Effect of migration and environmental heterogeneity on the maintenance of quantitative genetic variation: a simulation study

TEGAN KRISTA MCDONALD & SAM YEAMAN
Department of Biological Sciences, University of Calgary, Calgary, Canada

Keywords:
environmental heterogeneity; genetic variation; migration; mutation–selection balance; theory.

Abstract

The paradox of high genetic variation observed in traits under stabilizing selection is a long-standing problem in evolutionary theory, as mutation rates appear too low to explain observed levels of standing genetic variation under classic models of mutation–selection balance. Spatially or temporally heterogeneous environments can maintain more standing genetic variation within populations than homogeneous environments, but it is unclear whether such conditions can resolve the above discrepancy between theory and observation. Here, we use individual-based simulations to explore the effect of various types of environmental heterogeneity on the maintenance of genetic variation ($V_A$) for a quantitative trait under stabilizing selection. We find that $V_A$ is maximized at intermediate migration rates in spatially heterogeneous environments and that the observed patterns are robust to changes in population size. Spatial environmental heterogeneity increased variation by as much as 10-fold over mutation–selection balance alone, whereas pure temporal environmental heterogeneity increased variance by only 45% at max. Our results show that some combinations of spatial heterogeneity and migration can maintain considerably more variation than mutation–selection balance, potentially reconciling the discrepancy between theoretical predictions and empirical observations. However, given the narrow regions of parameter space required for this effect, this is unlikely to provide a general explanation for the maintenance of variation. Nonetheless, our results suggest that habitat fragmentation may affect the maintenance of $V_A$ and thereby reduce the adaptive capacity of populations.

Introduction

As genetic variation is the fundamental basis upon which evolution acts, it is important to understand how variation is maintained in order to provide a foundation for answering various questions in biology and related fields, such as the ‘missing heritability’ problem in medical genetics (Maher, 2008), conservation of biodiversity (Cook & Sgro, 2017) and population potential to respond to change (Houle, 1992). And yet, the relative importance of factors that influence variation and the mechanism(s) under which it is maintained are not well understood (Barton & Turelli, 1989; Mackay et al., 2009). Quantitative traits commonly experience stabilizing selection (Kingsolver et al., 2001), which in theory should erode genetic variation. However, high levels of standing variation and heritability of quantitative traits are consistently observed in nature (Johnson & Barton, 2005). This paradox – a high degree of genetic variation maintained in the face of stabilizing selection – remains a long-standing, unsolved problem in evolutionary biology and quantitative genetics theory.

The most widely studied explanation for this paradox is mutation–selection balance (henceforth referred to as MSB), the appeal of which lies in its intuitive logic: mutation, as the ultimate source of genetic variation, provides enough input to offset the eroding effect of selection, leading to a state of equilibrium. Under such models, stabilizing selection is assumed according to a Gaussian fitness function with parameter $V_S$ setting the
strength of selection on genotypes (where large values result in weaker selection). Multiple MSB models have been proposed, most notably the continuum-of-alleles model from Kimura & Crow (1964) of which two main approximations have been put forth: the Gaussian approximation [Kimura, 1965; later expanded by Lande (1976)] and the house-of-cards approximation (Turelli, 1984; Bürger et al., 1989). The continuum-of-alleles model makes the basic assumption that for a locus with a continuous distribution of possible alleles, mutation can result in a new allele with an effect different from the pre-existing one. In general, the two approximations differ in the way each handles this assumption. For example, for an arbitrary diploid locus \( i \) with an allele of effect \( x_i \) and a mutation with effect \( y \), under the Gaussian approximation the mutated allele would take on a new value which is conditional on the previous state \( (x_i + y) \). This results in a predicted equilibrium genetic variance \( V_G(G) \approx 2\mu \xi^2 V_s \), where \( \xi \) is the total number of loci that compose the quantitative trait, \( \mu \) the per locus mutation rate and \( \xi^2 \) the variance of the distribution of mutational effects. Under the house-of-cards approximation, \( x_i \) may take on any value independent of the previous value, and mutation results in the replacement of \( x_i \) by \( y \). In this case, the predicted variance at equilibrium, \( V_G(HC) \approx 4\mu\xi V_s \), is independent of the distribution of mutation effects. In contrast with the continuum-of-alleles model and its approximations, the diallelic model (Bulmer, 1972; Barton, 1986) assumes only two possible values at a locus with equal forward and backward mutation rates, such that mutation causes a flip between states (e.g. \( x_i = \{0,1\} \)). In this case, the resulting equilibrium variance has been shown to be generally comparable to the house-of-cards approximation \( V_G(HC) \) (Johnson & Barton, 2005).

Although these models have been extensively studied, conflicting evidence over their applicability in relation to realistic biological parameters has left debate open (Barton & Turelli, 1989; Johnson & Barton, 2005; Zhang & Hill, 2005). The assumptions and requirements of these models are unrealistic: primarily, they require extreme mutation rates or too many loci per trait (Johnson & Barton, 2005). This can be most simply illustrated by considering two empirical observations: heritabilities, \( h^2 \), for the majority of traits are relatively high \( (h^2 = 0.2–0.6; Mousseau & Roff, 1987) \), and stabilizing selection is typically relatively strongly in nature, as evidenced by the median value of estimates of reported quadratic selection gradient \( (\gamma = -0.1; \text{Turelli et al., 2001}) \). The value \( \gamma = -0.1 \) implies that the ratio of selection to phenotypic variation is \( V_R/V_P = 5 \) (Kingsolver et al., 2001; Johnson & Barton, 2005), which can be rearranged to account for typical estimates of heritability, yielding an expected empirical range of values for variation maintained in a natural population of approximately \( 0.04 < V_R/V_P < 0.12 \). Putting this in the context of Turelli’s (1984) \( V_G(HC) \) model under the assumption that \( V_G \approx V_A \), then \( 0.01 < n \mu < 0.03 \). If, for example, there are 100 loci underlying a given trait, then this would require per-locus mutation rates on the order of \( 10^{-4} \), which is 10 to 100 times higher than most estimates from quantitative genetic studies, typically thought to be in the range of \( 10^{-5} \) to \( 10^{-5} \) (Barton & Turelli, 1989; Bürger, 2000). Alternatively, given instead 1000 loci underlying a trait, then a per-locus mutation rate on the order of \( 10^{-5} \) satisfies the bounds, and this discrepancy is not an issue. However, this alternative scenario is also not realistic, as pleiotropy puts limits on the number of traits that could be independently under selection, so very few traits could maintain variation in this way (Johnson & Barton, 2005).

More complex extensions of these models have been investigated (Bürger et al., 1989; Zhang & Hill, 2002; Zhang et al., 2002, 2004; Bürger, 2014), namely by extending to multiple traits, such that ‘apparent stabilizing selection’ is generated through pleiotropic effects. Pure pleiotropy and joint effects models have been studied, as well as other extensions to include factors such as dominance or balancing selection. Nonetheless, these models have yet to provide a sufficient explanation for the patterns of variation maintained as suggested by empirical data (Johnson & Barton, 2005; but see Zhang & Hill, 2005). Another extension of mutation–selection balance models is to explore the interaction between spatial structure and genetic redundancy, which occurs when many different genotypes can yield the same phenotype (this is assumed to be operating when stabilizing selection occurs). When there is high genetic redundancy, drift can cause divergence among populations in the combinations of alleles that yield an optimal phenotype under uniform stabilizing selection, and migration can cause modest increases in variance, although the magnitude of this effect is also moderate and unlikely to provide a general explanation (Goldstein & Holsinger 1992; Phillips, 1996).

The above models all assume stabilizing selection to a constant environment, yet environments varying in space and time can also affect the maintenance of variation. If environmental heterogeneity can maintain sufficient differences in allele frequencies within a subdivided population, and migrant individuals introduce novel variants that can be maintained for some time, then migration can result in an increase in genetic variation within a population. For evolution in a quantitative trait, structured populations with limited amounts of gene flow have the potential to increase within-population variance (Slatkin, 1978; Lythgoe, 1997; Barton, 1999; Tufto, 2000; Spichtig & Kawecki, 2004; Huisman & Tufto, 2012; Bürger, 2014), and a temporally fluctuating environment has been shown in some cases to increase variance under particular regions of parameter space (Kondrashov & Yampolsky, 1996; Bürger & Gimelfarb, 2002; Gulisija & Kim, 2015). However, these studies have not explicitly framed results in
terms of the relative increase in variance over mutation–selection to address the problem of the discrepancy between MSB predictions and empirical observations, and the effect(s) of time and space have not been investigated simultaneously for quantitative traits. Results about the maintenance of variation in spatial and temporally heterogeneous environments can also be derived from single- and two-locus population genetic models (Levins, 1964; Felsenstein, 1976; Hedrick, 1986, 2006; Nagylaki & Lou, 2008), but such models do not give much insight into quantitative traits, due to the assumptions involved in extrapolating to many loci under phenotypic selection (as per Spichtig & Kawecki, 2004).

Here, we assess how different kinds of environmental heterogeneity affect the maintenance of genetic variation. We simulate spatial and temporal heterogeneity both independently and simultaneously to explore how the maintenance of variation is affected by variation in migration rate, population size, mutation rate, strength of selection and pattern of environmental variability (Table 1). In all cases, we explicitly focus on comparing heterogeneous vs. homogeneous environments to represent the increase in variance relative to that expected under mutation–selection balance. This provides some indication of how much more variance can be maintained under heterogeneous environments, which can be compared to the discrepancy between the mutation–selection balance predictions and empirical observations described above.

Materials and methods

Simulation set-up

A diploid Wright–Fisher two-patch scenario was modelled using the stochastic, individual-based simulation program Nemo (v2.3.45; Guillaume & Rougemont, 2006) under various parameter combinations of population size \( N \) (1000, 10 000), strength of selection \( V_S \) (2, 5, 10), mutation rate \( \mu \) (\( 10^{-6}, 10^{-5}, 10^{-4} \)) and backwards migration \( m \) (rate ranging from \( 10^{-7} \) to 0.1 per generation) between two equal-sized patches with some local optimum \( \theta \). See Table 1 for detailed reference of each parameter and range of values investigated.

In each simulation, relative fitness is determined by Gaussian stabilizing selection acting on a single quantitative trait of 50 unlinked loci, each with an independent allelic value \( a \) (\( n = 50; \ r = 0.5 \)). Nemo defines individual fitness of a quantitative trait according to the equation:

\[
W(z) = \exp \left[ -\frac{(z - \theta)^2}{2w^2} \right]
\]

where \( z \) is the phenotypic value of the individual, equal to \( \sum (a_{i,1} + a_{i,2}) \) for all \( i \) diploid loci, \( \theta \) the local optimum and \( w^2 \) the strength of selection on the phenotype (\( V_S = w^2 + V_E \)). As no effect of environmental variance was included (\( V_E = 0 \)), hereafter we use \( V_S \) to represent the strength of selection on the genotype to simplify comparison to theoretical models. At initialization of a simulation, each locus is assigned an allelic value \( a_0 \) drawn randomly from an exponential distribution (mean = 0.05). Subsequent mutation causes a change of \( ( -1 ) \cdot a_0 \), such that locus \( i \) has only two possible states, \( +a_0 \) or \( -a_0 \) for the entire simulation (this follows a pseudo-diallelic model with a house-of-cards mutation scheme; note that the diallelic model precludes the gradual evolution of genetic architecture that tends to occur over long periods of time under a continuous mutation model; Yeaman & Whitlock, 2011). A different set of mutation effect sizes is drawn for each simulation replicate within a given parameter set, and the same set of effect sizes was then used for all subsequent parameter sets, to minimize the effect of stochastic outlier alleles on differences among parameter sets. Loci have additive effects on the phenotype, with no dominance or epistasis and free recombination between adjacent loci (\( r = 0.5 \)).

A total of six scenarios were run, each differing in the relation of local optima between patches:

1 Homogeneous patches (\( \theta_1 = \theta_2 = 1 \)). The control set is used for comparison to the other scenarios; the local optimum is the same in each patch such that any differences that arise are due to drift rather than selection.

2 Pure spatial heterogeneity (\( \theta_1 = +1; \ \theta_2 = -1 \)). This set of simulations introduced spatial environmental heterogeneity with opposite local optima between each patch.

3 Pure temporal heterogeneity (\( \theta_{1,t} = \theta_{2,t} \)). To reflect a population inhabiting a changing environment, temporal fluctuations were modelled using an oscillating optimum defined by a sine wave centred about zero with amplitude of 1; \( \theta(t) = \sin(\frac{\pi t}{P}) \). Simulations were run with varying periodicity (\( P = 4, 16, 40, 100 \),

| Parameter | Definition | Value |
|-----------|------------|-------|
| \( N \)  | Population size | 1000, 10 000 |
| \( V_S \) | Selection strength | 2, 5, 10 |
| \( \mu \) | Mutation rate | \( 10^{-6}, 10^{-5}, 10^{-4} \) |
| \( m \)  | Backwards migration rate | 0.00001 – 0.1 |
| \( \theta \) | Local (patch) optimum | 4, 16, 40, 100, 1000 |
| \( P \)  | Periodicity, time (as the number of generations) | over which temporal oscillations complete one cycle of a sinusoidal wave |
1000). The main focus of the analysis considers a temporal oscillation over 100 generations, unless otherwise specified.

4 Combined (spatial & temporal) heterogeneity. The last set of simulations combined the set-up of (2) and (3). Three subsets were used to represent differing degrees of spatial variation: (A, $\theta_{1,1} = -\theta_{2,2}$) opposing optima reflected about the horizontal axis; (B, $\theta_{1,1} = \theta_{2,2} + 2$) in phase temporal fluctuations with one patch oscillating about $-1$ and the other about $+1$; and (C, $\theta_{1,1} = \theta_{2,2}^{(P/A)}$) a partial phase shift between optima functions, such that $\theta_{2,2}$ lags behind $\theta_{1,1}$ (see Fig. S1 for illustration).

All simulations were run for an initial period of 40 000 generations under these conditions, allowing the trait value to stabilize at an approximate equilibrium (assessed visually for the absence of change in trajectory of the trait value over time). Following selection and migration, a census of the population was taken every 100 generations ($t = 100$, or $t = 5$, for temporal simulations with $P \leq 100$ generations) for the mean trait value and genetic variation (in terms of $V_A$, the additive genetic variation) present in each patch. These data were averaged over the stable period of the last 10 000 generations (1000 for temporal simulations with $P \leq 100$) for 40 replicates. For consistency, a 40x50 matrix of allele values was generated such that each replicate had a different complement of mutation effect sizes, but that this complement was held constant across all simulation scenarios. Under the simulations that included temporal heterogeneity, preliminary results showed a cycling pattern emerge for both the trait value and $V_A$ over time; to use an estimate as representative as possible, the $V_A$ maintained within a cycle was sampled at 20 evenly spaced time points within a cycle, and the mean was calculated over the last ten cycles of the simulation. To obtain a measure of variation maintained within the simulated population that is comparable to empirical evidence, the ratio $V_A/V_S$ was calculated from the final mean $V_A$ estimate of each simulation set. All data processing and analysis were performed with R v3.3 (R Development Core Team, 2016).

**Results**

**Pure homogeneity, Set 1 ($\theta_1 = \theta_2 = 1$)**

We first describe patterns in the homogeneous environment, which acts as a control demonstrating the effect of finite population MSB processes in a two-patch model, to study the effects of mutation rate ($\mu$), selection strength ($V_A$) and population size ($N$), on the maintenance of variation (Fig. 1a). When compared to general MSB models (Turelli, 1984; Bürger et al., 1989), the observed genetic variance at equilibrium is on the same order of magnitude as model predictions (See Fig. S2). As expected from quantitative genetic theory (Falconer & Mackay, 1996), variance was somewhat higher in larger populations, although the difference is minimal under most parameters (Fig. S3). Also following expectations, $V_A$ increases with $\mu$, with the effect of mutation rate approximately linear and independent of selection. Similarly, $V_A$ increases with $V_S$ (selection weakens), which qualitatively matches predictions of Turelli (1984) and Bürger et al. (1989), although some quantitative deviations occur from the predictions, likely due to differences in the mutation model (see Figs S2.1 and S2.2 for simulated results under both a diallelic and continuous mutation model). Although from Fig. 1 there appears to be an interaction between selection and mutation rate, the relative effect of a 10-fold increase in mutation (from $\mu = 10^{-5}$ to $\mu = 10^{-4}$) is similar for $V_S = 2$ vs. $V_S = 10$ ($V_A$ changes by a factor of 8.87 vs. 8.59 for $V_S = 2$ vs. $V_S = 10$; $m = 0.1$).

Notably, these patterns are qualitatively independent of $N$ (Fig. S3), and so the results below are presented for the subset of $N = 1000$ only. There is minimal response to rate of migration between the two patches, except for a small increase at low-to-intermediate $m$, which is consistent with predictions of Goldstein & Holsinger (1992), and arises due to an interaction between effects of genetic redundancy and genetic drift (hereafter, referred to as the ‘redundancy-drift effect’; see Discussion).

**Pure spatial heterogeneity, Set 2 ($\theta_1 = +1; \theta_2 = -1$)**

With spatial environmental heterogeneity, two key differences arise when compared to the control case (Set 1): an overall large increase in the amount of $V_A$ maintained, and the appearance of a threshold at high migration, above which the divergence among populations and amount of $V_A$ maintained both decrease [which is qualitatively consistent with the critical migration threshold predicted by Yeaman & Otto (2011)]. Although the migration rate at which $V_A$ peaks varies as a function of $V_S$, the maximum $V_A$ reached under these parameters ($\approx 0.32$) is similar among levels of $V_S$. Given the number of loci, mutation effect sizes and mutation rates that we used, spatial environmental heterogeneity generates up to a ~10-fold increase in $V_A$ compared to the homogeneous case (for the same parameters, $V_A$[homogeneous] = 0.033), as shown in Fig. 1b. This demonstrates that the redundancy-drift effect described previously by Goldstein & Holsinger (1992) is small relative to the effect of migration and spatial heterogeneity, at least under these conditions. Additionally, the effect of mutation becomes relatively less pronounced under spatial heterogeneity, as most of the variation is maintained by migration in this case. Whereas a 10-fold change in $\mu$ (from $\mu = 10^{-5}$ to $\mu = 10^{-4}$) results in an increase by a factor of 8.87 in Fig. 1a, this increase only changes $V_A$ by a
factor of 1.09 under spatial heterogeneity ($V_S = 2, m = 0.1$).

**Pure temporal heterogeneity, Set 3 ($\theta_{t1} = \theta_{t2}$)**

In contrast to the case with spatial heterogeneity, under a temporally fluctuating optimum there is in general little change in $V_A$ in comparison with Set 1, the homogeneous case. As shown in Fig. 2a, a similar qualitative pattern is observed in response to migration rate, with an increase at low-to-intermediate $m$ and rapid drop with $m > 0.001$, akin to the redundancy–drift effect observed in the homogeneous case. The resulting maximum peak in $V_A$ however is approximately 45% greater than the homogeneous case ($V_S = 5, \mu = 10^{-3}, m = 2.5 \times 10^{-4}$), suggesting that temporal

---

**Fig. 1** Effect of migration on the maintenance of genetic variation ($V_A$) in a subdivided population ($N = 1000$) between (a) homogeneous patches and (b) spatially heterogeneous patches, under combinations of different mutation rate (line colour; $\mu = \{10^{-6}, 10^{-5}, 10^{-4}\}$) and selection strength (panels; $V_S = \{2, 5, 10\}$). The homogeneous environment type (broken line) represents the case of mutation–selection balance.
heterogeneity can only marginally increase the importance of this effect.

Different periods ($P$) for the temporal oscillation of the environment were simulated, and complex interactions were found (Fig. S4.1). Setting $P \leq 100$ resulted in a response to migration similar to the homogeneous case and an overall increase in $V_A$ with $\mu$, as expected. However, the behaviour for $P = 100$ deviates under very strong selection ($V_S = 2$), with the redundancy–drift effect causing a peak in $V_A$ at lower migration rates than for shorter periodicities (see Figs 2a and S4.1). Across changes in all parameter combinations, the redundancy–drift effect causes a maximum increase in $V_A$ of 28% relative to the $V_A$ maintained under $m = 0.1$, which was observed for $P = 100$ and $V_S = 10$. Under a very long cycle ($P = 1000$ generations), there

---

**Fig. 2** Effect of migration on the maintenance of genetic variation ($V_A$) in a subdivided population ($N = 1000$) between (a) temporal and (b) combined spatial and temporal environmentally heterogeneous patches under combinations of different mutation rate (line colour; $\mu = \{10^{-6}, 10^{-5}, 10^{-4}\}$) and selection strength (panels; $V_S = \{2, 5, 10\}$). The homogeneous environment type (broken line) represents the case of mutation–selection balance and the period = 100 generations in (a).
is minimal response to migration rate (i.e. no effect of redundancy–drift) and very little increase in variance when $\mu = 10^{-6}$ relative to simulations with shorter periods, but much more increases in variance under long periods at $\mu = 10^{-4}$ (see Fig. S4.1: $V_A$ under $P = 1000$ is 1.3-fold ~2.5-fold higher than $P = 100$ under $\mu = 10^{-4}$). The degree of increase in $V_A$ under temporal heterogeneity relative to homogeneity therefore depends on mutation, period and selection, but is typically considerably less than under spatial heterogeneity with intermediate migration.

**Combination of spatial and temporal heterogeneity, Set 4**

For spatial and temporal heterogeneity combined ($\theta_{t,1} = \theta_{t,2} + 2$), the amount of $V_A$ and patterns that arise are qualitatively similar to those in the case of pure spatial heterogeneity. Similar threshold behaviour is observed, and similar levels of $V_A$ are maintained, peaking between migration rates of 0.01–0.1, again depending on $V_S$ (Fig. 2b). Under this form of combined spatial and temporal heterogeneity, there is a max 31% increase in $V_A$ compared to pure spatial heterogeneity (at $V_S = 5, \mu = 10^{-8}, m = 0.001$), which is very small when compared to the 97.5-fold difference in $V_A$ for the pure spatial heterogeneity vs. homogeneity under the same parameters $V_S$ and $\mu$.

In contrast, the complex patterns involving periodicity described above that were observed in the purely temporal set do not appear to carry over under combined spatial and temporal heterogeneity, likely because they are swamped out by the relatively larger effect of migration, which is relatively consistent across different period lengths (see Supplementary Materials, S4.2). There is an overall increase in $V_A$ maintained, which appears more similar to the case of pure spatial heterogeneity than pure temporal heterogeneity or homogeneity. This result is perhaps unexpected, as spatial and temporal effects of heterogeneity might be expected to work synergistically.

**Comparing sets**

Figure 3 shows a comparison among the six different patterns of environmental variation for a subset of the total parameter combinations used. For a given parameter set, there is a general trend of increased $V_A$ maintained with an increase in heterogeneity, particularly with the degree of spatial heterogeneity between patches. Temporal heterogeneity on its own does little to increase $V_A$ compared to the homogeneous set. When temporal heterogeneity is combined with spatial heterogeneity, the increase in $V_A$ is most pronounced in sets where the temporal fluctuations in the two patches are most out of phase (Fig. 3; see Supplementary Materials, S1 for relation between $\theta_{s,1}$ and $\theta_{s,2}$). Of the three combinations investigated, subset 4B – where $\theta_{s,1} \neq \theta_{s,2}$ for any $t$ – maintains a considerably higher level of $V_A$ compared to subset 4A, which in turn maintained more $V_A$ than in 4C, which was most in phase between patches. This further supports the emphasis on spatial heterogeneity over temporal heterogeneity as a stronger force for maintaining variation.

To explore whether environmental heterogeneity could yield more realistic ratios of $V_A/V_S$ (i.e. 0.04 < $V_A/V_S$ < 0.12), we compared values for this ratio under homogeneous vs. heterogeneous simulation scenarios with all other parameters held constant (Fig. 4). Without any spatial heterogeneity, none of the parameter combinations result in $V_A$ that falls within this range for expected typical values (see Supplementary Materials, S5 and S6, for simulation sets that did not fall into range). However, with spatial heterogeneity, the combination of high migration ($m \geq 0.01$) and strong selection ($V_S = 2$ or 5) does result in $V_A$ within the typical range (i.e. an increase of ~10× relative to MSB). For the combination of spatial and temporal heterogeneity, there is little change for most parameter combinations compared to spatial heterogeneity alone (Fig. 4b); for those where there is a noticeable deviation, whether the parameter set falls in or out of the target range remains unchanged (with an exception at $V_S = 2, m = 0.001$).

Given migration between two patches with different optima, the genetic load observed due to deviation from the optimum (Fig. 5a) and due to genetic variance (Fig. 5b) increases considerably when $m \geq 0.01$ (load, $L = \frac{W_{\text{max}}}{W_{\text{load}}}$). This behaviour is expected, as the migrant individuals are introducing maladapted alleles, which increases both the variance and deviation from the optimum, contributing to load. As optimal local adaptation occurs when at least some alleles are strongly differentiated, this result is also reflected in the differences in allele frequency between patches, which decrease with increasing migration (See Fig. S7). For the homogeneous case, the observed load is minimal ($L$ [homogeneous] $\ll L$[heterogeneous]), and on the order of the genomic mutation rate, as expected from classical theory on mutation load (Bürger, 2000; Fig. S2B). We note that these results were qualitatively indistinguishable for load calculated with $W_{\text{max}} = 1$, because genetic variance is sufficiently high such that at least some nearly optimal genotypes are typically present in the population (See Fig. S8).

**Discussion**

The results of these simulations show that environmental heterogeneity can have a large impact on the maintenance of additive genetic variation ($V_A$) under a variety of parameter values. The direction and magnitude of this impact depend upon the rate of migration between subpopulations, with intermediate rates of
dispersal providing the largest effect. This effect is also dependent on the type of environmental heterogeneity – spatial, temporal or a combination of both all result in different magnitude and pattern of response to migration. Spatial heterogeneity has a much larger effect than temporal, and the effect of combined heterogeneity increases with the degree of spatial difference when temporal oscillations in each patch are more out of phase.

The redundancy–drift effect

For a range of migration rates (dependent on \( V_S \) and \( \mu \)), spatial environmental heterogeneity substantially increases the variation a population can maintain. However, migration alone, even between two environmentally similar patches, can increase the variation maintained at very low rates of migration. This result is due to the interaction of genetic redundancy and the stochastic force of genetic drift, as described by Goldstein & Holsinger (1992). With enough population structure (i.e. limited gene flow), the genotypes existing in each patch are expected to undergo different histories of mutation and drift, resulting in relatively independent fluctuations in allele frequency within each patch, and possibly local fixation or loss. Over time, this results in a different pool of genotypes present in each patch, even if the mean phenotype is equivalent. Migration between patches with different allele frequencies will then increase the variance within patches; movement of individuals is adapted to a given environment with one complement of alleles that result in a similarly fit phenotype, such that it can be maintained in the new population; then, this flux of allelic variants/combinations should increase the variation maintained in the population. This effect is mainly dependent on two factors: (i) the number of possible redundant genotypes, and (ii) the degree of population structure – patches must be sufficiently connected such that new genotypes can be introduced by migration at a high enough rate to affect genetic variance, but not overly connected such that drift is no longer acting (relatively) independently in each patch. This genetic redundancy–drift effect is evident
...in the control case, where there is an observable – albeit minimal – response to low migration, which persists regardless of population size, rate of mutation or strength of selection. This effect is somewhat increased under temporal heterogeneity, resulting in larger peaks over the same range of low dispersal as the homogeneous case, particularly under high mutation and strong selection \((\mu = 0.0001, V_S = 2)\). The increase of this effect may be due to the temporally variable selection within each patch causing different alleles to change in frequency more rapidly than migration homogenizes differences, resulting in a larger pool of differentiated loci than would be generated by drift alone.

**Environmental heterogeneity**

Nonetheless, this redundancy–drift effect is small relative to the effect of spatial heterogeneity. In this case, selection acts to pull each subpopulation towards its optima, resulting in different allele frequencies within each patch. With migration, the alleles that would otherwise be purged by selection are persistently reintroduced into the other patch, increasing the \(V_A\) maintained in the population. However, there is a threshold behaviour to this system – a critical migration rate beyond which polymorphism can no longer be maintained (Felsenstein, 1976; Yeaman & Otto, 2011). Past this point, the allele frequencies of each patch are

---

**Fig. 4** Comparison of the ratio of variance components, \(V_A/V_S\), between simulations of different pattern of environmental heterogeneity for all combinations of migration rate \((m; \text{point colour})\) and selection strength \((V_S = \{2, 5, 10\})\). For each combination of \(m\) and \(V_S\), simulations were run under three mutation rates \((\mu = \{10^{-5}, 10^{-4}, 10^{-3}\})\) to illustrate the full range of results shown in Figs 1–3, but note that these are not differentiated in the plot. The grey box represents the range of predicted ‘typical’ values expected based on empirical estimates of trait heritability \((h^2 = 0.2–0.6)\). Points that fall below the 1:1 line (dashed) are parameter combinations under which more \(V_A\) is maintained without (a) environmental heterogeneity or (b) addition of temporal heterogeneity \((\text{period} = 100)\). In all cases, \(N = 1000\).

**Fig. 5** Genetic load due to (a) deviation from the optimum and (b) genetic variance, observed for spatial heterogeneity (solid) and homogeneous patches (broken) given mutation rate. In all cases, \(\mu = \{10^{-5}, 10^{-4}\}; V_S = 5; N = 1000\).
similar enough that migrants are much less likely to re-introduce novel/differentiated alleles, and the effect of migration on \( V_A \) is reduced.

Most populations experience some form of temporally variable selection, be it seasonal fluctuations or some otherwise unstable environment. Due to this, there has been much interest around fluctuating selection and its effect(s) on the maintenance of variation (Kondrashov & Yampolsky, 1996; Bürger & Gimelfarb, 2002; Siepielski et al., 2009; Morrissey & Hadfield, 2012; Gulisija & Kim, 2015). Difficulties arise when attempting to draw comparisons across such studies due to a variety of different assumptions, such as the way in which the moving optimum is modelled and the nature of the underlying genetic architecture. A simple sinusoidal wave is commonly used, but varying parameters of this function (amplitude, centre, periodicity) appear to have complex interactions with each other as well as other parameters such as mutation rate and genetic redundancy.

Bürger & Gimelfarb (2002) found that in a diallelic multilocus model with recurrent mutation, the relative genetic variance increased with period length (typically \( \geq 24 \) generations), but the magnitude of change and pattern of response to period depended on the amplitude of the oscillations. They observed the highest levels of genetic variance when the amplitude was set such that the optimum cycled between the most extreme possible genotypes, and that greater amplitudes, where the optimum fell beyond the maximum possible genotypic values, resulted in a decline in \( V_A \) at high period lengths (\( \geq 52 \) generations). However, given the small number of loci (2–6) and scaling of allelic effects, this represents a scenario with highly nonredundant genetics. By contrast, Kondrashov & Yampolsky (1996) explored the maintenance of variation under temporal heterogeneity with high genetic redundancy. As empirical data become available for quantitative analyses of a population undergoing temporal variation, this will provide a realistic standard set of parameter values to be used in theory studies, allowing for a more quantitative assessment of the likely effects of temporal fluctuations and further evaluation of the relative importance of the above models.

### Spatial vs. temporal vs. combined heterogeneity

Interestingly, it appears that spatial heterogeneity has a much stronger effect than temporal, both in isolation and when combined. This is evidenced by the difference in magnitude of \( V_A \) maintained in Sets 2 vs. 3 (pure spatial and pure temporal heterogeneity, respectively), as well as the comparisons among the three scenarios of combined spatial and temporal heterogeneity (Set 4 A, B, C). The combination of spatial and temporal heterogeneity only marginally increased the variation maintained by spatial heterogeneity alone, and only in the case of complete spatial heterogeneity (Set 4B). Interestingly, among the three combined sets, \( V_A \) increased with the spatial component (in general, \( V_A[4C] \leq V_A[4A] \ll V_A[4B] \)), which was robust to changes in \( \mu \) and \( V_s \).

Thus, our results indicate that spatial heterogeneity is more important for the maintenance of variation than a temporally fluctuating environment, at least for the genetic architectures and assumptions about fitness used here. Additionally, from the ratio \( V_A/V_s \) (Figs 4 and 5), temporal fluctuations alone fail to result in any simulation maintaining sufficient variation such that it is within the range \( 0.04 \leq V_A/V_s \leq 0.12 \), whereas the scenarios incorporating spatial heterogeneity do result in a shift towards this range. This indicates that it is insufficient on its own as a mechanism to explain the discrepancy between mutation–selection balance models and empirical observations.

### Does heterogeneity maintain sufficient variation to explain empirical observations?

Given empirical estimates of selection and heritability, a realistic amount of variation maintained in a
population should fall within the range $0.04 \leq V_A/V_S \leq 0.12$, which would require either an unrealistically large number of loci or very high mutation rates under models of mutation–selection balance, as described in Introduction. To explore whether heterogeneity could result in more realistic ratios of $V_A/V_S$, we compared simulations run under homogeneous vs. heterogeneous conditions and examined the degree to which heterogeneity increased this ratio. The shift towards this window for simulations with spatial heterogeneity suggests that this is a viable mechanism for the maintenance of variation, but is limited to populations experiencing strong selection and high migration rates. As migration–selection balance is a spatially segregated form of negative frequency dependence, it seems likely that other mechanisms causing frequency dependence within a single population could also maintain considerable variation, such as sexually antagonistic selection (Turelli & Barton, 2004; Barson et al., 2015), nontransitive competition among genotypes (Lively & Sinervo, 1996) or fine-grained environmental heterogeneity (Sokolowski et al., 2007). This could be a particularly potent effect if strong selection acts on a few loci and indirectly protects linked loci from the expected erosion of variance under simple mutation–selection balance (e.g. Patten et al., 2010).

The results presented in Fig. 4 are intended to illustrate how the maintenance of realistic levels of $V_A$ is possible under environmental heterogeneity, rather than to represent the full span of parameter combinations that result in realistic behaviour. Little is known about the true values for many parameters involved, and changes in some values, such as the number of loci, are likely to alter the quantitative results. Rather, these results illustrate that for a given set of assumptions about the genetic basis of a trait, considerably more variation can be maintained by migration–mutation selection balance than mutation–selection balance alone. Moreover, the magnitude of this change can be sufficient to explain empirical observations in some cases, indicating environmental heterogeneity is a viable mechanism to consider for the maintenance of variation. However, we stress that given the sensitivity of this result to the balance between selection and migration and the narrowness of the region that maintains significantly more variance, it seems unlikely that environmental heterogeneity will generally resolve the ‘paradox of variation’.

Although these results are qualitatively quite clear, comparison to established analytical models of mutation–selection balance is somewhat complicated by the differences in mutation scheme. The majority of theory and analytic models assume either diallelic loci of equal effect or quantitative traits with a normally distributed continuum of alleles. Our simulations attempt to combine the two varying allelic effects across loci ($a_i$), with a single $|a|$ per locus, and an exponential distribution of possible values $a_i$. We chose this approach to reflect a (potentially) more biologically realistic scheme of mutation. For example, evidence suggests that the distribution of allelic effects at a locus is more likely to follow a gamma rather than a normal distribution, with a small number of loci of large effect and many loci of small effect (Orr, 2003). However, this leads to a less straightforward parameterization of $V_m$ or $\sigma^2$, which causes challenges when making comparisons to other analytical predictions (e.g. Supplementary Materials, Comparison of simulations to theory). Additionally, we note that altering the shape parameter of the gamma distribution did not substantially affect the maintenance of variation in our model, suggesting that the consequences of these assumptions are relatively small.

Given the relatively narrow range of parameter space over which we observe a significant effect of spatial heterogeneity on variance, it may often be difficult to evaluate whether natural populations commonly fall within this range, given the hurdles involved in quantifying selection and migration in nature (see Moore et al., 2007 for a comprehensive study). Evidence of migration maintaining quantitative genetic variation in spatially heterogeneous environments was found in lodgepole pine (Yeaman & Jarvis, 2006), but genomic signatures of this effect at individual loci may be difficult to detect. Presumably the increase in $V_A$ we observed was driven by increased heterozygosity at the underlying loci. However, given the difficulties involved in detecting signatures of local adaptation at individual loci using genome scans (De Mita et al., 2013; Lotterhos & Whitlock, 2015; Hoban et al., 2016), detecting the secondary effect of migration causing an increase in heterozygosity would likely be still more difficult. We did not explore the effect of migration–selection on heterozygosity at individual loci here, as a rigorous treatment would require extensive exploration of the effect of variation in the underlying genetic architecture and distribution of mutation effect sizes.

**Implications**

We have shown that spatial environmental heterogeneity and migration have the potential to increase genetic variance substantially, although temporal heterogeneity has a much more modest effect. As local adaptation is common (Hereford, 2009), this may be an important factor affecting evolvability. It has been noted that conservation management lacks proper consideration of evolutionary theory, implementation of which can improve overall efforts and long-term outcome (Cook & Sgrò, 2017). The described effect of migration suggests that the ecology underlying the maintenance of variation may be important to consider when planning conservation efforts with a focus on genetics.

For example, a major criterion for the IUCN (International Union for Conservation of Nature) Red List...
status is population size; however, there are examples of listed populations that have remained relatively stable, despite small $N_e$ (e.g. Florida Panthers, *Puma concolor coryi*: Benson *et al.*, 2011). The genetic consequences of population size have been emphasized due to [entirely legitimate] concerns about reduced diversity, which is often used as a proxy measure for the ‘health’ of a population of concern in conservation biology. Here, we showed that a 10-fold change in population size had less effect on $V_A$ than a reduction in migration, which suggests that migration may also be an important factor to consider when assessing and managing the genetic health of populations of concern. Indeed, the related concepts of gene flow and connectivity have been previously suggested for consideration and implementation in conservation management, albeit with some contention, as too much migration can result in a reduction in variability (Weeks *et al.*, 2012; Cook & Sgrò, 2017). Long-term maintenance of evolvability may depend in some cases on an intact and interconnected environment, rather than endogenous generation of variation (mutation-selection), which would depend less on environment. Therefore, a move towards improving connectivity between subpopulations and protection against habitat fragmentation may be a pertinent consideration for the maintenance of quantitative genetic variation, in addition to the well-recognized problems of reducing inbreeding and mitigating demographic fluctuations.

In conclusion, our results demonstrate a clear effect of migration on the maintenance of variation in all four investigated scenarios and highlight the potential for environmental heterogeneity to substantially increase $V_A$. However, it seems unlikely that a single mechanism best explains the maintenance of variation that we see in nature; rather, the many concepts put forth over the past decades may be viable explanations in only certain scenarios, just as has been demonstrated here with environmental heterogeneity. Further investigation into potential mechanisms, particularly into those scenarios under which they succeed and fail, and how biologically relevant successful scenarios may be, is needed to get a complete picture of the fundamental process that is the maintenance of variation.

**Acknowledgments**

We would like to thank F. Guillaume for help with the simulations and M. Turelli for clarifying a theoretical point. We wish to also thank Associate Editor C. Haag, N. Barton and two anonymous reviewers for their constructive feedback on this manuscript. This project was enabled in part by computational support provided by WestGrid and Compute Canada, and was funded by an AIHS grant to SY.

**References**

Barson, N.J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G.H., Fiske, P. *et al.* 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* 528: 405–408.

Barton, N.H. 1986. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genet. Res.* 47: 209–216.

Barton, N.H. 1999. Clines in polygenic traits. *Genet. Res.* 74: 223–236.

Barton, N. & Turelli, M. 1989. Evolutionary quantitative genetics: how little do we know. *Annu. Rev. Genet.* 23: 337–370.

Benson, J.F., Hostetler, J.A., Onorato, D.P., Johnson, W.E., Roelke, M.E., O’Brien, S.J. *et al.* 2011. Intentional genetic introgression influences survival of adults and subadults in a small, inbred felid population. *J. Anim. Ecol.* 80: 958–967.

Bulmer, M.G. 1972. The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genet. Res.* 19: 17–25.

Bürger, R. 2000. The Mathematical Theory of Selection, Recombination, and Mutation, Wiley, Chichester.

Bürger, R. 2014. A survey of migration-selection models in population genetics. *Discrete Continuous Dynamical Syst. Ser. B* 19: 883–959.

Bürger, R. & Gimelfarb, A. 2002. Fluctuating environments and the role of mutation in maintaining quantitative genetic variation. *Genet. Res.* 80: 31–46.

Bürger, R., Wagner, G.P. & Stettinger, F. 1989. How much heritable variance can be maintained in finite populations by mutation selection balance. *Evolution (N.Y.)* 43: 1748–1766.

Cook, C.N. & Sgrò, C.M. 2017. Aligning science and policy to achieve evolutionarily enlightened conservation. *Conserv. Biol.* 31: 501–512.

De Mita, S., Thuilliet, A.C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J. *et al.* 2013. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Mol. Ecol.* 22: 1383–1399.

Falconer, D.S. & Mackay, T.F.C. 1996, *Introduction to Quantitative Genetics*, 4th edn. Longman, Harlow, Essex.

Felsenstein, J. 1976. The theoretical population genetics of variable selection and migration. *Annu. Rev. Genet.* 10: 253–280.

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

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

Gulisija, D. & Kim, Y. 2015. Emergence of long-term balanced polymorphism under cyclic selection of spatially variable magnitude. *Evolution (N.Y.)* 69: 979–992.

Hedrick, P. 1986. Genetic polymorphism in heterogeneous environments: a decade later. *Annu. Rev. Ecol. Syst.* 17: 535–566.

Hedrick, P.W. 2006. Genetic polymorphism in heterogeneous environments: the age of genomics. *Annu. Rev. Ecol. Evol. Syst.* 37: 67–93.
Hereford, J. 2009. A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173**: 579–588.

Hoban, S., Kelley, J.L., Lotterhos, K.E., Antolini, M.F., Bradburd, G., Lowry, D.B. et al. 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am. Nat.* **188**: 379–397.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* **130**: 195–204.

Huisman, J. & Tufto, J. 2012. Comparison of non-Gaussian quantitative genetic models for migration and stabilizing selection. *Evolution* **66**: 3444–3461.

Johnson, T. & Barton, N. 2005. Theoretical models of selection and mutation on quantitative traits. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**: 1411–1425.

Kimura, M. 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. USA* **54**: 731–736.

Kimura, M. & Crow, J.F. 1964. The number of alleles that can be maintained in a finite population. *Genetics* **49**: 725–738.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M.M., Berrigan, D., Vignieri, S.N.N., Hill, C.E.E. et al. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* **157**: 245–261.

Kondrashov, A.S. & Yampolsky, L.Y. 1996. High genetic variability under the balance between symmetric mutation and fluctuating stabilizing selection. *Genet. Res.* **68**: 157–164.

Lande, R. 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**: 221–235.

Levins, R. 1964. Theory of fitness in a heterogeneous environment. *J. Theor. Biol.* **7**: 224–240.

Lively, C.M. & Sinervo, B. 1996. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380**: 240–243.

Lotterhos, K.E. & Whitlock, M.C. 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* **24**: 1031–1046.

Lythgoe, K.A. 1997. Consequences of gene flow in spatially structured populations. *Genet. Res.* **69**: 49–60.

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

Maher, B. 2008. Personal genomes: the case of the missing heritability. *Nature* **456**: 18–21.

Moore, J.-S., Gow, J.L., Taylor, E.B. & Hendry, A.P. 2007. Quantifying the constraining influence of gene flow on adaptive divergence in the lake-stream threespine stickleback system. *Evolution* **61**: 2015–2026.

Morrissey, M.B. & Hadfield, J.D. 2012. Directional selection in temporally replicated studies is remarkably consistent. *Evolution (N.Y.)* **66**: 435–442.

Mousseau, T.A. & Roff, D.A. 1987. Natural selection and the heritability of fitness components. *Heredity (Edinb.)* **59**: 181–197.

Nagylaki, T. & Lou, Y. 2008. *The Dynamics of Migration–Selection Models*, vol. **1922**, pp. 117–170. Springer, Berlin, Heidelberg.

Orr, H.A. 2003. The distribution of fitness effects among beneficial mutations. *Genetics* **163**: 1519–1526.
Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 The change in local optima (θ) over time for each patch in the three sub-scenarios (A, B, C) of simulations with combined spatial and temporal heterogeneity

Figure S2.1 Comparison of genetic variance observed in simulations (filled) to predictions of mutation-selection balance models (the house-of-cards approximation & the Stochastic house-of-cards)

Figure S2.2. Comparison of genetic variance (Vₐ) maintained under a continuous mutation scheme and diallelic mutation scheme (as used for the main results) under homogeneous scenario (θ₁ = θ₂) and spatially heterogeneous patches (θ₁ ≠ θ₂)

Figure S3 Effect of migration on the maintenance of genetic variation (Vₐ) for homogeneous patches (θ₁ = θ₂) with a population size of 1000 (solid line) or 10 000 (dashed line) for different mutation rates (µ) and strength of selection (Vₛ)

Figure S4.1 Effect of migration on the maintenance of genetic variation (Vₐ) for temporally heterogeneous patches

Figure S4.2 Effect of migration on the maintenance of genetic variation (Vₐ) for patches with combined spatial and temporal (θₜ₁ = θₜ₂ + 2) heterogeneity for various periodicities

Figure S5 Vₐ/Vₛ for simulations with temporal heterogeneity compared to simulations with homogeneous patches, for various rates of migration, mutation, and selection strength

Figure S6.1 Vₐ/Vₛ for simulations with combined temporal and partial spatial heterogeneity (set 4A; θₜ₁ = -θₜ₂) compared to simulations with homogeneous patches, for various rates of migration, mutation, and selection strength

Figure S6.2 Vₐ/Vₛ for simulations with combined temporal and partial spatial heterogeneity (set 4C; θₜ₁ = θₜ₂ + Pₜ₃) compared to simulations with homogeneous patches, for various rates of migration, mutation, and selection strength

Figure S7 Genetic Load due to (A) deviation from the optimum (B) genetic variance, assuming Wmax = 1 (i.e., the theoretical optimal phenotype) observed for spatial heterogeneity (solid) and homogeneous patches (broken) given mutation rate µ = {10⁻³, 10⁻⁴}; Vₛ = 5; N = 1000

Figure S8 Difference in frequencies of the ‘+’ allele between patch 1 and patch 2 across the range of effect sizes for increasing rate of migration (Left to Right: m = 0.0001, 0.001, 0.01)

Figure S9 Effect of ordering of lifecycle events on the observed genetic variance

Received 12 September 2017; revised 15 June 2018; accepted 19 June 2018