Research

Ecography 35: 446–457, 2012
doi: 10.1111/j.1600-0587.2011.07030.x

Subject Editor: Francisco Pugnaire. Accepted 17 June 2011

Genetic structure among arable populations of *Capsella bursa-pastoris* is linked to functional traits and in-field conditions

Graham S. Begg, Jane Wishart, Mark W. Young, Geoff R. Squire and Pietro P. M. Iannetta

G. S. Begg (graham.begg@hutton.ac.uk), J. Wishart, M. W. Young, G. R. Squire and P. P. M. Iannetta, The James Hutton Inst., Invergowrie, Dundee DD2 5DA, Scotland, UK.

Wild arable plants can be an economic burden but they also support diverse arable food webs and contribute to valuable ecosystem functions. These benefits may have been compromised over recent decades by declining weed diversity. The decline in wild arable plant diversity has been viewed predominantly in terms of species shifts a view that ignores the genetic and functional variation existing within species and the impact on ecological and evolutionary processes which this may have. To examine within-species diversity, ISSR markers were used in parallel with environmental and phenotypic characterisation, to investigate the population structure and diversity of *Capsella bursa-pastoris* (shepherd's purse) from arable fields in the UK. Analysis of 338 ISSR products for 109 individuals from 51 accessions obtained from the seed banks of 33 arable fields showed that in-field populations of shepherd's purse were genetically differentiated between individuals, and among accessions and fields. In addition, cluster analysis identified three genetically distinct regional-scale populations. Phenotypic variation was present at all scales of genetic differentiation, including the regional scale where populations differed in their key life-history traits: flowering time, fecundity and dormancy. Genetic drift is proposed as a contributor to differentiation among genetically isolated but locally co-occurring accessions. In addition, the genetic and phenotypic variation in shepherd's purse exhibited large scale, spatial trends and showed statistically significant associations with cropping intensity and soil-pH. These results suggest that adaptation as a result of selection by cropping practise and soil-pH has played a role in the ability of shepherd's purse to colonise and persist in arable fields.

The arable landscape is essential for food security (Rothstein 2007) and the arable wild plants in this landscape mediate essential ecological processes (Altieri 1999). Arable landscapes have nevertheless suffered some of the greatest reductions of wild plant biodiversity in the last 50 yr (Robinson and Sutherland 2002, Marshall et al. 2003). Many arable plant species appear to lack the capacity to adapt to the extreme selection pressures presented by the diverse array of agricultural practices, in particular, the increasing reliance upon efficient species-indiscriminate chemical herbicides and winter cropping (Marshall et al. 2003). However, some long established species appear well adapted as they remain in abundances while in addition a few new species have become commonplace, such as crop-derived volunteer weeds (Debeljak et al. 2008).

The ecological and economic implications of such changes in weed composition are still uncertain. They have been examined mainly through the unit of plant species, commonly by assigning an inferred ecological function to each species (Hawes et al. 2005). In some instances the inferred function is no more than general: for example, wild oat *Avena fatua* or blackgrass *Alopecurus myosuroides* is considered economically damaging and of little value in the arable food web, while shepherd's purse or redshank *Persicaria maculosa* support the food web by providing food and shelter to invertebrates without being economically damaging. In other instances the assignment of function is much more specific, based for example on life-history characteristics, nutritional quality and trophic links to invertebrates (Hawes et al. 2009). Such a species-based approach is nevertheless limited by ignoring the great genetic and functional variation that may exist within the commoner species (Hawes et al. 2005) and the impact on ecological and evolutionary processes this may have.

An understanding of within-species variation and its consequences is therefore needed for an assessment of the economic and ecological consequences of change in the arable flora. However, at present there is little knowledge of whether within-species variation in arable plants responds to selection pressures through management, soil or climate and whether such variation has ecological significance, for example by influencing food webs or crop-weed competition. Much of the research on causative factors that drive natural selection has been carried out in biodiverse tropical regions on animal, and especially insect species, due to their more obvious (compared to plants) phenotypic shifts (Silvertown 2004). There is limited and inconclusive evidence more generally that the functional attributes of arable plants segregate along environmental niche axes. Individual-based models suggest that they do, in that plants interacting by physiological traits give rise to assemblages having ecological properties (Pachepsky et al. 2001, Bown et al. 2007).

However, carefully constructed field experiments to test the hypotheses remain to be carried out.

It is possible that the diverse variety of environments in arable landscapes, often frequently and intensely disturbed, might themselves offer the conditions that allow examination of the hypothesis that plant variants are consequents of differences in locality and management. The occurrence of trait- or character-displacement under extreme selection pressures has been exemplified most dramatically to date by the increased prevalence of herbicide tolerant ecotypes (Holt and LeBaron 1990). Confirmation of trait-displacement among more general phenological attributes requires a set of potential plant variants whose genetic and phenotypic characteristics can be linked to factors of locality and management. Among the common and widespread arable plants, shepherd's purse is a suitable model, having a range of genetic variants that differ phenotypically in traits such as time to flowering, duration of the reproductive period and germination (Iannetta et al. 2007, Iannetta 2010). Its abundance and prevalence (assessed at the level of species) in the arable seedbank differs greatly with geographical location and the intensity of field management (Debeljak et al. 2008). A means is therefore needed to define the genetic differences among the variants of shepherd's purse and to examine their link to environment and phenotype.

In this respect, inter-simple sequence repeats (ISSRs) emerge as a suitable tool (Bornet and Branchard 2001). This technique may be applied with relative ease to low amounts of even crude genomic DNA extracts in a rapid and high throughput fashion, as no prior DNA manipulation, or 'discovery' approaches are necessary before the standard PCR amplification. These advantages allow easy application of the method to new species and high reproducibility across different laboratories that is not afforded by other dominant (and co-dominant), marker techniques (Wolfe and Liston 1998). This method is universal in its applicability to eukaryotic genomes as SSRs are present and abundant in a wide range of organisms. Furthermore, ISSRs are inherited in a Mendelian fashion, are hyper-variable (in length) and occur with sufficient frequency to allow a high mapping density that is representative of the euchromatic gene space (Munn et al. 2006). Consequently, ISSR markers have been used to discriminate crop cultivars (Charters et al. 1996), in hybridization studies of natural plant populations (Sica et al. 2005) and population genetic studies of the invasive weeds (Wang et al. 2008). The aim, in these studies has been to examine phylogenetic relationships, to describe population structure and diversity, to assess hybridization or find population- or cultivar-specific markers (Whitlock et al. 2007). Most recently, an ISSR approach has also been used to categorise accessions and to assess the relative abundance of types in grassland systems (Fridley and Grime 2007, Whitlock et al. 2007).

In this study, ISSR markers are used to describe the genetic differentiation and population structure of shepherd's purse from arable fields in the UK (Iannetta et al. 2007), to explore their phenotypic differences and environmental associations, and to examine the existence and possible implications of genetic diversity of these populations.

Material and methods

Plant material

Accessions of shepherd's purse were isolated from soil seedbank samples collected in 2002 from 33 (around 15%) of the arable fields used in the Farm Scale Evaluations (FSE; Squire et al. 2003). To obtain a wide range of conditions, fields were selected from those with high, medium and low baseline seedbanks, and from locations proportionate to the total, from the west (e.g. Gloucestershire, Warwickshire, Cheshire), south central and south east (Hertfordshire, Cambridgeshire, Norfolk), south of the Humber (Lincolnshire, Nottinghamshire), north-east England (Humberside, Yorkshire) and north-east Scotland. Data for each site on soil-carbon (%), nitrogen (%), pH and crop rotation history were described by Debeljak et al. (2008). As part of general seedbank sampling for each site, shepherd's purse seeds were germinated after sieving soil and spreading it in trays to encourage emergence (Debeljak et al. 2008). Of the shepherd's purse plants emerging from each site, a total of 51 individual plant accessions, up to three from each field-site, were transferred to pots and grown to maturity when seed was collected. Phenotypic characterisation was conducted on the progeny of accession plants to minimise bias introduced by differences in the environmental conditions to which the accessions had initially been exposed. In total, 109 plants were grown, comprised of typically, two plants (siblings) from the seed of each accession. Seeds were first processed using sterile culture. Samples of 2 mg of seed were surface sterilised in 1 ml of sterilising solution (1% (final concentration) sodium hypochlorite (NaOCl), with 0.01% [v/v] Tween20 for 5 min with shaking. After washing (3×) in 1 ml of sterile distilled water the seeds were germinated on 1% sterile distilled water agar [1% w/v] and stratified for 3 d at 4°C) and cultured at 21°C. One-week-old seedlings were transferred to small peat pots for 6 weeks before being transferred to large 15 litre peat containing pots which were positioned in a random plot design in a glasshouse. The plants were provided with water daily (150 ml) and 150 ml of liquid fertiliser (10% [w/v] '20:20:20 Sangral Soluble Fertiliser', William Sinclair Horticulture, Lincoln UK) every seventh day. Plants were provided with 18 h of daylight (>250 W m^{-2}) and a day/night temperature regime of 25/15°C. Plants were isolated using breathable clear plastic bags before flowering to ensure self-pollination and to enable seed collection. Plants were allowed to fully mature and senesce before seeds were harvested (cf. Iannetta et al. 2007). Phenotypic characters were scored for each of the 109 progeny plants.

DNA extraction

From each of the 109 plants representing 1 to 3 progeny from 51 accessions, two liquid-nitrogen frozen leaf-discs (ca 8 mm diameter) were collected in a 1.2 ml micro-tube (in a 96-well format; Qiagen, no. 19560) that contained a single sterile 3 mm diameter tungsten-carbide ball-bearing. The

−80°C frozen material was milled for 1 min at 30 Hz (×2) in a Retsch mixer-mill (no. MM301).

Total DNA was extracted from the disrupted tissue according to a modified method from Mogg and Bond (2003) by incubation in 400 µl extraction buffer (100 mM Tris., 50 mM EDTA, 500 mM NaCl, 0.7% [w/v] SDS, 50 µg ml^{-1} Proteinase K and 50 µg ml^{-1} RNase) at 37°C for 2 h. NaCl (5M, 260 µl), was then mixed with the extract before centrifugation at 4000 rpm for 5 min. The DNA-containing supernatant was transferred to a new microtube containing 700 µl of 85% [v/v] isopropanol, mixed well and incubated at −80°C for 1 h before re-centrifugation at 5000 rpm for 5 min.

The pellets were washed by addition of 300 µl 70% [v/v] ethanol, microfuged at 5000 rpm for 5 min and dried at 37°C for 15 min before re-suspension in 100 µl sterile distilled water (SDW). Concentrations were measured using NanoDrop (ND-1000, Nanodrop Technologies) and DNA diluted to ca 20 ng μl^{-1} (usually a 1:10 dilution).

ISSR amplification and fragment analysis

Fourteen ISSR primers (Charters et al. 1996) were synthesised with a 5'6-FAM label (Sigma-Aldrich) and tested in single-primer reactions using a range of published methods and optimised on a 50–62°C temperature gradient to obtain reaction conditions for each primer with the shepherd's purse DNA template. The six most successful of these ISSR primers were chosen for the final experiment (Table 1). The reaction conditions were described by Sica et al. (2005), with 10 ng DNA, 2 mM $MgSO_4$, 0.2 mM each dNTP, 0.6 µM primer, 0.5 U Platinum Taq DNA polymerase (Invitrogen), 1 × PCR buffer in a total volume of 12.5 µl. The PCR thermal profile was 1.5 min at 94°C; 35 cycles of 40 s at 94°C, 45 s at the primer annealing optimal temperature (Table 1), 1.5 min at 68°C, 45 s at 94°C, and 45 s at the annealing temperature before extension at 68°C for 5 min. Negative controls were included in each run where all reagents but DNA were added to the reaction mix. The total number of samples processed, including extraction replicates (ca 2–4 per individual), varied for each primer from between 198 and 242.

Four rows from each 96 well plate were checked on an agarose gel and visualised by SYBR Safe staining (Invitrogen). 895 µl Hi Di Formamide (Applied Biosystems) and 5 µl X-Rhodamine MapMarker 1000 (BioVentures) were mixed and 9 µl added to each well of a 96 well sequencing plate. 1 µl of each PCR product (diluted 1:10) was then added. Gel electrophoresis was performed on an ABI automated sequencer (Applied Biosystems '3730 DNA Analyzer' ABI3730, 36 cm capillary array (48 capillaries)) using 'pop7 performance optimized polymer'. All gels were run for 1800 s at 15 KV using Data Collection software ver. 3.0 (Applied Biosystems). Sample files were imported into GeneMapper Software ver. 3.7 (Applied Biosystems) for analysis. The single stranded DNA MapMarker with a single fluorophore allowed accurate sizing of fragments from 50 to 1000 bps. Projects were created for each ISSR primer and 'bins' allocated for each fragment size with peak fluorescence exceeding 200 Fluorescence Units (FU). The fluorescence of each resulting product was expressed as a proportion of the total fluorescence for that sample and expressed as Relative Fluorescence Units (RFUs). For individual plants the mean RFU was calculated for each product (averaged over the available extraction replicates) and the presence of an ISSR product was established where the mean RFU exceeded a threshold of 1% of the total RFU.

Plant measurements

Plant traits were quantified as described in Iannetta et al. (2007) and include, reproductive duration (d), flowering time (d), total flowering stem dry-weight (g), number of flowering stems, total seed weight (mg), seed number, mean seed weight (mg), rosette diameter (mm), number of leaves, leaf relative chlorophyll concentration (maximum of three Single Photon Avalanche Diode (SPAD) containing meter readings), seed germination rate (expressed as the time in days to 50% germination), and germination fraction (%). Treatment with gibberellic acid showed seed viability to exceed 99% so that the complement of germination fraction is subsequently considered as a measure of dormancy, i.e. the 'dormant fraction'.

Site measurements

Site data was collated as part of the Farm Scale Evaluations (Squire et al. 2003) and are publicly accessible (<www.farmscale.org.uk>). The field characteristics recorded were: Easting and Northing defined as distance east and north in km from an origin which is 100 km north and 400 km

Table 1. Details the five 5'-labelled 6-carboxyfluorescein (6-FAM) inter-simple sequence repeat (ISSR) primers and the (50 to 1000 bp) products they generated with shepherd's purse genomic DNA. Primer-sequences obtained from the Univ. of British Columbia (UBC, Biotechnology Laboratory and The James Hutton Inst.) (Charters et al. 1996). T_a, temperature of annealing. The total number of individuals assessed using each primer is noted alongside the total number of ISSR products obtained per primer across all the assessed plants. The proportion of ISSR products per individual (± SD) and the number of polymorphic ISSR products are also detailed.

Primer	Sequence	T_a	Individuals	ISSRS	Polymorphic ISSRs	% ISSRs $individual^{-1}$ (± SD)
UBC 880	6-FAM-VHV $(GT)_7$	53.4	109	73	71	30.0 (4.2)
UBC 886	6-FAM-VBV $(CT)_7$	53.4	104	79	77	36.2 (4.6)
UBC 888	6-FAM-BDB $(CT)_7$	60.0	101	34	30	38.1 (5.3)
UBC 889	6-FAM-DBD $(AC)_7$	60.0	105	54	53	25.2 (4.6)
SCRI 1417	6-FAM-BDT $(CA)_7$	57.2	105	53	48	29.3 (4.8)
SCRI 1424	6-FAM-BDB $(CAC)_5$	57.2	95	45	42	46.2 (6.6)

west of true origin (2°W, 49°N), current crop type, current crop intensity, previous crop intensity for the preceding 1 to 6 yr, soil-pH, soil-nitrogen (%), soil-carbon (%), soil texture (cf. Hodgson 1974), seedbank diversity (total number of species that emerged from the soil samples processed from a location, standardised across treatments), total seedbank density and shepherd's purse seedbank density (both densities expressed as emerged seedlings m^{-2} of the field-site, sampling depth being standardised across samples at 15 cm).

A rank of 'cropping intensity' values were used to characterise each particular crop (or land management) type as applied at the time of soil sampling (for seedbank analysis that provided the shepherd's purse accessions tested). The cropping intensity values were attributed on the basis of professional (farmers') opinion and justified scientifically on the basis of available data for total pesticide (fungicide, insecticide and herbicide), applications defined by their combined rate in kilogrammes of 'active substance applied per hectare' (a.s. kg ha^{-1}) of the total crop area (Garthwaite and Thomas 2006, Table 2). Cropping intensity values were categorical and ranged from 1 (least intense, e.g. untreated fallow) to 7 (most intense, e.g. potato). Values for swede, maize and beet were estimated from local data. Due to the wide variation in application rates, the average nitrogen application per crop, or land management, type could not be ascertained. Nitrogen application rates were therefore not used to ascertain the cropping intensity ranking.

Statistical analysis

A principal component analysis (PCA) based on the correlation matrix of crop intensity scores for the current (year 0) and previous years (years 1–6; Table 2) was used to establish whether the cropping intensity experienced by individual fields was consistent over this 7 yr period. To complete the field characterisation, a second PCA was then performed on the first two components obtained from the intensity analysis plus the remaining field characteristics.

Table 2. Cropping intensity scores associated with the various crop options employed in rotations at those FSE sites from which the soil samples yielding the shepherd's purse accessions tested here were obtained. Cropping intensity ranking was ascribed on the basis of professional (farmers) opinion. This table demonstrates that the ranking attributed is borne out by data for total pesticide (fungicide, insecticide and herbicide), defined by their combined 'application rate' in kg of active substance per hectare (a.s. kg ha^{-1}) of the total crop area (Garthwaite and Thomas 2006).

Crop	Cropping intensity	Pesticide (a.s. kg ha^{-1})
Fallow/rough grazing	1	0.01
Grass lay	2	0.04
Swede/turnip	3	0.71
Kale	3	0.80
Spring oil seed rape	3	0.80
Spring wheat	4	0.85
Spring barley	4	1.20
Peas	5	2.05
Winter barley	5	2.30
Winter oil seed rape	5	2.58
Winter wheat	5	2.59
Beet	6	3.00
Potatoes	7	4.26
Maize	7	4.50
Carrot	7	5.62

PCA was used to examine plant phenotype using the correlation matrix of plant traits calculated from individual plant measurements and repeated using the mean values for each accession. A random effects model was then fitted by restricted maximum likelihood to each of the principle components obtained from the first of these PCAs (the dependent variables) to estimate the phenotypic variance among the nested random effects of field-site, accession, and the individual progeny. Refitting by maximum likelihood, these models were extended to include the field-site characteristics independently of each other, with the fixed effects of Easting and Northing as a two-dimensional quadratic regression surface. The significance of these was determined by likelihood ratio test.

The ISSR data were analysed to assess the genetic structure and diversity of shepherd's purse populations using a variety of distance (band-based sensu Bonin et al. 2007) and allele frequency based methods. These were agglomerative hierarchical clustering (UGPMA) of ISSR haplotypes, Bayesian model-based clustering using the STRUCTURE software (Pritchard et al. 2000) and the estimation of hierarchical variance components and F-statistics by AMOVA (Excoffier et al. 1992) further details of which are provided below and in the Results section.

ISSRs are dominant markers and each product was treated as the phenotype at a single biallelic locus on the basis of the assumptions that both the presence and absence of a product are homologous and that length polymorphisms are absent so that non co-migrating products are independent. The validity of these assumption remains to be established, though presence homoplasy within species is thought to be limited (Bonin et al. 2007), while the distance methods, with the exception of the AMOVA, used Jaccard's coefficient which discounts the similarity between individuals resulting from shared product absence so that the results are insensitive to absence homoplasy. Even with the acceptance of these assumptions, ISSRs, like other dominant markers, are subject to an inherent ambiguity in the representation of genotype. Though tetraploid, shepherd's purse shows disomic inheritance (Hurka and Düring 1994) and the ambiguity lies in distinction between the homozygote dominant and heterozygous genotypes which are indistinguishable. This ambiguity was accounted for by the model based clustering methods employed but not by the distance based methods which treat the ISSR data as haplotypic.

The model-based clustering approach of STRUCTURE (Pritchard et al. 2000) assumes that sampled individuals are drawn from a fixed number of populations, K, that are characterised by a set of allele frequencies at each locus; allele frequencies and the population of origin of each allele are inferred from the data using a Markov Chain Monte Carlo algorithm. In all cases the 'admixture model' (Pritchard et al. 2000) was used, allowing the ancestry of individuals to be inferred by estimation of the admixture proportions, i.e. the fraction of an individual's genotype that originated from each of the K populations. In addition, field-site was used to inform the prior distribution of the admixture proportions (Hubisz et al. 2009), and population allele

frequencies were assumed to be correlated between the K clusters (Pritchard et al. 2000). Inference of population number was addressed heuristically by comparing the penalized log likelihood of models of differing K. For each model of K 1 to 6, five independent runs were completed. All simulations were run using a burn in period of 30 000 iterations with a post-burn in period of 10 000 iterations.

Following the model-based cluster analysis, individuals were assigned to one of several clusters on the basis of estimated admixture proportions (see Results section). Variation in cluster assignments in response to the field-site characteristics were analysed by mixed-effects logistic regression models in which field and accession were treated as random effects. The field-site characteristics were treated as fixed effects and their significance determined independently of each other by comparison with the null model using the likelihood ratio test. The exception to this was made were a significant spatial trend was identified, in this case the effect of field-site characteristics was also tested against the model inclusive of the spatial trend. The model-based cluster assignments were also the basis of an inter-cluster comparison of phenotype. This was carried out by multi-level ANOVA of the plant traits plus composites of these, i.e. the first three components derived from the plant trait PCA described above. The significance of differences between clusters was determined by conditional F-test.

Throughout the analysis of field-site, plant trait and genetic data, the suitability of the linear model formulation was assessed by graphical analysis of the model residuals.

Results

Site characteristics

Crop intensity was positively correlated across all 7 yr of the rotations indicating that fields experienced crop rotations of consistent intensity. This predominant pattern in crop intensity was captured by the first component of the PCA with positive loadings for the crop intensity rankings from each year (data not shown), greatest for years 1 and 2. This component, which explained 28% of the variation in the crop intensity data, was used as an additional composite descriptor of cropping intensity, termed CI, in subsequent analysis. The second component obtained from the principal component analysis explained 21% of the variance and was dominated by current crop intensity (year 0).

The correlations between the other field characteristics were less pronounced with just 30% of the variance being explained by the first two components. The biplot of these components shows that soil-pH and cropping intensity increased while the diversity and abundance of the seedbank declined from west to east and north to south (Fig. 1).

Plant traits

The PCA of plant traits yielded similar results independent of whether individual plant values or accession mean values were considered. Considering only the variation among individuals, 33, 19 and 12% of the variation was explained by the first three components. The loadings demonstrate that

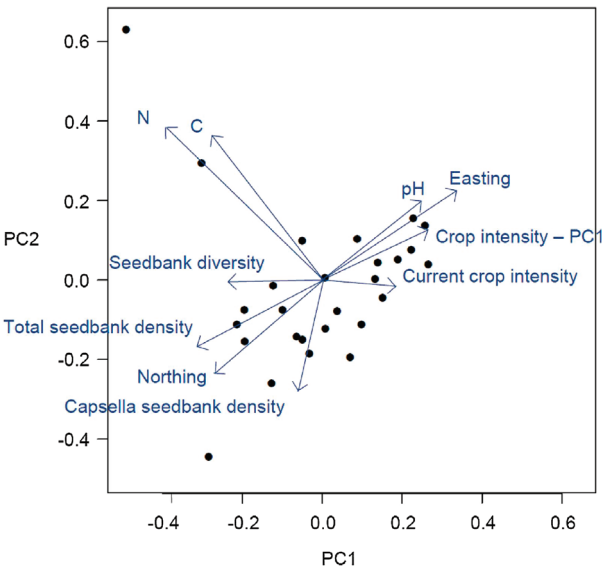

Figure 1. Biplot of first two components from the principal components analysis of field characteristics. Arrows indicate the loading of the field characteristics with respect to the components. Loading for soil texture classes have been suppressed to improve clarity.

the first 2 components represent the variation in traits associated with the reproductive and somatic status, respectively (Table 3), with PC1 being inversely related to reproductive duration and output as a consequence of a delayed time to flowering, and PC2 being negatively correlated with plant size. The loadings on the third component were greatest for the seed related variables of average seed size and germination rate with high PC3 scores relating to plants with small, slow germinating seeds.

Random effects models were fitted by restricted maximum likelihood to estimate the variance of the scores for the three principal components, i.e. PC1 to PC3 described above, among the field-sites, accessions and progeny. The variance estimates for the three PC scores establish that functional diversity, i.e. variation in reproductive and somatic

Table 3. Results of the principal components analysis of individual plant phenotypes based on the correlation matrix of plant traits showing the variance (standard deviation and the proportion of total variance) and loadings associated with the first three principal components.

	PC1 reproductive	PC2 somatic	PC3 seed
Standard deviation	2.08	1.58	1.27
Proportion of variance (%)	33.1	18.7	12.4
	loadings	loadings	loadings
Rosette diameter	<0.1	−0.434	<0.1
Number of rosette leaves	<0.1	−0.447	−0.115
Chlorophyll concentration	0.164	−0.447	0.195
Stem number	−0.324	−0.219	−0.242
Stem dry weight	0.272	−0.365	−0.267
Total seed weight	−0.428	<0.1	0.195
Total seed number	−0.429	<0.1	0.237
Average seed weight	0.195	<0.1	−0.586
Reproductive duration	−0.394	<0.1	<0.1
Flowering time	0.399	−0.231	0.126
Germination rate	<0.1	<0.1	−0.547
Dormant fraction	0.233	<0.1	<0.1

Table 4. Variation in plant function as represented by restricted maximum likelihood estimates of the standard deviation, with 95% confidence intervals in brackets, of the scores for the first three components of the principal components analysis.

Grouping	Reproductive	Somatic	Seed
Among field-site	1.27 (0.73, 2.22)	0.64 (0.26, 1.57)	0.39 (0.07, 2.18)
Among accession, within field-site	1.11 (0.65, 1.88)	0.92 (0.59, 1.45)	0.72 (0.42, 1.22)
Within accession	0.99 (0.80, 1.21)	0.83 (0.68, 1.01)	0.63 (0.52, 0.78)

status and seed characteristics, was present at each level and was broadly similar between these levels for each trait (Table 4). The subsequent inclusion of Easting and Northing as a two dimensional, quadratic spatial regression surface established the presence of a significant regional trend in reproductive status which declined to the south and east ($DF = 9$, $LR = 13.6035$, $p = 0.0183$). Consistent with this was a significant effect of soil-pH on reproductive status ($DF = 1$, $LR = 7.6347$, $p = 0.0057$). Neither somatic nor seed status showed a significant response to variation in the site characteristics. As a consequence of their potential importance to the life history strategy of annual plants, variation in flowering time and dormancy was compared against the field-site characteristics to establish if relationships existed but were masked by the composite nature of the somatic and seed status variables. The variation in flowering time was consistent with that already shown for reproductive status with a significant spatial trend ($DF = 3$, $LR = 8.3229$, $p = 0.0398$), though linear in this case, with an additional significant effect of pH ($DF = 1$, $LR = 8.9348$, $p = 0.0028$). Soil-pH was also significantly associated with variation in the dormant fraction ($DF = 1$, $LR = 5.9823$, $p = 0.0144$). Finally, the dormant fraction was positively associated with flowering time ($DF = 1$, $LR = 4.3860$, $p = 0.0362$).

Genetic structure and diversity

A total of 338 distinct ISSR products were identified from 109 individuals (51 different accessions), using data from the 6 different ISSR primers (Table 1). The total number of ISSR products per primer ranged from 34 (UBC 888) to 79 (UBC 886) with around 33% of these being present in any one individual (Table 1). Apparent polymorphism was high across the six primers used, ranging from 88 to 98% of ISSRs products per primer for the individual plants that were assessed (Table 1).

Table 5. Analysis of molecular variance (AMOVA) for in-field populations of shepherd's purse based on ISSR haplotypes obtained from all primers. P values for F-statistics obtained by permutation test based on 10 100 randomisations.

Effect	DF	Variance	Percent variance	F-statistic	p value
Among field-site	32	6.22	23.59	$F_{FT} = 0.2359$	< 0.0001
Among accession within field-site	18	5.07	19.22	$F_{MF} = 0.2513$	< 0.0001
Within accession	58	15.08	57.19		
Total	108	26.37			

Estimates of the hierarchical variance components and F-statistics (Table 5) are evidence of the structuring of shepherd's purse populations (Excoffier et al. 1992). Genetic subdivision is present at all levels with a distinction among fields ($F_{FT} = 0.2359$, $p < 0.0001$) and among accessions within fields ($F_{MF} = 0.2513$, $p < 0.0001$). However, the percentage variance in genotypes is not greatest at these levels but between individuals within accessions (Table 5). To test the contribution of random genotyping error to the estimate of within accession variation a further AMOVA was conducted independently on two subset of loci obtained using primers UBC 880 and UBC 886 (Table 6). In these analyses the presence of the products was determined for each extraction replicate rather than for the individual progeny plants after averaging the extraction replicate RFUs. This allowed the variation among replicates within progeny to be compared with that among progeny within accessions and also among accessions. The results obtained for the primer UBC 880 (Table 6) are consistent with the preceding results (Table 5) with a 60:40 split for the within versus among accession variation and show that for this subset of loci, 35% of the 60% of the variation present within accessions can be attributed to differences in the replicates. The results were somewhat different for the loci obtained by primer UBC 886 (Table 6). In this case, the within accession variation accounted for a greater proportion of the total genetic variation (75%) with 60% attributed to variation among replicates within progeny and 15% to variation among progeny within accessions (Table 6). Taken together these results suggest an approximately equal division of variation is present at the within individual, accession and field levels each accounting for around a fifth of the total, with the further two fifths resulting from genotyping error.

Despite the presence of random variation due to genotyping error the subsequent analyses confirm and extend the

Table 6. Estimates of hierarchical variance components obtained by analysis of molecular variance (AMOVA) for in-field populations of shepherd's purse based on ISSR haplotypes obtained from primers UBC 880 and UBC 886.

	UBC 880			UBC 886		
Effect	DF	Variance	Percent variance	DF	Variance	Percent variance
Among accessions	51	2.85	40.05	48	2.07	25.09
Among progeny within accession	57	1.76	24.69	54	1.20	14.53
Within progeny	128	2.51	35.26	137	4.98	60.38
Total	236	7.11		239	8.25	

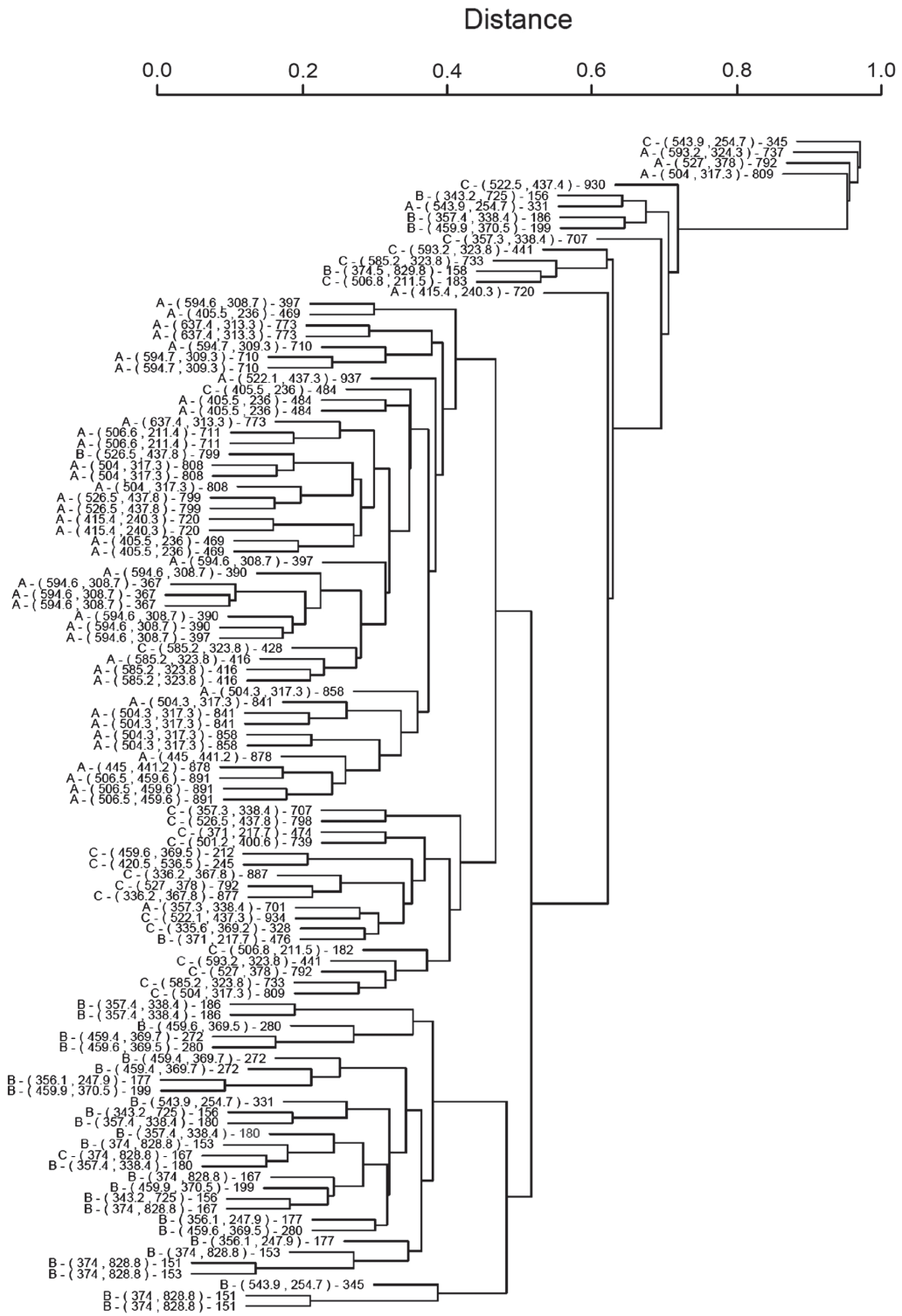

Figure 2. Dendrogram derived from the agglomerative hierarchical clustering of ISSR haplotypes using pairwise genetic distance (1-Jaccard's coefficient) and UGPMA. Information on the genetic and geographical origin of the individual plant is provided in the labels of the terminal nodes as follows: cluster – (Easting, Northing) – accession code.

patterns of population structure identified by the AMOVA. Agglomerative hierarchical clustering (UGPMA) also provided evidence of population structure (Fig. 2). Individuals, the distance between which typically ranged between 0.1 and 0.3, clustered first by accession and then by field-site. There was also an indication of clustering into 2–4 groups separated by a distance of 0.45–0.5 that corresponded to a regional, mostly east-west, segregation. In addition, there are a further 15, outlying individuals that show little similarity with other plants (Fig. 2).

The penalized log likelihood were highly consistent between replicate runs of the STRUCTURE algorithm

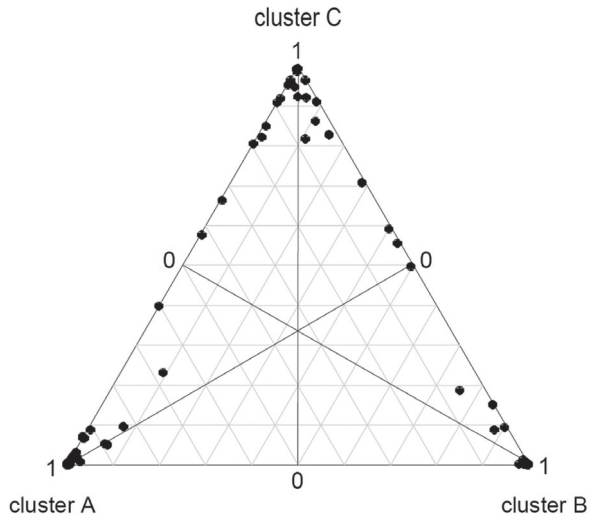

Figure 3. Summary of the clustering results showing the admixture proportions of individual plants (the proportion of an individual's genome from each of the 3 populations). The cluster admixture proportion for an individual is given by the distance from the side of the equilateral triangle as indicated by the axes scale from 0 to 1.

and approached a plateau at K = 3, although continuing to increase across all K. The assignment of individuals on the basis of admixture proportions estimated for the 3 cluster model was highly consistent with the results of agglomerative hierarchical clustering (Fig. 2). Furthermore the admixture proportions did not support the inclusion of a fourth cluster as no individuals had more than a minimal proportion of the genome (< 0.01) assigned to the fourth cluster. Selection of the simpler model can also be defended on the basis that shepherd's purse is predominantly self-fertilising and that multiple progeny were sampled from the majority of accessions obtained from sampling locations, both lead to departures from the model in a way that leads to an over estimate of K. The estimated admixture proportions, which themselves are thought to be relatively insensitive to these departures, are presented in Fig. 3 for a three cluster model. This shows that clusters A and B are distinct as admixed individuals combined genotype elements of a third cluster, C with cluster A or B, but not A with B.

Plants were assigned to the cluster for which the admixture proportion was greatest, in all cases this value exceeded 0.5 and in the majority of cases assignment was unambiguous (Fig. 3). The distribution of the admixture proportions for each of these clusters is an indication of their spatial segregation at the regional scale (Fig. 4). Cluster A is associated with a sub-population located predominantly to the south-east of mainland Britain. The distribution of the other clusters is less well defined but in line with the patterns of admixture (Fig. 3); individuals originating predominantly from cluster B lie to the north and west and do not coincide with cluster A, while a regional pattern in the assignment of individuals to cluster C was not evident and was overlapping with both cluster A and B.

The presence of regional patterns of segregation was confirmed by fitting a mixed-effects logistic regression model independently to each of the cluster assignments with a two-dimensional linear or quadratic spatial trend surface as a fixed effect and the inclusion of field-site and accession as nested random terms. Likelihood ratio test identified a significant linear trend in the assignment to cluster A, and quadratic trend in assignment to cluster B. There was no spatial trend, linear or quadratic, in the probability of assignment to the remaining sub-population, cluster C (Table 7).

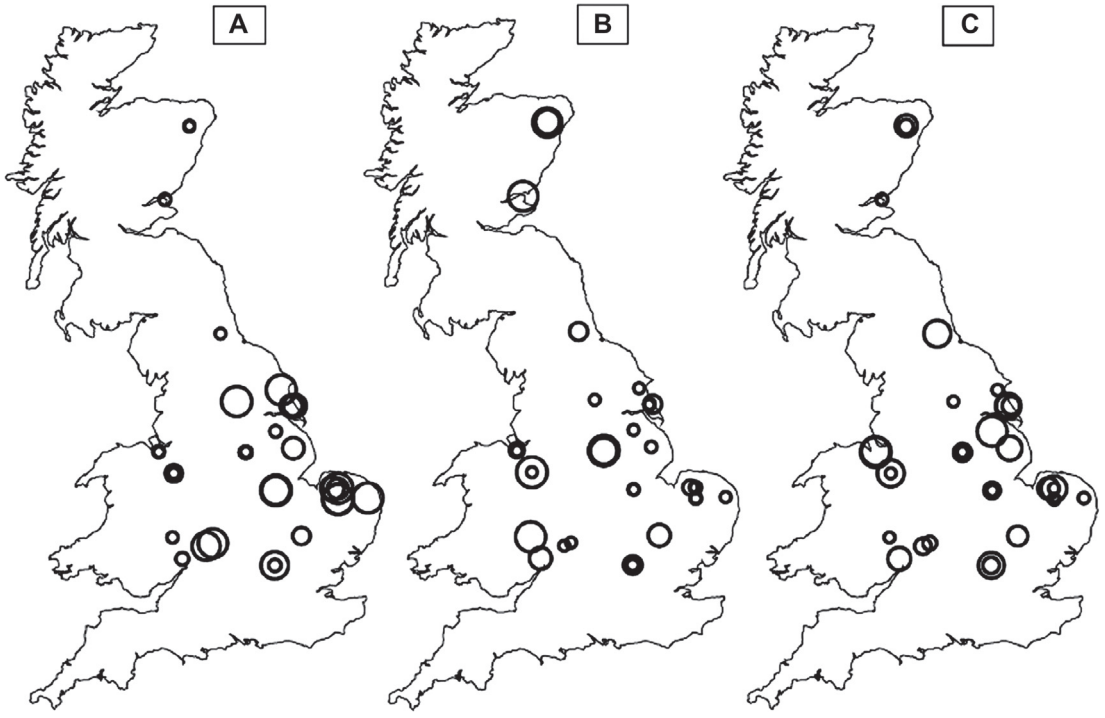

Figure 4. Spatial variation in the admixture proportions for cluster A, cluster B and cluster C. The location of each field-site is represented by a single circular symbol with circle diameter representing the mean admixture proportion for each site.

Table 7. Comparison between mixed-effects logistic regression models of cluster assignment. The significance of each field-site characteristic, including spatial trend surface[1] (EN), was tested by addition of terms to the null model, i.e. intercept only, and the spatial trend surface model if this was significant. Models were compared by change in AIC (ΔAIC) and likelihood ratio test. The direction of response is given for significant effects. The results for soil-C were omitted as they follow those for soil-N concentration with which it is highly correlated. Other field-site characteristics (see Material and methods: Site measurements for complete list) were excluded as no significant effects were observed across all tests. Significance codes are included for ease[2].

Model comparison	ΔAIC	LR	DF	p	Sig.[2]	Effect[1]
Cluster A						
null v EN	651.97	657.97	3	<0.0001	***	(+, −)
null v pH	647.70	649.70	1	<0.0001	***	+
null v nitrogen	646.01	648.01	1	<0.0001	***	−
null v intensity	644.96	640.96	1	<0.0001	***	−
EN v (EN + pH)	2.39	4.39	1	0.0340	*	+
EN v (EN + nitrogen)	1.54	3.54	1	0.0600	.	−
EN v (EN + intensity)	4.79	6.79	1	0.0092	**	−
Cluster B						
null v EN	21.18	37.18	8	<0.0001	***	(− , +)
null v pH	4.19	6.19	1	0.0129	*	−
null v nitrogen	−1.48	0.52	1	0.4706	NS	
null v intensity	4.53	6.52	1	0.0106	*	−
EN v (EN + pH)	−1.85	0.15	1	0.6978	NS	
EN v (EN + nitrogen)	−1.16	0.84	1	0.3582	NS	
EN v (EN + intensity)	−0.93	1.07	1	0.3015	NS	
Cluster C						
null v EN	−4.87	1.13	3	0.7702	NS	
null v pH	0.75	2.74	1	0.0979	.	+
null v nitrogen	1.81	3.80	1	0.0511		+
null v intensity	1.61	4.11	1	0.0427		+

[1]See text for description of spatial trend surface.
[2]Significance codes: *** <0.001> ** <0.01> * <0.05>. <0.1> NS.

The responses of the cluster assignments to the other field-site characteristics were tested by their addition to both null models and the trend surface models described above, where these were significant. With the spatial trend surface excluded, the probability with which plants were assigned to cluster A showed a highly significant response to soil-pH, soil-N, soil-C, and cropping intensity due to the strong regional variation in these characteristics (Table 7). Cropping intensity and soil-pH remained significant once the regional covariation had been accounted for but the response to soil-Nitrogen concentration was marginally non-significant (Table 7). The probability of assignment to cluster B appeared less dependent on the site characteristics measured, responding significantly only to soil-pH and cropping intensity and not even to these once the potential for regional covariation was accounted for (Table 7). The response of assignment to cluster C was inconclusive with results being either marginally significant (cropping intensity) or marginally non-significant (soil-pH, soil-N; Table 7).

The comparison of the phenotypic characteristics of the genetic clusters by multi-level ANOVA of the reproductive, somatic and seed status of individuals showed a significant difference in reproductive status among clusters based on the field-site averages ($F_{2,28} = 5.2759$, $p = 0.0114$). Further testing of the differences in plant traits among the clusters showed this to be consistent with a significant increase in reproductive duration in cluster A and a marginally non-significant reduction in flowering time (Table 8, significant when based on accession averages, $F_{2,16} = 4.0238$, $p = 0.0384$). Testing of the other plant traits identified differences between clusters in seed germination with, for example, a smaller fraction of dormant seeds in cluster A (Table 8).

Discussion

In this study a model approach is presented in which the parallel characterisation of environmental, molecular and ex situ phenotypic variation has been used to describe and understand the population structure of shepherd's purse, a representative arable weed. This approach confirmed that there are significant levels of intra-specific functional diversity in shepherd's purse (Hawes et al. 2005, Iannetta et al. 2007, 2010) and yielded evidence of structure in the in-field populations at multiple scales, with genetic and phenotypic differentiation evident between siblings resulting from self-fertilisation, accessions, field and regional populations. These patterns of differentiation were found to be partially related to the geophysical (location, soil-pH) and management (cropping intensity) properties of the fields.

Self-fertilisation in plants is associated with genetic differentiation and reduced genetic diversity, particularly within populations (Hamrick and Godt 1996, Glémin et al. 2006). Previous studies suggest that shepherd's purse conforms to this pattern with estimates of selfing rates based on isozyme data of 98% or more and genetic differentiation between populations, mostly non-arable, along transects of 2500 km (Neuffer 2010) and within a 2.5 km^2 area (Hameister et al. 2009).

Consequently, the high levels of self-fertilisation of shepherd's purse is likely to contribute to the genetic and functional differentiation observed in the present study among

Table 8. Estimates of the cluster specific means, with standard error of the mean in brackets, and ANOVA results for those phenotypic traits that differed significantly between clusters.

	Cluster					
Plant trait	A	B	C	DF	F	p
Flowering time (d)	74.70 (2.80)	88.03 (3.07)	84.87 (6.18)	2, 30	2.5967	0.0912
Reproductive duration (d)	82.35 (2.62)	71.52 (2.18)	73.18 (3.69)	2, 30	5.1838	0.0116
Dormant fraction (%)	25.21 (1.24)	33.27 (2.14)	31.49 (2.35)	2, 30	5.8268	0.0073
Germination rate (d)	18.41 (0.34)	20.09 (0.62)	17.25 (0.44)	2, 30	5.2590	0.0110

accessions within fields, and among field-sites. For example, the divergence among co-occurring but genetically isolated accessions may be the result of genetic drift as the accessions are exposed to common environmental conditions, so precluding divergent selection.

Genetic drift may also contribute to divergence between field populations, promoted by bottlenecks associated with founder events or extinction-colonisation dynamics at the field scale (Ceplitis et al. 2005, Glémin et al. 2006). However, the available evidence suggests differentiation at this scale resulted from rapid selection acting on existing variation in arable populations of shepherd's purse. For example, model based clustering, supported by the results of hierarchical clustering analysis, provided evidence for genetic differentiation at a scale larger than the field-site, and identified three clusters or sub-populations, two of which exhibited a degree of regional separation. The pattern of population differentiation is associated with varying environmental conditions (management intensity, soil-pH, and to a lesser extent soil-N) and trait divergence. One sub-population (cluster A) has, on average, an earlier flowering time and increased reproductive duration and output, with a higher proportion of seeds germinating. These results are supported by the findings of Toorop et al. (2008) who demonstrated that earlier time to flowering individuals have increased germination efficiency and a reduced capacity for secondary-dormancy. The results are not consistent with the notion of shepherd's purse as an 'ideal weed' that combines a general genotype with phenotypic plasticity to achieve successful colonisation (Baker 1974) but favour the role of selection in generating locally adapted and differentiated genotypes (Clements et al. 2004).

Disturbance by soil cultivation and herbicide treatment is a major feature of arable cropping and has an obvious impact on the mortality and fitness of weeds. Neuffer and Meyer-Walf (1996) observed an earlier flowering time in an arable population of shepherd's purse when compared to a non-arable population and it has been proposed that the increased frequency of disturbance under intensive cropping selects for early flowering variants (Iannetta et al. 2007). The present study found that early flowering was more common in the south-east of Britain, a region of more intense farming practices. However, the probability of an individual being assigned to cluster A, the sub-population that predominates in the south-east and which flowered on average 10 d earlier than any other sub-population, was actually lower in fields of high intensity cropping. The combination of traits expressed by these plants (i.e. early flowering, high fecundity, reduced dormancy) are generally associated with 'weediness' (Baker 1974) and are consistent with an r-selected life-history strategy that is well adapted to frequent disturbance. In this case, early flowering can be viewed as a conservative bet-hedging strategy employed to minimize the risk that individuals fail to reproduce before disturbance (Childs et al. 2010). However, dormancy also plays a pivotal role as a bet-hedging strategy in weeds, allowing populations to spread the risk of future disturbance (Childs et al. 2010). The balance of dormancy and early flowering between the cluster sub-populations suggests alternative life-history strategies are employed by populations depending on the degree of cropping intensity experienced. It appears that flowering time can be reduced sufficiently in shepherd's purse to reliably avoid disturbance at most levels of crop intensity. However, at the frequency of cultivation and pesticide use associated with the highest levels of crop intensity, the reduced flowering time strategy appears insufficient to minimise the risk of disturbance, thereby demanding greater recourse to the diversified bet-hedging strategy of elevated dormancy. The negative association between the capacity of a plant for early flowering and the extent of dormancy in their progeny indicates that there is a trade-off which precludes a single, universally successful, early flowering-high dormancy strategy, leading instead to selective divergence in the life-history strategies of shepherd's purse.

Shepherd's purse appears to respond to soil-pH, plants exhibit earlier flowering and increased dormancy at sites with more alkaline soil which is also associated with increased probability of finding cluster A genotypes. These results point toward soil-pH as a driver of population structure. The influence of pH on germination and dormancy (Footitt and Cohn 1992, Cosgrove 1997) provides a plausible route via which pH may act to select life-history strategies in this way though the existence of such mechanisms in shepherd's purse is yet to be established.

The population structure of arable weeds and the underpinning evolutionary processes have seldom been studied. However, there is evidence that weed populations can exhibit an adaptive response to selection pressure. The evolution of herbicide resistance illustrates an adaptive response to crop management (Neve 2007). Crop management and in particular, any form of weed control, can also lead to the rapid selection of life-history characteristics as in the case of selection for early flowering time in wild oats (Imam and Allard 1965) and *Arabidopsis thaliana* (thalecress; Jones 1971). As a further example, *Brassica rapa*, a ruderal and arable Brassicaceous species, has shown very rapid local adaptation in response to environmental conditions, reducing their time to flowering as a drought avoidance strategy (Franks and Sim 2007).

The patterns and processes of weed species assemblage are well established and suggest that environment and management influence the fitness of arable weeds. For example, Debeljak et al. (2008) found that the abundance in the UK of a newly introduced seedbank species, volunteer oilseed rape, and functionally similar weeds such as shepherd's purse and mayweed (*Matricaria* spp.), is highest in the north and west and lowest in the south east of England (Debeljak et al. 2008). Furthermore, the associations between cluster assignment and crop intensity and soil-pH observed in the present study are consistent with species-shifts of in-field weed communities which are most significantly affected by crop type or cover and soil-pH (Fried et al. 2008, Pinke et al. 2010).

The adaptive capacity of in-field populations of shepherd's purse is consistent with the global distribution of this species which encompasses very diverse niches and climates. Shepherd's purse also exhibits extensive variation in life-history, morphological and developmental traits, some of which appear to be in response to environmental conditions (Neuffer and Bartelheim 1989, Neuffer 2010) indicating adaptation to local environments. A frequently studied example of this is time to flowering which varies along latitudinal and altitudinal gradients (Neuffer and Bartelheim 1989, Neuffer 2010) and, associated with which, are differences in gene expression consistent with adaptation to photoperiod (Slotte et al. 2007).

In a review of the consequences of genetic diversity for ecological functioning, Hughes et al. (2008) established the influence of genetic diversity on ecological processes at population, community and ecosystem scales. For ecological processes to be affected in this way requires the genetic diversity to be associated with variation in phenotypic traits. In the case of shepherd's purse, the extensive genetic diversity within the arable populations is associated with phenotypic variation and, though based on neutral markers (ISSRs), the presence of character-displacement among genetically defined sub-populations establishes the correlation between genotype and phenotype and establishes the potential for the genetic and associated phenotypic diversity in shepherd's purse to influence the functioning and management of arable ecosystems.

Knowledge of the phenotypic variation allows us to speculate on the ecological significance of the genetic differentiation and diversity in arable populations of shepherd's purse. The phenology of weeds is key in determining their response to the frequency and timing of control measures. As a corollary to the selective pressure of cropping intensity, we can predict that the variation in time to reproduction (a function of flowering time and seed dormancy) will result in differential susceptibility to control measures across shepherd's purse populations and thereby influence population regulation and subsequent adaptation. It is with ecological (Mortensen et al. 2000) and evolutionary (Neve et al. 2009) understanding of the type gathered here that these effects might be predicted and control strategies devised that are effective and durable (Jordan and Jannink 1997). For example, intraspecific variation in weeds may also be relevant to crop protection. Tobacco rattle virus (TRV), an important pathogen of potato crops, infects weed species that are thought to be an important reservoirs of infection. Iannetta et al. (2010) have shown that susceptibility to TRV infection varies among some of the same accessions of shepherd's purse investigated here. The risks posed to crops by shepherd's purse may therefore vary as a consequence of the population structure described here. Such interactions also indicate that phenotypic variation in shepherd's purse has potential consequences for the arable food-web. The effect of ecotype exclusion from the arable food-web was addressed by Karley et al. (2008), who showed that aphid reproductive output decreased by 25% on shepherd's purse variants with high tissue water content, and low C:N ratio (*Aphis fabae*), and on variants with low phloem nitrogen (*Myzus persicae*). Furthermore, early flowering time variants produce more seed and have a higher density of leaf hairs (Karley et al. 2008) and so are likely to benefit seed feeders but not leaf-chewing insect herbivores. As weeds are disproportionately beneficial to the in-field food web compared to crop plants (Hawes et al. 2003) and shepherd's purse is an abundant weed in the UK arable system, a significant impact of shepherd's purse diversity is predicted as it propagates through the arable food web via trophic interactions such as those described above.

Conclusion

The patterns of genetic and phenotypic variation observed in this study suggest that adaptation as a result of selection by environmental conditions and cropping practise has played a role in the capacity of shepherd's purse to colonise and persist in an arable setting and has lead to the differentiation of shepherd's purse into a number of sub-populations. The evidence of past adaptation plus the existing extensive genetic and phenotypic diversity of the in-field populations also suggests the potential for further adaptation. The arable system will continue to face substantial pressure as a result of changing cropping practice, management, and change. Knowledge of the structure and diversity of weed populations and of the evolutionary processes at play are essential if we are to learn to manage the adaptive capacity of the weed flora so as to satisfy different ecosystem services such as the provision of yield and the support of the essential arable food web.

Acknowledgements – This work was funded by the Scottish Government. We thank Gavin Ramsay (JHI) and Peter Toorop (RBG Kew, Millennium Seed Bank), for their guidance on ISSR profiling and seed physiology (respectively). For his guidance in obtaining and calculating pesticides application rates we thank David Garthwaite (FERA, the Food and Environment Research Agency, formerly Central Science Laboratory).

References

Altieri, M. 1999. The ecological role of biodiversity in agroecosystems. – Agric. Ecosyst. Environ. 74: 19–31.

Baker, H. G. 1974. The evolution of weeds. – Annu. Rev. Ecol. Syst. 5: 1–24.

Bonin, A. et al. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. – Mol. Ecol. 16: 3737–3758.

Bornet, B. and Branchard, M. 2001. Nonanchored inter simple sequence repeats (ISSR) markers: reproducible and specific tools for genome fingerprinting. – Plant Mol. Biol. Rep. 19: 209–215.

Bown, J. L. et al. 2007. Consequences of intraspecific variation for the structure and function of ecological communities. Part 1: model development and predicted patterns of diversity. – Ecol. Model. 207: 264–276.

Ceplitis, A. et al. 2005. Bayesian inference of evolutionary history from chloroplast microsatellites in the cosmopolitan weed *Capsella bursa-pastoris* (Brassicaceae). – Mol. Ecol. 14: 4221–4233.

Charters, Y. M. et al. 1996. PCR analysis of oilseed rape cultivars *(Brassica napus* L. ssp. *oleifera)* using 5'-anchored simple sequence repeat (SSR) primers. – Theor. Appl. Genet. 92: 442–447.

Childs, D. Z. et al. 2010. Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. – Proc. R. Soc. B 277: 3055–3064.

Clements, D. R. et al. 2004. Adaptability of plants invading North American cropland. – Agric. Ecosyst. Environ. 104: 379–398.

Cosgrove, D. J. 1997. Relaxation in a high-stress environment: the molecular bases of extensible cell walls and cell enlargement. – Plant Cell 9: 1031–1041.

Debeljak, M. et al. 2008. Relation between the oilseed rape volunteer seedbank, and soil factors, weed functional groups and geographical location in the UK. – Ecol. Model. 212: 138–146.

Excoffier, L. et al. 1992. Analysis of molecular variance inferred for metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. – Genetics 131: 479–491.

Footitt, S. and Cohn, M. A. 1992. Seed dormancy in red rice VIII. Embryo acidification during dormancy-breaking and subsequent germination. – Plant Physiol. 100: 1196–1202.
Franks, S. J. and Sim, S. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. – Proc. Natl Acad. Sci. USA 104: 1278–1282.
Fridley, J. D. and Grime, J. P. 2007. Genetic identity of intraspecific neighbours mediates plant responses to competition and environmental variation in a species-rich grassland. – J. Ecol. 95: 908–915.
Fried, G. et al. 2008. Environmental and management factors determining weed species composition and diversity in France. – Agric. Ecosyst. Environ. 128: 68–76.
Garthwaite, D. G. and Thomas, M. R. 2006. Pesticide usage survey report 213: arable crops in Great Britain (including aerial applications 2003–2005). – Pesticide Usage Survey Team, FERA (previously Central Science Laboratory), Sand Hutton, York, UK.
Glémin, S. et al. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. – Proc. R. Soc. B 273: 3011–3019.
Hameister, S. et al. 2009. Genetic differentiation and reproductive isolation of a naturally occurring floral homeotic mutant within a wild-type population of *Capsella bursa-pastoris* (Brassicaceae). – Mol. Ecol. 18: 2659–2667.
Hamrick, J. L. and Godt, M. J. W. 1996. Effects of life history traits on genetic diversity in plant species. – Phil. Trans. R. Soc. B 351: 1291–1298.
Hawes, C. et al. 2003. Responses of plant and invertebrate trophic groups to contrasting herbicide regimes in the Farm Scale Evaluations of genetically-modified herbicide-tolerant crops. – Phil. Trans. R. Soc. B 358: 1899–1913.
Hawes, C. et al. 2005. Individuals as the basic accounting unit in studies of ecosystem function: functional diversity in *Capsella* (shepherd's purse). – Oikos 109: 521–534.
Hawes, C. et al. 2009. Functional approaches for assessing plant and invertebrate abundance patterns in arable systems. – Basic Appl. Ecol. 10: 34–42.
Hodgson, J. M. 1974. Soil survey field handbook. – Soil Survey of England and Wales, Tech. Monogr. no. 5, Harpenden Press, p. 99.
Holt, J. S. and LeBaron, H. M. 1990. Significance and distribution of herbicide resistance. – Weed Technol. 4: 141–149.
Hubisz, M. J. et al. 2009. Inferring weak population structure with the assistance of sample group information. – Mol. Ecol. Res. 9: 1322–1332.
Hughes, A. R. et al. 2008. Ecological consequences of genetic diversity. – Ecol. Lett. 11: 609–623.
Hurka, H. and Düring, S. 1994. Genetic control of plastidic l-glutamate dehydrogenase isozymes in the genus *Capsella* (Brassicaceae). – Heredity. 72: 126–131.
Iannetta, P. P. M. 2010. *Capsella*. – In: Kole, C. (ed.), Wild crop relatives: genomic and breeding resources – oilseeds. Springer, pp. 37–62.
Iannetta, P. P. M. et al. 2007. Variation in *Capsella* (shepherd's purse): an example of intraspecific functional diversity. – Physiol. Plant. 129: 542–554.
Iannetta, P. P. M. et al. 2010. Sustainable disease control using weeds as indicators: *Capsella bursa-pastoris* and tobacco rattle virus. – Weed Res. 50: 511–514.
Imam, A. G. and Allard, R. W. 1965. Population studies of predominantly self-pollinating species. VI. Genetic variability between and within natural populations of wild oats from differing habitats in California. – Genetics 53: 633–659.
Jones, M. E. 1971. The population genetics of *Arabidopsis thaliana.* II. Population structure. – Heredity 27: 51–58.
Jordan, N. R. and Jannink, J. L. 1997. Assessing the practical importance of weed evolution: a research agenda. – Weed Res. 37: 237–246.
Karley, A. J. et al. 2008. Intraspecific variation in *Capsella bursa-pastoris* in plant quality traits for insect herbivores. – Weed Res. 48: 147–156.
Marshall, E. J. P. et al. 2003. The role of weeds in supporting biological diversity within crop fields. – Weed Res. 43: 77–89.
Mogg, R. J. and Bond, J. M. 2003. A cheap, reliable and rapid method of extracting high-quality DNA from plants. – Mol. Ecol. Note 3: 666–668.
Mortensen, D. et al. 2000. The role of ecology in the development of weed management systems: an outlook. – Weed Res. 40: 49–62.
Munn, J.-H. et al. 2006. Distribution of microsatellites in the genome of *Medicago trunculatat*: a resource of genetic markers that integrate genetic and physical maps. – Genetics 172: 2541–2555.
Neuffer, B. 2010. Native range variation in *Capsella bursa-pastoris* (Brassicaceae) along a 2500 km latitudinal transect. – Flora doi: 10.1016/j.flora.2010.03.001
Neuffer, B. and Bartelheim, S. 1989. Genetic-ecology of *Capsella bursa-pastoris* from an altitudinal transect in the Alps. – Oecologia 81: 521–527.
Neuffer, B. and Meyer-Walf, M. 1996. Ecotypic variation in relation to man made habitats in *Capsella*: field and trampling area. – Flora 191: 49–57.
Neve, P. 2007. Challenges for herbicide resistance evolution and management: 50 years after Harper. – Weed Res. 47: 365–369.
Neve, P. et al. 2009. Evolutionary-thinking in agricultural weed management. – New Phytol. 184: 783–793.
Pachepsky, E. et al. 2001. Consequences of intraspecific variation for the structure and function of ecological communities. Part 2: linking diversity and function. – Ecol. Model. 207: 277–285.
Pinke, G. et al. 2010. Effects of environmental factors on weed species composition of cereal and stubble fields in western Hungary. – Cent. Eur. J. Biol. 5: 283–292.
Pritchard, J. K. et al. 2000. Inference of population structure using multilocus genotype data. – Genetics 155: 945–959.
Robinson, R. A. and Sutherland, W. J. 2002. Post-war changes in arable farming and biodiversity in Great Britain. – J. Appl. Ecol. 39: 157–176.
Rothstein, S. J. 2007. Returning to our roots: making plant biology research relevant to future challenges in agriculture. – Plant Cell 19: 2695–2699.
Sica, M. et al. 2005. ISSR markers show differentiation among Italian populations of *Asparagus acutifolius* L. – BMC Genet. 6: 17.
Silvertown, J. 2004. Plant coexistence and the niche. – Trends Ecol. Evol. 19: 605–611.
Slotte, T. et al. 2007. Differenial expression of genes important for adaptation in *Capsella bursa pastoris* (Brassicaceae). – Plant Physiol. 145: 160–173.
Squire, G. R. et al. 2003. On the rationale and interpretation of the farm-scale evaluations of genetically-modified herbicide-tolerant crops. – Phil. Trans. R. Soc. B 358: 1779–1800.
Toorop, P. E. et al. 2008. Germination characters and myxospermy of shepherds purse (*Capsella bursa-pastoris* (L.) Medic) time to flowering variants. – Polish J. Nat. Sci. 5: 101.
Wang, T. et al. 2008. Population genetic variation and structure of the invasive weed *Mikania micrantha* in southern China: consequences of rapid range expansion. – J. Hered. 99: 22–33.
Whitlock, R. et al. 2007. The role of genotypic diversity in determining grassland community structure under constant environmental conditions. – J. Ecol. 95: 985–907.
Wolfe, A. D. and Liston, A. 1998. Contributions of PCR-based methods to plant systematics and evolutionary biology. – In: Soltis, D. E. et al. (eds), Plant systematics II. Kluwer, pp. 43–86.