Natural selection maintains a single-locus leaf shape cline in Ivyleaf morning glory, *Ipomoea hederacea*

BRANDON E. CAMPITELLI* and JOHN R. STINCHCOMBE*†

*Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON Canada M5S 3B2,
†Center for the Analysis of Genome Evolution and Function, University of Toronto, 25 Willcocks Street, Toronto, ON M5S 3B2, Canada*

Abstract

Clines in phenotypic traits with an underlying genetic basis potentially implicate natural selection. However, neutral evolutionary processes such as random colonization, spatially restricted gene flow, and genetic drift could also result in similar spatial patterns, especially for single-locus traits because of their susceptibility to stochastic events. One way to distinguish between adaptive and neutral mechanisms is to compare the focal trait to neutral genetic loci to determine whether neutral loci demonstrate clinal variation (consistent with a neutral cline), or not. Ivyleaf morning glory, *Ipomoea hederacea*, exhibits a latitudinal cline for a Mendelian leaf shape polymorphism in eastern North America, such that lobed genotypes dominate northern populations and heart-shaped genotypes are restricted to southern populations. Here, we evaluate potential evolutionary mechanisms for this cline by first determining the allele frequencies at the leaf shape locus for 77 populations distributed throughout *I. hederacea's* range and then comparing the geographical pattern at this locus to neutral amplified fragment length polymorphism (AFLP) loci. We detected both significant clinal variation and high genetic differentiation at the leaf shape locus across all populations. In contrast, 99% of the putatively neutral loci do not display clinal variation, and *I. hederacea* populations show very little overall genetic differentiation, suggesting that there is a moderate level of gene flow. In addition, the leaf shape locus was identified as a major F_{ST} outlier experiencing divergent selection, relative to all the AFLP loci. Together, these data strongly suggest that the cline in leaf shape is being maintained by spatially varying natural selection.

Keywords: cline, *Ipomoea hederacea*, leaf shape, natural selection, single-locus polymorphism

Received 29 February 2012; revision received 13 July 2012; accepted 25 July 2012

Introduction

Clines are defined as a gradient in some measurable character over a series of adjacent populations (Huxley 1938; Haldane 1948; Endler 1973). For example, phenotypic clines often involve changes in an observable trait that could result from gradual or abrupt shifts in allelic or genotypic frequencies for the trait of interest (e.g. coat colour in deer mice, Mullen & Hoekstra 2008). While natural selection can produce clines, both theoretical and empirical evidence demonstrate that it is not

required: several neutral processes (e.g. genetic drift in peripheral populations, recent admixture of previously isolated populations, spatially restricted gene flow) can generate a cline, particularly if more than one occurs simultaneously (Endler 1973). For traits governed by one or few genetic loci, clines should be interpreted with caution because they are particularly susceptible to neutral processes (Vasemägi 2006). Accordingly, clines offer a unique opportunity to investigate the relative role of adaptive and neutral evolutionary mechanisms affecting the geographical distribution of traits and allele frequencies (Haldane 1948; Endler 1973). Here, we evaluate the role of selection versus stochastic processes underlying a cline for a leaf shape polymorphism in the

Correspondence: Brandon E. Campitelli, Fax: +1 416 978 5878;
E-mail: brandon.campitelli@utoronto.ca

Ivyleaf morning glory, *Ipomoea hederacea*, first documented by Bright (1998).

Leaf shape is arguably one of the most variable plant characteristics and has been shown to affect major physiological properties that enable plants to occupy a wide variety of environmental conditions, although evidence for this is typically correlational. Global patterns of leaf shape are well documented; for example, plants possessing broad, undissected leaves are consistently found in mesic habitats, whereas those with dissected and lobate leaves tend to dominate open and more arid habitats (Givnish & Vermeij 1976; Givnish 1979, 1988; Lambers *et al.* 2006). Similarly, intraspecific variation in leaf shape may be important for several different processes, including water relations (e.g. Wyatt & Antonovics 1981), thermoregulation (e.g. Gurevitch 1988) and defence against fungal enemies (e.g. Bright 1998) and herbivores (e.g. Bright 1998; Campitelli *et al.* 2008).

Bright (1998) described a latitudinal cline in *I. hederacea* for leaf shape, where northern populations are composed chiefly of lobed individuals, and southern populations are found to be either monomorphic for heart-shaped individuals, or are polymorphic (Fig. 1). Although the molecular genetic basis of this polymorphism is unknown, crosses between morphs have repeatedly revealed that leaf shape segregates in a Mendelian fashion and crosses within a morph never yield recombinants (Elmore 1986; Bright 1998; personal observation). Hence, leaf shape in this species behaves as a single-locus trait where the lobed allele (*L*) expresses partial dominance over the heart allele (*l*) (Elmore 1986; Bright 1998), and heterozygotes (*Ll*) displaying an intermediate phenotype (Fig. 1). Furthermore, Bright & Rausher (2008) measured selection on leaf shape in experimental populations of *I. hederacea* near the clinal boundary and found both directional selection favouring lobed individuals and heterozygote advantage. Despite the genetic variation for leaf shape in *I. hederacea*, its clinal distribution, and the noted importance of leaf form, interpreting the leaf shape cline in *I. hederacea*—especially the predominance of lobed genotypes in Northern populations—as the result of selection remains difficult because of its single-locus basis and the potential for stochastic forces to influence its distribution.

An increasingly common method to assess the adaptive significance of clines is to compare the geographical pattern for the trait or locus of interest to a set of putatively neutral loci (e.g. allozymes, microsatellites, AFLPs and SNPs). Stochastic evolutionary processes (e.g. genetic drift, founder events, limited gene flow and random migration) should affect all loci in the genome (neutral and selected), but selection should produce

locus-specific affects (Cavalli-Sforza 1966; Berry & Kreitman 1993; Vasemägi 2006). If a cline is neutral, neutral markers should exhibit a cline (or population subdivision) comparable to the trait of interest (Lewontin & Krakauer 1973). In contrast, if adjacent populations exhibit differentiation at the locus or trait of interest, but not at neutral loci, this suggests that natural selection on the locus of interest (or a tightly linked locus) is maintaining the observed differentiation (Lewontin & Krakauer 1973; Vasemägi 2006). Because patterns of divergence are not uniform across the genome (Kulathinal *et al.* 2009; Strasburg *et al.* 2009), it is important to compare clines to genome-wide distributions of differentiation (e.g. Kolaczkowski *et al.* 2011; Cheng *et al.* 2012). Consequently, although evolutionary ecological genomics has heavily focused on the molecular genetic basis of traits (Stinchcombe & Hoekstra 2008), an additional important development is that the availability of genome-wide markers in nonmodel systems facilitates evaluating the evolutionary significance of genetic and phenotypic clines.

Here, we compare the leaf shape cline in *I. hederacea* to genome-wide estimates of population structure and subdivision to evaluate the relative role of selective and stochastic forces producing it, by specifically asking: (i) Do allele and genotype frequencies at the leaf shape locus in *I. hederacea* exhibit a latitudinal cline congruent to what Bright (1998) described using phenotypic data?; (ii) When compared to presumably neutral loci, is the leaf shape locus an outlier?; and (iii), What is the pattern of population genetic diversity in *I. hederacea* populations in eastern North America?

Methods

Natural history

Ipomoea hederacea (L.) Jacquin (Convolvulaceae) is an annual vine with a range spanning the southern USA through the mid-Atlantic and mid-western USA. It is unclear whether *I. hederacea* was introduced to North America from tropical America (Strausbaugh & Core 1964; Long & Lakela 1971; Wunderlin 1982), or if it is native (Mohr 1901). Regardless, according to herbaria records, it has inhabited its current range for at least 150 years (Bright 1998). Furthermore, it shows a latitudinal cline in flowering time consistent with adaptive differentiation (Klingaman & Oliver 1996), suggesting it has occupied its current range long enough to experience and respond to selection. It is commonly found in disturbed areas such as roadside ditches and crop fields. It germinates between May and August and persists until the first hard frost, with flowering commencing ~4 weeks postgermination and fruit production
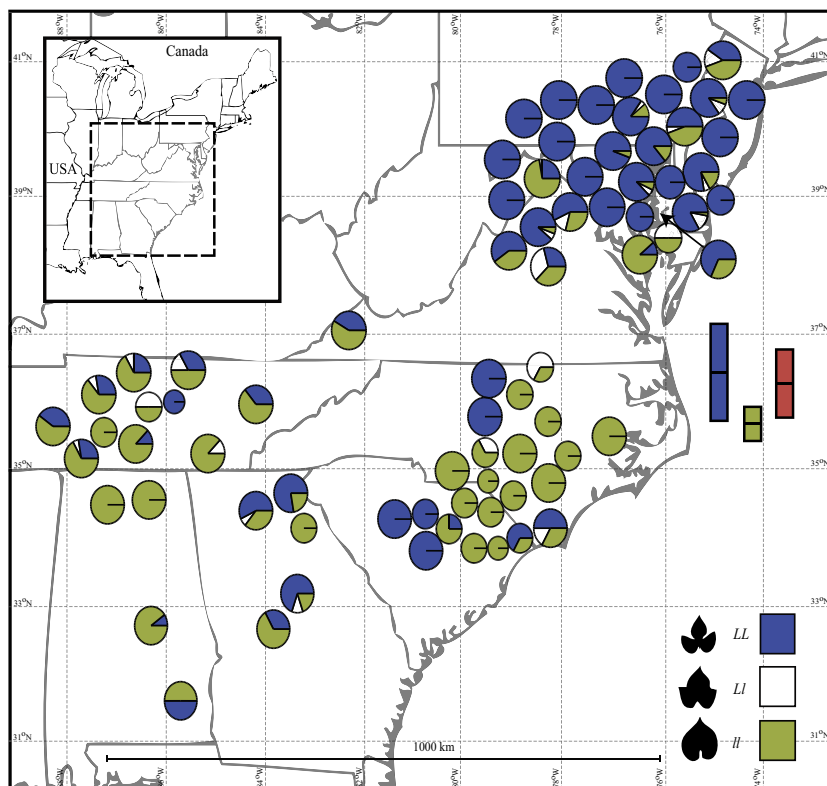
Fig. 1 Map showing the 77 populations sampled. Pies show the fraction of each population occupied by each genotype, and the pie's size indicates the relative number of individuals sampled. *LL* genotypes are homozygous-lobed, *Ll* are heterozygotes, and *ll* are homozygous heart. Boxes represent the centres and associated widths of the clines as determined by Analyse (see text and Fig. 2); far right box represents the allelic cline (both L and l), as they are mirror images of one another. Note: pies that are a single colour with a small horizontal line or internal tick indicate populations that are fixed for that genotype.

beginning shortly thereafter. Although natural seed dispersal is limited to gravity, contemporary dispersal and colonization is thought to occur primarily through agricultural activity (Epperson & Clegg 1986). *Ipomoea hederacea* has a mixed mating system, with populations exhibiting average selfing rates of 66% (Hull-Sanders *et al.* 2005) to as much as 93% (Ennos 1981).

Sampling

We sampled *I. hederacea* individuals from 77 populations (Fig. 1; Table S1) spanning the full clinal range of leaf shape found in Fig. S2 of Bright & Rausher (2008). Our collection consists of 452 maternal seed families we gathered ourselves and 353 seed families gathered and shared by colleagues. We collected seeds from individuals separated by at least 2 m to reduce the chance of collecting from immediate siblings and when possible instructed our colleagues to do the same. When fully matured seeds were not available, we harvested two whole leaves from individuals and preserved them in calcium sulphate desiccant; this totalled 164 individuals from 18 populations, constituting 5–100% (mean = 49%) of the individuals in these populations. Preliminary analyses indicated that the presence of clinal variation in leaf shape was robust to the inclusion or exclusion of these individuals, so they are included below.

Progeny tests to determine leaf shape genotype

To determine the leaf shape genotype of the plants we collected from in the field, we used a combination of progeny testing of up to 10 seeds per line (mean = 8.4) and the rules of Mendelian inheritance to determine genotypes. First, all heart-shaped individuals in the field were assigned *ll*, as genotype can be inferred directly from phenotype. Second, all lobed plants whose progeny included any *ll* individuals were determined to be *Ll*, as this is the only genotype of lobed individuals capable of producing *ll* progeny, regardless of selfing/ outcrossing and the local pollen pool. Third, lobed plants in the field that produced 100% *LL* progeny were assigned as *LL*. To guard against inadvertently categorizing *Ll* progeny as *LL*, we used a liberal visual screen based on the criteria described by Bright & Rausher (2008) to identify putative heterozygotes, followed by a 2nd round of progeny testing by allowing *all* putative heterozygote progeny to self; we selected a random sample of *LL* for further testing as a control. We found 100% *LL* offspring in our controls and used the presence or absence of full segregation in the 2nd round of progeny testing to unambiguously identify first-generation *Ll* and *LL* individuals and hence update our original genotype inferences. A small fraction of parental lines (23/805, or 2.9%) produced progeny (e.g. *Ll* and *LL*, *Ll* and *ll*,or only *Ll*) that did not allow parental

genotypes to be inferred because multiple, equally likely parental genotypes could produce such segregation. These were removed from the final data set, along with 45 families (5.6%) that failed to germinate in the greenhouse, leaving a total of 737 seed families. We provide a full description of specific greenhouse protocols (planting density, temperature, light conditions, watering treatments) in the Appendix S1.

Analysis of the leaf shape cline

We examined allele and genotype frequencies of populations using Analyse version 1.3 (Barton & Baird 1998). Analyse employs a Metropolis-Hastings algorithm to search for maximum-likelihood estimates of cline parameters. We chose a one-dimensional transect model, which fits a *tanh* sigmoidal curve to cline data based on a single coordinate axis, which is appropriate for data sets on single-locus traits (Barton & Baird 1998). We allowed four parameters to vary; cline *width* (distance over which the transition between extremes occurs, defined as the inverse of the maximum slope), cline *centre* (the point where the maximum slope occurs), *pmin* and *pmax* (average allele frequency on either side of the *centre*). To ensure robust model results, prior values for cline width and centre were purposely set to extreme values (i.e. combinations of large or small cline widths coupled with the centre set to the edges of sampling range), and the algorithm was allowed 5000 iterations to obtain the best estimate of the cline parameters (typically <2000 iterations were required for convergence). In total, we tested nine unique combinations of priors to ensure model convergence, and we did this for all three leaf shape genotypes and both alleles individually. We additionally analysed populations along both longitudinal and several angular (south-west to north-east) transects because the populations also span a significant longitudinal range (~13° longitude [1100 km]); there was no longitudinal pattern for any leaf shape genotype or allele, and all angular transects yielded less-significant clinal variation than latitude. Accordingly, we only present results for latitude.

Using width to assess the leaf shape cline

We first compared the estimated width of the cline to the overall sampling range. If the cline width is narrow relative to the sampling range, this suggests a steep transition zone. The opposite pattern (width > range) would indicate an absence of significant clinal variation in leaf shape.

We then used cline width to investigate the strength of selection that is necessary to maintain the observed cline in the face of gene flow using two models. The first model (*Stepped model*; Haldane 1948), assumes dispersal distances comparable to the cline width, and abrupt environmental gradients, and is written as:

$$w = \sigma/\sqrt{s} \qquad (1)$$

where w is cline width, σ is the standard deviation of dispersal distance (SDD) from the adult to offspring stage (i.e. seed dispersal) and s is the selection coefficient. The second model (*Gradient model*; May *et al.* 1975; Endler 1977; also see Fisher 1950) assumes a cline width greater than the potential dispersal distance, and a continuous gradient, and is written as:

$$b = l^2(2.4/w)^3 \qquad (2)$$

where b is the selection gradient, l is the SDD (as above) and w is the width of the cline.

Because we do not have dispersal distance data, we estimated the upper and lower bounds of directional selection (0.18 and 0.15 respectively) against the l allele from Bright & Rausher (2008, their Fig. 4) to approximate a range of SDDs that would be consistent with the maintenance of the cline in the face of the eroding effects of gene flow. We note that heterozygote advantage, as found by Bright & Rausher (2008), is expected to increase the width of clines (Slatkin 1973).

DNA extraction and AFLP genotyping

We extracted DNA from 666 of 737 of our seed lines (due to logistical constraints) representing 76 of the populations (mean = 8.8 individuals/population). We developed amplified fragment length polymorphism (AFLP) markers employing *Eco*RI and *Mse*I as restriction enzymes, following a modified protocol of Vos *et al.* (1995). We submitted all samples for fragment analysis to The Centre for Applied Genomics at Sick Kids Hospital (Toronto, ON, Canada), where they were run on an ABI PRISM® 3730XL Genetic Analyser (Applied Biosystems). For fully detailed protocols, see the Appendix S1 and Table S2 for a list of oligonucleotide sequences for adapters and primers.

AFLP profile generation and error rate calculation

To evaluate error rates, we re-genotyped AFLPs in 48 randomly selected samples. Using the program Raw-Geno 2 (Arrigo *et al.* 2009), we set bin width to 1–1.5 bp and analysed loci in the range of 100–550 bp with a minimum of 100 relative fluorescent units. By restricting our data set to polymorphic loci that were ≥90% reproducible, we retained 173 AFLP loci with a mismatch error rate (calculated as in Bonin *et al.* 2004) of 2.25% (see Appendix S1 for details).

Comparison of leaf shape locus to neutral loci

We compared the leaf shape cline to patterns of differentiation in AFLPs using several tests, differing in their implementation, assumed evolutionary models, and use of geographical information. First, we compared clinal patterns at the leaf shape locus to any clinal patterns in genotype frequencies of the presumably neutral AFLP loci. We calculated the frequency of the *recessive* (absence of a peak) genotypes for each AFLP locus and tested for clinal patterns using Analyse v1.3 (Barton & Baird 1998) using identical settings as were implemented for the leaf shape locus with the following adjustments; for our priors, we only tested a single combination of centre and width (~36°N and 2°latitude, respectively), which closely mimicked our findings for the leaf locus. For each AFLP locus, we compared the cline *centres* and associated *widths* to those of the leaf shape locus to determine whether either of them overlapped.

Second, we compared genetic differentiation at the leaf shape locus to a multilocus estimate for neutral loci to determine whether the leaf shape locus is more or less differentiated than expected under neutrality. To do this, we estimated global population genetic differentiation from the AFLP loci employing Hickory v1.1 (Holsinger *et al.* 2002). Weir & Cockerham's (1984) F_{ST} estimator is referred to as θ^{II} in Hickory; we also estimated θ^{II} at the leaf shape locus for comparison. We accepted default settings in Hickory and report results from the *full-model* because these yielded the best deviance information criterion values for both AFLPs and the leaf locus.

Third, we tested whether the leaf shape locus was an outlier relative to the rest of the genome using DFDIST (Beaumont & Nichols 1996) and BayeScan (Foll & Gaggiotti 2008), two programs designed to identify genomic markers under selection (e.g. Bonin *et al.* 2006; Wood *et al.* 2008; Galindo *et al.* 2009; Tice & Carlon 2011). DFDIST first estimates Weir & Cockerham's (1984) F_{ST} (β in DFDIST) for each locus and then uses a hierarchical Bayesian coalescent approach to estimate the confidence limits (CL) of β, conditional on the expected heterozygosity of each locus. In doing so, it generates a putatively neutral distribution of β and highlights outliers as candidate loci potentially experiencing selection. We instructed DFDIST to simulate 50 000 loci with a distribution of β that had a similar mean as the empirical estimate. We report results from a model with 99% CL, and in which, the absolute mean F_{ST} was set as the target value. Beaumont & Balding (2004) and Caballero *et al.* (2008) recommend implementing a trimmed mean to remove the influence of outlier loci in generating a neutral distribution, using a trimmed mean did not change the results.

BayeScan v2.1 (Foll & Gaggiotti 2008) employs a multinomial-Dirichlet Bayesian model to calculate the probability that each locus is subject to selection. It also allows a false discovery rate (FDR) to be used when determining which loci are deemed significant outliers. We set our parameters to a sample size of 10 000, FDR = 0.05, and a prior odds of a neutral model set to 10. We ran 20 pilot runs at a length of 5000, with an additional burn-in of 100 000 iterations.

For both DFDIST and BayeScan, we report results that include the leaf shape locus in the initial data file by treating each allele at this locus as an individual (*L* = dominant and *l* = recessive, with *Ll* individuals contributing one to each category). We also ran models where we coded *Ll* individuals as either dominant or recessive, or removed them altogether, and although F_{ST} and heterozygosity estimates differed, the patterns were identical. We ran two separate models in each program: (i) a 2-deme model by assigning populations to *north* or *south* according to their geographical location relative to the cline centre determined by Analyse, and (ii) a 76-deme model where each population was a deme.

Fourth, we used a principle coordinates analysis (PCoA) to determine whether more closely related populations shared a similar geographical location (and hence leaf shape). We used GenAlEx 6.4 (Peakall & Smouse 2006) to first calculate Nei's genetic distance between all pairwise combinations of populations and then performed a PCoA using the covariance-standardized option. We assigned populations to the geographical categories *north* or *south* (relative to clinal boundary) and statistically tested whether the northern and southern groups were different by performing a t-test on the principle coordinate scores for PCoA axes 1 and 2.

Because gene flow will affect the maintenance of clines and genetic differentiation among populations, we also used GenAlEx 6.4 to explore patterns of isolation-by-distance (IBD) amongst *I. hederacea* populations. We used the pairwise Nei's genetic distance from above, and geographical distances (GD), and then computed the Pearson product–moment correlation coefficient (R_{xy}) for these matrices (*Nei's D**GD). To test for statistical significance of IBD, we used both Mantel's test and spatial-autocorrelation analysis as implemented by GenAlEx; for the Mantel's test, we used 9999 permutations to estimate the null distribution of R_{xy}.

Genetic diversity of I. hederacea populations

We also tested whether northern populations (which are mostly *LL*) are generally lacking polymorphism or heterozygosity, as would be expected following postglacial

colonization, or whether the leaf shape locus is distinct. We used AFLPSurv v1.0 (Vekemans 2002) to estimate the per cent of polymorphic loci and expected heterozygosity for each population based on AFLP genotypes. We instructed AFLPSurv to employ a Bayesian approach for estimating population allele frequencies using prior knowledge of Wright's inbreeding coefficient (F_{IS}), which is 0.591 for our populations (estimated using Hickory; [Holsinger *et al.* 2002], see above). We then regressed these against both latitude and longitude.

Results

Geographical variation at the leaf shape locus

We found only a few southern populations were fixed for lobed individuals with most being dominated by, or fixed for, heart-shaped individuals. Northern populations were largely composed of lobed individuals (Fig. 1; Table S1). These results parallel those in Bright (1998, Fig. 2.4), which were estimated with herbarium and flora records.

Linear regressions revealed significant clinal variation at the leaf shape locus (Table 1); the frequency of homozygous lobed individuals and the *L* allele increased with latitude, while heart genotypes and the *l* allele decreased. The best fit cline model for each leaf shape genotype and both alleles at the leaf shape locus is illustrated in Fig. 2. Although there is some scatter in the data, these models likewise detected significant clinal variation for all genotypes except heterozygotes (Fig. 2; Table 1). Note that the cline *width* is broader for *LL* compared to *ll* genotypes, which results from there being more *LL* individuals in the south relative to *ll* individuals in the north. In Fig. 1, we show the cline *centre* and associated *width* of the *LL* and *ll* genotypes, as well as each allele, which reveals that the allelic cline is intermediate in centre and width to both genotypes.

Table 1 also provides the results for the estimated SDD for each cline model. Depending on the model, strength of selection and genotype or allele in question, SDD ranged from 21 to 150 km, suggesting that the cline can be maintained in the face of long-distance dispersal.

Comparison of leaf shape locus and neutral loci

A direct comparison of the cline *width* and *centre* of the leaf shape locus to that of each AFLP locus (i.e. AFLP
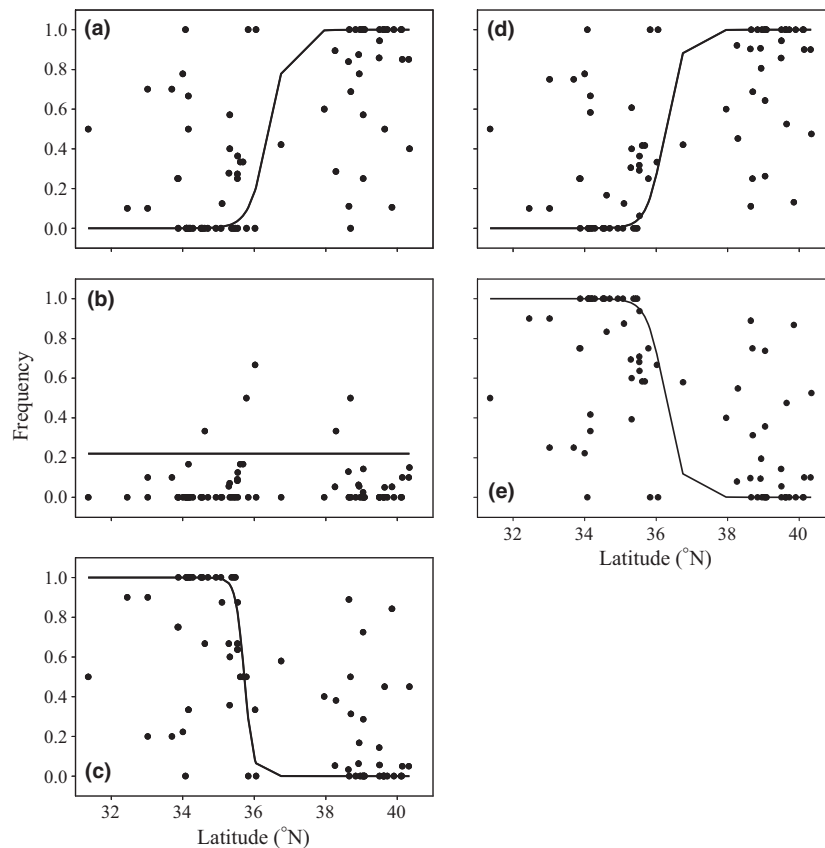
Fig. 2 Best fit cline model for the leaf shape locus: (a) lobed (*LL*), (b) heterozygotes (*Ll*), (c) heart (*ll*), (d) *L* allele, (e) *l* allele.

Table 1 Summary results for the latitudinal cline at the leaf shape locus in *Ipomoea hederacea*

| Trait	Linear R^2	Analyse					Stepped		Gradient	
		Width (km)	Centre (°N)	*LogL*	*Dir*		0.15	0.18	0.15	0.18
L allele	**0.41**	110.63	36.26	−222.8	+		42.84	46.92	121.16	132.72
l allele	**0.41**	110.63	36.26	−222.8	−		42.84	46.92	121.16	132.72
Lobe (*LL*)	**0.39**	120.13	36.41	−34.9	+		46.23	50.97	137.15	150.24
Heart (*ll*)	**0.41**	54.76	35.72	−31.63	−		21.21	23.23	42.21	46.24
Hetero (*Ll*)	0.03	3×10^{11}	-9×10^8	−10.83	N/A		N/A	N/A	N/A	N/A

R^2, linear regression of trait on latitude (boldface indicates $P \leq 0.05$); *width* and *centre,* cline width and geographical centre of cline estimated using Analyse (Barton & Baird 1998); *LogL*, log likelihood; *Dir*, directionality of cline from south to north where (+) is increase and (−) is decrease in frequency; *Stepped* and *Gradient,* estimated standard deviation of dispersal distance employing the given model (see text), where 0.15 and 0.18 represent the lower and upper bounds of the strength of selection on leaf shape (Bright & Rausher 2008).

loci with widths at least as narrow as the broadest estimate of width at the leaf shape locus [*LL* = 120.13 km] and that have a centre between 35.72 and 36.41°N) revealed that only 2 of 173 AFLP loci (1.2%) showed a latitudinal cline similar to the leaf shape locus. Of these two loci, only one (ay_150) shows a marked change in its estimated frequency across the cline (*pmin* and *pmax*; Table S3); however, this locus was not identified as an outlier (Fig. 3). Note that this comparison only considers *LL* and *ll* individuals, as well as each allele separately, because *Ll*s are not clinally distributed (Table 1). Full outputs from Analyse for each AFLP locus are presented in Table S3.

Global population differentiation averaged across all AFLP loci was low but significant (F_{ST} = 0.0047; 95% CI, 0.0029–0.0069). In contrast, genetic differentiation at the leaf shape locus was significantly greater (F_{ST} = 0.59, 95% CI = 0.49–0.68), suggesting appreciably more differentiation than at putatively neutral loci spread throughout the genome.

The leaf shape locus behaves as an outlier relative to the rest of the genome regardless of how we analyse it (Figs 3a,b and S1); it fell outside the bounds of neutral expectations when we split populations based on their *north* or *south* location relative to the leaf shape cline boundary (Fig. 3a,b), or when we employed a 76-deme model (Fig. S1). The leaf shape locus has a high F_{ST} compared to AFLP loci estimated in both DFDIST and BayeScan, and a *q*-value of 0.013, strongly indicating divergent selection. Note, WC's F_{ST} estimator (β) ranges from −1 to 1; however, points below the zero line in Fig. 3a result from sampling error and should be interpreted as showing no differentiation (Weir & Cockerham 1984; also see Galindo *et al*. 2009; Tice & Carlon 2011).

Our PCoA revealed two major clusters of populations separated along PC1 (37.9% of variation; clusters 1 and 2)
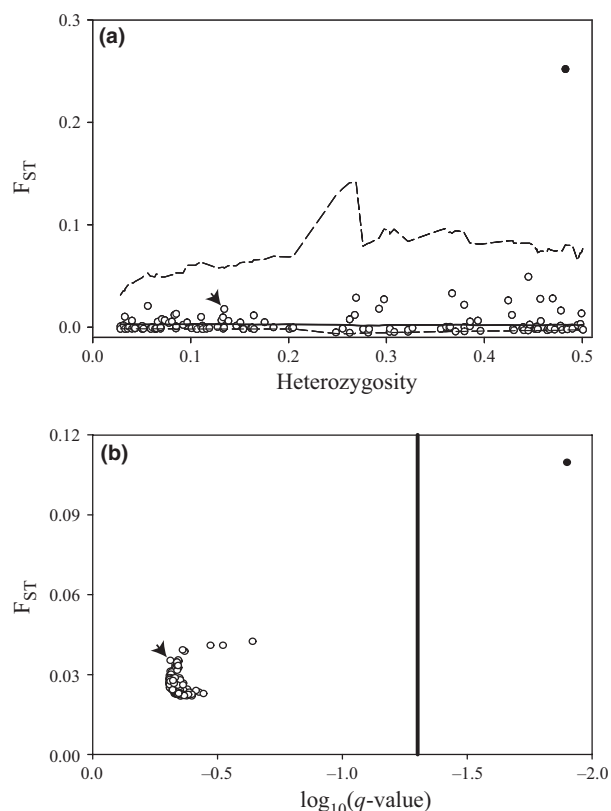
Fig. 3 Outputs of a 2-deme model for (a) DFDIST and (b) BayeScan, where populations are assigned to *north* or *south* based on their position relative to the cline boundary determined by Analyse. In (a), the solid line represents the mean $F_{ST,}$ and dashed lines represent the 99% CLs of the neutral distribution. In (b), the vertical line represents the $\log_{10}$ (*q*-value) corresponding to an false discovery rate (FDR) = 0.05, to the right of which represents potentially selected loci. Filled circles represent the leaf shape locus, and open circles represent amplified fragment length polymorphism (AFLP) loci. The arrowhead indicates AFLP locus ay_150.
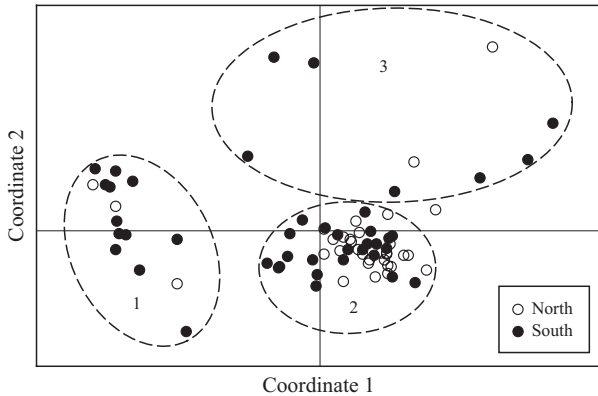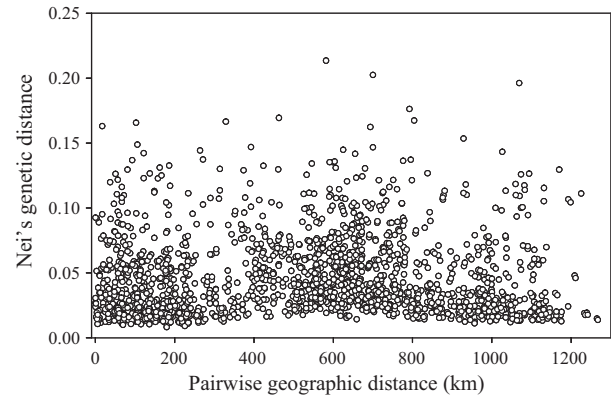
Fig. 4 Principle coordinate analysis of pairwise Nei's genetic distance. Populations are assigned to *north* or *south* based on their location relative to the leaf shape cline boundary.

Fig. 5 Isolation by distance as measured by pairwise Nei's genetic distance plotted against geographical distance.

and possibly a third cluster (cluster 3) separated along PC2 (20.2% of variation; Fig. 4). Although these clusters do not appear to correspond to populations from similar latitudes (*north* and *south* of the clinal boundary; Fig. 4), the mean scores for PC1 are significantly different between *north* and *south* groups ($t = -2.75$, $P = 0.008$), but not for PC2 ($t = 1.35$, $P = 0.18$). This indicates that geographical origin can predict genetic relatedness to some degree; however, supplemental Bayesian clustering analysis by Instruct (Gao *et al.* 2007) did not reveal any clustering or interpretable patterns (see Appendix S1 for model details and Fig. S2). The majority of *north* populations cluster tightly within cluster 2 (28 of 33 populations), whereas *south* populations show a wider spread, suggesting that northern populations may be more genetically similar to one another relative to southern populations. Despite our PCoA results, there was no evidence for isolation by distance in the AFLPs, using either Mantel's test (Fig. 5) or spatial autocorrelation analysis (Appendix S1; Fig. S3). Pairwise values of Nei's genetic distance (Fig. 5) did not show a significant relationship with geographical distance ($R_{xy} = 0.049$, $P = 0.15$), suggesting either widespread gene flow or a population bottleneck during colonization.

Genetic diversity

Neither the proportion of polymorphic loci nor expected heterozygosity show significant relationships with latitude ($R^2 = 0.021$; $P = 0.21$ and $R^2 = 0.024$; $P = 0.18$, respectively; Fig. 6) or longitude (data not shown). Furthermore, all populations display moderate to high levels of polymorphism, demonstrating that northern populations are not lacking polymorphism, as would be expected based on the leaf shape
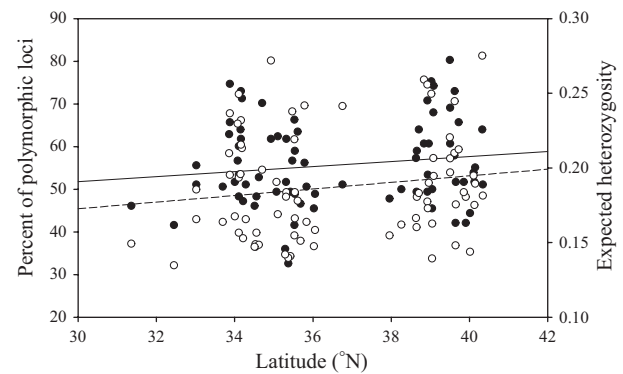
Fig. 6 Geographical patterns of genetic diversity as measured by per cent of polymorphic amplified fragment length polymorphism (AFLP) loci (filled circles and solid line; $R^2 = 0.021$, $P = 0.21$) and expected heterozygosity (open circles and dashed line; $R^2 = 0.024$, $P = 0.18$) for each population.

locus or a model of postglacial colonization (Soltis *et al.* 2006).

Discussion

Our comprehensive analysis of leaf shape genotype and allele frequencies and comparison to presumably neutral AFLP loci in *I. hederacea* revealed three major findings. First, we have shown that both allele and genotype frequencies (except heterozygotes, which are rare) at the leaf shape locus show significant clinal variation. Second, in contrast to patterns at the leaf locus, we found no signal of a latitudinal clines or differentiation in the overwhelming majority of putatively neutral markers we assessed. Third, no geographical pattern of isolation-by-distance (IBD) was found, and northern populations are not lacking genetic diversity. We discuss these results below.

Natural selection and the leaf shape cline

Our results demonstrate a significant latitudinal cline at the major gene governing the leaf shape polymorphism in 77 populations of *I. hederacea*; the frequency of the lobing allele (*L*) increases with latitude leading to northern populations primarily composed of lobed individuals, whereas the heart allele (*l*), and therefore heart shaped individuals, are mostly restricted to the south (Fig. 1). Our results confirm Bright's (1998) description of the cline from flora and herbaria samples, suggesting it has been maintained for at least several decades. Furthermore, the estimated cline *widths* are narrow relative to the overall range of *I. hederacea* (Fig. 2; Table 1), suggesting that natural selection may be maintaining a steep transition zone between leaf shapes despite the potential for gene flow to erode the cline. Classical models (i.e. Haldane 1948; Endler 1977) indicate that the cline can be maintained in the face of long-distance dispersal (up to 150 km).

The selective agents on leaf shape in *I. hederacea* are unknown; however, Bright & Rausher (2008) observed greater deer damage and fungal infection on *ll* genotypes in their experiments. In addition, we have previously shown that insect species can exhibit differential performance when feeding on *LL* and *ll* inbred lines (Campitelli *et al.* 2008). Combined, these studies suggest that leaf shape may affect herbivore and pathogen resistance. Furthermore, ecophysiological arguments (Lambers *et al.* 2006) predict that lobed leaves should remain warmer during cold nights, suggesting that *LL* individuals may dominate the cooler northern habitats because of an improved thermoregulatory ability; tests of these hypotheses are ongoing.

When comparing the leaf shape locus to putatively neutral loci, we discovered that only 2 of 173 AFLP markers exhibited a latitudinal cline similar to the leaf shape locus. Additionally, estimates of population genetic differentiation revealed that the leaf shape locus exceeds the average for all AFLP loci by more than 100-fold. Finally, a detailed analysis of genetic differentiation demonstrated that the leaf locus is consistently highlighted as a major outlier relative to neutral expectations (Fig. 3). Given that a neutral explanation for the leaf shape cline would predict that substantially more AFLP loci would display similar clinal behaviour, these data strongly support the hypothesis that natural selection has produced and/or maintained the leaf shape cline. Although gene-surfing (Klopfstein *et al.* 2006) can produce allele frequency clines, it is unclear why it would do so only for the leaf shape locus and not in AFLPs. Despite its simple genetic basis, the leaf shape cline does not appear to have arisen stochastically.

Several studies have compared clinally varying single-locus phenotypic traits or loci to putatively neutral loci, which are predicted to be particularly sensitive to drift and/or historical processes (Vasemägi 2006). A prevalent pattern across a wide variety of systems and types of markers is that focal loci (and the traits they affect) appear to be more differentiated than neutral makers. Examples include the mannose-6-phosphate isomerase locus in the acorn barnacle, *Semibalanus balanoides* (Schmidt & Rand 1999, 2001), the *MC1r* locus and coat colour in various species of mice (Hoekstra *et al.* 2004; Mullen & Hoekstra 2008), latitudinal clines in the *Clock* gene of several salmon species, (*Oncorhynchus*: O'Malley & Banks 2008; O'Malley *et al.* 2010), flower colour in *Linanthus parryae* (Schemske & Bierzychudek 2007), albumin locus in deer mice, *Peromyscus maniculatus* (Storz & Dubach 2004), various genetic clines in *Drosophila melanogaster* (Berry & Kreitman 1993; Anderson *et al.* 2005; McKechnie *et al.* 2010; Paaby *et al.* 2010) and *D. subobscura* (Prevosti 1964; Prevosti *et al.* 1985). Hence, our results fit into an emerging pattern suggesting either that selection producing single-locus clines is typically quite strong or that stochastic processes producing such clines are rare (or that we have a propensity for choosing loci and traits that experience [or have experienced] selection). For several of these studies, the remaining challenge is to connect clinally varying markers to organismal phenotypes and the ecological mechanisms producing selection, which has been done a handful of times (e.g. Mullen & Hoekstra 2008). For *I. hederacea*, the link between locus and phenotype is clear (because it behaves as a Mendelian polymorphism), although the molecular variant underlying it and selective mechanism(s) remain unknown.

Although our evidence supports natural selection acting on leaf shape, we cannot exclude the possibility that selection is acting on a different locus in linkage disequilibrium (or physical linkage) with the leaf shape locus. Field studies such as those by Bright & Rausher (2008) using recombinant progeny (F_2, F_3, RILs) to eliminate associations between the leaf shape locus and other loci (except those in tight linkage) would allow an evaluation of the fitness advantage of the different genotypes. Experiments north of the clinal boundary would be ideal, because there is a clear prediction that *LL*s should have greater fitness than *ll*s. To date, existing field studies have either used inbred lines (Stinchcombe & Rausher 2001; Simonsen & Stinchcombe 2010) or have been carried out near the clinal boundary (Bright & Rausher 2008) where the expected fitnesses of different genotypes based on the cline are more difficult to predict.

Genetic diversity of I. hederacea in North American populations

One striking result was the lack of population genetic structure; F_{ST} across AFLP loci was quite low. The average selfing rate in *I. hederacea* populations in southern USA was found to be 66% (Hull-Sanders *et al.* 2005), with earlier studies (Ennos 1981) suggesting selfing rates as high as 93%, which leads to a prediction of higher neutral genetic differentiation and significant population structure.

One possible explanation for the lack of structure is that there is a high degree of gene flow leading to admixed populations. Given the potential prevalence of selfing (Hull-Sanders *et al.* 2005), gene flow is likely to be achieved primarily through seed dispersal. Our results indicate that the observed cline can be maintained despite dispersal events of 20–150 km. Because *I. hederacea* inhabits recently disturbed soils and is an agricultural weed, seed dispersal via human-mediated activity may result in long-range dispersal (Epperson & Clegg 1986; Hull-Sanders *et al.* 2005). Provided some outcrossing occurs between plants within populations (after colonization via seed), this process could impede the build-up of genetic structure. In the closely related congener, *I. purpurea*, which shares a very similar growth habitat, range in eastern USA and seed dispersal capacity, Epperson & Clegg (1986) found no evidence for increased IBD when comparing populations separated by metres to those separated by up to 500 km, implying that there must be substantial gene flow, which they attributed to agricultural activity. Our analysis of IBD shows similar findings (Fig. 5).

A second potential explanation for the lack of genetic structure is that *I. hederacea* populations in North America may have emerged from a recent bottleneck, and hence, populations are equally genetically depauperate. The bottleneck interpretation would be consistent with the hypothesis that *I. hederacea* is a released ornamental species that originated in tropical America (Strausbaugh & Core 1964; Long & Lakela 1971; Wunderlin 1982) or that it is native to the southern USA and subsequently expanded northward. Our results do not support these hypotheses: the populations we sampled exhibit high neutral polymorphism suggesting they are not genetically depauperate (Fig. 6), which is not consistent with a pattern expected under a northward colonization model (reviewed by Soltis *et al.* 2006).

A third possible explanation is the inherent limitations of AFLPs: (i) They are dominant markers, and hence, allele frequencies cannot be determined without some form of estimation (Mueller & Wolfenbarger 1999; Bonin *et al.* 2004; Mba & Tohme 2005). (ii) Restriction enzyme–based markers suffer the possibility that fragments from different regions of the genome share the exact same size (i.e. size homoplasy, Arrigo *et al.* 2009), therefore artificially inflating the genetic similarity between individuals. (iii) AFLPs may experience an unusually high mutation rate driving an artificially large within-population variance component of the F_{ST} calculation (Mba & Tohme 2005). Although we controlled for these, the technical problems of AFLPs are difficult to avoid altogether. However, further technological advancements have facilitated the development of newer markers such as RADseq (Peterson *et al.* 2012) and SNP-arrays (Bourret *et al.* 2012), which are more reliable and numerous and show great promise for ecological and population genomics.

Geographical patterns of leaf shape

Intraspecific clinal variation in leaf shape has been demonstrated in several other species; however, the patterns appear to defy a single mechanistic explanation. For example, Wyatt & Antonovics (1981) investigated leaf shape in the Butterflyweed, *Asclepias tuberosa*, which displays a longitudinal cline. They showed that leaf shape gradually shifts from an ovate (broad base tapering towards the tip) morph to an obovate (narrow base widening towards the tip) morph and that this cline had been maintained for at least 30 years. In contrast, we did not detect any signal of a longitudinal cline. In another example, Yamaguchi (1987) demonstrated that the recessive entire morph of the Japanese wild radish, *Raphanuss ativus*, showed a significant increase in frequency with latitude. Again, in contrast, we found that the entire morph (heart-shaped genotypes) for *I. hederacea* decreased with latitude. However, in wild radish, the entire morph appears to possess more serrations (tooth-like projections; Yamaguchi 1987), which may have functional similarity to lobing in *I. hederacea*. Gurevitch (1988) found that common yarrow, *Achillea millefolium*, populations from low altitudes had more open and highly dissected leaves, whereas high-altitude individuals produced more compact, smaller leaves. She posited that this differentiation resulted from alternative strategies for temperature regulation; highly dissected leaves enable lowland plants to maintain cooler leaf temperatures, thereby avoiding overheating, while compact small leaves enable upland plants to achieve leaf temperatures higher than ambient and thus avoid cold damage (Gurevitch 1988). Again, in contrast to Gurevitch's (1988) findings, *I. hederacea* displays increased dissection (lobing) in northern cooler environments, with the heart leaves (potentially able to achieve warmer temperatures) found in southern habitats (Bright 1998; this study).

Given these results, are there likely to be general trends or mechanisms affecting the microevolution and ecology of leaf shape? Existing evidence suggests that it is unlikely. The evolution of different leaf shapes is a function of both phylogenetic history and adaptation to contemporary environmental conditions (Givnish 1987; Nicotra *et al.* 2011). Furthermore, leaf shape likely experiences trade-offs with several other aspects of plant physiology (Nicotra *et al.* 2011), potentially confounding simple predictions. For example, leaf lobing has been shown to be beneficial in both hot daytime and cool nighttime conditions, because of an increased thermoregulatory efficiency relative to their entire margined counterparts. Hence, lobed leaves should dissipate heat more effectively under hot sunny conditions and gain heat from the surrounding air more efficiently under cold nighttime conditions (Givnish 1988; Lambers *et al.* 2006). Based on thermal arguments, lobed genotypes of *I. hederacea* should dominate in both the south (where hot days are prevalent during summer) and the north, a prediction neither Bright's (1998) nor our study supports. However, heat-transfer limitations can be offset by other leaf features such as changes in stomata or trichome density.

Leaf shape can interact with several different environmental forces, such as temperature (Vogel 1968, 1970, 2009; Parkhurst & Loucks 1972), precipitation (Givnish 1979; Wyatt & Antonovics 1981), herbivory (Rausher 1978; reviewed by Brown *et al.* 1991; Campitelli *et al.* 2008), fungal infection (Bright 1998) and many others. It will therefore be difficult to formulate a universal hypothesis that describes most geographical patterns in leaf shape, and that is amenable to all conditions and combinations thereof (reviewed by Nicotra *et al.* 2011).

Conclusion

The leaf shape polymorphism in *I. hederacea* appears to follow an adaptive latitudinal cline. Further ecological work is necessary to measure the extent of fitness advantages for leaf-lobing and ecological mechanisms producing them, especially in populations where lobed genotypes predominate. Potential mechanisms include nighttime temperature regulation, pathogen resistance, herbivory or combinations of these factors. Future experimental field work evaluating the role of biotic interactions and ecophysiological function will be necessary to confirm the agents of selection producing the cline described here.

Acknowledgements

References

Anderson AR, Hoffmann AA, McKechnie SW, Umina PA, Weeks AR (2005) The latitudinal cline in the In (3R) Payne inversion polymorphism has shifted in the last 20 years in Australian *Drosophila melanogaster* populations. *Molecular Ecology*, **14**, 851–858.

Arrigo N, Tuszynski JW, Ehrich D, Gerdes T, Alvarez N (2009) Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. *BMC Bioinformatics*, **10**, 33.

Barton NH, Baird SJE (1998) Analyse: an application for analysing hybrid zones. http://www.biology.ed.ac.uk/research/groups/Mac/Analyse/index.html.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Beaumont MA, Nichols RA (1996) Evaluating loci for the use in the genetic analysis of population structure. *Proceedings of The Royal Society of London. Series B, Biological Sciences*, **263**, 1619–1626.

Berry A, Kreitman M (1993) Molecular analysis of an allozyme cline: alcohol dehydrogenase in *Drosophila melanogaster* on the east coast of North America. *Genetics*, **134**, 869–893.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Ranatemporaria*). *Molecular Biology and Evolution*, **23**, 773–783.

Bourret V, Kent MP, Primmer CR *et al.* (2012) SNP-array reveals genome wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). *Molecular Ecology*, doi: 10.1111/mec.12003.

Bright KL (1998) *Geographical variation and natural selection on a leaf shape polymorphism in the Ivyleaf morning glory (Ipomoea hederacea)*. PhD Thesis, Duke University, USA.

Bright KL, Rausher MD (2008) Natural selection on a leaf-shape polymorphism in the Ivyleaf morning glory (*Ipomoea hederacea*). *Evolution*, **62**, 1978–1990.

Brown VK, Lawton JH, Grubb PJ (1991) Herbivory and the evolution of leaf size and shape [and Discussion]. *Philosophical Transactions: Biological Sciences*, **333**, 265–272.

Caballero A, Quesada H, Rolán-Alvarez E (2008) Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics*, **179**, 539–554.

Campitelli BE, Simonsen AK, Rico Wolf A, Manson JS, Stinchcombe JR (2008) Leaf shape variation and herbivore consumption and performance: a case study with *Ipomoea hederacea* and three generalists. *Arthropod-Plant Interactions*, **2**, 9–19.

Cavalli-Sforza LL (1966) Population structure and human evolution. *Proceedings of The Royal Society of London. Series B, Biological Sciences*, **164**, 362–379.

Cheng C, White BJ, Kamdem C *et al.* (2012) Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach. *Genetics*, **190**, 1417–1432.

Elmore CD (1986) Mode of reproduction and inheritance of leaf shape in *Ipomoea hederacea*. *Weed Science*, **34**, 391–395.

Endler JA (1973) Gene flow and population differentiation. *Science*, **179**, 243–250.

Endler JA (1977) Geographic Variation, Speciation, and Clines. Princeton University Press, Princeton, New Jersey.

Ennos RA (1981) Quantitative studies of the mating system in two sympatric species of *Ipomoea* (Convolvulaceae). *Genetica*, **57**, 93–98.

Epperson BK, Clegg MT (1986) Spatial-autocorrelation analysis of flower color polymorphisms within substructured populations of morning glory (*Ipomoea purpurea*). *The American Naturalist*, **128**, 840–858.

Fisher RA (1950) Gene frequencies in a cline determined by selection and diffusion. *Biometrics*, **6**, 353–361.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Galindo J, Morán P, Rolán-Alvarez E (2009) Comparing geographical genetic differentiation between candidate and non-candidate loci for adaptation strengthens support for parallel ecological divergence in the marine snail *Littorinasaxatilis*. *Molecular Ecology*, **18**, 919–930.

Gao H, Williamson S, Bustamante CD (2007) A Markov chain Montecarlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics*, **176**, 1635–1651.

Givnish TJ (1979) On The Adaptive Significance of Leaf Form. In Topics in Plant Population Biology. Columbia Univ. Press, New York.

Givnish TJ (1987) Comparative studies of leaf form: assessing the relative roles of selective pressures and phylogenetic constraints. *New Phytologist*, **106**, 131–160.

Givnish TJ (1988) Adaptation to sun and shade: a whole-plant perspective. *Australian Journal of Plant Physiology*, **15**, 63.

Givnish TJ, Vermeij GJ (1976) Sizes and shapes of liane leaves. *The American Naturalist*, **110**, 743–778.

Gurevitch J (1988) Variation in leaf dissection and leaf energy budgets among populations of *Achillea* from an altitudinal gradient. *American Journal of Botany*, **75**, 1298–1306.

Haldane JBS (1948) The theory of a cline. *Genetics*, **48**, 177–284.

Hoekstra HE, Drumm KE, Nachman MW (2004) Ecological genetics of adaptive color polymorphism in pocket mice: geographic variation in selected and neutral genes. *Evolution*, **58**, 1329–1341.

Holsinger KE, Lewis PO, Dey DK (2002) A Bayesian approach to inferring population structure from dominant markers. *Molecular ecology*, **11**, 1157–1164.

Hull-Sanders HM, Eubanks MD, Carr DE (2005) Inbreeding depression and selfing rate of *Ipomoea hederacea* var. *integriuscula* (Convolvulaceae). *American Journal of Botany*, **92**, 1871–1877.

Huxley J (1938) Clines: an auxiliary taxanomic principle. *Nature*, **142**, 219–220.

Klingaman TE, Oliver LR (1996) Existence of ecotypes among populations of Entireleaf morning glory (*Ipomoea hederacea* var. *integriuscula*). *Weed Science*, **44**, 540–544.

Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, **23**, 482–490.

Kolaczkowski B, Kern AD, Holloway AK, Begun DJ (2011) Genomic differentiation between temperate and tropical Australian populations of *Drosophila melanogaster*. *Genetics*, **187**, 245–260.

Kulathinal RJ, Stevison LS, Noor MAF (2009) The genomics of speciation in *Drosophila*: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genetics*, **5**, e1000550.

Lambers H, Chapin III FS, Pons TL (2006) Plant Physiological Ecology. Springer, New York.

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.

Long RW, Lakela O (1971) A Flora of Tropical Florida. University of Miami Press, Coral Gables.

May RM, Endler JA, McMurtrie RE (1975) Gene frequency clines in the presence of selection opposed by gene flow. *The American Naturalist*, **109**, 659–676.

Mba C, Tohme J (2005) Use of AFLP markers in surveys of plant diversity. *Methods in Enzymology*, **395**, 177–201.

McKechnie SW, Blacket MJ, Song SV *et al.* (2010) A clinally varying promoter polymorphism associated with adaptive variation in wing size in *Drosophila*. *Molecular Ecology*, **19**, 775–784.

Mohr C (1901) Plant Life of Alabama, 6th edn, Brown Print Co., Montgomery, Alabama.

Mueller U, Wolfenbarger L (1999) AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution*, **14**, 389–394.

Mullen LM, Hoekstra HE (2008) Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution*, **62**, 1555–1570.

Nicotra AB, Leigh A, Boyce KC *et al.* (2011) The evolution and functional significance of leaf shape in the angiosperms. *Functional Plant Biology*, **38**, 535–552.

O'Malley KG, Banks MA (2008) A latitudinal cline in the Chinook salmon (*Oncorhynchus tshawytscha*) Clock gene: evidence for selection on PolyQ length variants. *Proceedings of The Royal Society of London. Series B, Biological Sciences*, **275**, 2813–2821.

O'Malley KG, Ford MJ, Hard JJ (2010) Clock polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of The Royal Society of London. Series B, Biological Sciences*, **277**, 3703–3714.

Paaby AB, Blacket MJ, Hoffmann AA, Schmidt PS (2010) Identification of a candidate adaptive polymorphism for *Drosophila* life history by parallel independent clines on two continents. *Molecular Ecology*, **19**, 760–774.

Parkhurst DF, Loucks OL (1972) Optimal leaf size in relation to environment. *The Journal of Ecology*, **60**, 505–537.

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE*, **7**, e37135.

Prevosti A (1964) Chromosomal polymorphism in *Drosophila subobscura* populations from Barcelona (Spain). *Genetics Research*, **5**, 27–38.

Prevosti A, Serra L, Ribo G *et al.* (1985) The Colonization of *Drosophila subobscura* in Chile. II. Clines in the chromosomal arrangements. *Evolution*, **39**, 838–844.

Rausher MD (1978) Search image for leaf shape in a butterfly. *Science*, **200**, 1071–1073.

Schemske DW, Bierzychudek P (2007) Spatial differentiation for flower color in the desert annual *Linanthus parryae*: was Wright right? *Evolution*, **61**, 2528–2543.

Schmidt PS, Rand DM (1999) Intertidal microhabitat and selection at MPI: interlocus contrasts in the northern Acorn barnacle, *Semibalanus balanoides*. *Evolution*, **53**, 135–146.

Schmidt PS, Rand DM (2001) Adaptive maintenance of genetic polymorphism in an intertidal barnacle: habitat- and life-stage-specific survivorship of *Mpi* genotypes. *Evolution*, **55**, 1336–1344.

Simonsen AK, Stinchcombe JR (2010) Quantifying evolutionary genetic constraints in the Ivyleaf morning glory, *Ipomoea hederacea*. *International Journal of Plant Sciences*, **171**, 972–986.

Slatkin M (1973) Gene flow and selection in a cline. *Genetics*, **75**, 733–756.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.

Stinchcombe JR, Rausher MD (2001) Diffuse selection on resistance to deer herbivory in the ivyleaf morning glory, *Ipomoea hederacea*. *The American Naturalist*, **158**, 376–388.

Storz JF, Dubach JM (2004) Natural selection drives altitudinal divergence at the albumin locus in deer mice, *Peromyscus maniculatus*. *Evolution*, **58**, 1342–1352.

Strasburg JL, Scotti-Saintagne C, Scotti I, Lai Z, Rieseberg LH (2009) Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. *Molecular Biology and Evolution*, **26**, 1341–1355.

Strausbaugh PD, Core EL (1964) Flora of West Virginia. West Virginia University Bulletin, Morgantown.

Tice KA, Carlon DB (2011) Can AFLP genome scans detect small islands of differentiation? The case of shell sculpture variation in the periwinkle Echinolittorina hawaiiensis. *Journal of Evolutionary Biology*, **24**, 1814–1825.

Vasemägi A (2006) The adaptive hypothesis of clinal variation revisited: single-locus clines as a result of spatially restricted gene flow. *Genetics*, **173**, 2411–2414.

Vekemans X (2002) AFLP-Surv Version 1.0. Distributed by Author. Laboratoire de Génétique et EcologieVégétale, UniversitéLibre de Bruxelles, Belgium.

Vogel S (1968) "Sun leaves" and "shade leaves": differences in convective heat dissipation. *Ecology*, **49**, 1203–1204.

Vogel S (1970) Convective cooling at low airspeeds and the shapes of broad leaves. *Journal of Experimental Botany*, **21**, 91–101.

Vogel S (2009) Leaves in the lowest and highest winds: temperature, force and shape. *The New Phytologist*, **183**, 13–26.

Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wood HM, Grahame JW, Humphray S, Rogers J, Butlin RK (2008) Sequence differentiation in regions identified by a genome scan for local adaptation. *Molecular Ecology*, **17**, 3123–3135.

Wunderlin RP (1982) Guide to the Vascular Plants of Central Florida, Universty Presses of Florida, Tampa.

Wyatt R, Antonovics J (1981) Butterflyweed re-revisited: spatial and temporal patterns of leaf shape variation in *Asclepias tuberosa*. *Evolution*, **35**, 529–542.

Yamaguchi H (1987) Latitudinal cline and intrapopulational differentiation in leaf shape of wild radish in Japan. *Japan Journal of Breeding*, **37**, 54–65.

B.E.C. is a PhD student studying the ecological, physiological, and population genetics of a leaf shape cline in morning glory. J.R.S.'s research focuses on plant ecological and evolutionary genetics.

Data accessibility

Population locations, allele and genotype frequencies, and sampling details are available in Table S1. Progeny test results are available in Table S4; AFLP genotypes and leaf shape genotypes are available in Table S5.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Summary of population location, genotype and allele statistics for all sampled populations.

Table S2 Oligonucleotide sequences used for AFLP generation.

Table S3 Summary of cline statistics for AFLP loci.

Table S4 Progeny test results.

Table S5 AFLP and leaf shape genotypes.

Fig. S1 Outputs of a 76-deme model for (a) DFDIST and (b) BayeScan. Filled circles represent the leaf shape locus, open circles represent AFLP loci.

Fig. S2 Posterior probabilities of InStruct outputs for $K = 2$–4.

Fig. S3 Spatial autocorrelation coefficients (r) ±95% CLs for the leaf shape locus (solid line and filled circles) and for all AFLP loci (black dashed line and empty circles) across 10 distance size classes.

Appendix S1 Detailed methods for the greenhouse experiment, DNA extraction, AFLP marker development, Instruct and spatial autocorrelation analysis.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.