

Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine

Sam Yeaman^{1,2,*} and Andy Jarvis^{2,3}

¹*Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia V6T 1Z4, Canada*

²*International Plant Genetic Resources Institute, and* ³*International Centre for Tropical Agriculture, AA 6713 Cali, Colombia*

Genetic variation is of fundamental importance to biological evolution, yet we still know very little about how it is maintained in nature. Because many species inhabit heterogeneous environments and have pronounced local adaptations, gene flow between differently adapted populations may be a persistent source of genetic variation within populations. If this migration–selection balance is biologically important then there should be strong correlations between genetic variance within populations and the amount of heterogeneity in the environment surrounding them. Here, we use data from a long-term study of 142 populations of lodgepole pine (*Pinus contorta*) to compare levels of genetic variation in growth response with measures of climatic heterogeneity in the surrounding region. We find that regional heterogeneity explains at least 20% of the variation in genetic variance, suggesting that gene flow and heterogeneous selection may play an important role in maintaining the high levels of genetic variation found within natural populations.

Keywords: genetic diversity; gene flow; heterogeneous selection; cline; lodgepole pine; climate change

1. INTRODUCTION

The maintenance of genetic variation within populations remains a central question in the study of evolutionary biology (Johnson & Barton 2005). Strongly selected traits are expected to have low levels of genetic variance and low heritabilities, due to the purging of non-optimal alleles (Fisher 1930). Empirical evidence, however, shows that variation is ubiquitous. Both natural and experimental populations often have high heritabilities for many traits (Mousseau & Roff 1987; Roff & Mousseau 1987), and traits that are closely associated with fitness tend to have higher levels of genetic variance (but lower heritability) than weakly selected traits (Houle 1992). Understanding the population genetic processes that maintain such high levels of variation is central to the study of evolution.

Numerous theoretical approaches have been taken to address this problem, including models of mutation–selection balance (Lande 1975; Turelli 1984; Bulmer 1989), genotype–environment interactions (Via & Lande 1987; Gillespie & Turelli 1989; Turelli & Barton 2004), population structure and stabilizing selection (Goldstein & Holsinger 1992), temporally fluctuating selection pressures (Waxman & Peck 1999; Burger & Gimelfarb 2002) and pleiotropic overdominance (Bulmer 1973; Gillespie 1984; Turelli & Barton 2004). Despite extensive theoretical research, however, there is still no consensus about which processes are most important and only limited evidence to evaluate this question empirically (Johnson & Barton 2005).

One of the more intuitive mechanisms proposed to maintain genetic variation is through spatially varying selection and gene flow. When populations inhabit environments with different local optima, the reduction

in variation caused by selection within each population can be opposed by gene flow between them. As long as gene flow is insufficient to overwhelm local adaptations and prevent divergence among populations, it will increase the standing genetic variance within populations. While numerous models have examined this migration–selection balance in single locus traits (reviewed in Felsenstein 1976), there have been comparatively few treatments of polygenic traits. Slatkin (1978) showed how gene flow through an environmental gradient increases the additive genetic variance within populations inhabiting intermediate regions of the cline, due to the mixing of alleles from populations with different means. Although Slatkin suggested that this effect would only be important under strong selection, Barton (1999) has shown that gene flow through a cline can result in substantial increases in variance under weak selection as well. Other models based on single populations experiencing stabilizing selection and immigration of non-locally adapted alleles (Tufto 2000) or two populations inhabiting different environments connected by gene flow (Spichtig & Kawecki 2004) have made similar qualitative conclusions, showing that gene flow can maintain genetic variation within populations, but that this effect will only occur under limited migration and strong selection.

Although these theoretical treatments have demonstrated the possibility that gene flow and spatially heterogeneous selection can maintain variation within populations, it is unclear whether this effect makes a substantial contribution to levels of variation in nature. Several studies have shown that gene flow in clines can maintain heterozygosity in single locus traits (Barton & Hewitt 1985; Lenormand & Raymond 2000), but there have been fewer demonstrations of this effect in

* Author for correspondence (yeaman@zoology.ubc.ca).

quantitative traits. Sgro & Blows (2003) examined genetic variance in a cline for development time in wild populations of *Drosophila serrata* and found that non-additive genetic variance was highest at the region with the sharpest change in trait mean. While this evidence conforms to the predictions of Barton's analysis of polygenic clines (1999), it is unclear why patterns of additive genetic variance did not also increase in this transition zone, as predicted by theory. Other studies of quantitative genetic clines have reported ambiguous patterns in genetic variance (van't Land *et al.* 1999; Magiafoglou *et al.* 2002) or have focused on changes in mean rather than variance (e.g. James *et al.* 1995; Hoffmann *et al.* 2002; Palo *et al.* 2003). Part of the reason for this lack of evidence may stem from the difficulty of testing predictions based on genetic variance; the theoretical expectations for patterns of genetic structure in polygenic traits change depending upon the strength of selection and the shape of the change in optimum. For example, under a gentle linear gradient, genetic variance should be relatively constant across the cline, whereas for a steep linear gradient, genetic variance should peak at the centre of the cline (Barton 1999). Because the shape of the environmental gradient is often unknown and because similar patterns could be expected under alternative processes such as localized mutation–selection balance, it is difficult to test hypotheses based on clines in polygenic traits. In a slightly different context, empirical studies of maladaptation caused by gene flow (Riechert 1993; Bossart & Scriber 1995; King & Lawson 1995; Hendry & Taylor 2004) suggest that gene flow can be a persistent source of variance within populations, yet there is still considerable disagreement about the strength of this effect (Hendry & Taylor 2004). Thus, while theoretical research suggests that gene flow through heterogeneous environments may maintain genetic variance, there is little empirical evidence to evaluate the importance of this process.

A general prediction about the effect of gene flow that is more easily tested is that genetic variance within populations should be correlated to regional environmental heterogeneity. Assuming homogenous migration rates and dispersal distances across the range of a species, the variance in trait values carried by alleles flowing into any given population should be proportional to the regional variance in trait means, which should in turn be positively related to the regional variance in environment. This prediction is a direct extrapolation from cline theory, as variance is expected to be higher in populations inhabiting steep clines with more heterogeneity within the region defined by dispersal distance (Slatkin 1978; Barton 1999). Importantly, this prediction will only hold when there is sufficient regional variation in environment to maintain local adaptations among populations (analogous to Slatkin's 'characteristic length' threshold in a linear cline).

To test this prediction, we studied patterns of genetic variation within populations of lodgepole pine (*Pinus contorta*) using previously collected data from a long-term common garden experiment (Rehfeldt *et al.* 1999) and compared these with quantitative estimates of environmental heterogeneity. Lodgepole pine is a long-lived, highly outcrossed conifer (Yeh & Layton 1979) with an extensive range spanning climatically heterogeneous mountain environments and relatively homogenous valleys

and plateaus from Baja California to the Yukon Territories and Alaska. Adaptation to the climatic variations across this range requires tradeoffs between maximizing growth rate and tailoring phenology to avoid exposing sensitive growing tissues to extreme environmental conditions (Howe *et al.* 2003). Like many tree species, neutral genetic variation in lodgepole pine is mostly partitioned within populations (Yang *et al.* 1996; Hamrick 2004). Studies of growth response, however, have found significant patterns of local adaptation and divergence in trait means among populations (Yang *et al.* 1996; Rehfeldt *et al.* 1999; Wu & Ying 2004). Because the environment is sufficiently heterogeneous to maintain divergence between populations in the presence of gene flow, lodgepole pine conforms to the assumptions of the gene flow/environmental heterogeneity hypothesis described above. Thus, if gene flow through heterogeneous environments is an important process maintaining variation, we should see significant correlations between genetic variance for growth response within populations and environmental heterogeneity in the surrounding region. We tested this prediction by comparing levels of genetic variance within populations with quantitative measures of heterogeneity in climatic conditions in their surrounding region.

2. MATERIAL AND METHODS

(a) Genetic variance

Genetic variance was estimated using previously collected data from a common garden experiment testing growth response to climate in 142 populations of lodgepole pine. This experiment was established by Keith Illingworth and maintained by the British Columbia Forest Service; the details are described more thoroughly in other papers (Rehfeldt *et al.* 1999; Wu & Ying 2004). Basic experimental design was as follows: seeds were collected from *ca* 15 different trees in each population (over an area of *ca* 1 km²) and bulked according to their population of origin (Alvin Yanchuk 2005, personal communication). Seeds from each population were then planted in several common gardens across a range of climatic conditions, with two randomized blocks of nine individuals in each site (Rehfeldt *et al.* 1999). Due to the limitations of such an extensive study, populations were not planted in the full set of 60 common gardens used for testing, but were planted in *ca* 33 different sites on average. Following 20 years of growth, the heights of all surviving individuals were recorded.

For the analysis in this study, we excluded any trials with fewer than 10 surviving individuals (out of 18) to minimize the error in estimating variance in small samples. Similarly, we also excluded populations that had been planted in fewer than five locations. Following these restrictions, 103 populations remained with an average of 28 planting sites per population (s.d. = 13.9). For each population, we calculated the mean and variance in 20-year growth response using LSMEANS and VARCOMP in SAS with a restricted maximum-likelihood model to control for block effects in the planting design. These measures of variance showed a weak correlation with the population means ($r=0.29$, $p<0.01$, d.f. = 97). In attempts to correct for this correlation, we tried three different transformations: log-transforming the raw height data, standardizing the final variances by the population means and standardizing the standard deviations by the population means. Each of these transformations, however, yielded estimates of variance that

had stronger, negative correlations with the population means ($r_{\log} = -0.68$; $r_{\sigma^2/\mu} = -0.61$; $r_{\sigma/\mu} = -0.87$), so for all subsequent analysis we used the original untransformed measures of variance.

A major limitation facing studies investigating the maintenance of variation is the difficulty of estimating genetic variance for a large number of populations. Here, we use phenotypic variance (V_P) measured from multiple common garden trials as a proxy for genetic variance (V_G) within each population. While imperfect, this approach may provide a reasonable estimation of genetic variance, assuming that environmental variance (V_E) is either positively correlated with V_G or is roughly constant among populations planted within the same test environment. Randomized planting design and the large number of replicate test sites across a broad range of environmental conditions (mean = 33) should reduce the likelihood of consistent differences in V_E due to experimental design and/or genotype \times environment interactions.

(b) Environmental heterogeneity

To represent patterns of spatial environmental variation, we constructed geographical information system (GIS)-based maps of precipitation and temperature (table 1). Each of these datasets consists of a square lattice with a resolution of approximately 1 km² per cell, covering a total area of roughly 4000 \times 4000 km. This spans the entire range of lodgepole pine, providing an approximate picture of the spatial variation in climatic conditions experienced by the species. All variables were created from 1 km GIS-based rasters in the WORLDCLIM global climatic datasets (Hijmans *et al.* 2004). The original WORLDCLIM dataset consists of maps showing mean monthly estimates of precipitation and minimum and maximum temperature, created by using thin plate splines to interpolate from a global network of weather stations based on distance and altitude. Weather station data typically represent 30 year averages from the period of 1960 to 1990. The WORLDCLIM data do not account for rain shadow effects or other complicated meteorological phenomena but the data should be sufficiently accurate to capture medium- and large-scale climatic patterns and approximate regional heterogeneity. As these climatic variables are continuous representations of the abiotic conditions rather than biologically based models of stress (e.g. incorporating the importance of the transition from freezing to non-freezing temperatures), they are only approximate representations of selection pressure. To increase the biological relevance of these maps and avoid the inclusion of climatic data from uninhabited alpine areas and lowland valleys, we eliminated all areas that are currently non-forested. We identified non-forested areas using maps of coniferous forest from 1992 (Defries *et al.* 2000), excluding any area with less than 10% cover.

For any given cell in the climatic maps, we calculated heterogeneity in the surrounding region using a weighted measure of variance:

$$\sigma^2 = \frac{\sum_{ij} (x_{ij} - \bar{x})^2 m_{ij}}{\sum_{ij} m_{ij}}, \quad (2.1)$$

where i, j are the geographical coordinates of the cells being considered, x_{ij} is the environmental condition at cell $[i, j]$, \bar{x} is the weighted mean for the region and m_{ij} is the weighting at cell $[i, j]$. In this study, the region of analysis is a 201 \times 201 km square surrounding the focal cell and the weights represent the relative probability of gene flow from the surrounding

Table 1. Climatic variables.

variable name	description
<i>drought</i>	number of 1/4 month periods with less than 25 mm of precipitation
<i>precip</i>	mean annual precipitation
<i>cold</i>	coldest mean monthly temperature
<i>temp</i>	mean annual temperature

cells to the focal cell. In lodgepole pine, gene flow is mainly the product of wind-mediated pollen dispersal, with seeds typically moving only short distances (Ennos 1994). While we could find no quantitative data on pollen dispersal curves in lodgepole pine, studies of other wind-pollinated conifers have generally found that pollen flow decreases with distance and is typically leptokurtic (Schuster & Mitton 2000; Robledo-Arnuncio & Gil 2005). Following Austerlitz *et al.* (2004), we used a Weibull distribution to calculate the weights for each cell, as this was the best fitting of all distributions they tested and is also flexible in its shape:

$$m_{ij} = \frac{\beta}{\alpha} \left(\frac{d}{\alpha}\right)^{\beta-1} e^{-(d/\alpha)^\beta}, \quad (2.2)$$

where d is the distance between the focal cell and cell $[i, j]$, α is the scale parameter and β is the shape parameter. For this distribution, α corresponds to the mean distance of gene flow and β describes the degree of leptokurtosis (decreasing values of β result in increasing leptokurtosis). It is important to note that, in this application, the mean distance of gene flow does not correspond exactly to the dispersal distance of pollen, due to the possibility of multi-generational stepping stone migration. As an initial parameter set, we measured heterogeneity under $\alpha = 0.5$ km and $\beta = 1$ (which corresponds to an exponential distribution with mean = 0.5). To examine the effect of different parameters describing the dispersal curve, we applied this method to calculate heterogeneity under a range of values of α (200 m–10 km) and β (0.2–2.0), using an analysis window of 201 \times 201 km. This simple representation of gene flow will fail to account for the influence of physical and phenological barriers to pollen flow and may be less accurate in mountainous areas where differences in flowering time can present effective barriers to gene flow (Schuster *et al.* 1989); it should, however, provide a coarse approximation.

In all applications of the above methods, we calculated heterogeneity for the cell corresponding to the geo-referenced point of origin for each provenance. Four of the provenance geo-reference points corresponded to cells that lay more than 2 km outside of currently forested areas, and these were excluded from the analysis. Following these restrictions, 99 populations had sufficient data for estimates of both genetic diversity and environmental heterogeneity. All variables except *drought* were log-transformed prior to the calculation of weighted variance to reduce the effect of large ranges in magnitude. To facilitate comparisons among variables, all measurements of heterogeneity were standardized to an index between 0 and 1 by subtracting the minimum value from all scores and dividing by the range. To evaluate the relationship between variance within populations and regional heterogeneity, we performed multiple linear regressions and tested Pearson correlations using 'lm' and 'cor.test' in the statistical package R. As a test of an alternative hypothesis, we calculated correlations between genetic variance within populations and

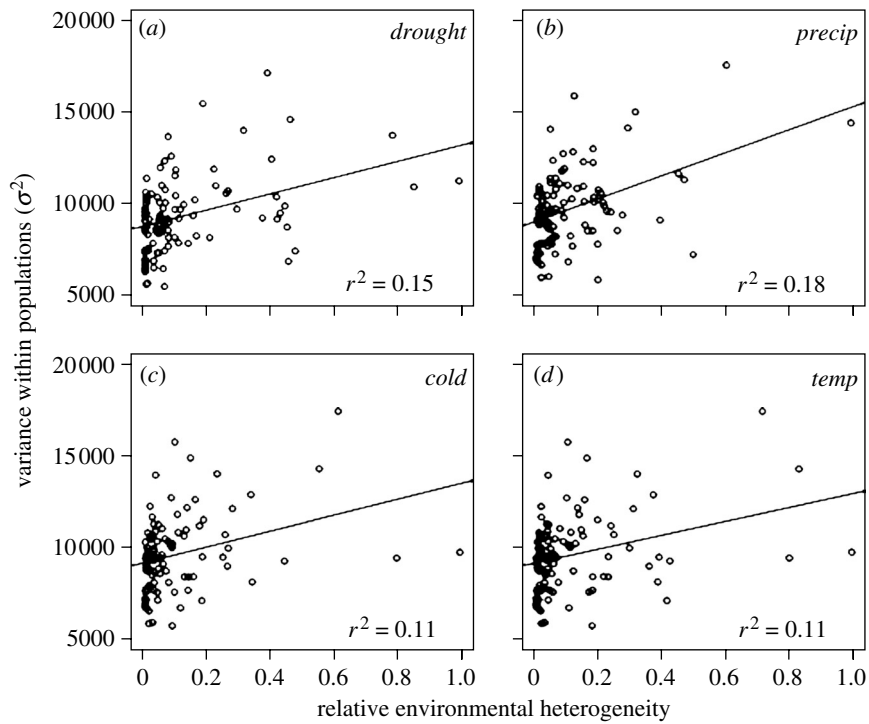


Figure 1. Relationship between variance within populations and regional heterogeneity in the four climatic variables (a) *drought*, (b) *precip*, (c) *cold* and (d) *temp*, calculated using $\alpha = 500$ m, $\beta = 1$ for the Weibull distribution representing gene flow.

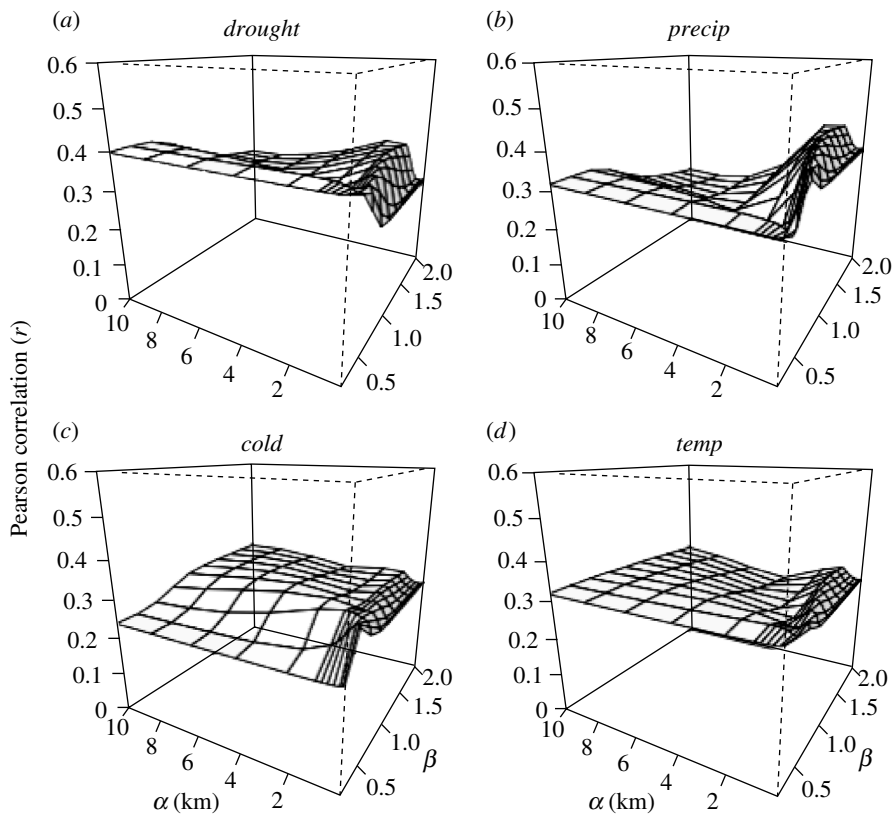


Figure 2. Pearson correlations between variance within populations and heterogeneity in the four climatic variables (a) *drought*, (b) *precip*, (c) *cold* and (d) *temp*, calculated under a range of values of α and β for the Weibull distribution representing gene flow.

both latitude and distance from the centre of the range (defining the centre as the mean of all latitudinal and longitudinal coordinates).

3. RESULTS

Genetic variance in 20-year growth responses within populations was correlated with regional heterogeneity in

all four environmental variables (figure 1). Surprisingly, these correlations were relatively insensitive to variations in the parameters defining the dispersal distance under the Weibull distribution (figure 2). Pearson correlations (r) calculated under all parameter combinations were significantly different from 0 ($p < 0.05$, d.f. = 97) but 95% confidence limits overlapped for all dispersal functions considered (not shown). Thus, although correlations

tended to be somewhat lower under certain parameter combinations (e.g. $\alpha < 1$ and $\beta > 1$ for *drought*, *cold* and *temp*), these were not statistically different from correlations under other combinations.

By calculation of r^2 , regional heterogeneity in *drought* explained 8–20% of the variation in genetic diversity, depending upon the values of α and β (figure 2). Heterogeneity in the other variables explained 7–14% (*cold*), 7–13% (*temp*) and 7–19% (*precip*) of the variation in genetic diversity within populations. Although some variables explained more of the variation than others and all were significantly different from zero ($p < 0.05$, d.f. = 97), differences between them were not statistically significant (all estimates lie within 95% confidence limits of other estimates). Thus, it was not possible to evaluate the relative importance of the different climatic variables used here. Multiple linear regression of genetic variance on heterogeneity in all four variables showed that they could collectively explain between 17 and 28% of the variation in genetic diversity within populations, depending upon the values of α and β .

To examine other possible explanations for patterns in genetic variance, we used linear regression to test whether variance within populations was correlated to either their latitude or their distance from the geographic centre of their range. Both variables showed significant inverse relationships, with latitude explaining 5% ($p < 0.05$, d.f. = 97) and distance explaining 10% of the variation ($p < 0.01$, d.f. = 97) in genetic variance. To test whether these two terms covaried with any of the heterogeneity measurements, we calculated the residuals of the single factor regressions of diversity on latitude and distance and then calculated the correlations between these residuals and the measurements of heterogeneity. In almost all cases, the correlations between genetic variance and heterogeneity were still significant, with heterogeneity explaining 9–18% (*drought*), 4–12% (*cold*), 6–10% (*temp*) and 6–19% (*precip*) of the variation in residuals of the distance regression, and 7–16% (*drought*), 6–12% (*cold*), 7–11% (*temp*) and 4–18% (*precip*) of the variation in residuals of the latitude regression, depending upon the values of α and β . Of these, all correlations were significant ($p < 0.05$, d.f. = 97) except for 6 of the 120 parameter combinations in *cold* versus distance residuals and 3 of the 120 parameter combinations in *precip* versus latitude residuals. We then calculated the residuals from the multiple linear regression of diversity on both latitude and distance. Multiple linear regression of these combined residuals on heterogeneity in all four variables showed that they could collectively explain ca 20% of the variation in genetic diversity within populations ($p < 0.001$, d.f. = 97; for $\alpha = 500$ m and $\beta = 1$). Thus, while latitude and distance from the centre of the range are significantly correlated to genetic variance, for the most part, they explain different portions of the variance from regional environmental heterogeneity.

4. DISCUSSION

Evolutionary theory suggests that quantitative genetic variation within populations can be maintained by spatially varying selection and gene flow (Slatkin 1978; Barton 1999; Tufto 2000; Spichtig & Kawecki 2004). Although local adaptation is pervasive (Hedrick *et al.* 1976;

Linhart & Grant 1996), it is unclear whether gene flow is strong enough to maintain the high levels of genetic variation found in nature. Other processes, such as mutation–selection balance or population bottlenecks, could have a greater effect on levels of genetic variation within populations, obscuring any contribution by gene flow. Evidence from this study shows strong correlation between regional heterogeneity and genetic variance in lodgepole pine ($r^2 \sim 20\%$), suggesting that gene flow and heterogeneous selection are making significant contributions to levels of genetic variation within populations.

As this inference is based on correlation, however, there are two possible alternative explanations that should be addressed. First, it is possible that *temporal* heterogeneity is responsible for maintaining diversity within populations (as per Burger & Gimelfarb 2002). If areas that are more spatially heterogeneous are also more *temporally* heterogeneous, then it would be impossible to evaluate which of these factors is maintaining diversity with the correlation-based approach used here. There is good reason to suspect that *temporal* and spatial heterogeneity in climate would not be well correlated, because other factors such as continentality and oceanic currents can influence *temporal* variations in climate without correlation to spatial influences such as altitude. At the present time, however, there are no long-term climatic records available at a sufficiently fine spatial scale, so this alternative hypothesis cannot be conclusively rejected.

As a second explanation, it is possible that environments within populations are also heterogeneous and that variance within populations is the product of micro-environmental adaptation in their immediate environment. If areas that are regionally heterogeneous also tend to be more heterogeneous at this small scale within populations, then correlations with regional heterogeneity could be an artefact, as described above with respect to *temporal* heterogeneity. Because the datasets we used are interpolated on a coarse 1×1 km scale, they are inappropriate for estimating fine-scale heterogeneity and testing this alternate hypothesis. While local adaptation in Douglas fir (*Pseudotsuga menziesii*) has been found over changes in altitude of only a few hundred metres (Campbell 1979), sampling for the establishment of common gardens in this experiment was conducted over small areas (ca 1 km^2) and care was taken to avoid sampling over obvious sources of environmental variation within a population. Thus, although it was not possible to conclusively test that variance was not being maintained by micro-environmental adaptation within populations, this seems an unlikely explanation for the correlations found in this study.

While theoretically possible, these alternative explanations are less likely than the suggestion that gene flow and regional heterogeneity maintain diversity within populations. Local adaptation has been demonstrated (Yang *et al.* 1996; Rehfeldt *et al.* 1999; Wu & Ying 2004), and gene flow in lodgepole pine is extensive (Perry 1978), so some effect of gene flow on levels of variance is expected. Assuming that our interpretation is correct, gene flow and regional heterogeneity explain approximately 20% of the variation in diversity within populations of lodgepole pine. In fact, due to the potential for errors in the heterogeneity modelling and estimation of genetic variance, the true correlation may be considerably stronger.

This evidence suggests that gene flow and heterogeneity play an important role in maintaining genetic variance, but how general is this effect? Under what conditions would we expect to see maintenance of variation by environmental heterogeneity and gene flow in other species? First, natural selection must be strong enough to maintain variation in trait means between populations, in spite of gene flow. Second, for the increase in variance to be significant, there must be considerable spatial variation in the optimum trait within the effective range of gene flow. Linear cline models show how the environment must change over a characteristic length defined by the strength of selection and distance of gene flow in order for selection to maintain localized adaptations (Slatkin 1978; Barton 1999). Assuming that there is some analogous threshold in more complex heterogeneous environments, there will be some scale of change in environment below which local adaptations are not maintained and genetic variance is unaffected by migration–selection balance. Lodgepole pine is an ideal species in which to detect such an effect, as it has extensive gene flow and inhabits both heterogeneous mountain environments and homogenous plateaus. It remains to be seen whether migration–selection balance plays a significant role in other species, where gene flow is limited and/or environments are more uniform. Unfortunately, the method used here is only applicable when there is substantial variation in the regional heterogeneity; tests based on correlation are powerless to detect an effect when all populations experience equal conditions. It is worth noting that if migration–selection balance plays a significant role in driving levels of diversity, the common statistical assumption of homogeneity of variances across populations may not be warranted when they inhabit environments with different levels of heterogeneity.

Most studies of the impact of gene flow on genetic structure within populations have focused on its maladaptive consequences. Empirical studies have described the impacts of migration load (e.g. Storfer & Sih 1998), gene swamping (e.g. Raymond & Marquine 1994) and outbreeding depression (e.g. Price & Waser 1979), while theoretical studies have noted the limits to local adaptation (e.g. Kirkpatrick & Barton 1997). Here, we have found evidence suggesting that variance within populations can be maintained by gene flow without homogenizing local adaptations and eliminating diversity between populations. While populations of lodgepole pine are not always perfectly locally adapted (Wu & Ying 2004), they are able to persist under this genetic load and even maintain ecological dominance. Although genetic load is always maladaptive in environments that do not change over time, genetic variance is necessary for any response to selection (Fisher 1930); thus, genetic load may be beneficial in times of rapid environmental change. While both mutation and migration can increase variance, migrant alleles from neighbouring populations with slightly different environments are more likely to be adaptive to slight temporal variations in environment than random mutations generated *de novo*. For example, if two populations inhabit wet and dry environments, respectively, and climatic change causes these habitats to reverse their characteristics, allelic variation maintained by gene flow would be inherently adaptive to the novel change, whereas alleles generated by mutation would

often be maladaptive. If migration–selection balance proves to be a widespread and significant factor maintaining variation in quantitative traits, it will be important to consider environmental heterogeneity and gene flow when evaluating conservation and management options, especially when considering adaptation to climatic change. Generally speaking, conserving both heterogeneous landscapes and historical levels of gene flow should maintain diversity within and between populations.

We wish to thank Mike Whitlock for his comments and encouragement and Gerry Rehfeldt for providing the raw data used for the estimates of genetic variance. Throughout the research process, many more than we can list have offered comments and suggestions, but we extend special thanks to SOWD, the Aitken lab, S. Aitken, S. Cook, M. Slatkin, A. Yanchuk, A. Gerstein, and M. Vellend. This work was funded by strategic funds available in the International Centre for Tropical Agriculture (CIAT), the BMZ-funded Sustainable use of Forest Genetic Resources in the Americas project led by the International Plant Genetic Resources Institute (IPGRI) and an NSERC grant to M. Whitlock.

REFERENCES

- Austerlitz, F., Dick, C. W., Dutech, C., Klein, E. K., Oddou-Muratorio, S., Smouse, P. E. & Sork, V. L. 2004 Using genetic markers to estimate the pollen dispersal curve. *Mol. Ecol.* **13**, 937–954. (doi:10.1111/j.1365-294X.2004.02100.x)
- Barton, N. H. 1999 Clines in polygenic traits. *Genet. Res. Camb.* **74**, 223–236.
- Barton, N. H. & Hewitt, G. M. 1985 Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* **16**, 113–148. (doi:10.1146/annurev.es.16.110185.000553)
- Bossart, J. L. & Scriber, J. M. 1995 Maintenance of ecologically significant genetic variation in the tiger swallowtail butterfly through differential selection and gene flow. *Evolution* **49**, 1163–1171.
- Bulmer, M. G. 1973 Maintenance of genetic variability of polygenic characters by heterozygous advantage. *Genet. Res. Camb.* **22**, 9–12.
- Bulmer, M. G. 1989 Maintenance of genetic variability by mutation–selection balance. *Genome* **31**, 761–767.
- Burger, R. & Gimelfarb, A. 2002 Fluctuating environments and the role of mutation in maintaining quantitative genetic variation. *Genet. Res. Camb.* **80**, 31–46.
- Campbell, R. K. 1979 Genecology of Douglas-fir in a watershed in the Oregon cascades. *Ecology* **60**, 1036–1050.
- Defries, R. S., Hansen, M. C., Townshend, J. R. G., Janetos, A. C. & Loveland, T. R. 2000 A new global 1-km dataset of percentage tree cover derived from remote sensing. *Glob. Change Biol.* **6**, 247–254. (doi:10.1046/j.1365-2486.2000.00296.x)
- Ennos, R. A. 1994 Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**, 250–259.
- Felsenstein, J. 1976 The theoretical population genetics of variable selection and migration. *Annu. Rev. Genet.* **10**, 253–280. (doi:10.1146/annurev.ge.10.120176.001345)
- Fisher, R. A. 1930 *The genetical theory of natural selection*. Oxford, UK: Clarendon Press.
- Gillespie, J. H. 1984 Pleiotropic overdominance and the maintenance of genetic variation in polygenic characters. *Genetics* **107**, 321–330.
- Gillespie, J. H. & Turelli, M. 1989 Genotype–environment interactions and the maintenance of polygenic variation. *Genetics* **121**, 129–138.

- 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.
- Hamrick, J. L. 2004 Response of forest trees to global environmental changes. *Forest Ecol. Manag.* **197**, 323–335. (doi:10.1016/j.foreco.2004.05.023)
- Hedrick, P. W., Ginevan, M. E. & Ewing, E. P. 1976 Genetic polymorphism in heterogeneous environments. *Annu. Rev. Ecol. Syst.* **7**, 1–32. (doi:10.1146/annurev.es.07.110176.000245)
- Hendry, A. P. & Taylor, E. B. 2004 How much of the variation in adaptive divergence can be explained by gene flow?—an evaluation using lake-stream stickleback pairs. *Evolution* **58**, 2319–2331.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. 2004 The WorldClim interpolated global terrestrial climate surfaces, v. 1.3. <http://biogeo.berkeley.edu/>.
- Hoffmann, A. A., Anderson, A. & Hallas, R. 2002 Opposing clines for high and low temperature resistance in *Drosophila melanogaster*. *Ecol. Lett.* **5**, 614–618. (doi:10.1046/j.1461-0248.2002.00367.x)
- Houle, D. 1992 Comparing evolvability and variability of quantitative traits. *Genetics* **130**, 195–204.
- Howe, G. T., Aitken, S. N., Neale, D. B., Jermstad, K. D., Wheeler, N. C. & Chen, T. H. H. 2003 From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Can. J. Bot.* **81**, 1247–1266. (doi:10.1139/b03-141)
- James, A. C., Azevedo, R. B. R. & Partridge, L. 1995 Cellular basis and developmental timing in a size cline of *Drosophila melanogaster*. *Genetics* **140**, 659–666.
- Johnson, T. & Barton, N. 2005 Theoretical models of selection and mutation on quantitative traits. *Phil. Trans. R. Soc. B* **360**, 1411–1425. (doi:10.1098/rstb.2005.1667)
- King, R. B. & Lawson, R. 1995 Color pattern variation in Lake Erie water snakes—the role of gene flow. *Evolution* **49**, 885–896.
- Kirkpatrick, M. & Barton, N. H. 1997 Evolution of a species' range. *Am. Nat.* **150**, 1–23. (doi:10.1086/286054)
- Lande, R. 1975 Maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res. Camb.* **26**, 221–235.
- Lenormand, T. & Raymond, M. 2000 Analysis of clines with variable selection and variable migration. *Am. Nat.* **155**, 70–82. (doi:10.1086/303295)
- Linhart, Y. B. & Grant, M. C. 1996 Evolutionary significance of local genetic differentiation in plants. *Annu. Rev. Ecol. Syst.* **27**, 237–277. (doi:10.1146/annurev.ecolsys.27.1.237)
- Magiafoglou, A., Carew, M. E. & Hoffmann, A. A. 2002 Shifting clinal patterns and microsatellite variation in *Drosophila serrata* populations: a comparison of populations near the southern border of the species range. *J. Evol. Biol.* **15**, 763–774. (doi:10.1046/j.1420-9101.2002.00439.x)
- Mousseau, T. A. & Roff, D. A. 1987 Natural selection and the heritability of fitness components. *Heredity* **59**, 181–197.
- Palo, J. U., O'Hara, R. B., Laugen, A. T., Laurila, A., Primmer, C. R. & Merila, J. 2003 Latitudinal divergence of common frog (*Rana temporaria*) life history traits by natural selection: evidence from a comparison of molecular and quantitative genetic data. *Mol. Ecol.* **12**, 1963–1978. (doi:10.1046/j.1365-294X.2003.01865.x)
- Perry, D. A. 1978 Estimate of effective range of pollen dispersal in lodgepole pine (*Pinus contorta* Dougl.). *Ann. Bot.* **42**, 1001–1002.
- Price, M. V. & Waser, N. M. 1979 Pollen dispersal and optimal outcrossing in *Delphinium nelsoni*. *Nature* **277**, 294–297. (doi:10.1038/277294a0)
- Raymond, M. & Marquine, M. 1994 Evolution of insecticide resistance in *Culex pipiens* populations—the Corsican paradox. *J. Evol. Biol.* **7**, 315–337. (doi:10.1046/j.1420-9101.1994.7030315.x)
- Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L. & Hamilton Jr, D. A. 1999 Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecol. Monogr.* **69**, 375–407.
- Riechert, S. E. 1993 Investigation of potential gene flow limitation of behavioral adaptation in an aridlands spider. *Behav. Ecol. Sociobiol.* **32**, 355–363.
- Robledo-Arnuncio, J. J. & Gil, L. 2005 Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. *Heredity* **94**, 13–22. (doi:10.1038/sj.hdy.6800542)
- Roff, D. A. & Mousseau, T. A. 1987 Quantitative genetics and fitness—lessons from *Drosophila*. *Heredity* **58**, 103–118.
- Schuster, W. S. F. & Mitton, J. B. 2000 Paternity and gene dispersal in limber pine (*Pinus flexilis* James). *Heredity* **84**, 348–361. (doi:10.1046/j.1365-2540.2000.00684.x)
- Schuster, W. S., Alles, D. L. & Mitton, J. B. 1989 Gene flow in limber pine: evidence from pollination phenology and genetic differentiation along an elevational transect. *Am. J. Bot.* **76**, 1395–1403.
- Sgro, C. M. & Blows, M. W. 2003 Evolution of additive and nonadditive genetic variance in development time along a cline in *Drosophila serrata*. *Evolution* **57**, 1846–1851.
- Slatkin, M. 1978 Spatial patterns in the distribution of polygenic characters. *J. Theor. Biol.* **70**, 213–228. (doi:10.1016/0022-5193(78)90348-X)
- Spichtig, M. & Kawecki, T. J. 2004 The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *Am. Nat.* **164**, 70–84. (doi:10.1086/421335)
- Storfer, A. & Sih, A. 1998 Gene flow and ineffective antipredator behavior in a stream-breeding salamander. *Evolution* **52**, 558–565.
- Tufto, J. 2000 Quantitative genetic models for the balance between migration and stabilizing selection. *Genet. Res. Camb.* **76**, 285–293.
- Turelli, M. 1984 Heritable genetic variation via mutation selection balance—Lerch's zeta meets the abdominal bristle. *Theor. Popul. Biol.* **25**, 138–193. (doi:10.1016/0040-5809(84)90017-0)
- Turelli, M. & Barton, N. H. 2004 Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and G×E interactions. *Genetics* **166**, 1053–1079. (doi:10.1534/genetics.166.2.1053)
- van't Land, J., van Putten, P., Zwaan, B., Kamping, A. & van Delden, W. 1999 Latitudinal variation in wild populations of *Drosophila melanogaster*: heritabilities and reaction norms. *J. Evol. Biol.* **12**, 222–232. (doi:10.1046/j.1420-9101.1999.00029.x)
- Via, S. & Lande, R. 1987 Evolution of genetic variability in a spatially heterogeneous environment: effects of genotype–environment interaction. *Genet. Res. Camb.* **49**, 147–156.
- Waxman, D. & Peck, J. R. 1999 Sex and adaptation in a changing environment. *Genetics* **153**, 1041–1053.
- Wu, H. X. & Ying, C. C. 2004 Geographic pattern of local optimality in natural populations of lodgepole pine. *Forest Ecol. Manag.* **194**, 177–198. (doi:10.1016/j.foreco.2004.02.017)
- Yang, R. C., Yeh, F. C. & Yanchuk, A. D. 1996 A comparison of isozyme and quantitative genetic variation in *Pinus contorta* ssp. *latifolia* by F–ST. *Genetics* **142**, 1045–1052.
- Yeh, F. C. & Layton, C. 1979 Organization of genetic variability in central and marginal populations of lodgepole pine (*Pinus contorta* spp. *latifolia*). *Can. J. Genet. Cytol.* **21**, 487–503.