ancestral allele which pushes modifier
415 frequency back to zero (solid curve in fig. 4c), with diverged sub-populations now being
416 maintained at α_{anc} itself.

417 This sort of ‘resident strikes back’ dynamics also emerges in other evolutionary models
418 with multiple attractor states (Mylius and Diekmann 2001), and in the present model,
419 occurs when the ancestral population has intermediate α_{anc} (chequer-shaded region of
420 PIP in fig. 4a) for which divergence levels show a bistable dependence on V_0 (the initial
421 variance of the population before equilibration). The invading high-assortment modifier
422 merely increases population variance, causing the bistable system to switch to the alter-
423 native (bimodal) state in which the ancestral allele has an advantage due to its tendency
424 to associate with relatively fit hybrids (fig. 5c). The range of α_{anc} allowing for the tran-
425 sient invasion of high-assortment modifiers shrinks with increasing V_0 , and vanishes (fig.
426 4b) when V_0 is so high that there is no distinction between the assortment level required
427 to induce or maintain divergence in the ancestral population.

428 Figure 4c illustrates the variety of dynamical behaviours that can occur when assort-
429 ment modifiers invade an ancestral population with weak assortment ($\alpha_{anc} > \alpha_c$) and low
430 V_0 (before equilibration). Modifiers that further widen preference range are eliminated,
431 while modifiers that reduce α typically invade the population in the short run, but suffer a
432 long-term fate that depends sensitively on the values of α_{anc} and α_{mod} . Thus, $\alpha_{mod} < \alpha_{anc}$
433 modifiers either fix (corresponding to the filled region in fig. 4a), or decline towards zero
434 frequency (typically after invading transiently and inducing divergence, chequer-shaded
435 region in fig. 4a) or form a dimorphism with the ancestral allele (diagonally shaded
436 region).

437 **Assortment ESS under different migration-selection scenarios.** Having ana-
438 lyzed modifier evolution in detail for one set of parameters, we now ask: how does the
439 ES assortment level vary with migration and selection strengths, and is it always high
440 enough to induce divergence, while simultaneously being too low to complete reproduc-
441 tive isolation between diverged sub-populations (as in fig. 4b and accompanying text)?

442
443 Figure 5a shows how the scaled ES assortment strength, given by $(\alpha_*/2\mu)^{-2}$, and the
444 divergence threshold, $(\alpha_c/2\mu)^{-2}$, vary with 2μ , the phenotypic distance between the se-
445 lection optima, for two different migration rates. As selection across demes becomes
446 more divergent (2μ increases), the ES assortment level changes in a rather complex way,
447 revealing *four* qualitatively different regimes of assortment evolution. For small μ , ran-
448 dom mating is the ESS, with $(\alpha_*/2\mu)^{-2} \sim 0$. For intermediate μ , the population evolves
449 non-zero assortment, which is still lower than the threshold for divergence (dashed line),
450 so that the populations are hybridized at the ESS. For larger μ , i.e., in the falling part
451 of the $(\alpha_*/2\mu)^{-2}$ vs. 2μ plot in fig. 5a, assortment evolves exactly to the level required
452 to induce divergence but no further, while for very large μ , the ES assortment strength
453 is much higher than the divergence threshold $(\alpha_c/2\mu)^{-2}$, which is not even well-defined
454 when μ is very large and divergence is driven primarily by ecological selection.

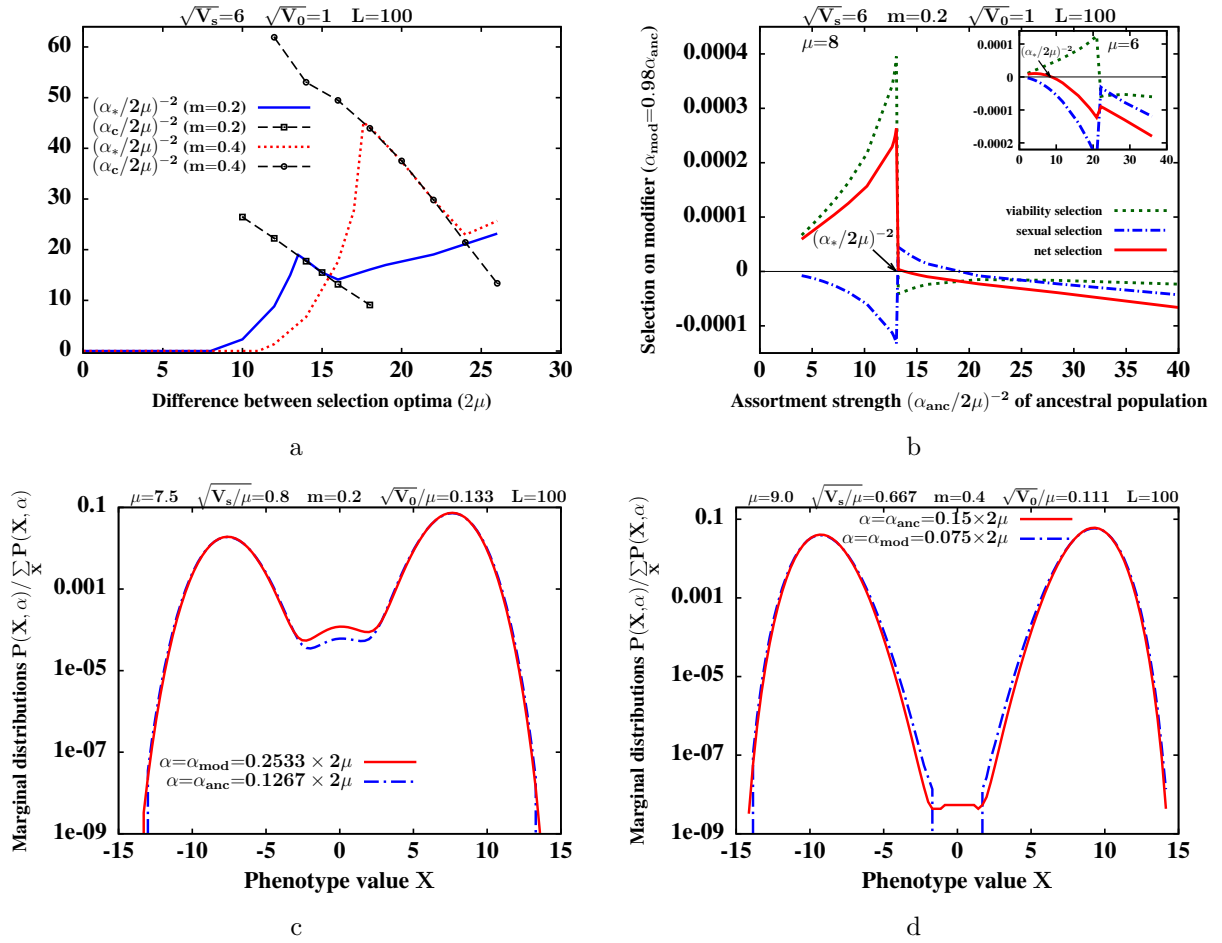


Figure 5: (a) The scaled ES assortment strength $(\alpha_*/2\mu)^{-2}$ and scaled threshold assortment strength for divergence, $(\alpha_c/2\mu)^{-2}$, vs. 2μ , the difference between the selection optima, for two values of m . (b) Coefficients of viability, sexual and net selection on a modifier that reduces preference range ($\alpha_{mod}=0.98\alpha_{anc}$) vs. assortment level $(\alpha_{anc}/2\mu)^{-2}$ within the ancestral population, for $\mu=8$ (main plot) and $\mu=6$ (inset). Selection coefficients are measured 2000 generations after introduction of modifier at initial frequency $p_0=0.1$. Net selection on modifier changes sign at $\alpha=\alpha_*$. For $\mu=6$ (inset), negative sexual selection on modifiers prevents assortment strength from evolving beyond the ES level, while for $\mu=8$, negative viability selection prevents a further increase. (c)-(d) Marginal phenotypic distributions $P_\alpha(X)=P(X, \alpha)/\sum_X P(X, \alpha)$ associated with an intermediate-assortment ($\alpha \sim \alpha_*$) and a high-assortment ($\alpha \sim 0.5\alpha_*$) allele segregating at the modifier locus in one of the demes (with optimum at $+\mu$), 100 generations after the high-assortment modifier is introduced at frequency 0.1 in the ancestral population close to its ESS. Lower marginal fitness of the high-assortment modifier in a bimodal population may be due to the lower frequency of relatively fit hybrids and higher frequency of unfit migrant phenotypes in the associated phenotypic distribution (fig. 5(c)), or due to the higher frequency of outliers in the resident and migrant sub-distributions associated with the modifier (fig. 5(d)). All plots show HM predictions.

455 To understand why assortment strength cannot evolve to a level sufficient for divergence
 456 when μ is small, while evolving up to or beyond the divergence threshold for larger
 457 μ , consider how assortment modifiers fare in hybridized (unimodal) populations. In
 458 such populations, modifiers that reduce α tend to associate with outlier or ‘extreme’
 459 phenotypes— this generates indirect selection on the modifier, if outliers either have high
 460 viability (are close to the selection optimum) or low mating success (have phenotypes
 461 far from those of most individuals). We distinguish between these two possibilities by
 462 separately tracking the viability and sexual selection/mating success components of the
 463 net (indirect) selection acting on the modifier (see Appendix S1). Note that for the
 464 symmetric model, indirect selection is the sole driver of modifier evolution (neglecting
 465 drift).

466 For small μ (corresponding to the weak selection regime in fig. 2b), outliers are not
 467 favoured by viability selection since they are less fit than the generalist ($X \sim 0$) phenotype.
 468 Moreover, outliers are also chosen for mating with a probability less than their frequency,
 469 as they are farther from the majority. Thus, in this regime both viability and sexual
 470 selection disfavour modifiers that increase assortment, resulting in $(\alpha_*/2\mu)^{-2} \sim 0$ at the
 471 ESS.

472 For intermediate or large μ (corresponding to the assortment-dependent divergence
 473 regime in fig. 2b), outlier phenotypes associated with the modifier are fitter than phe-
 474 notypes near $X \sim 0$, resulting in positive viability selection on the modifier. However, as
 475 long as the population is unimodal, the modifier still undergoes negative sexual selection
 476 due to the reduced mating success of outliers. Thus, natural and sexual selection drive
 477 modifier evolution in opposite directions, and the ensuing ES assortment level depends
 478 on the relative strength of the two, which depends on selection and migration parameters.
 479 To clarify this, we explicitly track, for two different selection parameters, how the coef-
 480 ficients of viability and sexual selection on a modifier of small effect ($\alpha_{mod} = (1 - \delta)\alpha_{anc}$,
 481 $\delta = 0.02$) vary with the assortment level, $(\alpha_{anc}/2\mu)^{-2}$, in the ancestral populations (fig.

482 5b).

483 For intermediate μ , outliers enjoy a rather modest fitness advantage over hybrids. Thus,
484 a modifier that increases discrimination in a randomly mating population experiences
485 weak viability selection, which is just slightly higher than the sexual selection acting
486 against it (inset, fig. 5b). In fact, as assortment builds up, i.e., with increasing $1/\alpha_{anc}^2$,
487 sexual selection against modifiers (that reduce α) becomes stronger than positive viability
488 selection, inhibiting a further increase in assortment towards the divergence threshold
489 (inset, fig. 5b). For larger μ , phenotypes close to the deme optimum are much fitter than
490 hybrids clustered around $X=0$ (*strong* selection for specialists over generalists). Thus,
491 viability selection on modifiers that lower α is much stronger, and prevails over negative
492 sexual selection (main plot, fig. 5b), driving assortment to a level sufficient for divergence.

493 Within diverged populations, assortment evolution is governed by somewhat different
494 effects. Close to α_c , when populations have just become bimodal, it is possible for highly
495 unfit phenotypes that migrate into the deme to produce relatively fit hybrid offspring
496 by mating disassortatively, i.e., with well-adapted residents. By contrast, modifiers with
497 stronger assortment ($\alpha_{mod} < \alpha_c$) allow for almost no hybridization between phenotypes
498 at the two optima and are hence associated with fewer hybrids and a higher fraction of
499 phenotypes near the immigrant optimum (fig. 5c). While this may give rise to a weak
500 sexual selection advantage for such modifiers (since migrant phenotypes have a mating
501 advantage with respect to hybrids in bimodal populations, being more numerous), it also
502 results in lower marginal fitness of the $\alpha < \alpha_c$ modifier. Thus high-assortment modifiers
503 undergo negative viability selection within bimodal populations (main plot, fig. 5b),
504 which prevents any further increase in assortment beyond the divergence threshold, at
505 least when μ is not extremely large (falling part of the curve in fig. 5a).

506 As μ increases further, the divergence threshold $(\alpha_c/2\mu)^{-2}$ decreases, resulting in a fall
507 in the ES assortment strength, $(\alpha_*/2\mu)^{-2}$, with μ (fig. 5a). For very large μ , populations

508 approach the strong selection regime in fig. 2b, where divergence is driven more by
509 ecological selection than assortment, and is thus accompanied by extensive hybridization
510 between resident and migrant phenotypes (note, for instance, the trimodal distribution
511 in fig. 3a). The high rate of hybridization pulls the means of the distributions around the
512 two optima closer to $X=0$, resulting in somewhat reduced local adaptation. Modifiers
513 that further increase assortment in such a population do associate more strongly with
514 unfit migrant phenotypes than with fitter hybrids, but they also tend to form associations
515 with better adapted (closer to optimal) phenotypes within the resident pool. Due to
516 these two opposing effects, such modifiers may actually undergo positive or very weakly
517 negative viability selection, causing assortment strength to evolve beyond the very low
518 level required to split the unimodal population (very large μ regime in fig. 5a).

519 Figure 5a also illustrates how migration influences the ES assortment strength. An
520 increase in m shifts the $(\alpha_*/2\mu)^{-2}$ vs. 2μ curve towards larger μ , while also increasing
521 the ES assortment level attained at large μ . This is explained by noting that the range of μ
522 for which populations exist in the weak selection regime and consequently have random
523 mating ESS, is larger for $m=0.4$ than for $m=0.2$ (fig. 2b). Even when assortment-
524 dependent divergence becomes possible, high-assortment modifiers are less effective in
525 generating increased phenotypic variance and skew towards the selection optimum in
526 populations with high m , because of the homogenizing effects of gene flow. Thus, weak
527 viability selection on modifiers is easily canceled out by negative sexual selection even
528 for fairly large μ , resulting in ES assortment levels that are insufficient for divergence.
529 Only when μ is quite large does the indirect fitness advantage of such modifiers become
530 strong enough to drive assortment to the threshold for divergence. Interestingly, for
531 high m , this assortment threshold can be so high as to produce complete reproductive
532 isolation (zero hybridization) between populations at the onset of divergence itself, in
533 contrast to the moderate hybridization observed at the ESS for lower migration rates.
534 A modifier that further increases assortment does not significantly reduce the (already

535 negligible) hybrid frequency. Instead, it tends to associate with outliers *within* resident
536 and migrant sub-populations in each deme (see fig. 5d). These outliers, especially within
537 the resident sub-population, are both less fit and less likely to be selected for mating than
538 phenotypes at the selection optimum, resulting in negative viability and sexual selection
539 against modifiers that increase assortment strength beyond the divergence threshold.

540 **Assortment evolution in individual-based simulations.** The preceding discussion
541 highlights the sensitivity of modifier evolution to genetic variation within demes. Finite
542 populations subject to stochastic fluctuations may, however, exhibit much less variation
543 than predicted by the HM (see fig. 3). How does loss of variation affect modifier dynamics,
544 and does the resultant ESS differ from the HM prediction?

545 Consider a situation with $\alpha_* \sim \alpha_c$ (as predicted by the HM). Figure 6 shows that the
546 HM accurately predicts modifier dynamics in individual-based simulations when the an-
547 cestral population in which the modifier appears is weakly assortative and hybridized,
548 i.e., $\alpha_{anc} > \alpha_{mod} > \alpha_* \simeq \alpha_c$ (fig. 6a), but not within diverged ancestral populations with
549 strong assortment, i.e., for $\alpha_{anc} < \alpha_{mod} < \alpha_*$ (fig. 6b). Diverged populations tend to lose
550 phenotypic variation and congeal around the selection optima (fig 3a); thus the effects
551 that drove modifier evolution under the HM (involving selection for a specific level of
552 hybridization or against outliers), no longer operate in the absence of variation. Conse-
553 quently, while weakly assortative populations in individual-based simulations can evolve
554 towards the divergence threshold $\alpha_c \sim \alpha_*$ by fixing small-effect modifiers that reduce α
555 (in agreement with the HM), once in the diverged state, modifiers that reduce α even
556 further, evolve neutrally and are not selected against (fig. 6b).

557 The analysis so far assumes no mutation. However, mutation contributes substan-
558 tially to polygenic variability (Barton and Keightley 2002). To test whether assortment
559 evolution in populations with stable genetic variation agrees qualitatively with HM pre-
560 dictions, we perform individual-based simulations allowing for recurrent mutation at the

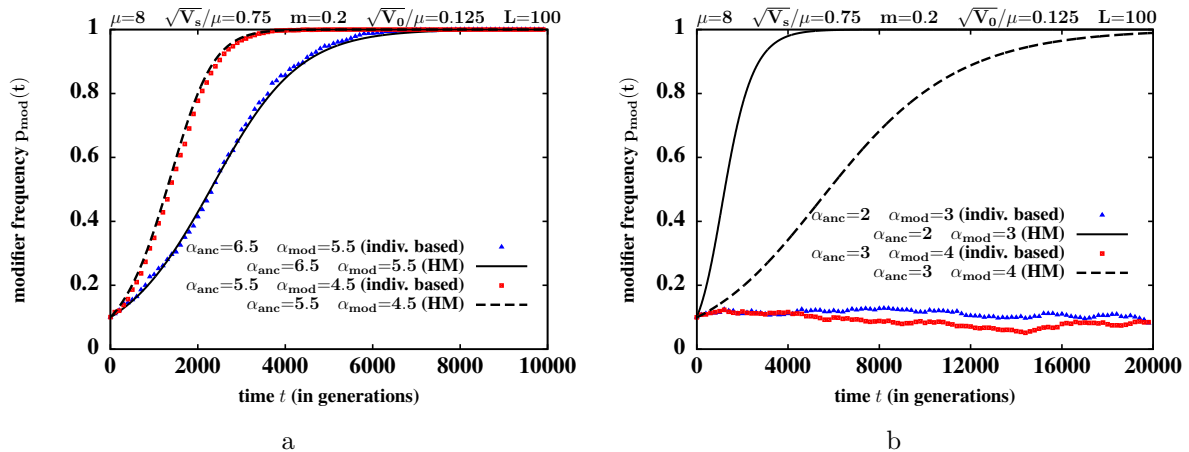


Figure 6: Modifier dynamics from individual-based simulations of populations with $N=10^6$ (points) vs. HM predictions (lines). Modifier frequency vs. time t is accurately predicted by HM for $\alpha_{\text{anc}} > \alpha_*$, i.e., when the ancestral population is hybridized (fig. 6a) but not for $\alpha_{\text{anc}} < \alpha_*$, i.e., when the ancestral population is diverged (fig. 6b). Modifiers that change α in finite populations with $\alpha_{\text{anc}} < \alpha_*$ undergo nearly neutral evolution (see data points in fig. 6b, obtained by averaging over 5 replicates), and not the positive selection predicted by the HM.

561 trait loci (Appendix S6). These show that large, finite populations with mutation have
 562 a well-defined ES level of assortment, as long as the mutation rate is not too low, with
 563 modifiers that take the population towards this ESS undergoing positive selection in both
 564 strongly assortative and weakly assortative populations (fig. S5, Appendix S6).

565 **Asymmetric models**

566 Several theoretical studies have argued that reinforcement is strongly affected by the
 567 directionality of gene flow, and occurs under more stringent conditions for continent-island
 568 migration than with reciprocal introgression (Servedio and Kirkpatrick 1997; Servedio
 569 2000). To test this, we study our model with $m_{12} \neq m_{21}$, and find that while asymmetric
 570 gene flow does make it more difficult for assortment-increasing modifiers to invade, it
 571 also facilitates higher divergence at any given assortment strength (see Appendix S7 and
 572 Discussion).

573 Discussion

574 Our results highlight a number of subtle effects that arise in simple magic-trait scenarios
575 of speciation when the trait is polygenic and can sustain high genetic variation. The core
576 question addressed in the paper is: How does the interplay of gene flow, natural and sexual
577 selection shape the evolution of such a trait along with the *degree* of mate discrimination
578 for the trait? Studying the co-evolution of mate discrimination with the assortment trait
579 on which it acts, sheds light on the evolutionary accessibility and stability of diverged
580 states, thus providing a window into long-term divergence and speciation.

581 Established assortment

582 A striking feature of the onset of divergence in the model is its highly non-linear charac-
583 ter, with small changes in the female preference range at α_c triggering a steep increase
584 in trait differences between the two demes (figs. 1b, 2a). The existence of sharp di-
585 vergence thresholds has important implications for reverse speciation (Seehausen 2006),
586 suggesting that the collapse of diverged species into a hybrid swarm due to reduced mate
587 discrimination may be an abrupt and unpredictable event, especially if the assortment
588 trait is polygenic.

589 Threshold effects involving a precipitous rise in divergence over time are common in
590 speciation models (Gavrilets 2004). The abrupt transition in this model at α_c has
591 the same underlying cause, namely, the synergistic changes in allele frequency that occur
592 when genome-wide LD builds up to a critical value. In fact, stronger assortment generates
593 higher LD (fig. 1b), thus inflating the variance of polygenic traits despite sexual selection
594 against phenotypic outliers. This effect contributes significantly to the divergence of
595 polygenic traits determined by unlinked or weakly linked loci. However, when trait loci
596 are tightly linked, assortative mating has a qualitatively different effect, with stronger
597 assortment reducing rather than increasing (the already high) LD and trait variance
598 (Kirkpatrick and Nuismer 2004; Bürger and Schneider 2006). Thus, extending the

599 model to include arbitrary linkage, as in de Cara et al. (2008), could provide a useful
600 perspective on the conclusions of this study.

601 Sharp transitions to a diverged state only occur when selection and migration are
602 competing evolutionary forces and divergence is assortment-dependent. As in the model
603 of van Doorn et al. (2000), which considers preference-trait mating in a similar two-deme
604 setting, this model also exhibits an ecological speciation (low V_s, m) and a weak selection
605 (high V_s, m) regime, where assortment has little effect and populations exhibit high or
606 low divergence, independently of α (fig. 2b). Weak selection thresholds emerge quite
607 generally in two-deme models, even with single-locus and AD assumptions (Brown and
608 Pavlovic 1992; Meszéna et al 1997), and reflect a switch in the topology of the fitness
609 landscape (from disruptive to stabilizing selection) near the generalist phenotype.

610 Unlike the AD framework, which is mutation-limited, both the HM and the infinitesimal
611 model assume high standing genetic variation— this allows for generation of new
612 phenotypes via recombination, and long-term phenotypic change under selection, far be-
613 yond the initial phenotypic range of the population. Thus, as long as assortment is strong,
614 divergence occurs even if populations have limited initial phenotypic variation ($\sqrt{V_0} \ll \mu$)
615 and do not include optimal phenotypes at the outset. Further, the long-term state of
616 populations is largely insensitive to where they start out in phenotypic space. For in-
617 stance, under strongly divergent selection (large μ) and low migration, sub-populations
618 (with a modest segregation variance) can always evolve towards their respective optima
619 via shifts in their phenotypic distributions, irrespective of whether they start out as gen-
620 eralist populations or being locally adapted to one of the optima or clustered around any
621 intermediate phenotype (Appendix S2). This contrasts with asexual populations in AD
622 models (Meszéna et al 1997), which undergo evolutionary branching under similar eco-
623 logical conditions only if they start out exactly at the generalist strategy (or are invaded
624 by mutants with very large phenotypic effects).

625 Both the HM and the more widely applicable infinitesimal model make qualitatively
626 similar predictions (see Appendix S3), in particular, that phenotypic variance always
627 increases with increasing assortment. However, in individual-based simulations with finite
628 L , increasing assortment may collapse variance by destabilizing polymorphisms. This ‘loss
629 of polymorphism’ regime has been highlighted in several oligogenic models (Kirkpatrick
630 and Nuismer 2004; Pennings et al. 2008; Rettelbach et al. 2013), and becomes
631 important in the present model under certain conditions (see fig. 3), e.g., for unimodal
632 populations subject to stabilizing or weakly divergent selection across demes (resulting
633 in reduced variation around the generalist phenotype), or for diverged populations which
634 hybridize weakly (and undergo loss of variation about the selection optima). However,
635 even under these conditions, as L increases, selection coefficients responsible for loss of
636 polymorphism at individual loci become very small, thus recovering the infinitesimal limit,
637 in which polymorphic variation is more or less stable over other relevant timescales, e.g.,
638 the time required for assortment modifiers to fix within a population. Moreover, even
639 for moderate values of L , mutation (with rate $U \sim 0.1$) can maintain enough polymorphic
640 variation for HM predictions to be valid over long timescales, at least in large populations
641 (Appendix S6).

642 **Evolving assortment**

643 In the evolving assortment scenario, we explore how the evolutionarily stable (ES)
644 assortment strength varies with selection and migration parameters, in order to determine
645 when divergence and complete reproductive isolation occur at the ESS.

646 We identify several qualitatively different regimes of assortment evolution. With weak
647 selection, divergence is not possible at any assortment strength, and randomly mating
648 populations are immune to invasion by modifiers that increase assortment. Such random
649 mating ESS also arise in oligogenic models (Pennings et al. 2008; Rettelbach et al.
650 2013) and simply reflect the absence of disruptive selection on the whole population.

651 Even when divergence is possible above a certain assortment threshold, the diverged
652 state need not be an evolutionary attractor, in that the ES assortment may be lower than
653 this threshold. This is typically the case when migration rates are high and/or selection
654 across demes only moderately divergent, so that the gain in fitness due to local adaptation
655 is offset by the loss in mating success that outliers suffer. Under these conditions, even
656 though assortative mating increases trait variance, the accompanying sexual selection
657 acts against assortment modifiers associated with the higher variance, thus constraining
658 evolution of strong assortment and preventing divergence.

659 This sort of antagonism between natural and sexual selection also plays a role in diploid,
660 single-locus models of competitive speciation, where assortment evolution stops at low
661 values despite disruptive selection, precisely because of positive sexual selection on het-
662 erozygotes when they are widespread (Matessi et al. 2001; Otto et al. 2008; Pennings
663 et al. 2008)—an effect similar to the one described above. Interestingly, this weak-
664 assortment state is a global attractor in the polygenic model (with even highly diverged
665 populations evolving towards it through the substitution of low-assortment modifiers, at
666 least for the HM) but not in one-locus models.

667 As selection across demes becomes more divergent, positive viability selection on mod-
668 ifiers becomes strong enough to drive assortment levels up to the divergence threshold,
669 so that populations are bimodal at the ESS. Significantly, there is effective stabilizing
670 selection on the strength of assortment even within bimodal populations, with high-
671 discrimination modifiers ($\alpha \ll \alpha_c \sim \alpha_*$) undergoing negative selection due to their associa-
672 tion with certain phenotypes, at least in the HM and infinitesimal model. We identify
673 two distinct explanations for this, which apply in different parameter ranges.

674 First, very high assortment depletes hybrids while simultaneously increasing the fre-
675 quency of ill-adapted migrants within each deme (fig. 5c). The resultant selection against
676 high assortment is essentially positive selection for moderate hybridization between di-

677 verged populations, which however must not be so high as to significantly reduce lo-
678 cal adaptation. We verify that this kind of selection for hybridization is not sensitive
679 to the ratio of hybrid to migrant fitness in each deme by also considering alternative
680 (quadratic/stretched exponential) selection functions (results not shown).

681 Second, strong assortment increases the frequency of outliers within resident and mi-
682 grant sub-populations (fig. 5d); the lower fitness and reduced mating success of such
683 outliers selects against the associated high-assortment modifier.

684 The first explanation for selection against $\alpha < \alpha_*$ modifiers applies when there is sig-
685 nificant hybridization between diverged populations, while the second is more relevant
686 if there is complete reproductive isolation at the divergence threshold itself, i.e., if α_c
687 is so low that the switch to bimodality is enough to ensure a close to zero probability
688 of heterospecific matings. This is typically the case when both 2μ and m are high (fig.
689 5a); high migration, then, actually aids stronger reproductive isolation, by inducing fe-
690 males to evolve very restrictive preference ranges at the ESS (see Rettelbach et al.
691 (2013) for a somewhat similar effect in a one-locus model). Thus selection for inter-
692 mediate assortment in this situation does not reflect incomplete reproductive isolation
693 between diverged populations, as found by Servedio (2011), but arises instead due to
694 selection against assortment-driven increase in sub-population variance, as also observed
695 in polygenic models of competitive speciation (Doebeli 1996).

696 In finite populations with zero or low mutational variance, complete reproductive isola-
697 tion can also evolve via the action of stabilizing selection on diverged resident and migrant
698 sub-populations in each deme. Stabilizing selection depletes standing variation within
699 the sub-populations; as the phenotypic range of both migrants and residents shrinks, hy-
700 bridization between the two becomes weaker, which further accelerates loss of variation,
701 finally leading to the emergence of two distinct phenotypes, narrowly clustered around
702 the optima, which are too ‘far apart’ to hybridize (fig. 3a). This finding has two general

703 implications. First, hybridization between incipient species may not be as much due to
704 mating between the prototypical (fittest or most abundant) phenotypes, but is likely to
705 involve outliers within each population. Thus the extent of hybridization between pop-
706 ulations is intimately connected to the degree of phenotypic variation they harbour, or
707 have harboured in the past. Second, contrary to the reinforcement view, in this situation,
708 it is selection for increased discrimination that drives initial divergence, and stabilizing
709 selection that causes populations to evolve complete reproductive isolation over longer
710 timescales.

711 To assess the robustness of our results, we also analyze the model with asymmetric
712 migration and unequal selection strengths in the two demes. Most qualitative conclu-
713 sions remain unaltered in the presence of these asymmetries. However, asymmetric gene
714 flow inhibits the evolution of assortment over a wider parameter range than symmetric
715 migration (fig. S6b, Appendix S7), as is consistent with earlier studies which suggest
716 that indirect, LD-mediated selection on modifiers can be washed out by net influx of a
717 foreign allele at the modifier locus (Servedio and Kirkpatrick 1997; Servedio 2000). This
718 is, however, partially offset by the fact that for a fixed strength of assortment, divergence
719 is higher with more asymmetric migration (fig. S6a, Appendix S7). Thus, island popula-
720 tions subject to one-way introgression may evolve weaker assortment, but similar levels
721 of divergence as populations with reciprocal introgression.

722 Many of our results depend on the fact that populations experience stabilizing selection
723 around well-defined optima. Thus it would be useful to consider how these conclusions
724 change when selection in each deme is directional (Cotto et al. in preparation). A
725 limitation of the model is that the preference range is assumed to be determined by a
726 single locus. Modeling the preference range as a polygenic trait with similar amounts
727 of standing variation as the assortment trait could be another interesting direction for
728 future work.

729 In conclusion, divergence and assortment evolution based on highly polygenic traits is
730 well-described by phenotypic approaches such as the infinitesimal and the hypergeometric
731 model, at least over short timescales. Polygenic traits can maintain substantial genetic
732 and phenotypic variation in spite of stabilizing selection when there is gene flow between
733 demes and if density regulation occurs independently in each deme. Assortative mating
734 further amplifies phenotypic variation by building up associations between loci, thus
735 facilitating a strong response to divergent selection. The extent of genetic variation also
736 affects the degree of hybridization between diverged populations and the evolution of
737 discrimination within these, and is thus a crucial determinant of the fate of incipient
738 species.

739 **Author Contributions**

740 H.S. and N.H.B. designed the study, H.S. did the calculations, H.S. wrote the manuscript
741 with input from N.H.B. .

742 **Acknowledgements**

743 The research leading to these results has received funding from the People Programme
744 (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-
745 2013) under REA grant agreement Nr. 291734 (H.S.) and ERC grant SelectionInforma-
746 tion 250152 (N.B.). We thank Maria Servedio, Olivier Cotto and Christelle Fraisse for
747 useful comments on the manuscript, and two anonymous reviewers for detailed and in-
748 sightful reviews.

749 **References**

- 750 Barton, N.H. 1992. On the spread of new gene combinations in the third phase of Wrights
751 shifting balance. *Evolution* 46:551-557.
- 752 Barton, N. H., and M. A. R. de Cara. 2009. The evolution of strong reproductive isolation.
753 *Evolution* 63:1171-1190.

- 754 Barton, N.H., A.M. Etheridge, and A. Véber. 2016. The infinitesimal model. bioRxiv doi:
755 <http://dx.doi.org/10.1101/039768>.
- 756 Barton, N.H., and P.D. Keightley. 2002. Understanding quantitative genetic variation.
- 757 Barton N. H., and M. Shpak. 2000. The stability of symmetrical solutions to polygenic
758 models. *Theor. Popul. Biol.* 57:249-264.
- 759 Bolnick, D.I., and B.M. Fitzpatrick. 2007. Sympatric speciation: models and empirical
760 evidence. *Annual Review of Ecology, Evolution, and Systematics*, 38, 459-487.
- 761 Brown, J. S., and N. B. Pavlovic. 1992. Evolution in heterogeneous environments: effects
762 of migration on habitat specialization. *Evol. Ecol.* 6:360-382.
- 763 Bulmer, M.G. 1980. *The mathematical theory of quantitative genetics*. Oxford University
764 Press.
- 765 Bürger, R., and K. Schneider. 2006. Intraspecific competitive divergence and convergence
766 under assortative mating. *American Naturalist* 167:190-205.
- 767 Butlin, R. R. 1993. The variability of mating signals and preferences in the brown plan-
768 thopper, *Nikzparuata lugens* (Homoptera: Delphacidae). *J. Insect Behav.* 6:125-140.
- 769 Cotto, O., et al, in preparation.
- 770 Coyne, J. A, and H. A. Orr. 2004. *Speciation*. Sinauer, Sunderland, MA.
- 771 de Cara, M. A. R., N. H. Barton, and M. Kirkpatrick. 2008. A model for the evolution
772 of assortative mating. *Am. Nat.* 171:580-596.
- 773 Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *Am. Nat.* 74:312-
774 321.
- 775 Doebeli, M. 1996. A quantitative genetic competition model for sympatric speciation, *J.*
776 *Evol. Biol.* 9, 893-910.

- 777 Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of
778 animals? *Evolution*, 35, 124-138.
- 779 Gavrillets, S. 2004. *Fitness landscapes and the origin of species*. Princeton Univ. Press,
780 Princeton, NJ ; Oxford, England.
- 781 Geritz, S. A. H., É. Kisdi, G. Meszéna, and J. A. J. Metz. 1998. Evolutionary singular
782 strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.*
783 12:35-57.
- 784 Hey, J. 2006. Recent advances in assessing gene flow between diverging populations and
785 species. *Current Opinion in Genetics and Development*, 16, 592-596. *Nature Reviews*
786 *Genetics*, 3, 11-21.
- 787 Kelly, J. K., and M. Noor. 1996. Speciation by reinforcement: a model derived from
788 studies of *Drosophila*. *Genetics* 143:1485-1497.
- 789 Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. *Proc. R.*
790 *Soc. London Ser. B* 267:1649-55.
- 791 Kirkpatrick, M., and N.H. Barton. 1997. The strength of indirect selection on female
792 mating preferences. *Proc Natl Acad Sci USA* 94(4):1282-1286.
- 793 Kirkpatrick, M., and S. L. Nuismer. 2004. Sexual selection can constrain sympatric spe-
794 ciation. *Proceedings of the Royal Society of London, B* 271:687-693.
- 795 Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models
796 and experiments. *Am. Nat.*, 159, S22-S35.
- 797 Kondrashov, A. S. (1984). On the intensity of selection for reproductive isolation at the
798 beginnings of sympatric speciation. *Genetika*, 20:408-415.
- 799 Kondrashov A.S., and M. Shpak. 1998. On the origin of species by means of assortative
800 mating. *Proc. R. Soc. London Ser. B* 265:2273-2278

- 801 Kopp, M., and J. Hermisson. 2008. Competitive speciation and costs of choosiness. *J.*
802 *Evol. Biol.*, 21, 1005-1023.
- 803 Matessi, C., A. Gimelfarb, and S. Gavrilets. 2001. Long-term buildup of reproductive
804 isolation promoted by disruptive selection: How far does it go? *Selection* 2:41-64.
- 805 Meszéna, G., I. Czibula, and S. A. H. Geritz. 1997. Adaptive dynamics in a 2-patch
806 environment: a toy model for allopatric and parapatric speciation. *J. Biol. Syst.* 5:265-
807 284.
- 808 Mylius, S.D., and O. Dieckmann. 2001. The resident strikes back: invader-induced switch-
809 ing of resident attractor. *J. Theor. Biol.*, 211, 297-311.
- 810 Niemiller, M.L., B.M. Fitzpatrick, and B.T. Miller. 2008. Recent divergence-with-gene-
811 flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene
812 genealogies. *Molecular Ecology*, 17, 2258-2275.
- 813 Ortiz-Barrientos, D., and M. A. Noor. 2005. Evidence for a one-allele assortative mating
814 locus. *Science* 310:1467.
- 815 Otto, S.P., M.R. Servedio, and S.L. Nuismer. 2008. Frequency-dependent selection and
816 the evolution of assortative mating. *Genetics* 179:2091-2112.
- 817 Pennings, P. S., M. Kopp, G. Meszéna, U. Dieckmann, and J. Hermisson. 2008. An
818 analytically tractable model for competitive speciation. *Am. Nat.* 171:E44-E71.
- 819 Polechová, J., and N. H. Barton. 2005. Speciation through competition: a critical review.
820 *Evolution* 59:1194-1210.
- 821 Rettelbach, A., M. Kopp, U. Dieckmann, and J. Hermisson. 2013. Three models of adap-
822 tive speciation in spatially structured populations. *Am. Nat.* 182:E215-E234.
- 823 Ripa, J. 2009. When is sympatric speciation truly adaptive? An analysis of the joint
824 evolution of resource utilization and assortative mating. *Evol. Ecol.* 23:31-52.

- 825 Seehausen, O. 2006. Conservation: losing biodiversity by reverse speciation *Curr. Biol.*,
826 16 , pp. 334-337.
- 827 Servedio, M. R. 2000. Reinforcement and the genetics of nonrandom mating. *Evolution*
828 54:21-29.
- 829 Servedio, M.R. 2011. Limits to the evolution of assortative mating by female choice under
830 restricted gene flow. *Proc. R. Soc. Lond. B Biol. Sci.* 278:179-187.
- 831 Servedio, M. R., and M. Kirkpatrick 1997. The effects of gene flow on reinforcement.
832 *Evolution* 51:1764-1772.
- 833 Servedio, M., and M. Noor. 2003. The role of reinforcement in speciation: theory and
834 data. *Ann. Rev. Ecol. Evol. Syst.* 34:339-364.
- 835 Servedio, M. R., G. S. van Doorn, M. Kopp, A. M. Frame, and P. Nosil. 2011. Magic
836 traits in speciation: magic but not rare? *Trends Ecol. Evol.* 26:389-397.
- 837 van Doorn, G. S., P. Edelaar, and F. J. Weissing. 2009. On the origin of species by natural
838 and sexual selection. *Science* 326:1704-1707.
- 839 Wang, R.L., J. Wakeley J, and J. Hey. 1997. Gene flow and natural selection in the origin
840 of *Drosophila pseudoobscura* and close relatives. *Genetics*, 147, 1091-1106.