The American Naturalist

Local adaptation by small-effect alleles

--Manuscript Draft--

Manuscript Number:	55813
Full Title:	Local adaptation by small-effect alleles
Short Title:	Local adaptation by small-effect alleles
Article Type:	Symposium
Abstract:	How does migration-selection balance shape the genetic architecture of local adaptation? At first glance, the predictions of population and quantitative genetic models of migration-selection balance appear to contradict each other. Population genetic models predict that small-effect alleles can be swamped by migration, and are therefore unlikely to contribute to local adaptation. Quantitative genetic models are based on the assumption that all alleles are of small effect, and predict that local adaptation will evolve if there is sufficient standing genetic variation (VG). Empirically, both VG and local adaptation are common, and most VG seems due to alleles of small effect. But if small alleles are prone to swamping, how does local adaptation evolve? Here, I review how this seeming contradiction arises and compare the predictions of population and quantitative genetic models. I then use individual-based simulations to explore the evolution of genetic architecture when alleles are small and prone to swamping, illustrating how population and quantitative genetic predictions complement, rather than contradict each other. Traits made up of many small, swamping prone alleles can yield considerable local adaptation, but have low FST, high allelic covariance, and very transient architectures, where no locus makes an important contribution for very long. 24 pages, 6 figures

LOCAL ADAPTATION BY SMALL-EFFECT ALLELES

Sam Yeaman

Department of Forest and Conservation Sciences, 2424 Main Mall, University of British Columbia, Vancouver, Canada, V6T 1Z4.

Department of Botany, 6270 University Blvd., University of British Columbia, Vancouver, Canada, V6T 1Z4.

yeaman@zoology.ubc.ca

Abstract

How does migration-selection balance shape the genetic architecture of local adaptation? At first glance, the predictions of population and quantitative genetic models of migration-selection balance appear to contradict each other. Population genetic models predict that small-effect alleles can be swamped by migration, and are therefore unlikely to contribute to local adaptation. Quantitative genetic models are based on the assumption that all alleles are of small effect, and predict that local adaptation will evolve if there is sufficient standing genetic variation (V_G). Empirically, both V_G and local adaptation are common, and most V_G seems due to alleles of small effect. But if small alleles are prone to swamping, how does local adaptation evolve? Here, I review how this seeming contradiction arises and compare the predictions of population and quantitative genetic models. I then use individual-based simulations to explore the evolution of genetic architecture when alleles are small and prone to swamping, illustrating how population and quantitative genetic predictions complement, rather than contradict each other. Traits made up of many small, swamping prone alleles can yield considerable local adaptation,

29 but have low F_{ST}, high allelic covariance, and very transient architectures, where no locus
30 makes an important contribution for very long.
31
32 **Introduction**
33 What is the genetic basis of adaptation? How many variants contribute to a given
34 adaptive phenotype? How big are their effects? Where are they in the genome? The
35 answers to these questions are currently the subject of intense empirical scrutiny and
36 theoretical discussion (Stinchcombe and Hoekstra 2008; Rockman 2011; Martin and
37 Orgogozo 2013; Savolainen et al. 2013; Lee et al. 2014). With recent advances in
38 sequencing technology, it is now possible to explore associations between millions of
39 SNPs and the phenotypes and environments involved with local adaptation (e.g.,
40 Fournier-level et al. 2011; Hohenlohe et al. 2011, Soria-Corrasco et al. 2014). Beyond
41 our mechanistic interest in connecting genotype to phenotype, characterizing these
42 variants allows us to answer questions about how evolution shapes a species. In
43 particular, the way that a species inhabits a landscape can have profound effects on the
44 expected genetic architecture of adaptive traits. As individuals disperse among regions,
45 they may encounter different environments or competitors, which can result in natural
46 selection acting at different strengths or in different directions in trait space. Depending
47 on the relationship between spatial variation in selection and the rate of migration across
48 a landscape, heterogeneous environments can lead to local adaptation (Felsenstein 1976;
49 Hedrick et al. 1976; Hedrick 1986; Linhart and Grant 1996; Hereford 2009), affect the
50 maintenance of standing genetic variance (Slatkin 1978; Lytghoe 1997; Tufto 2000;
51 Yeaman and Jarvis 2006), and result in speciation (Schluter 2000; Nosil 2012). By

formulating clear predictions about how migration, selection, and drift combine to shape genetic architecture and testing these predictions with empirical data, we can learn much about how landscape shapes these fundamental aspects of evolution.

Population genetic models have provided the most tractable approach to generating testable predictions about the evolution of genetic architecture under the balance between divergent selection and migration. Haldane (1930) showed that a locally favored allele would be lost from a single island population when the rate of immigration (m) of a maladapted genotype exceeds the strength of selection (s) against it. Subsequent studies have identified similar critical thresholds for the maintenance of locally adapted polymorphisms in models with a wide range of demographics and patterns of environmental variation (reviewed in Felsenstein 1976; Bürger 2014). Conceptually, the results of these models can be understood as follows: the mean fitness of an allele depends on the proportion of evolutionary time (over multiple generations) that it spends in each type of environment and the fitness it experiences there. At higher migration rates, a locally adapted allele will spend an larger proportion of its evolutionary time in environments where it is disfavored, and above the critical migration rate, it will have a lower mean fitness than an intermediate, non-locally adapted allele. At this point, the maintenance of polymorphism will be deterministically disfavored, although loss of polymorphism is also possible due to drift when the mean fitness of the locally adapted alleles is close to that of an intermediate allele (Yeaman and Otto 2011).

The generality of the critical migration threshold result implies that large-effect alleles will be more likely to contribute to local adaptation, as they are more able to overcome the homogenizing effect of migration, or ‘gene swamping’ (Lenormand 2002).

Even when migration rates are below the critical threshold, larger alleles also have higher probabilities of establishment and longer persistence times, further suggesting that architectures evolving under migration-selection balance should be enriched for large-effect alleles (Yeaman and Otto 2011; Yeaman and Whitlock 2011; Aeschbacher and Bürger 2014). Simulation studies show that the distribution of allele effect sizes underlying a locally adapted trait tends to be shifted towards larger alleles when there is persistent migration-selection balance (Griswold 2006; Yeaman and Whitlock 2011). Thus, we can test for the importance of migration in shaping adaptive evolution by studying the genetic architecture of traits in comparisons of allopatric vs. parapatric pairs of populations inhabiting contrasting environments (as per Renaut et al. 2013).

If we observe a given rate of migration between contrasting environments, we might predict that adaptation would be effectively prevented by swamping if all genetic variants are of small effect relative to the migration rate. Hereafter, I will use the term 'swamping-prone alleles', or SPAs to denote alleles with selection coefficients that fall below the critical threshold in the relevant model, analogous to $s < m$ in Haldane's continent-island model, and 'swamping-resistant alleles' or SRAs for the converse. It is important to note that the actual value of the critical threshold can differ substantially according to model assumptions, and the transition from SRA to SPA with can involve a gradual transition in dynamics (*e.g.*, gradual changes in the persistence time of a polymorphism or mean allelic divergence at equilibrium; Yeaman and Otto 2011). While SPAs have non-zero probabilities of establishing and contributing to adaptation, they do not persist for long once established (Yeaman and Otto 2011), suggesting adaptation would be transient and unstable if all variation was due to SPAs. Alternatively, if

available mutations occur in a large range of sizes, single-locus models predict that we should fail to observe alleles contributing to local adaptation with sizes below the minimum effect size that can resist swamping.

The effect of linkage disequilibrium

When extrapolating to the whole genome scale, the predictions of single-locus models need to be modified to account for the effect of linkage disequilibrium (LD) between locally adapted alleles (Slatkin 1975; Barton 1983; Yeaman and Whitlock 2011; Bürger and Akerman 2011). LD can alter the mean fitness of an allele through associations with other locally adapted alleles that arise due to either physical linkage when the alleles are on the same chromosome or statistical association when the alleles are on different chromosomes. In either case, the magnitude of fitness increase for a focal allele depends on the number of other alleles under selection, their frequencies, and their linkage relationships (Barton and Bengtsson 1986; Bürger and Akerman 2011). This increased fitness results in a decreased rate of gene flow at the focal locus relative to the migration rate, which has been termed the 'effective migration rate' (Petry 1983; Bengtsson 1985; Bengtsson and Barton 1986), and has also been discussed in terms of reduced rates of between-population recombination (Via 2012; Feder et al. 2013). If this LD-mediated indirect effect of divergent selection is sufficiently large, it may cause a new locally adapted mutation that would be otherwise prone to swamping to instead be resistant to swamping (Yeaman and Whitlock 2011), as has been elegantly shown using two-locus analytical models (Bürger and Akerman 2011; Akerman and Bürger 2014).

120 Two broad predictions emerge from the extensive body of theoretical work on the
121 effects of LD on adaptation (reviewed in Smadja and Butlin 2011; Feder et al. 2012a):
122 physical linkage should most strongly affect the mean fitness of a focal allele when the
123 strength of divergent selection on a linked background locus is large relative to the
124 recombination rate between them ($s > r$; Petry 1983; Barton and Bengtsson 1986;
125 Yeaman and Whitlock 2011; Bürger and Akerman 2011; Akerman and Bürger 2014), and
126 statistical associations with locally adapted alleles that are physically unlinked will have
127 substantial effects on mean fitness at a focal allele only when the fitness of immigrants is
128 very strongly reduced (Bengtsson 1985; Barton and Bengtsson 1986; Feder et al. 2012b;
129 Flaxman et al. 2013). The effects of LD for physically linked and unlinked loci have been
130 discussed in terms of 'divergence hitchhiking' and 'genomic hitchhiking', respectively,
131 but these terms have been applied to both causal processes and resulting patterns in the
132 literature, so it is not clear how they should be applied to discuss swamping. Taken
133 together, the results of the above studies suggest that if LD facilitates local adaptation via
134 SPAs, it will likely be restricted to clustered regions in the genome, unless the relative
135 fitness of migrants and their offspring is very low. As the scope for LD to facilitate
136 divergence would be especially limited at the outset of an adaptive walk, swamping
137 presents an intriguing and potentially significant theoretical hurdle to local adaptation.
138
139 *Does swamping limit local adaptation in quantitative traits?*
140 Despite the predicted importance of swamping, it is not immediately clear how to
141 reconcile this prediction with the behavior of quantitative genetic models of local
142 adaptation. Quantitative genetic models come from a very different perspective than

143 population genetic models, beginning from the assumption that standing genetic variance
144 is due to many alleles of vanishingly small effect. This allows the details of genetic
145 architecture to be subsumed into parameters representing the mean, variance, and higher
146 moments of the distribution of genetic values (Turelli and Barton 1994; Falconer and
147 Mackay 1996). Models of local adaptation predict the trait divergence at equilibrium by
148 finding the point where the response to divergent selection is balanced by the
149 homogenizing effect of migration (*e.g.*, Tufto 2000; Hendry et al. 2001; Lopez et al.
150 2008; Yeaman and Guillaume 2009; Huisman and Tufto 2012). The behavior of
151 quantitative genetic models can be seen most easily by examining the model of Hendry,
152 Day, and Taylor (2001):
153
154 $$D = \frac{D_\Theta V_G}{V_G(1-m)+(\omega^2+V_P)m} \qquad (1)$$
155
156 where the equilibrium trait divergence (D) is a function of the difference in
157 environmental trait optima (D_θ), stabilizing selection (ω^2), migration rate (m), and genetic
158 (V_G) and phenotypic variance (V_P). Because D scales with V_G, and V_G is a parameter
159 rather than a result of the evolutionary dynamics, divergence cannot be predicted *a priori*
160 from these models without knowing what determines V_G, which can itself be affected by
161 migration-selection balance. If we examine a simplified case where two populations
162 experience different environments (Θ = -1 vs. +1), m = 0.005, ω^2 = 25, and $V_P = V_G$, this
163 model predicts equilibrium divergence to be $D/D_\theta = 8V_G / (8V_G + 1)$.
164 By contrast, population genetic models of migration-selection balance are
165 explicitly focused on the importance of allele effect size, and predictions about swamping

can be made for individual alleles underlying a quantitative trait. Using α to represent the effect size of a new mutation and assuming that populations begin from an undifferentiated state where all phenotypes = 0, the fitness of the genotypes can be calculated from the appropriate fitness function (see supplementary materials for details), and substituted into equation 11 from Yeaman and Otto (2011) to calculate the critical migration rate (m_{crit}) above which the new allele will be prone to swamping:

$$m_{crit} = \left(\frac{w_{1,Bb}}{w_{1,Bb} - w_{1,BB}\left(1+\frac{1}{4N}\right)} - \frac{w_{2,Bb}}{w_{2,BB}\left(1+\frac{1}{4N}\right) - w_{2,Bb}} \right)^{-1}, \quad (2)$$

where $w_{1,xy}$ and $w_{2,xy}$ are the relative fitness values of genotype xy in patch 1 and patch 2, N is the patch size, allele b is favored in patch 1, and B is favored in patch 2. The two-patch model with symmetrical selection and migration has $m_{crit} = 0.5$ for mutations of any effect size when $N \rightarrow \infty$ (Bulmer 1972b), but much lower values of m_{crit} for asymmetrical selection, dominance, or migration or for finite populations (Gavrilets and Gibson 2002; Yeaman and Otto 2011). Using the Gaussian fitness function with $\omega^2 = 25$ to calculate $w_{1,xy}$ and $w_{2,xy}$ for a mutation of size α, equation (2) can be rearranged to solve for α_{crit} when $m = 0.005$, giving $\alpha_{crit} = 0.0215$ (by setting $N = 1000$, to simplify comparison to the symmetrical model in (1); see supplementary materials for details). For the same parameters used above with equation 1, this predicts that local adaptation will not be maintained at the allelic level when $\alpha << 0.0215$, which is about 1% of the difference in environmental optima ($\Theta_2 - \Theta_1 = 2$).

188 At first glance, these predictions appear to contradict each other across wide
189 ranges of parameter space. For the above example, if mutations are all small ($\alpha <<$
190 0.0215) equation 2 predicts that local adaptation should not evolve, but if there is any
191 standing genetic variation, equation 1 would still predict local adaptation. Standing
192 genetic variation is nearly ubiquitous in quantitative traits (Houle 1992; Falconer and
193 Mackay 1996) and in most cases appears to be due to alleles of small effect (Mackay et al
194 2009; Rockman 2011), suggesting that this problematic region of parameter space is quite
195 realistic. It is important to stress that these models do not actually contradict each other
196 on purely theoretical grounds, because (1) will yield the same prediction as (2) when it is
197 known that $V_G = 0$. However, the contradiction that arises in parameterizing them
198 suggests that we have an incomplete understanding of how local adaptation evolves when
199 alleles are individually prone to swamping. If most alleles are of small effect, how does
200 local adaptation persist? Is swamping a useful concept to understand constraints to local
201 adaptation in quantitative traits? Does the maintenance of variation at the whole genome
202 scale transcend the prediction that swamping at individual alleles would prevent
203 divergence? Here, I will use individual-based simulations with explicit genetic
204 architecture to examine some situations where population genetic models predict
205 swamping, but substantial local adaptation can be maintained under mutation-migration-
206 selection balance. I show that allele-level swamping may not present a significant barrier
207 to adaptation for traits with highly polygenic architectures, but that it does affect the
208 underlying architecture of this adaptation.
209
210 **Simulation model**

To explore the impact of mutation on adaptation under migration-selection balance, I used a modified version of Nemo (Guillaume and Rougemont 2006) to simulate a scenario with two patches connected by migration (m), similar to that described in Yeaman and Whitlock (2011). Each generation, new individuals are drawn and survive with probability proportional to their fitness until the carrying capacity of the local patch (N) is reached. Individual fitness is determined by the function $w = 1 - \phi (|\Theta - Z|/2)^{\gamma}$, where Z is the individual phenotype, Θ is the locally optimal phenotype, ϕ is the strength of selection on the phenotype, and γ is the curvature of the fitness function. Unless otherwise indicated, the optima are set to $\Theta_1 = -1$ and $\Theta_2 = 1$, and $\gamma = 1$, resulting in linear stabilizing selection, so that an individual that is perfectly adapted to one patch experiences a fitness cost of ϕ in the other patch.

To facilitate comparison to single-locus population genetic models, the recombination rate between loci is set to $r = 0.5$ and mutations are all diallelic, with phenotypic effects of either + or -α. Individual phenotypes are determined by the additive effects of all mutations present in their genome, with the total number of diploid loci equal to n_{tot}. In most cases, simulations are run using $\alpha = 0.01$ and $\phi = 0.1$, which results in a critical migration threshold of $m_{crit} = 0.0021$, when $N = 1000$ (from equation 2). While this prediction should approximate the true threshold for multiple unlinked mutations of the same size, due to $\gamma = 1$, complications arise with multiple loci (Spichtig and Kawecki 2004; see supplementary materials) and LD and indirect selection would cause slight deviations (Akerman and Bürger 2014). In any case, this prediction should be approximately accurate for multi-locus traits, suggesting that adaptation should be effectively prevented by swamping for $\{\alpha = 0.01; \phi = 0.1\}$ when $m >> 0.0021$.

234 Because of the assumption of stabilizing selection, the selection coefficients (s)
235 experienced by individual mutations will depend upon the genetic background and
236 architecture underlying the trait. In the above case where all $Z = 0$, then $s \sim \phi\alpha$. When a
237 phenotype is at a local optimum, however, any further mutations with the same sign as
238 the optimum become deleterious. This highlights a critical assumption that arises in
239 modeling quantitative traits: when n_{tot} is small enough that mutations at all loci are
240 required to yield the locally fittest phenotype (*i.e.* $2\sum_i \alpha_i \leq \Theta$), then all loci will
241 experience only directional selection within each patch. By contrast, when the number of
242 loci is larger than the number of alleles required to yield a locally optimal phenotype (*i.e.*
243 when $2\sum_i \alpha_i > \Theta$), the fitness effect of an allele depends upon its genetic background.
244 This phenomenon is known as genetic redundancy (Goldstein and Holsinger 1992;
245 Nowak 1997)—there are more loci contributing to the phenotype than are needed to
246 achieve the optimum phenotype. The number of loci required to yield a locally optimal
247 phenotype here is simply $n_{opt} = \Theta/(2\alpha)$ due to the diallelic and equal-effect-size
248 assumptions, but this quantity would have a less simple definition for more realistic
249 distributions of effect size. It is worth noting that the simulations used in Feder et al.
250 (2012a) and Flaxman et al. (2013; 2014), and other recent papers assume pure directional
251 selection, while those of Le Corre and Kremer (2003, 2011, 2012) and others assume
252 spatially variable stabilizing selection.
253
254 **Results**
255 In some regions of parameter space, transitions between low to high trait divergence are
256 consistent with predictions from population genetic models. When the migration rate and

257 total strength of selection on the phenotype (ϕ) are held constant and only α and n_{opt} are
258 allowed to vary (holding constant $2\alpha n_{opt} = \theta$), there is considerable divergence in the
259 mean trait values of the populations (D) when $\{n_{opt} = 1$ and $\alpha = 0.5\}$ and very little
260 divergence when $\{n_{opt} = 50$ and $\alpha = 0.01\}$ (Figure 1A). These results are consistent with
261 the swamping predictions from single-locus population genetic theory: individual
262 selection coefficients are weaker with small α and therefore the critical migration
263 thresholds are lower. As predicted by equation 2, little divergence is observed when the
264 critical migration thresholds for the alleles fall below the simulation migration rate
265 (Figure 1A, orange lines), even though the strength of selection on the phenotype is high.
266 However, in other cases, simulation behavior appears inconsistent with population
267 genetic predictions, at least a the phenotypic level: considerable trait divergence can be
268 maintained when swamping is predicted for individual alleles, especially when the trait is
269 highly polygenic and there is high genetic redundancy. If we hold constant $\{n_{opt} = 50$ and
270 $\alpha = 0.01\}$ but increase n_{tot}, considerable divergence can evolve when $n_{tot} >> n_{opt}$ and
271 mutation rates are high (Figure 1B), even though single-locus models predict that these
272 alleles should be prone to swamping. Similar amounts of divergence after 50,000
273 generations are observed whether the populations were initialized with no genetic
274 variation and no divergence between populations or initialized with individuals optimally
275 adapted to their local environments and high heterozygosity at all loci (Supplementary
276 materials, Figure S1A).
277 How is this stable adaptive phenotypic divergence maintained when swamping is
278 predicted for individual loci? One possible explanation is that the conditions for the
279 maintenance of polymorphism are relaxed when there is enough standing genetic

280 variation that some individual's phenotypes overshoot the local optimum, resulting in
281 stabilizing selection, which only occurs when $n_{tot} >> n_{opt}$. If this is critical to the
282 establishment of divergence, then we should see little to no divergence until standing
283 variation builds up and a substantial proportion of individuals have phenotypes beyond
284 the local optima. In parameter sets where substantial adaptive divergence evolved, there
285 was usually enough standing genetic variation at the end of the simulations that
286 stabilizing selection was occurring (Figure 2A: proportion of individuals with $Z > |\Theta|$).
287 However, by examining the time-course of evolution for one parameter set that results in
288 particularly high adaptive divergence and standing variation, it is clear that only pure
289 directional selection was acting during the time that the initial adaptive divergence was
290 evolving (Figure 2B, prior to red arrow). Thus, while the transition to stabilizing selection
291 likely alters the dynamics underlying the maintenance of polymorphism, it is not a
292 necessary prerequisite for the maintenance of local adaptation when $n_{tot} >> n_{opt}$.
293 Another possible explanation for persistent trait divergence despite the prediction
294 of allelic swamping is that different alleles are constantly establishing and being lost, so
295 that individual alleles make only fleeting contributions to phenotypic divergence. As long
296 as sufficient genetic variance exists within the population and phenotypic selection is
297 strong relative to migration, it may be possible for phenotypic divergence to stably persist
298 despite constant flux in the underlying architecture. For the additive diallelic loci
299 simulated here, the contribution of a locus to phenotypic divergence in the direction of
300 the difference in optima (d) can be calculated from the frequency of the $+\alpha$ allele in patch
301 1 (p_1) and patch 2 (p_2): $d = (\alpha p_2 - \alpha(1 - p_2)) - (\alpha p_1 - \alpha(1 - p_1)) = 2\alpha(p_2 - p_1)$,
302 so that the total divergence in mean phenotypes is then $D = \sum 2d$. In the highly polygenic

303 simulations, d varies considerably among loci, with many contributions to divergence
304 occurring in the direction opposite the difference in environmental optima (Figure 3B).
305 Alleles making large contributions to divergence, shown in the tails of the distribution in
306 Figure 3B, usually do so for only a few hundred generations, and the architecture of
307 allelic divergence is constantly shifting (Figure 3C; also see supplementary materials,
308 Figure S2 for a more detailed plot). High flux in architecture occurs for all of the
309 simulations plotted in Figure 1B: a bout of high divergence for a given allele, where the
310 value of d is in the top 20% of the distribution, lasts less than 300 generations on average
311 (Figure S3).

312 By contrast, for simulations where $\phi = 0.5$ and $\alpha = 0.1$, when selection on
313 individual alleles is strong relative to migration and swamping is not predicted, there are
314 very few loci with $d << 0$ (Figure 3E; note log10 scale) and architectures are temporally
315 stable with large positive contributions to divergence that typically last many thousands
316 of generations (Figure 3F). Interestingly, high flux in architecture can also occur when
317 alleles are resistant to swamping, if $n_{tot} >> n_{opt}$ and mutation rates are high (Figure 4), a
318 finding also reported by Yeaman and Whitlock (2011; their Figure S2). In this case, there
319 is a dramatic change in the stability of the genetic architecture underlying divergence
320 with just an order of magnitude of change in the per-locus mutation rate (Figure 4A vs.
321 B), which does not occur for less redundant architectures (Figure 4 C&D). This likely
322 occurs because there are many alleles with equivalent fitness effects when there is high
323 genetic redundancy and high mutation rate, and these alleles can then replace each other
324 due to drift. Because of the importance of drift, this effect would be most important in
325 small or demographically unstable populations.

When all alleles are prone to swamping, there is a strong and nearly linear relationship between standing genetic variance (V_G) and adaptive trait divergence (Figure S1B), and V_G also increases with μ and n_{tot} (Figure 1C). Because V_G can be inflated by migration among diverged populations above baseline expectations for mutation-selection balance (Slatkin 1978; Lythgoe 1997), it is not immediately clear if this increase in V_G is attributable to mutation alone. Rerunning these simulations with a homogeneous environment ($\Theta_1 = \Theta_2 = 0$) and high migration ($m = 0.5$) does not substantially affect the levels of V_G (Figure 1C), which shows that maintenance of V_G under these parameters is mainly driven by mutation, not migration. By contrast, when alleles are resistant to swamping, a much larger amount of V_G is maintained in heterogeneous environments, relative to homogeneous environments (Figure 1D). The linear relationship between V_G and trait divergence (Figure S1B) is consistent with the behavior of quantitative genetic models (equation 1): when there is very little genetic variance within populations, there is little adaptive divergence between them.

When traits are highly polygenic, Latta (1998, 2003) and Le Corre and Kremer (2003, 2011, 2012), showed that considerable trait divergence can evolve with limited allelic divergence (F_{ST}) but high between-population covariance in allelic effects, represented by θ_B. However, it was not clear from their work whether these results would extend to traits made up of swamping prone alleles. The results observed here show that this covariance-based mechanism does extend to SPAs: comparing the results of simulations with $n_{tot} = 5000$ and considerable trait divergence, to those with $n_{tot} = 50$ and little trait divergence, both result in similar levels of F_{ST}, proportion of polymorphic loci, and expected heterozygosity (Figure 5A & B). On the other hand, simulations where high

349 n_{tot} results in high D have much higher between-population allelic covariances (Figure
350 5C). The relative insensitivity of V_G to migration when adaptation is due to SPAs, shown
351 in Figure 1C, occurs because there is very limited allelic differentiation between
352 populations, and phenotypic divergence is mainly driven by positive between-population
353 allelic covariances. As a result, migration does not substantially impact allele frequencies
354 or inflate genetic variance within populations.
355 Because the amount of divergence when alleles are prone to swamping depends
356 on V_G, predicting the potential for local adaptation in this context becomes a problem of
357 understanding the maintenance of genetic variation. Genetic variance can be greatly
358 impacted by drift, and much less divergence is seen in simulations with smaller
359 population size (Figure S1C), as per Blanquart et al. (2012). While theoretical models of
360 mutation-selection balance generally predict that V_G should increase linearly with $n_{tot}\mu$
361 (Latter 1960; Bulmer 1972a; Turelli 1984; Bürger et al. 1989), both V_G and trait
362 divergence increase more with n_{tot} than with μ (*e.g.*, $\{n_{tot} = 500 \ \& \ \mu = 10^{-4}\}$ vs. $\{n_{tot} =$
363 $5000 \ \& \ \mu = 10^{-5}\}$ (Figure 1B). This occurs because expected heterozygosity becomes
364 saturated at high mutation rates (Figure 5B), which limits the potential for increases in V_G
365 with further increases in mutation.
366
367 **Discussion**
368 The results presented here build on a large body of previous research from
369 population and quantitative genetics, using a single simulation framework to illustrate
370 how genetic and ecological constraints can shape the architecture of adaptive divergence
371 in polygenic traits. As predicted by single-locus population genetic models (reviewed in

372 Felsenstein 1976 and Bürger 2014; Yeaman and Otto 2011), stable locally adapted
373 polymorphisms do not persist when the critical migration threshold for a given allele falls
374 below the migration rate (Figure 1A). However, the swamping of adaptive divergence at
375 the level of individual alleles does not necessarily prevent local adaptation and
376 divergence at the phenotypic level. If all alleles are prone to swamping, substantial trait
377 divergence can still evolve if sufficient genetic variance can be maintained within
378 populations under mutation-selection-drift balance (Figure 1), consistent with quantitative
379 genetic models. When this occurs, allelic divergence tends to be limited (Figure 5) and
380 transient (Figure 3), and trait divergence tends to be generated more by between-
381 population allelic covariances (Figure 5C), as per Latta (1998; 2003) and Le Corre and
382 Kremer (2003; 2011; 2012). By contrast, when alleles are large enough to be highly
383 resistant to swamping, the maintenance of adaptive allelic divergence does not depend on
384 V_G or recurring mutation (Figure 1D; Yeaman and Otto 2011; Aeschbacher and Bürger
385 2014), and architectures are more temporally stable (Figure 3F). The concept of
386 swamping is therefore very useful for understanding how genetic and ecological
387 constraints shape the underlying genetic architecture of divergence, but of limited utility
388 for predicting whether overall phenotypic local adaptation will occur, as allelic
389 constraints can be overcome by covariance-mediated quantitative genetic adaptation.
390 Because the potential for covariance-mediated evolution depends upon the maintenance
391 of V_G, it is less likely to occur when there are few loci or low mutation rates (Figure 1B),
392 or when demography results in small effective population size (Figure S1C). This
393 suggests that there may be a strong relationship between effective population size or
394 species range and the extent of local adaptation. Taken together, these results show that

predictions about the architecture of local adaptation can be broadly characterized in terms of the susceptibility of alleles to swamping and the factors affecting the maintenance of baseline standing genetic variation (*i.e.*, independently of migration; Figure 6).

Where do real species fall along the continua illustrated in Figure 6? In cases with very low migration between populations, most alleles would be effectively resistant to swamping, while the reverse would be true when migration rates are very high. With intermediate migration rates, allele effect sizes underlying a single trait may span the range from swamping-resistant to -prone. In such cases, adaptation may involve a combination of covariance-mediated responses via SPAs (where sufficient genetic variance exists) and stable allelic divergence via SRAs, depending upon allele frequencies and effect size distributions. Over long periods of time, if traits have high genetic redundancy ($n_{tot} >> n_{opt}$), mutations that are larger or more highly clustered will tend to outcompete and replace smaller alleles (Yeaman and Whitlock 2011; Yeaman 2013) and the gradual establishment of more SRAs would likely reduce the proportion of divergence contributed by allelic covariances among transient SPAs. While such dynamics were not explicitly investigated here, the high genetic variance associated with adaptation due to SPAs would involve considerable genetic load (Blanquart et al. 2012). It would be interesting to see if small genomic rearrangements yielding clustering (as per Yeaman 2013) would evolve for loci with alleles that are typically swamping-prone.

Although the simulations explored here are useful for building intuition about how local adaptation evolves, it is difficult to use them to make predictions about whether we should expect to see local adaptation via SPAs in empirical contexts.

418 Beyond the sensitivity of simulation results to parameters about which we know little, the
419 very assumptions about evolutionary dynamics represented in these simulations may not
420 accurately represent how genetic variation is maintained. Classical models of mutation-
421 selection balance (Latter 1960; Bulmer 1972a; Turelli 1984; Bürger et al. 1989) seem
422 unable to explain observed levels of genetic variation (Barton and Turelli 1989; Johnson
423 and Barton 2005), and because the simulations used here follow similar assumptions,
424 their results should not be considered quantitatively representative of what we should
425 expect to observe in nature. Nonetheless, it is reasonable to expect that traits governed by
426 very few loci are unlikely to harbor high levels of V_G and would therefore be unlikely
427 candidates for adaptation via SPAs. Conversely, since most quantitative traits have
428 substantial levels of V_G (Houle 1992; Falconer and Mackay 1996), local adaptation via
429 SPAs may be quite common in highly polygenic traits. While genetic redundancy is not
430 critical for adaptation via SPAs, traits with limited redundancy are necessarily less
431 polygenic, and therefore less likely to harbor high levels of V_G. Ultimately, empirical
432 data will be required to evaluate whether local adaptation via SPAs commonly occurs in
433 nature.
434
435 *Swamping and statistical linkage disequilibrium*
436 In some situations, the effects of statistical LD can greatly facilitate adaptive divergence,
437 even when all mutations are small enough that they would be individually prone to
438 swamping. Flaxman and colleagues (2014) showed that when divergently selected alleles
439 are present at a large number of loci, populations can go through a rapid transition from
440 being relatively undifferentiated to a state where there is high LD among alleles and they

441 are highly diverged between populations. This transition occurred earlier when physical
442 linkage was allowed, but still occurred when genome architecture precluded physical
443 linkage, indicating the importance of transient LD between many physically unlinked
444 alleles. Because this effect was always observed after divergently selected
445 polymorphisms had accumulated at a large number of loci, and it occurred along with a
446 substantial drop in the relative fitness of migrants, it seems consistent with the predictions
447 of Petry (1983) and Bengtsson and Barton (1985, 1986): LD alters the mean fitness of
448 alleles and therefore, their susceptibility to swamping. The results of Flaxman et al.
449 (2014) therefore represent an interesting example of how extreme divergent adaptation
450 can occur via LD among numerous unlinked small-effect alleles. While it is not clear
451 whether the mutations used in Flaxman et al. were in fact prone to swamping in the strict
452 sense (see supplementary materials for further discussion), it seems that they were not
453 very resistant to swamping. Regardless, the rapid transition from a state of low to high
454 differentiation illustrates how LD can cause a shift from a state where divergence is
455 determined mainly on a locus-by-locus basis, to one where alleles gain fitness from
456 statistical associations with each other.

457 There are some important differences between the type of adaptive divergence
458 described by Flaxman et al. and that shown here. Whereas divergence in Flaxman et al.
459 occurred with a substantial reduction in effective migration rate (effectively resulting in
460 speciation), divergence shown in Figure 1B occurs with very limited reduction in
461 effective migration rate, due to the relatively weaker strength of selection with $\phi = 0.1$.
462 This weaker selection would limit the potential for LD-mediated dynamics of the type
463 described by Petry (1983) and Barton and Bengtsson (1986). Also, allelic differentiation

observed here with SPAs is transient, whereas in the simulations of Flaxman et al., allelic differentiation is presumably very stable following the transition. These differences in simulation behavior arise because of differences in selection regime (directional vs. stabilizing selection) and how traits are modeled (multiplicative vs. additive effects). Also, the mutation rate in Flaxman et al., where one mutation occurred per generation in the entire meta-population, was much lower than the rates that resulted in local adaptation in Figure 1B, where many mutations occurred per generation. Allelic covariances are equivalent to LD, but the effects of these associations are much more stable and pronounced in the simulations of Flaxman et al (2014). Real traits and environments will likely vary in how closely they resemble the assumptions of these different approaches to modeling divergent selection, and both types of adaptation via small alleles may occur in nature.

Genomic signatures of local adaptation

We know that local adaptation is common at the phenotypic level (Hedrick et al. 1976; Hedrick 1986; Linhart and Grant 1996; Hereford 2009) and we are rapidly accumulating evidence about adaptation at the genomic level (reviewed in Mitchell-Olds et al. 2007; Stapley 2010; Feder et al. 2013; Savolainen et al. 2013), but we are still far from a comprehensive understanding of where real populations fall within the spectrum of Figure 6. A now frequently used approach to study the architecture of local adaptation is to scan the genome for SNPs or genomic regions that have strong statistical associations with a locally adapted phenotype or environmental variable of interest or extreme values of F_{ST}, beyond the patterns in allele frequency due to demography (Storz 2005; Schoville

et al. 2012). These approaches have identified signatures consistent with local adaptation in a range of species (e.g., Fournier-level et al. 2010; Hohenlohe et al. 2011, Soria-Corrasco et al. 2014), however it is not always clear how well they avoid false positives due to the confounding effects of demography (Vilas et al. 2012; Schoville et al. 2012; de Mita et al. 2013; Lotterhos and Whitlock 2014). Furthermore, it has been argued that many examples of clusters of loci with extreme F_{ST} may in fact be false positives driven by reductions in heterozygosity within populations due to selective sweeps or background selection, rather than migration-selection balance (Cruikshank and Hahn 2014). Beyond the problem of identifying false positives, it isn't clear how well such tools can identify true positives, because alleles under relatively weak selection, even when resistant to swamping, will not have large differences in allele frequency compared to putatively neutral markers (Latta 1998; Le Corre and Kremer 2003; Storz 2005). Le Corre and Kremer (2012) showed that with increasing numbers of loci underlying a trait, values of F_{ST} at causal loci were increasingly difficult to distinguish from F_{ST} at neutral markers. With SPAs, these problems are likely to be so severe that it will be very difficult to identify causal loci without also generating a large proportion of false positives. Further simulation studies that explicitly compare patterns at selected vs. neutral loci will be required to evaluate the feasibility of applying genome scan methods to detect a genomic signature of local adaptation via SPAs.

Even with perfect knowledge of which loci contribute to a given quantitative trait, when traits are highly genetically redundant and divergence is due to SPAs, the contributions of individual alleles tend to fluctuate dramatically over a few hundred generations (Figure 3C). These fluctuations would likely occur more gradually at larger

population sizes, but would be exacerbated by rapid changes in population size. Fluctuation in architecture is not confined to SPAs, and can occur when alleles are resistant to swamping, which may be relatively rapid if there is sufficient genetic redundancy and a high enough rate of mutation (Figure 4; Yeaman and Whitlock 2011). Such extreme flux in the allelic composition of divergence challenges the conception of local adaptation as a well-defined and stable genetic strategy for responding to environmental heterogeneity. Similarly, because adaptation involving highly redundant, polygenic architectures is associated with high genetic variance within populations, there are many different ways to build an optimal phenotype present within a population at any given time. If genetically redundant, SPA-mediated adaptation actually happens in nature, this flux could result in very different genetic architectures of divergence in pairs of populations spanning similar environmental gradients in different parts of a species range, if the pairs of populations were connected by sufficiently low migration rates. Thus, while range-wide population genomic studies have greater power to detect signals of parallel adaptation along multiple transects, such parallelism would not even be expected for adaptation via SPAs or for traits with high genetic redundancy (given sufficient isolation-by-distance between transects), reducing the relative power of range-wide studies.

An interesting alternative to the locus-by-locus approaches of genome scans recently proposed by Berg and Coop (2014) is to use the results of a genome wide association study (GWAS) to estimate genotypic values within populations as the frequency-weighted sum of the individual SNP effects on a trait of interest. It is then possible to test whether variance among populations in estimated genotypic values

exceeds background expectations, estimated using putatively neutral markers, which is effectively a genomic analogue of Q_{ST} vs. F_{ST} tests. The genotypic values used in this test are analogous to *D* here, as they represent the sum of all locus effects (*d*), and include the effects of both allelic differentiation and covariance. This test may be especially useful for detecting a signature of adaptation via SPAs, which tend to contribute to divergence mainly through covariances. However, because the small-effect alleles likely to be prone to swamping are also difficult to detect using GWAS, and because of the difficulties of controlling for population structure, this approach might require very large sample sizes and careful design to get sufficient power (Berg and Coop 2014).

At the phenotypic level, when local adaptation is highly polygenic and driven primarily by allelic covariances, migration does not substantially inflate standing genetic variance within populations (Figure 1C). Thus, finding higher genetic variance within populations in regionally heterogeneous environments (Yeaman and Jarvis 2006) or in populations experiencing higher migration might suggest that local adaptation is at least partly due to substantial changes in allele frequency and SRAs. Also, if sufficiently powered genome scans are unable to detect a strong signature of local adaptation in a given species, despite observed phenotypic divergence, this is suggestive of the importance of SPAs. However, such negative evidence is hardly very convincing, and direct positive associations of alleles or covariances with divergence are much more appealing, and hopefully attainable.

Conclusions

555 There is a long history of using single locus population genetic models to understand
556 local adaptation. These models provide useful predictions about the conditions where
557 stable allelic divergence should be expected to contribute to local adaptation. However,
558 not all local adaptation involves stable allelic divergence, and these models therefore
559 should not be used to make predictions about limits to overall divergence for highly
560 polygenic traits. The genetic redundancy of a trait has very important effects on the
561 stability of the genetic architecture underlying local adaptation, and considerable
562 phenotypic divergence may be maintained with only transient allelic divergence.
563 Population genomic studies contrasting the genetic architecture of standing variation vs.
564 population differentiation may help understand how real species respond to
565 heterogeneous environments. However, statistical challenges may limit the potential for
566 confident inference and much remains to be learned about the power of these analytical
567 methods, especially in the non-equilibrium conditions found in nature.
568

569 **Acknowledgments**

570 I would like to thank Mike Whitlock for many stimulating discussions that led up to this
571 manuscript. Kay Hodgins, Katie Lotterhos, Jason Holliday, Sally Aitken, Loren
572 Rieseberg, and the Kassen lab also provided helpful comments and discussion. This work
573 is part of the AdapTree Project funded by the Genome Canada Large Scale Applied
574 Research Project program, with co-funding from Genome BC and the Forest Genetics
575 Council of BC (co-Project Leaders S. N. Aitken and A. Hamann).

576
577
578 **References**
579

580 Aeschbacher, S., R. Bürger. 2014. The effect of linkage on establishment and survival of
581 locally beneficial mutations. Genetics 197:317-336.

582 Akerman, A., R. Bürger. 2014. The consequences of gene flow for local adaptation and
583 differentiation: a two-locus two-deme model. J. Math. Biol. 68:1135-1198.

584 Barton, N. H. 1983. Multilocus clines. Evolution 37:454–471.

585 Barton, N. H., and B. O. Bengtsson. 1986. The barrier to genetic exchange between
586 hybridizing populations. Heredity 57:357–376.

587 Barton, N. H., and M. Turelli. 1989. Evolutionary quantitative genetics: how little do we
588 know? Annu. Rev. Genet. 23:337–370.

589 Bengtsson, B. O. 1985. The flow of genes through a genetic barrier. Pp. 31–42. *in* J. J.
590 Greenwood, P. H. Harvey, and M. Slatkin, eds. Evolution essays in honor of John
591 Maynard Smith. Cambridge Univ. Press, Cambridge.

592 Berg, J. J., G. Coop. 2014. A population genetic signal of polygenic adaptation. PLoS
593 Genetics 10:e1004412.

594 Blanquart, F., S. Gandon, S. L. Nuismer. 2012. The effects of migration and drift on local
595 adaptation to a heterogeneous environment. J. Evol. Biol. 25:1351-1363.

596 Bulmer, M. G. 1972a. The genetic variability of polygenic characters under optimising
597 selection, mutation and drift. Genetical Research 49:157-173.

598 Bulmer, M. G. 1972b. Multiple niche polymorphism. Am. Nat. 106:254– 257.

599 Bürger, R., G. P. Wagner, F. Stettinger. 1989.How much heritable variation can be
600 maintained in finite populations by mutation-selection balance? Evolution 43:1748–1766.

601 Bürger, R., A. Akerman. 2011. The effects of linkage and gene flow on local adaptation:
602 a two-locus continent-island model. Theor. Popul. Biol. 80:272-288.

603 Bürger, R. 2014. A survey of migration-selection models in population genetics. Discrete
604 and continuous dynamical systems B 19:883-959.

605 Cruickshank, T. E., M. W. Hahn. 2014. Reanalysis suggests that genomic islands of
606 speciation are due to reduced diversity, not reduced gene flow. Mol. Ecol. 3133-3157.

607 De Mita, S., A. C. Thuillet, L. Gay, et al. 2013. Detecting selection along environmental
608 gradients: analysis of eight methods and their effectiveness for outbreeding and selfing
609 populations. Molecular Ecology 22, 1383–1399.

610 Falconer, D. S., T. F. C. Mackay. 1996. Introduction to quantitative genetics. Pearson
611 Education Ltd., Essex.

612 Feder, J. L., S. P. Egan, P. Nosil. 2012a. The genomics of speciation-with-gene-flow.

613 Trends Genet. 28:342-350.

614 Feder, J. L., R. Geijji, S. Yeaman, P. Nosil. 2012b. Establishment of new mutations under
615 divergence and genome hitchhiking. Phil. Trans. R. Soc. B 367:461-474.

616 Feder, J. L., S. Flaxman, S.P. Egan, A.A. Comeault, P. Nosil. 2013. Geographic mode of
617 speciation and genomic divergence. Ann. Rev. Ecol. Syst. 44:73-97.

618 Felsenstein, J. 1976. The theoretical population genetics of variable selection and
619 migration. Ann. Rev. Genet. 10:253-280.

620 Flaxman, S. M., J. L. Feder, P. Nosil. 2013. Genetic hitchhiking and the dynamic buildup
621 of genomic divergence during speciation with gene flow. Evolution 67:2577-2591.

622 Flaxman, S. M., A.C. Wacholder, J. L. Feder, P. Nosil. 2014. Theoretical models of the
623 influence of genomic architecture on the dynamics of speciation. Mol. Ecol. 23:4074-
624 4088.

625 Fournier-Level, A., A. Korte, M.D. Cooper, M. Nordborg, J. Schmitt, A.M. Wilczek.
626 2011. A map of local adaptation in *Arabidopsis thaliana*. Science 334:86-90.

627 Goldstein, D. B., Holsinger, K. E. 1992. Maintenance of polygenic variation in spatially
628 structured populations: roles for local mating and genetic redundancy. Evolution 46:412-
629 429.

630 Griswold, C. K. 2006. Gene flow's effect on the genetic architecture of a local adaptation
631 and it's consequences for QTL analyses. Heredity 96:445–453.

632 Guillaume, F., and J. Rougemont. 2006. Nemo: an evolutionary and population genetics
633 programming framework. Bioinformatics 22:2556–2557.

634 Haldane, J. B. S. 1930. A mathematical theory of natural and artificial selection. Part VI.
635 Isolation. Proc. Cambridge Philos. Soc. 26:220–230.

636 Hedrick, P.W., M. E. Ginevan, and E.P. Ewing. 1976. Genetic polymorphism in
637 heterogeneous environments. Annu. Rev. Ecol. Syst. 7:1–32.

638 Hedrick, P. W. 1986. Genetic polymorphism in heterogeneous environments: a decade
639 later. Ann. Rev Ecol. Syst. 17:535-566.

640 Hendry, A. P., T. Day, E. B. Taylor. 2001. Population mixing and the adaptive
641 divergence of quantitative traits in discrete populations: a theoretical framework for
642 empirical tests. Evolution 55:459–466.

643 Hereford, J. 2009. A quantitative survey of local adaptation and fitness tradeoffs. Am.
644 Nat. 173:579–588.

645 Hohenlohe, P. A., S. Bassham, P. D. Etter, N. Stiffler, E.A. Johnson, W. A. Cresko.

646 Population genomics of parallel adaptation in threespine stickleback using sequenced
647 RAD tags. PLoS Genetics 6:e1000862.

648 Houle, D. Comparing evolvability and variability of quantitative traits. Genetics 130:195-
649 204.

650 Huisman, J., J. Tufto. 2012. Quantitative genetic models for migration and stabilizing
651 selection. Evolution 66:3444-3461.

652 Johnson, T., and N. H. Barton. 2005. Theoretical models of selection and mutation on
653 quantitative traits. Phil. Trans. R. Soc. B. 360:1411–1425.

654 Kremer, A., V. Le Corre. 2011. Decoupling of differentiation between traits and their
655 underlying genes in response to divergent selection. Heredity 108:375-385.

656 Latta, R. G. 1998. Differentiation of allelic frequencies at quantitative trait loci affecting
657 locally adaptive traits. Am. Nat. 151:283-292.

658 Latta, R. G. 2003. Gene flow, adaptive population divergence and comparative
659 population structure across loci. New Phytologist 161:51-58.

660 Latter, B. D. H. 1960. Natural selection for an intermediate optimum. Aust. J. Biol. Sci.
661 13:30-35.

662 Le Corre, V., A. Kremer. 2003. Genetic variability at neutral markers, quantitative trait
663 loci, and trait in a subdivided population under selection. Genetics 164:1205-1219.

664 Le Corre, V., A. Kremer. 2012. The genetic differentiation at quantitative trait loci under
665 local adaptation. Mol. Ecol. 21:1548-1566.

666 Lee, Y. W., B. A. Gould, J. R. Stinchcombe. 2014. Identifying the genes underlying
667 quantitative traits: a rationale for the QTN programme. AoB Plants: plu004.

668 Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends Ecol. Evol.
669 17:183–189.

670 Lewontin, R.C. 1974. The Genetic Basis of Evolutionary Change. Columbia University
671 Press, New York.

672 Linhart, Y. B., and M. C. Grant. 1996. Evolutionary significance of local genetic
673 differentiation in plants. Annu. Rev. Ecol. Syst. 27:237– 277.

674 Lopez, S., R. Rousset, F. H. Shaw, R. G. Shaw, O. Ronce. 2008. Migration load in plants:
675 role of pollen and seed dispersal in heterogeneous landscapes. J. Evol. Biol. 21:294–309.

676 Lotterhos, K. E., M. C. Whitlock. 2014. Evaluation of demographic history and neutral
677 parameterization on the performance of F_{ST} outlier tests. Mol. Ecol. 23:2178-2192.

678 Lythgoe, K. A. 1997. Consequences of gene flow in spatially structured populations.

679 Genet. Res. 69:49–60.

680 Mackay, T. F. C, E. A. Stone, J. F. Ayroles. 2009. The genetics of quantitative traits:
681 challenges and prospects. Nat. Rev. Genet. 10:565-577.

682 Martin, A., V. Orgogozo. 2013. The loci of repeated evolution: a catalog of genetic
683 hotspots of phenotypic variation. Evolution 67:1235-1250.

684 Mitchell-Olds, T., J. H. Willis, and D. B. Goldstein. 2007. Which evolutionary processes
685 influence natural genetic variation for phenotypic traits? Nat. Rev. Genet. 8:845–856.

686 Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and
687 heterogeneous genomic divergence. Mol. Ecol. 18:375–402.

688 Nosil, P. 2012. Ecological speciation. Oxford University Press, Oxford.

689 Nowak, M.A., M. C. Boerlijst, J. Cooke, J. Maynard Smith. 1997. Evolution of genetic
690 redundancy. Nature 388:167-171.

691 Petry, D. 1983. The effect on neutral gene flow of selection at a linked locus. Theor.
692 Popul. Biol. 23:300–313.

693 Renaut S., C. J. Grassa, S. Yeaman, B. T. Moyers, Z. Lai, et al. 2013. Genomic islands of
694 divergence are not affected by geography of speciation in sunflowers. Nature
695 Communications. 4:1827.

696 Rockman, M. V. 2011. The QTN program and the alleles that matter for evolution: all
697 that's gold does not glitter. Evolution 61:1-17.

698 Savolainen, O., M. Lascoux, J. Merilä. 2013. Ecological genomics of local adaptation.
699 Nat. Rev. Genet. 14:807-820.

700 Schluter D. 2000. The ecology of adaptive radiation. Oxford University Press, Oxford.

701 Schoville, S. D., A. Bonin, O. François, S. Lobreaux, C. Melodelima, S. Manel. 2012.
702 Adaptive genetic variation on the landscape: Methods and cases. Ann. Rev. Ecol. Evol.
703 Syst. 43:23-43.

704 Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, et al. 2014.
705 Genomics and the origin of species. Nat. Rev. Genet. 15:176-192.

706 Slatkin, M. 1975. Gene flow and selection in a two-locus system. Genetics 81:878-802.

707 Slatkin, M. 1978. Spatial patterns in the distributions of polygenic characters. J. Theor.
708 Biol. 70:213-228.

709 Smadja, C. M., R. K. Butlin. 2011. A framework for comparing processes of speciation in
710 the presence of gene flow. Mol. Ecol. 20:5123-5140.

711 Soria-Carrasco, V., Z. Gompert, A. A. Comeault, T. E. Farkas, T. Parchman, et al. 2014.
712 Stick insect genomes reveal natural selection's role in parallel speciation. Science
713 344:738-742.

714 Spichtig, M., and T. J. Kawecki. 2004. The maintenance (or not) of polygenic variation
715 by soft selection in heterogeneous environments. Am. Nat. 164:70–84.

716 Stapley, J., J. Reger, P. G. D. Feulner, C. Smadja, J. Galindo, et al. Adaptation genomics:
717 the next generation. Trends Ecol Evol. 25:705-712.

718 Stinchcombe, J. R., H. E. Hoekstra, 2008. Combining population genomics and
719 quantitative genetics: finding the genes underlying ecologically important traits. Heredity
720 100:158-170.

721 Storz, J.F. 2005. Using genome scans of DNA polymorphism to infer adaptive population
722 divergence. Mol. Ecol. 14:671-688.

723 Tufto, J. 2000. Quantitative genetic models for the balance between migration and
724 stabilizing selection. Genet. Res. 76:285–293.

725 Turelli, M., N. H. Barton. 1994. Genetic and statistical analyses of strong selection on
726 polygenic traits: what me normal? Genetics 138:913–941.

727 Via, S. 2012. Divergence hitchhiking and the spread of genomic isola- tion during
728 ecological speciation-with-gene-flow. Philos. Trans. R. Soc. Lond. B 367:451–460.

729 Vilas, A., A. Perez-Figueroa, A. Caballero. 2012. A simulation study on the performance
730 of differentiation-based methods to detect selected loci using linked neutral markers.
731 Journal of Evolutionary Biology 25:1364–1376.

732 Yeaman, S., A. Jarvis. 2006. Regional heterogeneity and gene flow maintain variance in a
733 quantitative trait within populations of lodgepole pine. Proc. R. Soc. B. 273:1587–1593.

734 Yeaman, S., F. Guillaume. 2009. Predicting adaptation under migration load: the role of
735 genetic skew. Evolution 63:2926–2938.

736 Yeaman, S. S. P. Otto. 2011. Establishment and maintenance of adaptive genetic
737 divergence under migration, selection, and drift. Evolution 65:2123-2129.

738 Yeaman, S. M. C. Whitlock 2011. The genetic architecture of adaptation under
739 migration-selection balance. Evolution 65:1897-1911.

740 Yeaman, S. 2013. Genomic rearrangements and the evolution of clusters of locally
741 adaptive loci. Proc. Natl. Acad. Sci. USA. 110:E1743-E1751.

742

743

744 **Figure Legends**

745 Figure 1. Divergence in trait means when allele effect sizes vary from SPA to SRA and
746 $n_{tot} = n_{opt}$ (A) vs. when all alleles are prone to swamping ($\phi = 0.1$, $\alpha = 0.01$) and $n_{tot} \geq n_{opt}$
747 (B). The predicted critical migration thresholds for alleles of different sizes are plotted in
748 orange in (A), with the dashed line indicating the threshold for $m = 0.005$. For the
749 simulations shown in (B), the genetic variance within populations is hardly affected by
750 migration and environmental heterogeneity, as similar V_G is maintained for homogenous
751 ($\Theta_1 = 0$; $\Theta_2 = 0$) environments when $m = 0.5$ or $m = 0.005$, and in heterogeneous
752 environments ($\Theta_1 = -1$; $\Theta_2 = +1$) when $m = 0.005$ (C). By contrast, when alleles are
753 resistant to swamping ($\phi = 0.5$, $\alpha = 0.1$), there are large effects of migration and
754 environmental heterogeneity on standing genetic variance (D). In all cases $N = 1000$ and
755 $m = 0.005$.

756

757 Figure 2. Proportion of individuals with phenotypes that fall beyond the local optimum (Z
758 $> |\Theta|$), shown at the end of the simulations (A) and in the first 5000 generations of the
759 simulations run with $n_{tot} = 5000$, N = 1000, $\mu = 0.00001$. The red arrow in (B) indicates
760 the last timestep at which there were no individuals with extreme phenotypes. Divergence
761 in trait means shown in both panels for reference; all other parameters are as in Figure
762 1B.

763

764 Figure 3. Comparison between genetic architectures when alleles are prone to swamping
765 (A-C; $\phi = 0.1$ and $\alpha = 0.01$) or resistant to swamping (D-F; $\phi = 0.5$ and $\alpha = 0.1$). A and D
766 show the divergence in phenotypic trait means over the simulation and the last 2000
767 generations over which the data for the remaining panels was collected. B and E show the
768 distribution of contributions to divergence (d) made by individual loci (note the log10
769 scale in E), with red and blue colorations indicating those loci with larger values of d
770 (cutoffs set as +/- $\alpha/2$ and +/- $\alpha/4$). C and F show the values of d from figures B and E,
771 plotted over time at each of the 500 loci in the simulations; white spaces indicate values
772 of $d < |\alpha/4|$. In both simulations, $N = 1000$; in A-C, $\mu = 10^{-4}$; in D-F, $\mu = 10^{-5}$.

773

774 Figure 4. Contributions to divergence (d) for traits with swamping resistant alleles ($\phi =$
775 0.5; $\alpha = 0.1$); Panel A reproduces the results shown in Figure 3F; other panels show
776 results with different values of n_{tot} and μ; all other parameters are as in Figure 3F.
777

778 Figure 5. Effect of n_{tot} and μ on mean and 95th percentile values of F_{ST} (panel A),
779 expected heterozygosity (H_S) and proportion of non-polymorphic loci (panel B), and
780 scaled genetic covariances within (θ_W) and between (θ_W) populations (panel C). All other
781 parameters are as in Figure 1B.

782

783 Figure 6. Schematic representation of predictions about genetic architecture based on the
784 susceptibility to swamping (y-axis) and three factors that can each independently affect
785 mutation-selection-drift balance and the maintenance of genetic variation within
786 populations: number of loci (n_{tot}), effective population size (N_e), and per-locus mutation
787 rate (μ). Ideally, this schematic would separate each of these factors onto its own axis,
788 and include axes for genetic redundancy and LD, but graphical constraints prevent this.

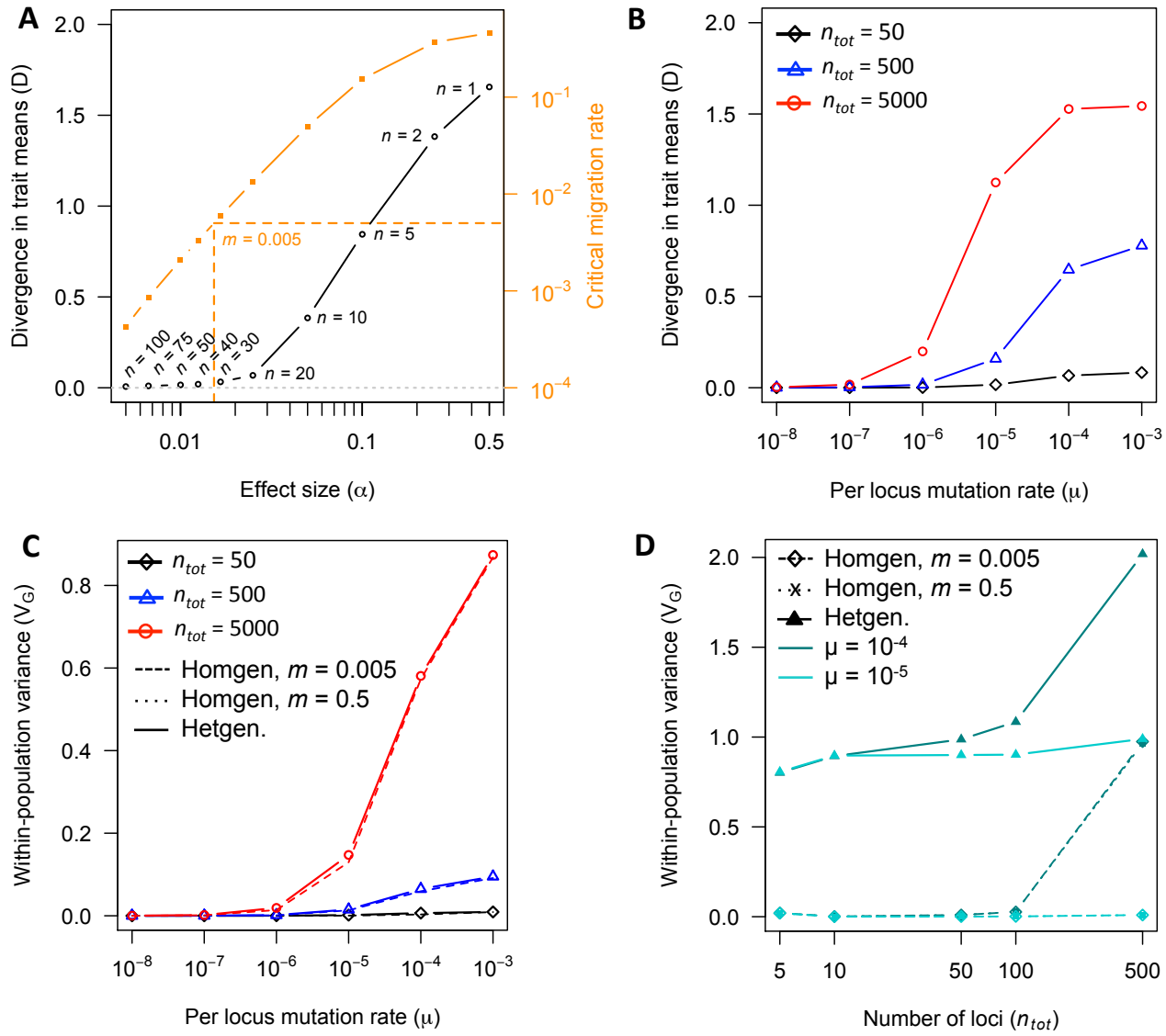
789

790

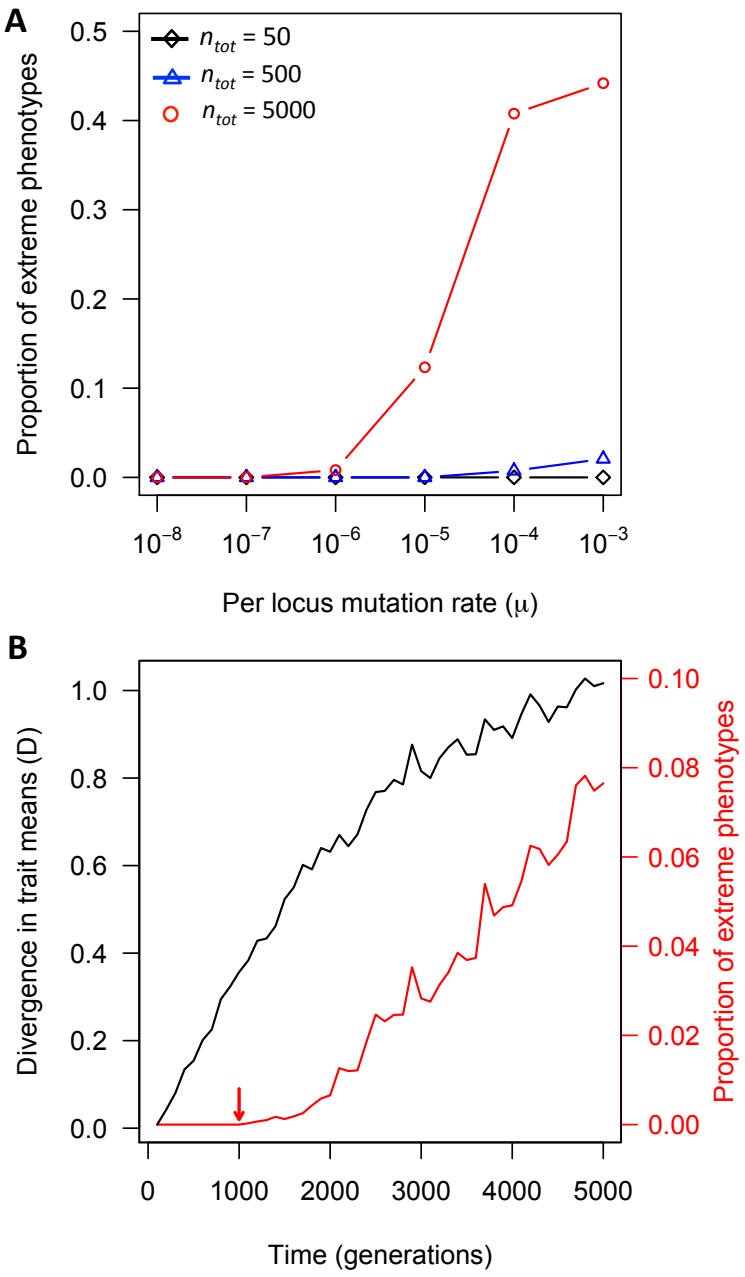
791

792

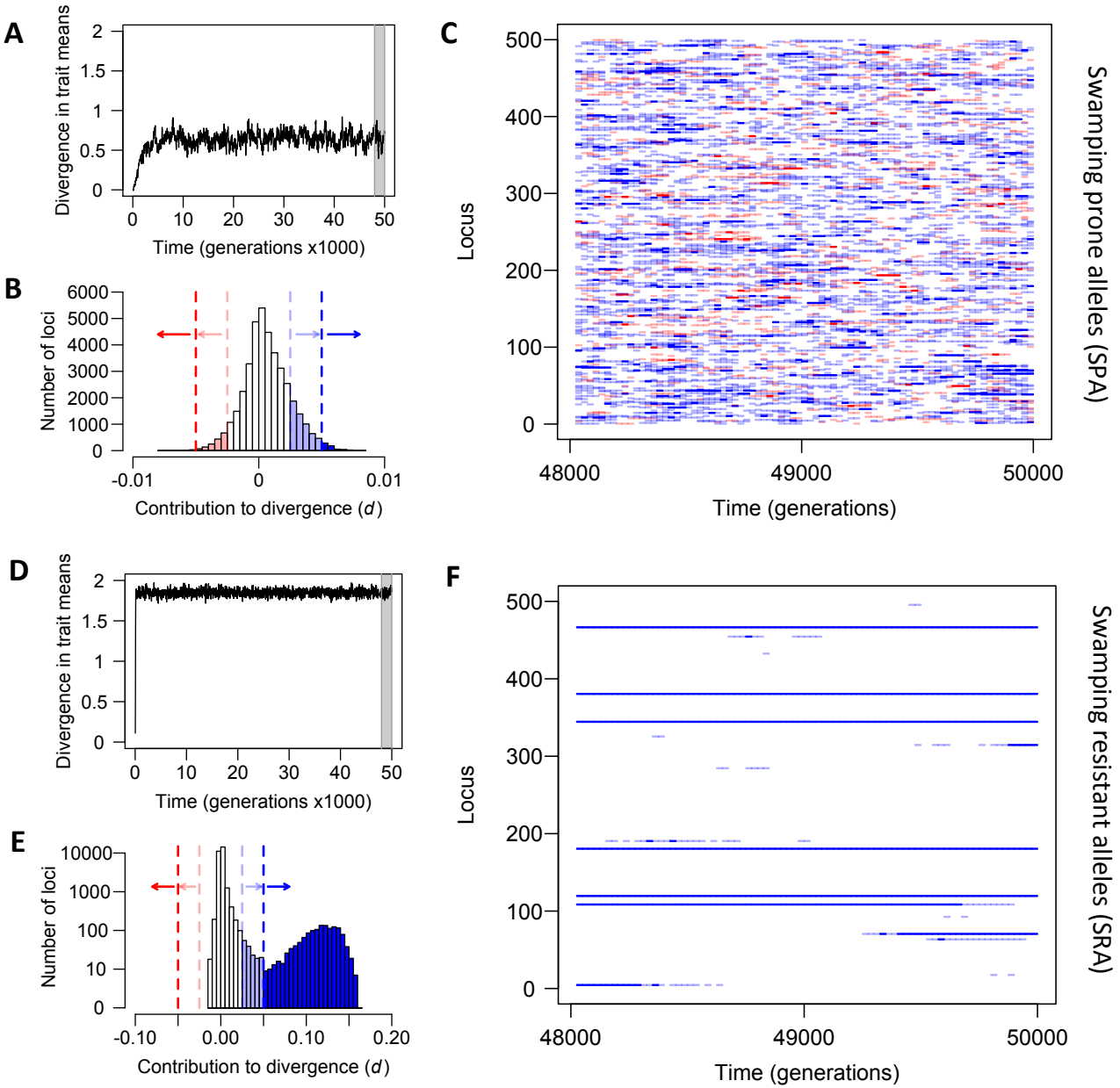
Figure


A
Divergence in trait means (D)
Critical migration rate
Effect size (α)
n = 1
n = 2
n = 5
n = 10
n = 20
n = 30
n = 40
n = 50
n = 75
n = 100
m = 0.005
B
Divergence in trait means (D)
Per locus mutation rate (μ)
n_tot = 50
n_tot = 500
n_tot = 5000
C
Within-population variance (V_G)
Per locus mutation rate (μ)
n_tot = 50
n_tot = 500
n_tot = 5000
Homgen, m = 0.005
Homgen, m = 0.5
Hetgen.
D
Within-population variance (V_G)
Number of loci (n_tot)
Homgen, m = 0.005
Homgen, m = 0.5
Hetgen.
μ = 10^-4
μ = 10^-5

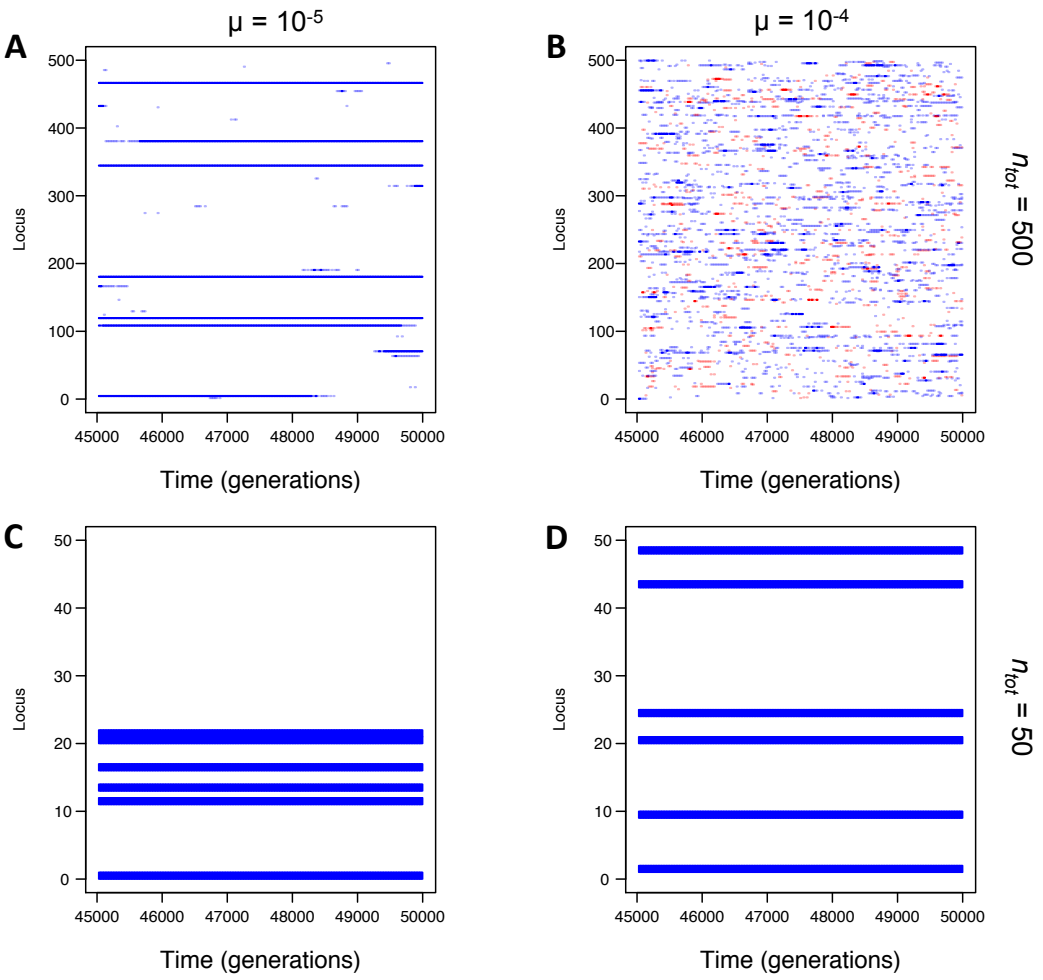
Figure


A
n_{tot} = 50
n_{tot} = 500
n_{tot} = 5000
Proportion of extreme phenotypes
0.0
0.1
0.2
0.3
0.4
0.5
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
Per locus mutation rate (μ)
B
Divergence in trait means (D)
0.0
0.2
0.4
0.6
0.8
1.0
Proportion of extreme phenotypes
0.00
0.02
0.04
0.06
0.08
0.10
0
1000
2000
3000
4000
5000
Time (generations)

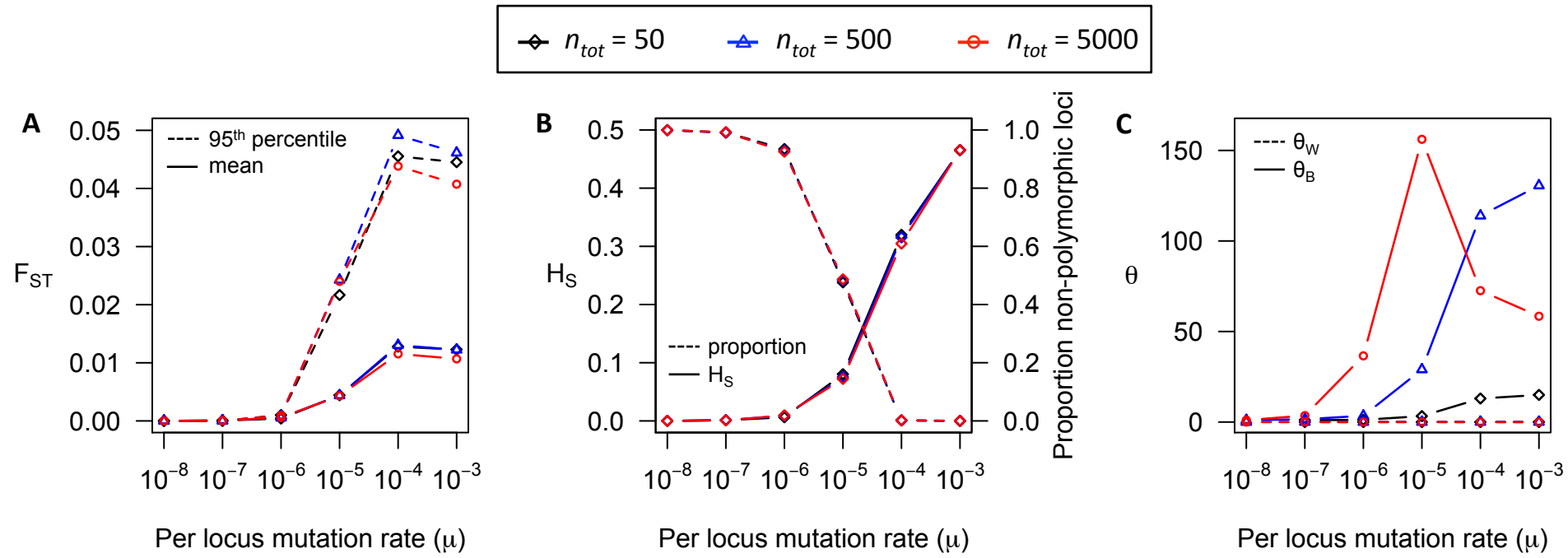
Figure


A
Divergence in trait means
Time (generations x1000)
B
Number of loci
Contribution to divergence (d)
C
Locus
Time (generations)
Swamping prone alleles (SPA)
D
Divergence in trait means
Time (generations x1000)
E
Number of loci
Contribution to divergence (d)
F
Locus
Time (generations)
Swamping resistant alleles (SRA)

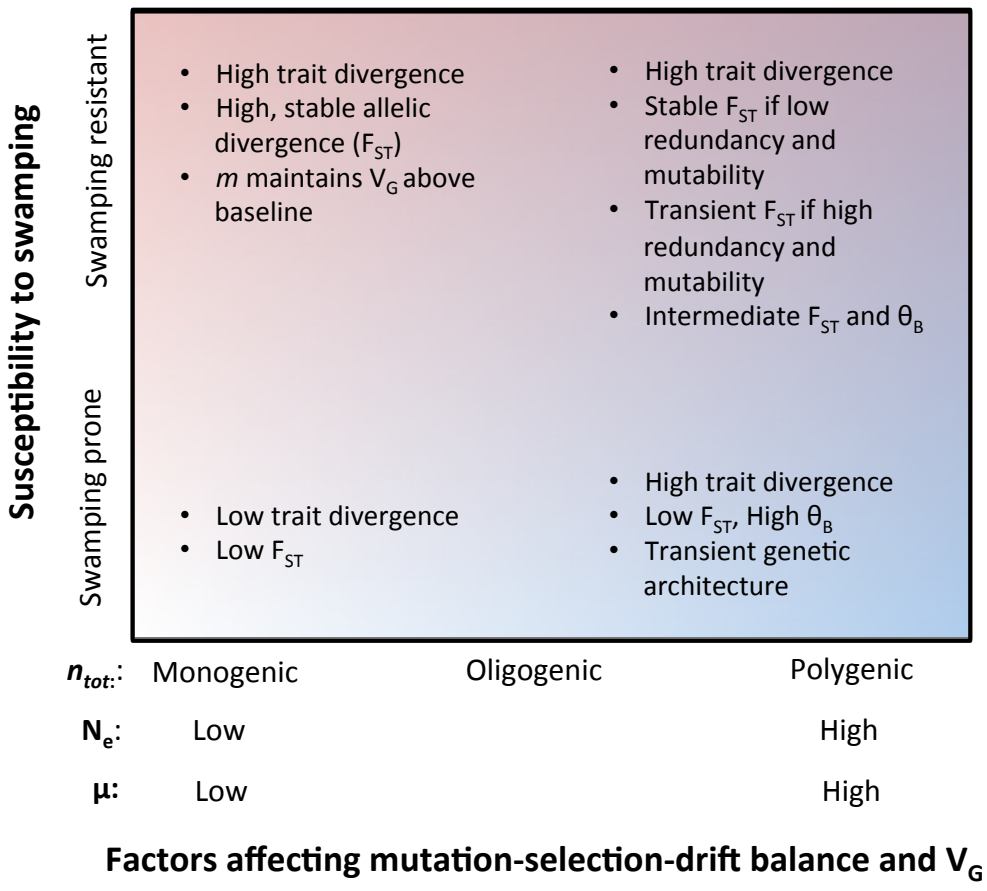
Figure


A
μ = 10-5
B
μ = 10-4
Locus
Time (generations)
ntot = 500
C
D
ntot = 50
0
100
200
300
400
500
10
20
30
40
50
45000
46000
47000
48000
49000
50000

Figure


n_{tot} = 50
n_{tot} = 500
n_{tot} = 5000
A
0.05
0.04
0.03
0.02
0.01
0.00
F_{ST}
---- 95th percentile
— mean
10^{-8} 10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3}
Per locus mutation rate (μ)
B
0.5
0.4
0.3
0.2
0.1
0.0
H_S
1.0
0.8
0.6
0.4
0.2
0.0
Proportion non-polymorphic loci
---- proportion
— H_S
10^{-8} 10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3}
Per locus mutation rate (μ)
C
150
100
50
0
θ
---- θ_W
— θ_B
10^{-8} 10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3}
Per locus mutation rate (μ)

Figure


Susceptibility to swamping
Swamping resistant
Swamping prone
• High trait divergence
• High, stable allelic divergence (F_{ST})
• *m* maintains V_G above baseline
• High trait divergence
• Stable F_{ST} if low redundancy and mutability
• Transient F_{ST} if high redundancy and mutability
• Intermediate F_{ST} and θ_B
• Low trait divergence
• Low F_{ST}
• High trait divergence
• Low F_{ST}, High θ_B
• Transient genetic architecture
n_{tot}: Monogenic Oligogenic Polygenic
N_e: Low High
μ: Low High
Factors affecting mutation-selection-drift balance and V_G