PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2009) **276**, 1495–1506
doi:10.1098/rspb.2008.1767
Published online 25 February 2009

Variation at range margins across multiple spatial scales: environmental temperature, population genetics and metabolomic phenotype

William E. Kunin[1,*], Philippine Vergeer[1], Tanaka Kenta[2,3], Matthew P. Davey[2], Terry Burke[2], F. Ian Woodward[2], Paul Quick[2], Maria-Elena Mannarelli[2], Nathan S. Watson-Haigh[2,4] and Roger Butlin[2]

[1]*Faculty of Biological Sciences, Institute for Integrative and Comparative Biology, University of Leeds, Leeds LS2 9JT, UK*
[2]*Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK*
[3]*Sugadaira Montane Experimental Center, University of Tsukuba, Ueda 386-2204, Japan*
[4]*CSIRO Livestock Industries, Queensland Bioscience Precinct, St Lucia, Queensland 4067, Australia*

Range margins are spatially complex, with environmental, genetic and phenotypic variations occurring across a range of spatial scales. We examine variation in temperature, genes and metabolomic profiles within and between populations of the subalpine perennial plant *Arabidopsis lyrata* ssp. *petraea* from across its northwest European range. Our surveys cover a gradient of fragmentation from largely continuous populations in Iceland, through more fragmented Scandinavian populations, to increasingly widely scattered populations at the range margin in Scotland, Wales and Ireland. Temperature regimes vary substantially within some populations, but within-population variation represents a larger fraction of genetic and especially metabolomic variances. Both physical distance and temperature differences between sites are found to be associated with genetic profiles, but not metabolomic profiles, and no relationship was found between genetic and metabolomic population structures in any region. Genetic similarity between plants within populations is the highest in the fragmented populations at the range margin, but differentiation across space is the highest there as well, suggesting that regional patterns of genetic diversity may be scale dependent.

Keywords: marginal populations; microclimate; isolation by distance; metabolomics; postgenomics; spatial structure

1. INTRODUCTION

The study of species distributions and their margins has become a topic of increasing interest in recent decades, because of both its intrinsic interest (Gaston 2003; Antonovics *et al.* 2006) and the potential importance of such margins in understanding population responses to anthropogenic environmental change (Mandák *et al.* 2005). Populations of a species living at the margin of its climatic or other environmental tolerances may provide vital information regarding the processes that determine species distributions. They may also harbour local adaptations that could be fundamental to the performance of the species as a whole under future climatic conditions (Jump & Penuelas 2005; Bridle & Vines 2007).

A major aspect of the debate over species margins concerns the evolutionary process of adaptation to local conditions. If a species expands its range to the limits of its environmental tolerances, as is commonly assumed, populations at range margins should be under substantial selective pressure to adapt to local conditions. If such adaptation is achieved, the limits of the species' environmental tolerances should be relaxed, allowing the species' distribution to expand. In principle, this process might continue indefinitely, so that in time all species would come to live everywhere (c.f. Willis 1922). Yet species show distinctive distributional properties and climatic limits, which may remain stable over very long periods (e.g. Coope 1995). How can we reconcile theory with observation?

Many explanations have been offered for such apparent niche conservatism (Lennon *et al.* 1997; Hochberg & Ives 1999; Gaston 2003; Holt & Keitt 2005). One type of explanation that has gained considerable attention in recent years concerns the effects of gene flow on local adaptation (Holt & Gomulkiewicz 1997; Kirkpatrick & Barton 1997; Butlin *et al.* 2003; Filin *et al.* 2008). In this school of thought, populations at the margins of ranges are considered less well adapted to local conditions than are populations in core areas of a species' distribution. The resulting higher population sizes (e.g. Brown *et al.* 1995, but see Sagarin & Gaines 2002) and greater reproductive output of core populations relative to those at range margins (e.g. Carey *et al.* 1995, but see Gaston 2003; Angert 2006) create a net flux of propagules—and of genes—from the core of a distribution to more marginal zones. In the extreme, core populations may be net population sources, whereas some marginal populations may be effectively demographic sinks, dependent on

* Author for correspondence (w.e.kunin@leeds.ac.uk).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2008.1767 or via http://rspb.royalsocietypublishing.org.

One contribution of 17 to a Special Issue 'Geographic range limits of species'.

Received 1 December 2008
Accepted 7 January 2009

the immigration for population persistence (Mandák *et al.* 2005). The net flow of genes from core to margin may be sufficient to prevent or severely limit local adaptation within marginal populations (Kirkpatrick & Barton 1997; Case & Taper 2000; Lenormand 2002). However, gene flow also increases the genetic variance for adaptive traits in marginal populations and this increases their ability to respond to selection, potentially outweighing the swamping effect (Barton 2001; Garant *et al.* 2007).

While gene flow-based models are increasingly popular, they face serious difficulties. The level of gene flow required to swamp selection seems far in excess of what is plausible between the core and the margin of a geographic range in many systems. Very little gene flow is required to prevent population differentiation by genetic drift (approx. one immigrant genome per generation, regardless of the population size, Felsenstein 1983), but much higher levels are required to counteract the effects of selection (Holt 2003; Alleaume-Benharira *et al.* 2006). Thus, in species with only moderate dispersal abilities, any effects of gene flow in preventing adaptation are more likely to be important across local environmental gradients than across species' ranges. Our main focus here is on plant populations, where mean seed dispersal distances are typically quite short (e.g. Venable *et al.* 2008: <1 m), and even pollen dispersal drops off sharply with distance (e.g. Cresswell 1997). The 'fat tails' of long-distance dispersal that may occur in both of these processes (e.g. Fenart *et al.* 2007; Soons & Bullock 2008) may be important for colonization, for overcoming genetic drift or to maintain a supply of genetic variation upon which selection can act, but they are far too infrequent to invoke as forces overpowering the effects of selection.

One way out of this conundrum is to rethink the nature of the range margin itself. Most analyses of distributional margins consider these phenomena across vast, continental scales, but similar distributional limits are displayed at much finer spatial scales as well. Species ranges are not continuously distributed over smooth environmental gradients; there are margin-like environments even in the core parts of a species' range, and fragments of core-like environments may be found near range margins (e.g. Roy & Thomas 2003). Within a region, for example, many species display upper and/or lower altitudinal limits, which may well be determined by environmental tolerances similar to those influencing geographic range limits (e.g. Merrill *et al.* 2008). At even finer scales, populations are typically patchily distributed in space, displaying very local-scale distributional margins (e.g. Kunin 1998; Wilson *et al.* 2002). These margins too may reflect (micro-) environmental limits to species persistence. Even if gene flow is insufficient to block adaptation when considered at biogeographic scales, much higher levels of gene flow would be anticipated when considering finer scale components of the species' distributional limits. In addition to dispersal between occupied patches, dispersal at moderate-to-fine spatial scales may also play a key role in the recolonization of empty sites, which can be important for both (meta-) population persistence (Lennon *et al.* 1997; Wilson *et al.* 2002) and patterns of genetic variation (Whitlock 2003). Because it may be common for marginal populations also to be fragmented, it is difficult to separate these various effects.

In this study, we focus on these fine-scale components of fragmented populations at range margins. We will be able neither to assess directly the factors limiting species distributions at these scales, nor to test the evolutionary processes giving rise to such limits. Nonetheless, as a first step towards addressing such issues, we will document the levels of climatic, genetic and phenotypic diversity within and between populations in different portions of a species' range. Local environmental variation is a prerequisite for local adaptation to occur, and local variation in genotypes and phenotypes can provide information on the degree of gene flow and of local adaptation present at different scales. We will focus on the western European distribution of a single plant species, *Arabidopsis lyrata* spp. *petraea* (hereafter *A. l. petraea*). We have published research elsewhere comparing the metabolic and growth properties of selected populations found in different regions (countries) of this distribution (Davey *et al.* 2008, in press; Vergeer *et al.* 2008). However, here we will focus on the variation at finer spatial scales, specifically on the variation within and between populations (sites) inside each region. There are a number of reasons to focus on such variation patterns.

Firstly, we need to quantify the degree of microclimatic variation across local populations (on scales of tens or hundreds of metres), and compare that with the variation found between populations and across space at coarser scales of resolution (tens or hundreds of km). Secondly, we need to quantify the spatial scaling of genetic differentiation within and between populations, within each region. The highly fragmented populations found in marginal areas of a distribution are generally expected to contain reduced levels of within-population genetic diversity (Honnay & Jacquemyn 2007; Kark *et al.* 2008), but fragmentation might also increase genetic diversification between populations (Cohan 1984). Thirdly, we need to quantify the levels of variation in putatively adaptive traits, as they too may be expected to change at range margins. A vast array of morphological traits can be measured to describe a phenotype (Pigliucci *et al.* 1999). It is likely that no single trait is critical at range margins, but trade-offs between traits may impose limits on adaptability (Angert *et al.* 2008). We opt here to focus on molecular phenotypes, as they arguably sum across a wide range of expressed variations. Morphological traits involved in plant growth are controlled by a variety of metabolic networks, allowing the protection and repair of plant cells in order to provide an appropriate response to changing environmental and resource conditions (Vinocur & Altman 2005; Meyer *et al.* 2007); and thus metabolic traits should be considered as a vital phenotype in wild plant populations. Metabolic phenotypes (or chemotypes) are now being used for environmental genomics research to identify ecologically important genes and traits (Benfey & Mitchell-Olds 2008).

Variation *within* populations may be reduced if isolated marginal populations have lost genetic diversity through inbreeding or genetic drift, or if diversity has been exhausted by the response to local selection. At the same time, variation *among* populations may also be lower in marginal regions if they are less able to respond to local selection than more diverse and connected populations in core regions of the distribution, or alternatively it may be increased owing to isolation and founder effects.

To test these predictions, we have examined within- and among-population variations in a component of climatic variation (temperature), neutral genetic diversity and metabolic profile in four regions stretching from the core to the extreme southern margin of the species' western European distribution. We first compare variability in environmental temperatures among and within regions. We then test the expectation that neutral genetic variation will be lower within populations, but more strongly structured between populations, in marginal than core regions. We then examine the patterns of metabolomic variation, and its relationship with both environment and genetics. We propose that variations in metabolic profiles among populations, and correlation between metabolic profiles and environmental conditions, are signatures of local adaptation. If so, we expect higher metabolomic variation and stronger correlations in core than marginal regions, if we see the predicted patterns of neutral genetic variation and consistent levels of environmental variation among sites.

2. MATERIAL AND METHODS

(a) *Focal system:* A. lyrata *ssp.* petraea

Our work centres on the northern rock cress (*A. lyrata* spp. *petraea* (L.) O'Kane & Al-Shehbaz), a small rosette-forming crucifer native to much of northwest and central Europe. Molecular phylogenetic information (Mitchell-Olds 2001; Clauss & Koch 2006; Clauss & Mitchell-Olds 2006) confirms a close relationship between the species and the model plant species, *Arabidopsis thaliana*, with sufficiently high genetic similarity between the two to allow the use of genomic and post-genomic resources that would not be available for most other plants. Despite this high level of relatedness, the ecology of *A. l. petraea* is very different from its better-known congener. The focal plant is a subarctic/subalpine perennial, found in a range of open natural habitats, ranging from rock faces to scree slopes and gravel bars, to coastal zones (Clauss & Koch 2006). Natural populations appear intolerant of competition or grazing, but may endure extremely stressful abiotic conditions including ultramafic (serpentine) substrates, salt spray, drought and inundation. Plants may form solitary rosettes, but are capable of vegetative reproduction, producing clonal patches in some sites, although clones can generally be differentiated owing to variability in leaf shape (Jonsell *et al.* 1995). The plant is also an obligate outcrosser, suggesting very different patterns of genetic diversity than those found in the largely self-pollinating *A. thaliana*.

(b) *Site selection, surveying and temperature recording*

To examine the genetic and phenotypic patterns across *A. l. petraea*'s geographic range, we focused on the species' northwest European distribution, stretching from Iceland to Scandinavia and south into northern and western regions of the British Isles (BI; Jalas & Suominen 1994). The species varies greatly in its abundance across this area. In Iceland, the species is very widely distributed, being one of the most common plants in the country (Kristinsson 1995), and it is reasonably common in both south-central Norway and in coastal zones of central Sweden, although in each case restricted to specific habitats such as scree, rock and riverine gravel bars in Norway, coastal rock and shingle in Sweden. Populations are much more widely dispersed in Scotland, whereas only a few scattered populations persist in Wales and Ireland. Thus, our focal regions can be considered as a geographic transect from nearly continuous populations (Iceland), through regions with more subdivided habitat-restricted patches (Norway, Sweden), to the progressively more fragmented and isolated populations at the southern margin of the species' distribution (Scotland, Wales and Ireland, collectively referred to as the 'British Isles' populations). The northern margin of the distribution is less easily studied, as it is fairly abrupt and often associated with the limits of available habitat.

Within each focal region, we acquired the best available data on the distribution of the species (see electronic supplementary materials). Using these records, we compiled a matrix of inter-population distances within each focal region. To examine the environmental and population differentiation at multiple scales, we adopted a multi-scale sampling protocol, which involved sampling plants within a hierarchy of nested scales both within and between populations. Details of the protocol are provided in electronic supplementary materials (ESM1). The resulting sample included eight focal populations in each region, with the exception of two each in Ireland and Wales (the only known extant populations in each). We surveyed each population to an equal level of intensity, and then selected 10 focal quadrats (50×50 cm) within each site: four at patch margins (the furthest north, south, east and west found) and six randomly chosen from the remainder of quadrats identified. We placed temperature monitoring probes (Thermochron iButtons Model DS1922L; Maxim Dallas Inc.) buried 1 cm below the soil surface adjacent to four of these quadrats, two chosen subjectively to be the warmest and coldest microsites of the set (based largely on slope and aspect) and the other two chosen at random from the remainder. Temperature recordings were taken every 120 min over a period of 2 years (August 2005–2007). These temperature records were analysed by month and year to produce a total of 124 variables (see electronic supplementary material), which were analysed by principal component analysis (PCA) to provide a description of 'temperature space' across the study areas. The first five axes of this PCA, which captured 89.3 per cent of the variance in the original dataset, will be used in our analyses (see electronic supplementary material), allowing distances in climate space to be calculated between sites. As only four iButtons were placed in each population, direct temperature data are available only for four focal plants at each site, and so only this subset of plants (141 in total) is used in our analyses here.

(c) *Measuring genetic similarity/diversity*

As part of our ongoing research, we have sampled genetic material from over 1000 *A. l. petraea* plants from across the focal regions. This includes 114 of the focal plants for which detailed microclimate recordings are available. Genomic DNA was extracted from silica gel-dried leaves of each plant following Whitlock *et al.* (2008). Genotypes were scored for 25 SNPs following the protocol of Multiplex SNP-SCALE (Kenta *et al.* 2008), a cost-effective medium throughput method based on allele-specific PCR, using Applied Biosystems 3730 DNA Analyser and GeneMapper 3.7 software. These SNPs were newly developed (T. Kenta, N. Watson-Haigh, M. Mannarelli, J. Slate, R. K. Butlin and T. Burke 2008 unpublished data) in two sorts of randomly chosen genes: (i) strictly randomly chosen genes from the

A. thaliana genome (Schmid *et al.* 2005) and (ii) randomly chosen genes that have GenBank registration for *A. thaliana* and at least one other species within the Brassicaceae. We assumed that genetic similarity/diversity patterns derived from those random genes more likely represent stochastic effects such as genetic drift rather than selection. Gene codes of those SNPs are provided in the electronic supplementary materials (electronic supplementary material, table A4). A panel of 10 plants that represent the whole study area was used to discover SNPs in these genes to avoid ascertainment bias. The Loiselle *et al.* (1995) kinship coefficient was calculated as an index of genetic similarity using SPAGeDi 1.2 (Hardy & Vekemans 2002) for all plant pairs available within each region. Genetic distances between populations were estimated using F_{ST}.

(d) *Measuring metabolomic profiles*

(i) *Growth*

Seeds of *A. l. petraea* were collected from 28 populations during 2005–2006 (see electronic supplementary materials, table A1). Seeds of each population were established and, after 5–8 days, 20 germinated seedlings from each population were transferred to a Steill (Stiell Facilities, Glasgow, UK) controlled-environment growth cabinet set to a 16/8 h day/night cycle; 20/15°C day/night. When the plants developed approximately 12 leaves (another 12–18 days), and after 5–7 h into the daylight period, the foliage was excised and immersed in liquid nitrogen (see electronic supplementary material, ESM1 for full details).

(ii) *Metabolite extraction and analyses*

Metabolites were extracted and analysed as described in Davey *et al.* (2008). Briefly, approximately 50 mg leaf tissue per plant was extracted using MeOH/$CHCl_3$/H_2O followed by MeOH/$CHCl_3$ and H_2O. The aqueous phase (MeOH and H_2O) was directly injected into a LCT mass spectrometer (Waters Ltd. Manchester, UK) in the negative ionization mode (50–800 *m*/*z*). For metabolite fingerprinting, raw centroid mass/charge (*m*/*z*) ratios were combined into 0.2 mass unit 'bins'. Binned *m*/*z* and per cent total ion count (%TIC) values were Pareto scaled and explored by PCA using SIMCA-P v. 12 (Umetrics, Sweden). For our metabolomic data, the first PCA axis displayed a strong batch effect, and so axes 2–6 were used here, as being the most informative (see electronic supplementary material, ESM2).

(e) *Statistical methods*

As noted above, the results obtained from most of our observations (of temperature and metabolomics) were subjected to principal component analyses. In each of these five-dimensional principal component spaces, the Euclidean distances could be calculated between any two samples by the use of simple geometry. Our genetic analyses are measured in terms of pairwise similarity (rather than distance). Each of these, in turn, can be examined as a function of the geographic distance separating the points being compared. Given the multiplicative scale of distances considered, these spatial distances were converted to logarithmic (base 10) scales before analysis. Variation in principal component scores was also partitioned between regional, population and individual variations, using the minimum norm quadratic unbiased estimator (MINQUE) algorithm, implemented in SPSS (v. 14.0), and then summed across axes. For genetic data, a similar partitioning was done using analysis of molecular variance (AMOVA) executed by genetic mixture analysis (GMA) (Lewis & Zaykin 2001).

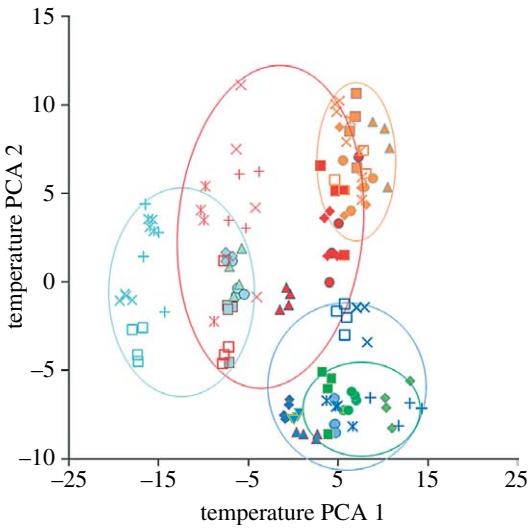

Figure 1. PCA ordination of temperature data from our field sites. This figure shows axes 1 and 2, which together represent 74.5% of the total variation. Ovals indicate variation in each region (turquoise, Iceland (circles, Ice_1; filled squares, Ice_2; uptriangles, Ice_3; diamonds, Ice_4; asterisks, Ice_5; pluses, Ice_6; open squares, Ice_7; crosses, Ice_8); light green, Ireland (circles, Ire_1; diamonds, Ire_2); red, Norway (circles, Nor_1; filled squares, Nor_2; triangles, Nor_3; diamonds, Nor_4; asterisks, Nor_5; pluses, Nor_6; open squares, Nor_7; crosses, Nor_8); dark blue, Scotland (circles, Sco_1; uptriangles, Sco_2; downtriangles, Sco_3; diamonds, Sco_4; asterisks, Sco_5; pluses, Sco_6; squares, Sco_7; crosses, Sco_8); orange, Sweden (circles, Swe_1; filled squares, Swe_2; uptriangles, Swe_3; diamonds, Swe_4; asterisks, Swe_5; pluses, Swe_6; open squares, Swe_7; crosses, Swe_8); dark green, Wales (circles, Wal_1; squares, Wal_2)). Details are given in the electronic supplementary materials.

We consider the following sets of regression analyses here:

(i) temperature difference × spatial distance and elevational difference,
(ii) genetic similarity × spatial distance, elevational and temperature differences,
(iii) metabolomic difference × spatial distance and temperature differences, and
(iv) metabolomic difference × genetic similarity.

We used matrix permutation to test the departure of the regression slopes from the null hypotheses of zero, using a custom script in R 2.7.0 (R Development Core Team 2008). We held the matrix for the dependent variable constant and permuted each of the independent variable matrices separately 1000 times. The reported estimate of the *p*-value is calculated from the number of permutations for which the absolute value of the regression coefficient exceeded the absolute value from the real data. This approach is similar to a Mantel test, but it used regression coefficients, rather than correlation, as statistics to test the effect of each explanatory variable separately.

As metabolomic data were measured in the laboratory from individuals grown from field-collected seeds of unknown paternity, the precise spatial position of each plant

Table 1. (*a*) Partitioning of variation between regional, site-level and within-population components. Partitioning was carried out for PCA data using MINQUE, and for genetic data using AMOVA methods. (*b*) Partitioning of variation within each region. BI, British Isles.

(*a*) data set	between regions (%)	between sites (%)	within populations (%)
temperature	68.34	24.50	7.16
genetics	50.2	35.1	14.7
metabolomics	30.47	6.64	62.89

(*b*)	temperature		genetic variation		metabolic profile	
	total variance	among-site component (%)	Hs	F_{ST}	total variance	among-site component (%)
core						
Iceland	56.88	91.6	0.198	0.084	11.71	6.6
Norway	68.45	79.8	0.204	0.086	13.83	12.5
Sweden	12.36	42.7	0.224	0.118	10.03	5.5
margin						
BI	30.85	84.0	0.201	0.243	12.71	8.9

could not be ascertained with any certainty. As a consequence, metabolomic data were considered only at spatial scales from the site scale and above. Mean metabolomic PCA values for each site were examined as a function of mean site position and mean temperature PCA scores.

3. RESULTS

(a) *Temperature variation within and between sites*

Substantial variation in temperature profile was found both between and within focal regions. The temperature PCA (figure 1) shows a roughly triangular scatter of points with Norwegian and lowland Icelandic populations in the centre; Ireland, Wales and parts of Scotland (with mild winters and cool summers) are placed in one corner; upland areas of Iceland (with cold winters and cool summers) are in another corner; and Sweden (with cold winters and hot summers) in the third corner. There is substantial variation within regions, largely associated with elevational differences; lowland areas of Iceland are quite close in temperature profiles to upland Norway, whereas lowland Norway approaches Sweden (where all populations are near sea level) and upland Scotland. Wales and Ireland cluster together with lowland Scottish sites.

The range of microclimatic variation found within sites was sometimes considerable, but nonetheless most variations in temperature were associated with region (68.34%) or site (24.50%) levels of analysis (table 1). While within-population differences in temperature accounted for only 7.16 per cent of the total variation, they were nonetheless substantial in places. The within-site temperature range (the difference between the two most dissimilar iButtons at each site) varied between regions, with averages ranging from 22 to 57 per cent of the maximum difference found within each region (electronic supplementary material, table A3). In a few sites (e.g. Iceland site 6, Norway site 8), local variation among iButtons was comparable to that found across the entire regional sample. More commonly, however, local variation in temperature was more limited, but there was still substantial overlap found between samples collected at sites tens or even hundreds of km apart. Interestingly, the within-site variation was much higher for PCA axis 2 (associated with summer heat) than for axis 1 (reflecting winter cold). Regions differed in their total variation in temperature profile and in the proportion of variation that was among sites (table 1*b*).

To test for spatial effects more explicitly, we used multiple regression analysis to compare differences between sample points in temperature PCA space (the Euclidean distance, calculated across the five highest PCA axes) with a matrix of between-sample geographic distances and elevational differences (table 2*a*). There was a significant effect of geographic distance on temperature differences in three of the regions (Sweden, Norway and Scotland), and a significant effect of elevational differences in four (all but Sweden, where very little elevational range was present). Overall, the explanatory power of these relationships was high in most regions, together explaining between 45 and 89 per cent of variance in all sites except Sweden (<4 per cent of variance explained) where there was, overall, substantially less variation than in the other regions (table 1*b*).

(b) *Genetic similarity as a function of distance and climate*

A. l. petraea populations were strongly genetically differentiated both among regions and among populations within regions (table 1*a*). The fraction of residual variance within populations (14.7%) was twice that found in temperature analyses and the proportion of variation among sites was also higher. Within-site genetic diversity was remarkably consistent across regions, contrary to our prediction (table 1*b*). However, in line with the expectations, very different spatial patterns of genetic similarity were displayed in different regions of the studied distribution (figure 2). There was no significant spatial structure displayed in genetic similarity within populations ($\log D<3$) in any region. In the more continuously distributed portions of the range (Iceland, Norway and Sweden), there was little evidence of genetic differentiation with distance even between populations; although Iceland showed significant genetic responses to climate and elevation, the explanatory power of these relationships

Table 2. Summary of regression coefficients and partial Mantel tests relating (*a*) temperature differences to log (distance) between points and differences in elevation; (*b*) genetic similarity as a function of log (distance), temperature difference and elevational differences, (*c*) metabolomic differences as a function of log (distance), temperature difference and elevational differences and (*d*) metabolomic differences as a function of genetic similarity. Italics and annotations indicate the levels of statistical significance, as indicated by the number of simulations (out of 1000) providing slopes steeper than the observed: $+0.1>p>0.05$, $*0.05>p>0.005$, $**p<0.005$.

region	intercept	log (distance)	Δ elevation	Δ temperature	R^2
(*a*) Δ temperature					
Iceland	1.777**	0.309	*0.0158***		0.892
Sweden	3.146**	*0.342***	0.0275		0.0358
Norway	1.767**	*1.290***	*0.00630***		0.472
Scotland	−0.873**	*1.181***	*0.0150***		0.454
Ireland and Wales	2.134*	0.122	*0.00895***		0.710
(*b*) genetic similarity					
Iceland	0.0919	−0.0274+	*−0.000124**	*0.00815***	0.015
Sweden	0.0924+	−0.0199	0.00190	−0.00488	0.0165
Norway	−0.0207	−0.000438	0.0000142	0.000510	0.00237
Scotland	0.249**	*−0.0562***	0.0000542	−0.00110	0.1226
Ireland and Wales	0.213*	*−0.104***	0.0000664	0.0285+	0.32562
(*c*) metabolomic differences					
Iceland	1.391	0.216		−0.0160	0.0259
Sweden	1.169	0.220		−0.0167	0.0300
Norway	1.522	0.342		−0.00212	0.0183
Scotland, Wales and Ireland	1.324	0.0781		0.109	0.1211

region	intercept	genetic similarity	R^2
(*d*) metabolomic differences versus genetic similarity			
Iceland	1.533	1.086	0.024
Sweden	2.590	2.590	0.052
Norway	1.111	4.987	0.101
Scotland, Wales and Ireland	2.814	−4.031	0.517

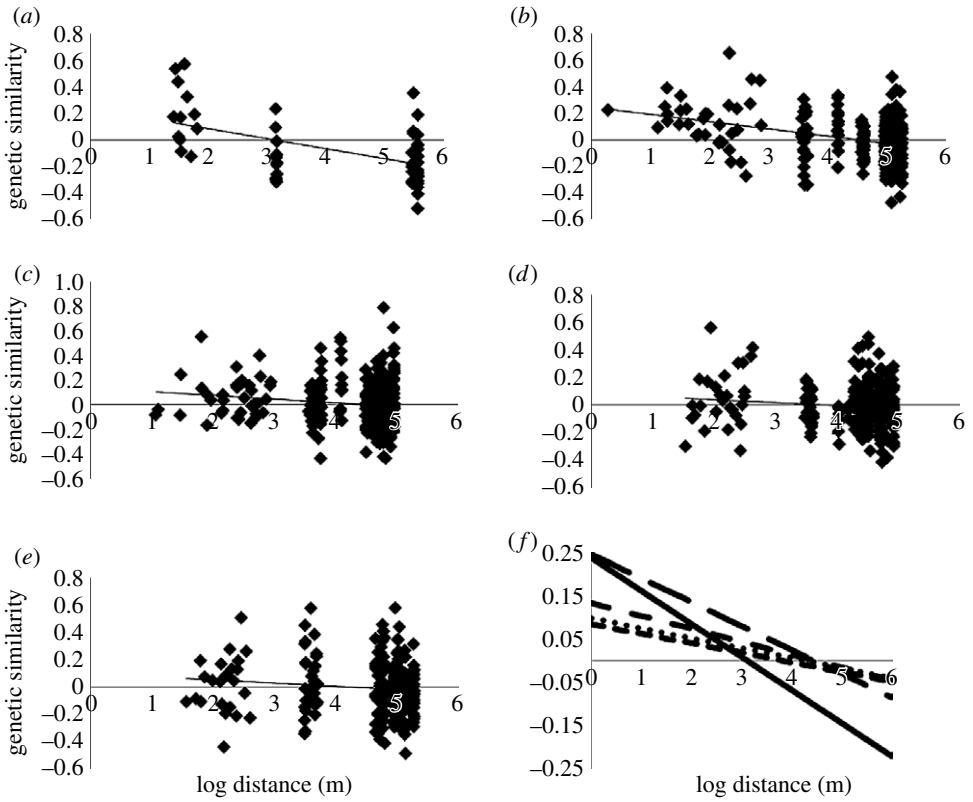

Figure 2. Decline in pairwise genetic similarity (Loiselle kinship coefficient) with distance, in each of the focal regions considered. (*a*) Ireland and Wales, (*b*) Scotland, (*c*) Norway, (*d*) Sweden and (*e*) Iceland. (*f*) The fitted relationships (Mantel tests, genetic similarity × log (distance)) of the regions together for comparison are provided (solid line, Ireland and Wales; long-dashed line, Scotland; medium-dashed line, Norway; small-dashed line, Sweden; dotted line, Iceland).

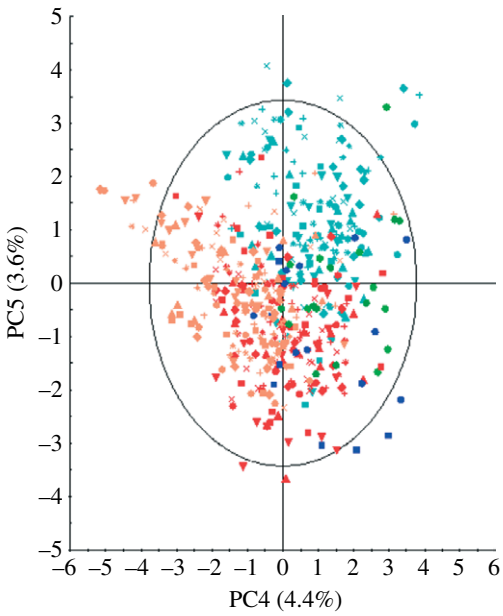

Figure 3. Score scatter plot from PCA (PC axes 4 and 5) of *m*/*z* values (binned to 0.2 Da) obtained by metabolic fingerprinting of *A. lyrata* spp. *petraea* populations from Iceland, Sweden, Norway, Ireland, Scotland and Wales. Fingerprints were obtained from direct injection mass spectrometry of the aqueous phase (methanol : water) in negative ionization. All data were Pareto scaled prior to PCA (turquoise, Iceland (circles, Ice_1; squares, Ice_2; uptriangles, Ice_3; diamonds, Ice_4; asterisks, Ice_5; pluses, Ice_6; downtriangles, Ice_7; crosses, Ice_8); light green, Ireland (circles, Ire_2); red, Norway (circles, Nor_1; squares, Nor_2; uptriangles, Nor_3; diamonds, Nor_4; asterisks, Nor_5; pluses, Nor_6; downtriangles, Nor_7; crosses, Nor_8); dark blue, Scotland (circles, Sco_5; squares, Sco_7); orange, Sweden (circles, Swe_1; squares, Swe_2; uptriangles, Swe_3; diamonds, Swe_4; asterisks, Swe_5; pluses, Swe_6; downtriangles, Swe_7; crosses, Swe_8) dark green, Wales (circles, Wal_1)).

was very low (table 2*b*). Overall F_{ST} values were low (table 1*b*). However, the more marginal and fragmented the populations, the higher was the association of genetic differentiation with distance, with the strongest spatial genetic patterns recorded in the most marginal British Isles populations (table 1*b* and figure 2). Within the British Isles, both the slope of decline in genetic similarity with distance and the explanatory power of the relationship grew as the extreme range margin (Wales and Ireland) was approached (Scotland: $\beta = -0.056$, $R^2 = 0.123$; Ireland/Wales: $\beta = -0.104$, $R^2 = 0.326$). As a consequence of this steep decline, in the most fragmented region (Wales and Ireland), levels of genetic similarity at the greatest distances considered (e.g. those more than 10 km apart) were substantially lower than those found at equivalent distances within the other, less fragmented, regions of the distribution.

(c) *Metabolomic differences between sites*

A high proportion (62.89%) of the observed variations in metabolic fingerprints in our samples occurred within populations (table 1*a*). Conversely, site-level variation within regions had very little explanatory power (6.6% of variance), such that there was no distinct clustering in the metabolic fingerprints of populations *within* any one region on any of the calculated principal components. Even though all plants were grown in standard laboratory conditions, there were distinct clusters at a regional scale, with PC4 separating Sweden and the British Isles, PC5 separating Iceland from all other regions and PC6 separating the British Isles populations from all other regions (figure 3). There was some overlap between the Swedish and Norwegian regions.

To test our predictions for spatial effects within regions, we used MANOVA to test for site-level effects and expressed the among-site component of variation as a proportion of total variation (table 1*b*). The site-level variation was significantly greater than zero in all cases ($p < 0.001$, except for the British Isles where $p = 0.036$). There was little among-site variation in Sweden, consistent with the low environmental variation observed there (table 2*b*). Among-site variation in the British Isles was not markedly different from levels in the core regions. Variation among families within sites was also significant in Norway and Sweden, and when regions were combined, indicating that genetic variation for metabolic profile exists within populations as well as among populations. Multiple regressions were used to compare differences between sample points in metabolomic PCA space (the Euclidean distance, calculated across PCA axes 2–6) with a matrix of mean between-sample geographic, temperature and genetic distances for each source population within a region. Although there were weakly positive slopes indicating greater metabolic dissimilarity with increasing distance within a region or, for the British Isles, temperature within a region (table 2*c*), these correlations were not statistically significant ($p > 0.05$). There were no significant relationships between metabolomic differences and genetic similarity in any of the focal regions (table 2*d*).

4. DISCUSSION

Our findings provide a wide-ranging overview of the scaling of environmental, genotypic and phenotypic variations in a focal species across a large fraction of its European range. We will discuss these three topics in the sections below, bringing together some general points raised by our findings.

(a) *The scaling of environmental variation*

While most models and many discussions of range margins have tacitly assumed that environments slide gradually and smoothly from core to marginal conditions, even a casual perusal of the natural environment suggests that much rougher and noisier gradients are typical. There is a growing literature (e.g. Pelletier 1997; Turcotte 1997; Halley *et al.* 2004; Scanlon *et al.* 2007) suggesting that substantial environmental variation can be found across a wide range of spatial scales, potentially producing spatially complex, fractal-like fitness surfaces for species across space. Species distributions, as well, typically show multi-scale patchiness (Erickson 1945; Kunin 1998; Wilson *et al.* 2002). The presence of such environmental and biotic variations may greatly complicate the interpretation of range margins, as it may allow core-like environments and populations to be found near the margins of a species' distribution. If core and marginal habitats are intimately interwoven, the necessary conditions for the scale of gene flow required to prevent

local adaptation could exist in many parts of a species range, as discussed in the introduction.

We measured only one aspect of environmental variation, near surface soil temperature, but this is likely to be important for the survival and reproduction of *A. l. petraea* in itself (P. Vergeer unpublished data) and to be correlated with other critical environmental variables. Our datasets provide evidence of substantial environmental variation within most populations, with some populations containing strikingly different microsites, and yet the great majority of variations were between sites and regions. In part, this may be due to the small number of samples taken at each site, but nonetheless we had endeavoured to capture some of the most extreme hot and cold microsites available within each population. Alternatively, the relatively modest variation in conditions sampled within many of our sites could itself be the evidence of micro-environmental habitat selection by our focal plant. If the wider landscape displays a considerable range of microclimates (e.g. Bennie *et al.* 2008), the fact that our populations inhabited only a limited range of temperature conditions in any specific area may signify the action of environmental constraints.

Elevation was a major factor explaining the among-site variation within regions and this partly explains why environmental variation in Sweden was lower than in other regions. Because the environment was so much more consistent in Sweden than other areas, it provides a poor test of our predictions about metabolomic variation among sites. However, the most marginal region, Britain and Ireland, has similar among-site environmental variation to the core regions, and yet did not display the predicted reduction in local phenotypic diversity.

(b) *Spatial structure of genetic diversity in core and margins of range*

We expected to find reduced genetic diversity within populations and greater differentiation among populations in marginal regions compared with core regions. Surprisingly, sites in marginal regions harboured levels of neutral genetic variation very similar to those in core regions, suggesting that local effective population sizes remain large on average, although clearly some individual demes are small. On the other hand, between-population spatial structure was stronger among the fragmented populations in these marginal areas. F_{ST} was three times higher in the British Isles region than the core regions. Genetic differentiation among regions and the marked difference among regions in the level of between-site variation could result, at least in part, from historical effects related to postglacial colonization rather than reflecting a drift–gene flow balance with restricted dispersal. However, recent phylogeographic work (Koch & Matschinger 2007) has suggested that *A. l. petraea* survived the glaciations in large populations in permafrost regions north of the central European ice sheets. This may explain why the present northwest European populations are genetically diverse and not structured by independent colonizations from different southern refugia. This suggests that the patterns of differentiation that we currently see are, indeed, a result of a present-day balance between dispersal and drift.

The striking pattern of high F_{ST} in marginal areas is reinforced by the shifting relationship between genetic similarity and distance as we move across the species' distribution. As we move from relatively densely occupied core landscapes (in Iceland) into more sparsely occupied landscapes (in Norway and Sweden) and ultimately to the extreme margins of the species' distribution (in Scotland, Wales and Ireland), we find increasingly pronounced isolation by distance relationships (figure 2). Fragmentation and isolation in marginal regions result in reduced gene flow and allow greater levels of genetic differentiation across space (Lesica & Allendorf 1995). Eckstein *et al.* (2006) also found steeper declines in genetic similarity with distance at range margins in their study of three *Viola* spp, and in one species (*Viola elatior*) differentiation was greater (although not significantly so) for widely spaced populations at the periphery than it was at the range centre, as in our findings. This pattern, if it is replicated in other biological systems, may help to explain an outstanding puzzle in spatial population genetics. There is a long-standing body of theory and data suggesting that genetic diversity falls as one moves towards the margins of a species' distribution (e.g. Carson 1959; Lewontin 1974; Lesica & Allendorf 1995). Conversely, others have argued that precisely the opposite pattern holds, with greater diversity at range margins (Fisher 1930; Burger 1988; Nevo 1988; Hoffmann & Parsons 1991; Parsons 1991), and still others suggest a hump-shaped pattern of increasing and then decreasing diversity as the margin is approached (Kark *et al.* 2008). Our results suggest that the dispute may ultimately come down to different scale perspectives on the same phenomenon: distributional margins may be both genetically depauperate (at within-population scales) and genetically rich (at between-population scales). Of course, observations of a single species across a single gradient are hardly conclusive, but our results suggest a potentially fruitful avenue for future research.

The question remains whether fragmentation in marginal regions, such as Scotland, Ireland and Wales in this case, results in poor local adaptation. Neutral genetic diversity may not be a good predictor of genetic variation for adaptive traits (Reed & Frankham 2001). Nevertheless, we can make two alternative predictions: protection from gene flow may allow local adaptation at range margins or alternatively small isolated populations may lack the genetic variation needed to adapt. In our system, neutral markers suggest that local populations are relatively more free of the restraining effects of gene flow from differently adapted populations in the marginal regions than those in the core regions, predicting greater phenotypic differentiation. They also suggest that the lack of variation is unlikely to constrain response to selection. However, populations may lack adaptive variation even when neutral diversity remains high, perhaps as a result of past selective pressures. The small number of surviving populations of *A. l. petraea* in Ireland and Wales and its sporadic distribution in Scotland (despite the widespread availability of apparently suitable habitat) suggest that adaptation is limited in these regions compared with more densely occupied areas such as Iceland and Norway. Our measurements of metabolic differentiation under common garden conditions allow us the possibility to test whether the variation among populations in Britain and Ireland is greater than that in the core regions, as predicted by

the pattern of environmental variation and neutral genetic variation.

(c) *Metabolomic phenotype*

Natural selection acts on the variation in phenotypes. Here, we considered a composite phenotype, the metabolic fingerprint, which provides a potentially much wider perspective than a single trait. These fingerprints can be used to identify ecologically important genes and pathways (Jackson *et al.* 2002; Hoffmann 2005; Benfey & Mitchell-Olds 2008). Because they integrate many possible responses to the environment, differences among populations may well be adaptive, although this has not been demonstrated directly. We hypothesized that this post-genomic phenotype would vary in different ways at range margins compared with the core. Variation among environmentally different sites may be reduced if marginal populations are unable to adapt, but may be greater if fragmentation and reduced gene flow remove constraints on adaptation. In fact, the metabolic phenotype, measured as variation in metabolic fingerprints obtained using direct injection mass spectrometry, did vary between geographically and climatically isolated populations within regions. Variation was markedly lower among the environmentally consistent Swedish localities than in more heterogeneous regions. However, it did not vary more or less strongly in the marginal and fragmented populations of the British Isles than among the large and connected populations of the core regions (Iceland and Norway), therefore fulfilling the predictions of neither of our two hypotheses.

Only a fraction of the metabolome was fingerprinted, and key components may not have been included. The weak patterns for metabolic phenotypes could be due to the laboratory rearing of plants, all grown under the same control conditions, if plasticity is the dominant mode of response to environmental conditions. Whether this is true remains to be tested and may help to establish the optimal metabolic phenotype for each population, which also remains unknown (Jackson *et al.* 2002). However, our finding is important regarding the spatial scaling of metabolic phenotypes. The low inter-population variability of metabolic fingerprints despite divergence at neutral loci may suggest stabilizing selection, despite the evidence for fine-scale variation in climate in some sites. The metabolic signal of adaptation for this species may not be measurable on such a fine scale (tens of km apart), where plastic responses to the fine-scale environmental variation may be dominant, but instead may only be detectable at a regional scale of hundreds of km as found using metabolic fingerprinting comparison across regions (Davey *et al.* 2008, in press) and targeted metabolic profiling of glucosinolates (Windsor *et al.* 2005). The degree of metabolic variation with spatial scale and genotype may vary greatly between species. For example, Petrakis *et al.* (2008) found that they could classify the geographic origin of olive oils within three regions in southern Greece at 87 per cent predictability, but this was reduced to 74 per cent when classified at a finer geographic scale. However, Ossipov *et al.* (2008), using metabolic profiling, were able to discriminate the different genotypes of birch (*Betula pendula*) and were able to discriminate which field trees were grown in. Such interesting results showing the influence that genetics, environment and space have on the metabolome need to be investigated further.

(d) *Local and regional adaptation*

As noted in the introduction, this paper falls short of testing the importance of gene flow in limiting local adaptation at fine-scale distributional limits, but it nonetheless provides evidence relevant to such ideas. We have demonstrated substantial variation in microclimates (specifically, in temperature profiles) between microsites within populations, sometimes comparable in scale to those found between populations separated by many kilometres, and similarly within-region differences that are commonly as great as those found between regions (figure 1). These differences appear to be great enough that local climatic adaptations between populations within regions, and indeed between microsites within populations, might be favoured by selection. The absence of spatial genetic structure within sites, however, suggests that gene flow at such fine scales may be quite strong, as would be expected in an obligately outcrossing animal-pollinated plant such as *A. l. petraea*. Such gene flow may help to explain the lack of spatial structure in metabolomic profiles, even between populations. On the other hand, when variance was partitioned between spatial scales (table 1*a*), greater proportions of genetic and (especially) metabolomic variances were found within populations than was found for temperature, holding open the possibility of substantial levels of local adaptation, even at these fine scales. Further research, including common garden performance trials or other experimental tests, would be required before definitive answers can be given as to the importance of local adaptation, and of gene flow, at fine spatial scales. Indeed, we are presently analysing the results of such experimental tests, which show clear evidence of local differentiation between populations within some regions (Vergeer *et al.* unpublished data). However, even with such results in hand, a study focusing on a single plant species could provide at best only tantalizing hints, rather than general answers. In *A. l. petraea*'s range (as in those of many other species), it is difficult to differentiate the effects of marginality from those of population fragmentation; our most marginal sites are also our most fragmented. Only with the accumulation of a wide range of studies, each addressing a specific case over a substantial range of spatial scales, will general patterns of gene flow and adaptation across scales become evident.

This research was funded by the Natural Environment Research Council Post-Genomics and Proteomics programme (NE/C507837/1). We thank B. Palmer, J. Edwards and T. Kubo for their analytical help and statistical advice, numerous field assistants for their efforts and Naomi Kingston of the Irish National Parks and Wildlife Service (Duchas), and Hywel Roberts and Barbara Jones of the Countryside Council for Wales, for their assistance with research permits. We also thank Kevin Gaston and two anonymous referees for their helpful comments.

REFERENCES

Alleaume-Benharira, M., Pen, I. R. & Ronce, O. 2006 Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *J. Evol. Biol.* **19**, 203–215. (doi:10.1111/j.1420-9101.2005.00976.x)

Angert, A. L. 2006 Demography of central and marginal populations of monkeyflowers (*Mimulus cardinalis* and *M. lewisii*). *Ecology* **87**, 2014–2025. (doi:10.1890/0012-9658(2006)87[2014:DOCAMP]2.0.CO;2)

Angert, A. L., Bradshaw, H. D. & Schemske, D. W. 2008 Using experimental evolution to investigate geographic range limits in monkeyflowers. *Evolution* **62**, 2660–2675. (doi:10.1111/j.1558-5646.2008.00471.x)

Antonovics, J., McKane, A. J. & Newman, T. J. 2006 Spatiotemporal dynamics in marginal populations. *Am. Nat.* **167**, 16–27. (doi:10.1086/498539)

Barton, N. H. 2001 Adaptation at the edge of a species' range. In *Integrating ecology and evolution in a spatial context* (eds J. Silvertown & J. Antonovics), pp. 365–392. Oxford, UK: Blackwell.

Benfey, P. N. & Mitchell-Olds, T. 2008 Perspective–from genotype to phenotype: systems biology meets natural variation. *Science* **320**, 495–497. (doi:10.1126/science.1153716)

Bennie, J., Huntley, B., Wiltshire, A., Hill, M. O. & Baxter, R. 2008 Slope, aspect and climate: spatially explicit and implicit models of topographic microclimate in chalk grassland. *Ecol. Model.* **216**, 47–59. (doi:10.1016/j.ecolmodel.2008.04.010)

Bridle, J. R. & Vines, T. H. 2007 Limits to evolution at range margins: when and why does adaptation fail? *Trends Ecol. Evol.* **22**, 140–147. (doi:10.1016/j.tree.2006.11.002)

Brown, J. H., Mehlman, D. W. & Stevens, G. C. 1995 Spatial variation in abundance. *Ecology* **76**, 2028–2043. (doi:10.2307/1941678)

Burger, R. 1988 The maintenance of genetic variation: a functional analytic approach to quantitative genetic models. In *Population, genetics and evolution* (ed. G. de Jong), pp. 63–72. Berlin, Germany: Springer-Verlag.

Butlin, R. K., Bridle, J. R. & Kawata, M. 2003 Genetics and the boundaries of species' distributions. In *Macroecology* (eds T. Blackburn & K. Gaston), pp. 274–295. Oxford, UK: Blackwell Science.

Carey, P. D., Watkinson, A. R. & Gerard, F. F. O. 1995 The determinants of the distribution and abundance of the winter annual grass Vulpia ciliate sspambigua. *J. Ecol.* **83**, 177–187. (doi:10.2307/2261556)

Carson, H. L. 1959 Genetic conditions that promote or retard the formation of species. *Cold Spr. Harb. Symp. Quant. Biol.* **24**, 87–103.

Case, T. J. & Taper, M. L. 2000 Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. *Am. Naturalist* **155**, 583–605. (doi:10.1086/303351)

Clauss, M. J. & Koch, M. A. 2006 Poorly known relatives of *Arabidopsis thaliana*. *Trends Plant Sci.* **11**, 449–459. (doi:10.1016/j.tplants.2006.07.005)

Clauss, M. J. & Mitchell-Olds, T. 2006 Population genetic structure of *Arabidopsis lyrata* in Europe. *Mol. Ecol.* **15**, 2753–2766. (doi:10.1111/j.1365-294X.2006.02973.x)

Cohan, F. M. 1984 Can uniform selection retard genetic divergence between isolated conspecific populations? *Evolution* **38**, 495–504. (doi:10.2307/2408699)

Coope, G. R. 1995 Insect faunas in ice age environments: why so little extinction? In *Extinction rates* (eds J. H. Lawton & R. M. May), pp. 55–74. Oxford, UK: Oxford University Press.

Cresswell, J. E. 1997 Spatial heterogeneity, pollinator behaviour and pollinator-mediated gene flow: bumblebee movements in variously aggregated rows of oil–seed rape. *Oikos* **78**, 546–556. (doi:10.2307/3545616)

Davey, M. P., Burrell, M. M., Woodward, F. I. & Quick, W. P. 2008 Population specific metabolic phenotypes of *Arabidopsis lyrata* ssp. *petraea*. *New Phytol.* **177**, 380–388. (doi:10.1111/j.1469-8137.2007.02282.x)

Davey, M. P., Woodward, F. I. & Quick, W. P. In press Intraspecfic variation in cold-temperature metabolic phenotypes of *Arabidopsis lyrata* ssp. *petraea*. *Metabolomics*. (doi:10.1007/s11306-008-0127-1)

Eckstein, R. L., O'Neill, R. A., Danihelka, J., Otte, A. & Kohler, W. 2006 Genetic structure among and within peripheral and central populations of three endangered floodplain violets. *Mol. Ecol.* **15**, 2367–2379. (doi:10.1111/j.1365-294X.2006.02944.x)

Erickson, R. O. 1945 The Clematis fremontii var. riehlii population in the Ozarks. *Ann. Mo. Bot. Gard.* **32**, 413–460. (doi:10.2307/2394445)

Felsenstein, J. 1983 *Theoretical population genetics*. Seattle, WA: University of Washington.

Fenart, S., Austerlitz, F., Cuguen, J. & Arnaud, J. F. 2007 Long distance pollen-mediated gene flow at a landscape level: the weed beet as a case study. *Mol. Ecol.* **16**, 3801–3813. (doi:10.1111/j.1365-294X.2007.03448.x)

Filin, I., Holt, R. D. & Barfield, M. 2008 The relation of density regulation to habitat specialization, evolution of a species' range, and the dynamics of biological invasions. *Am. Nat.* **172**, 233–247. (doi:10.1086/589459)

Fisher, R. A. 1930 *The genetical theory of natural selection*. Oxford, UK: Clarendon Press.

Garant, D., Forde, S. E. & Hendry, A. P. 2007 The multifarious effects of dispersal and gene flow on contemporary adaptation. *Funct. Ecol.* **21**, 434–443. (doi:10.1111/j.1365-2435.2006.01228.x)

Gaston, K. J. 2003 *The structure and dynamics of geographic ranges*. Oxford series in ecology and evolution. Oxford, UK: Oxford University Press.

Halley, J. M., Hartley, S., Kallimanis, A. S., Kunin, W. E., Lennon, J. J. & Sgardelis, S. P. 2004 Uses and abuses of fractal methodology in ecology. *Ecol. Lett.* **7**, 254–271. (doi:10.1111/j.1461-0248.2004.00568.x)

Hardy, O. J. & Vekemans, X. 2002 SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* **2**, 618–620. (doi:10.1046/j.1471-8286.2002.00305.x)

Hochberg, M. E. & Ives, A. R. 1999 Can natural enemies enforce geographical range limits? *Ecography* **22**, 268–276. (doi:10.1111/j.1600-0587.1999.tb00502.x)

Hoffmann, M. H. 2005 Evolution of the realized climatic niche in the genus *Arabidopsis* (Brassicaceae). *Evolution* **59**, 1425–1436. (doi:10.1111/j.0014-3820.2005.tb.01793.x)

Hoffmann, A. A. & Parsons, P. A. 1991 *Evolutionary genetics and environmental stress*. New York, NY: Oxford University Press.

Holt, R. D. 2003 On the evolutionary ecology of species' ranges. *Evol. Ecol. Res.* **5**, 159–178.

Holt, R. D. & Gomulkiewicz, R. 1997 How does immigration influence local adaptation? A reexamination of a familiar paradigm. *Am. Nat.* **149**, 563–572. (doi:10.1086/286005)

Holt, R. D. & Keitt, T. H. 2005 Species' borders: a unifying theme in ecology. *Oikos* **108**, 3–6. (doi:10.1111/j.0030-1299.2005.13145.x)

Honnay, O. & Jacquemyn, H. 2007 Scuseptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conserv. Biol.* **21**, 823–831. (doi:10.1111/j.1523-1739.2006.00646.x)

Jackson, R. B., Linder, C. R., Lynch, M., Purugganan, M., Somerville, S. & Thayer, S. S. 2002 Linking molecular insight and ecological research. *Trends Ecol. Evol.* **17**, 409–414. (doi:10.1016/S0169-5347(02)02571-5)

Jalas, J. & Suominen, J. 1994 *Atlas Florae Europaeae* vol. 10 Cruciferae (Sysymbrium to Aubrieta). Helsinki Univ. Publishing House.

Jonsell, B., Kustås, K. & Nordal, I. 1995 Genetic variation in *Arabis petraea*, a disjunct species in northern Europe. *Ecography* **18**, 321–332. (doi:10.1111/j.1600-0587.1995.tb00135.x)

Jump, A. S. & Penuelas, J. 2005 Running to stand still: adaptation and the response of plants to rapid climate change. *Ecol. Lett.* **8**, 1010–1020. (doi:10.1111/j.1461-0248.2005.00796.x)

Kark, S., Hadany, L., Safriel, U. N., Noy-Meir, I., Niles Eldredge, N., Tabarroni, C. & Randi, E. 2008 How does genetic diversity change towards the range periphery? An empirical and theoretical test. *Evol. Ecol. Res.* **10**, 391–414.

Kenta, T., Gratten, J., Haigh, N. S., Hinten, G. N., Slate, J., Butlin, R. K. & Burke, T. 2008 SNP-SCALE: a cost-effective medium-throughput single nucleotide polymorphism genotyping method. *Mol. Ecol. Resour.* **8**, 1230–1238.

Kirkpatrick, M. & Barton, N. H. 1997 Evolution of a species' range. *Am. Naturalist* **150**, 1–23. (doi:10.1086/286054)

Koch, M. A. & Matschinger, M. 2007 Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **104**, 6272–6277. (doi:10.1073/pnas.0701338104)

Kristinsson, H. 1995 *A Guide to the Flowering Plants and Ferns of Iceland*. Reykjavik, Iceland: Mál og menning.

Kunin, W. E. 1998 Extrapolating species abundance across spatial scales. *Science* **281**, 1513–1515. (doi:10.1126/science.281.5382.1513)

Lennon, J. J., Turner, J. R. G. & Connell, D. 1997 A metapopulation model of species boundaries. *Oikos* **78**, 486–502. (doi:10.2307/3545610)

Lenormand, T. 2002 Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17**, 183–189. (doi:10.1016/S0169-5347(02)02497-7)

Lesica, P. & Allendorf, F. W. 1995 When are peripheral populations valuable for conservation? *Conserv. Biol.* **9**, 753–760. (doi:10.1046/j.1523-1739.1995.09040753.x)

Lewis, P. O. & Zaykin, D. 2001 Genetic data analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). Available from the authors at: http://lewis.eeb.uconn.edu/lewishome/software.html.

Lewontin, R. C. 1974 *The genetic basis of evolutionary change*. New York, NY: Colombia University Press.

Loiselle, B. A., SorkV, L., Nason, J. & Graham, C. 1995 Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* **82**, 1420–1425. (doi:10.2307/2445869)

Mandák, B., Bímova, K., Plačková, I., Mahelka, V. & Chrtek, J. 2005 Loss of genetic variation in geographically marginal populations of Atriplex tatarica (Chenopodiaceae). *Ann. Bot.* **96**, 901–912. (doi:10.1093/aob/mci242)

Merrill, R. M., Gutierrez, D., Lewis, O. T., Gutierrez, J., Diez, S. B. & Wilson, R. J. 2008 Combined effects of climate and biotic interactions on the elevational range of a phytophagous insect. *J. Anim. Ecol.* 77, 145–155. (doi:10.1111/j.1365-2656.2007.01303.x)

Meyer, R. C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Törjeék, O., Fiehn, O., Eckardt, Ä. . , Willmitzer, L., Selbig, J. & Altmann, T. 2007 The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **104**, 4759–4764. (doi:10.1073/pnas.0609709104)

Mitchell-Olds, T. 2001 *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends Ecol. Evol.* **16**, 693–700. (doi:10.1016/S0169-5347(01)02291-1)

Nevo, E. 1988 Genetic diversity in nature–patterns and theory. In *Evolutionary biology* (eds M. K. Hecht & B. Wallace), pp. 217–246. New York, NY: Plenum Press.

Ossipov, V., Ossipova, S., Bykov, V., Oksanen, E., Koricheva, J. & Haukioja, E. 2008 Application of metabolomics to genotype and phenotype discrimination of birch trees grown in a long-term open-field experiment. *Metabolomics* **4**, 39–51. (doi:10.1007/s11306-007-0097-8)

Parsons, P. A. 1991 Evolutionary rates: stress and species boundaries. *Ann. Rev. Ecol. Syst.* **22**, 1–18. (doi:10.1146/annurev.es.22.110191.000245)

Pigliucci, M., Cammell, K. & Schmitt, J. 1999 Evolution of phenotypic plasticity a comparative approach in the phylogenetic neighbourhood of *Arabidopsis thaliana*. *J. Evol. Biol.* **12**, 779–791. (doi:10.1046/j.1420-9101.1999.00074.x)

Pelletier, J. D. 1997 Analysis and modeling of the natural variability of climate. *J. Climate* **10**, 1331–1342. (doi:10.1175/1520-0442(1997)010<1331:AAMOTN>2.0.CO;2)

Petrakis, P. V., Agiomyrgianaki, A., Christophoridou, S., Spyros, A. & Dais, P. 2008 Geographical characterization of Greek virgin olive oils (cv. Koroneiki) using H-1 and P-31 NMR fingerprinting with canonical dascriminant analysis and classification binary trees. *J. Agric. Food Chem.* **56**, 3200–3207. (doi:10.1021/jf072957s)

R Development Core Team 2008 R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

Reed, D. H. & Frankham, R. 2001 How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* **55**, 1095–1103. (doi:10.1554/0014-3820(2001)055[1095:HCCAMA]2.0.CO;2)

Roy, D. B. & Thomas, J. A. 2003 Seasonal variation in the niche, habitat availability and population fluctuations of a bivoltine thermophilous insect near its range margin. *Oecologia* **134**, 439–444. (doi:10.1007/s00442-002-1121-3)

Sagarin, R. D. & Gaines, S. D. 2002 The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecol. Lett.* **5**, 137–147. (doi:10.1046/j.1461-0248.2002.00297.x)

Scanlon, T. M., Caylor, K. K., Levin, S. A. & Rodriguez-Iturbe, I. 2007 Positive feedbacks promote power-law clustering of Kalahari vegetation. *Nature* **449**, 209–212. (doi:10.1038/nature06060)

Schmid, K. J., Ramos-Onsins, S., Ringys-Beckstein, H., Weisshaar, B. & Mitchell-Olds, T. 2005 A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. *Genetics* **169**, 1601–1615. (doi:10.1534/genetics.104.033795)

Soons, M. B. & Bullock, J. M. 2008 Non-random seed abscission, long-distance wind dispersal and plant migration rates. *J. Ecol.* **96**, 581–590. (doi:10.1111/j.1365-2745.2008.01370.x)

Turcotte, D. L. 1997 *Fractals and chaos in geology and geophysics*. Cambridge, UK: CUP.

Venable, D. L., Flores-Martinez, A., Muller-Landau, H. C., Barron-Gafford, G. & Becerra, J. X. 2008 Seed dispersal of desert annuals. *Ecology* **89**, 2218–2227. (doi:10.1890/07-0386.1)

Vergeer, P., Van den Berg, L. J. L., Bulling, M. T., Ashmore, M. R. & Kunin, W. E. 2008 Geographical variation in the response to nitrogen deposition in *Arabidopsis lyrata petraea*. *New Phytol.* **179**, 129–141. (doi:10.1111/j.1469-8137.2008.02445.x)

Vinocur, B. & Altman, A. 2005 Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Curr. Opin. Biotechnol.* **16**, 123–132. (doi:10.1016/j.copbio.2005.02.001)

Whitlock, M. C. 2003 Fixation probability and time in subdivided populations. *Genetics* **164**, 767–779. (doi:10.2307/2408196)

Whitlock, R., Hipperson, H., Mannarelli, M. & Burke, T. 2008 A high-throughput protocol for extracting high-purity genomic DNA from plants and animals. *Mol. Ecol. Res.* **8**, 736–741. (doi:10.1111/j.1755-0998.2007.02074.x)

Willis, J. C. 1922 *Age and area*. Cambridge, UK: Cambridge University Press.

Wilson, R. J., Ellis, S., Baker, J. S., Lineham, M. E., Whitehead, R. W. & Thomas, C. D. 2002 Large-scale patterns of distribution and persistence at the range margins of a butterfly. *Ecology* **83**, 3357–3368. (doi:10.1890/0012-9658(2002)083[3357:LSPODA]2.0.CO;2)

Windsor, A. J., Reichelt, M., Figuth, A., Svatos, A., Kroymann, J., Kliebenstein, D. J., Gershenzon, J. & Mitchell-Olds, T. 2005 Geographic and evolutionary diversification of glucosinolates among near relatives of *Arabidopsis thaliana* (Brassicaceae). *Phytochemistry* **66**, 1321–1333. (doi:10.1016/j.phytochem.2005.04.016)