Molecular Ecology (2004) **13**, 2959–2967 doi: 10.1111/j.1365-294X.2004.02314.x

Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species

FABIENNE VAN ROSSUM,* ISABELLE BONNIN,† STÉPHANE FÉNART, MAXIME PAUWELS, DANIEL PETIT and PIERRE SAUMITOU-LAPRADE

Laboratoire de Génétique et Evolution des Populations Végétales, UMR CNRS 8016, FR CNRS 1818, Université de Lille 1, Bâtiment SN2, F-59655 Villeneuve d'Ascq Cedex, France

Abstract

***Arabidopsis halleri*, a close wild relative of *A. thaliana*, is a clonal, insect-pollinated herb tolerant to heavy metals (Zn, Pd, Cd) and a hyperaccumulator of Zn and Cd. It is of particular interest in the study of evolutionary processes and phytoremediation. However, little is known about its population gene flow patterns and the structure of its genetic diversity. We used five microsatellite loci to investigate the genetic structure at a fine spatial scale (10 cm to 500 m) in a metallicolous population of *A. halleri*. We also studied the contributions made by clonal propagation and sexual reproduction (seed and pollen dispersal) to the genetic patterns. Clonal diversity was high (D_G > 0.9). Clonal spread occurs only at short distances (< 1 m). Both clonal spread and limited dispersal, associated with sexual reproduction, contribute to the significant spatial genetic structure revealed by spatial autocorrelation analysis. The shape of the autocorrelogram suggests that seed dispersal is restricted and pollen flow extensive, which may be related to intense activity by insect pollinators. Clonal spread was more extensive in the lowly polluted zone than in the highly polluted zone. This cannot be interpreted as a strategy for promoting the propagation of adapted genotypes under the harshest ecological constraints (highest heavy metal concentrations). The higher fine-scale spatial genetic structure found in the lowly polluted zone can be ascribed to plant densities that were lower than in the highly polluted zone. No evidence of genetic divergence due to spatial heavy metal heterogeneity was found between lowly and highly polluted zones.**

Keywords: *Arabidopsis halleri*, clonal spread, heavy metal tolerance, pollen dispersal, seed dispersal, spatial autocorrelation

Received 9 March 2004; revision received 7 June 2004; accepted 6 July 2004

Introduction

Contamination of soils by heavy metals is an important environmental and health problem because of the high toxicity of heavy metals (e.g. Salt *et al*. 1998). Some plant species can grow and reproduce on these soils; they are defined as heavy metal tolerant (Macnair 1987). These plants are of great interest for phytoremediation, a technique that uses heavy-metal-tolerant plants to extract metals from polluted soils or to limit the erosion of toxic elements (Salt *et al*. 1998). Because heavy metals represent highly constraining selective pressures, heavy metal tolerance also provides an excellent model system to study recent adaptive evolutionary processes (e.g. Macnair 1987). Metallicolous populations occurring in areas polluted by heavy metals due to industrial human activities are of particular interest; the sites often constitute contaminated islands in noncontaminated areas, and the amount of heavy metals in the soil is very high, which means there are very strong ecological constraints (e.g. Macnair 1987).

Correspondence: Pierre Saumitou-Laprade. Fax: +33 320 436 979; E-mail: pierre.saumitou-laprade@univ-lille1.fr

*Present address: Plant Science and Nature Management, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium.

†Present address: UMR de Génétique Végétale (CNRS-INRA-UPS-INAPG), Ferme du Moulon, F-91190 Gif sur Yvette, France.

Molecular tools and genomics offer new perspectives to study evolutionary processes, including the function and evolution of genes. But because of the considerable investments they require, such investigations have been carried out on only a few model species (Mitchell-Olds 2001). In this context, *Arabidopsis halleri* (Brassicaceae) is a particularly interesting species; this clonal perennial herb is a close wild relative of the model species *Arabidopsis thaliana*, for which a vast array of genomic tools and molecular markers is available (e.g. Mitchell-Olds 2001; Clauss *et al*. 2002). Unlike *A. thaliana*, *A. halleri* is tolerant to high soil concentrations of Zn, Cd and Pb, and hyperaccumulates Zn and Cd. It has been shown that tolerance to Zn is constitutive of the species, but variance for the tolerance has been observed among populations and individuals; metallicolous populations demonstrate increased tolerance to Zn compared with nonmetallicolous populations (Bert *et al*. 2000; Fournier *et al*. unpublished). It is, therefore, a good candidate species for studying adaptive processes, but also for phytoextraction (Bert *et al*. 2002). However, little is known about gene flow patterns and the structure of genetic diversity in *A. halleri* (Clauss *et al*. 2002). To our knowledge, very few studies have explored gene flow patterns within metallicolous populations (e.g. McNeilly 1967; Antonovics & Bradshaw 1970; Ducousso *et al*. 1990). Also, genetic structure at a fine spatial scale has not been investigated for metallicolous populations of insect-pollinated plant species.

Spatial genetic structure within plant populations is primarily determined by the balance and interplay between local genetic drift and gene flow through sexual reproduction, i.e. through seed and pollen dispersal (e.g. Wright 1943; Heywood 1991). The magnitude of structuring depends on the relative contribution of pollen and seed dispersal to gene flow, the level of self-fertilization and the availability of compatible mates, plant density, and pollinator abundance and foraging behaviour (e.g. Williams 1994; Richards 1997; Vekemans & Hardy 2004). For instance, a decrease in pollinator numbers may contribute to the disruption of pollination processes, and therefore reduce intrapopulational gene flow but increase local genetic drift (e.g. Kwak *et al*. 1998). If habitat deterioration due to human activities has been reported to negatively affect pollinators (e.g. Kwak *et al*. 1998; Kevan 1999), little is known about the effects of heavy metal contamination of soils on pollinator abundance (Westerbergh & Saura 1994; Leita *et al*. 1996). Clonal propagation can also contribute to the patterns of genetic structure. A nonrandom spatial arrangement of clones can substantially increase geitonogamy and/or local consanguineous matings, local genetic drift and isolation-by-distance (Heywood 1991; Charpentier 2002). Spatial genetic structure may also be influenced by the spatial variation of selection through environmental heterogeneity (Linhart & Grant 1996). Several studies on plant populations reveal a fine-scale, intrapopulation genetic structure based on neutral genetic markers that is associated with habitat variation, e.g. with edaphic factors for the wind-pollinated *Festuca ovina* (e.g. Prentice *et al*. 2000), with plant community composition for the outcrossing *Gypsophila fastigiata* (Lönn *et al*. 1996) and with aridity for *Triticum dicoccoides* (Li *et al*. 2000). Genetic adaptation to heterogeneous environments may result in the formation of distinct ecotypes, which may or may not continue to exchange effective migrants.

In this study, we use microsatellite markers to investigate the spatial genetic structure at a microgeographical scale (ranging from 10 cm to 500 m) in a metallicolous population of *A. halleri*. This large, continuous population is located in an anthropogenic calaminary site surrounded by industrial and urban areas. The soil shows a gradient of heavy metal concentrations (mainly Zn, Pb, Cd). We address the following questions:

1 How is genetic diversity structured within the whole population, and at a very fine spatial scale?
2 What are the relative contributions of sexual reproduction and clonal propagation to the genetic pattern?

These results are analysed and discussed in light of spatial environmental heterogeneity of the site, i.e. by comparing zones that vary in the amounts of heavy metals in the soil.

Materials and methods

The species

Arabidopsis halleri (Brassicaceae) is a diploid, self-incompatible, insect-pollinated, perennial, rosette-forming herb, which propagates by stolons. Plants produce large numbers of seeds. In Europe, it is distributed among continental and mountainous areas, occurring on acidic, fresh and oligotrophic soils, but also on soils with high heavy metals content (Clapham & Akeroyd 1993; Bert *et al*. 2002). In France, it is non-native and only found in heavy metal-contaminated sites due to human activities (e.g. Berton 1946).

Study site and sampling

'Bois des Asturies' in the town of Auby in northern France, comprises a very large (several thousands of individuals), continuous population of *A. halleri*, described since 1944 and probably introduced in 1920–25 (Berton 1946). This site was downwind of a zinc smelter and also served as a waste deposit site for the metallurgical factory Umicor. A gradient of heavy metal concentrations (mainly Zn, Pb and Cd) can be found in the soil, increasing from a lowly polluted (LP) zone (extractable [Zn] < 3000 μg/g, [Pb] < 1000 μg/g, [Cd] < 100 μg/g) to a highly polluted (HP)

zone (extractable [Zn] > 30 000 μg/g, [Pb] > 5000 μg/g, [Cd] > 300 μg/g) (Van Rossum *et al*. unpublished).

Leaf material was sampled from 612 individual rosettes (= ramets) of *A. halleri*, following two patterns: (i) a minimal distance of every 3 m along a transect covering the population area (232 ramets); and (ii) near the transect (0.5–3 m), using a grid, every 10 cm within four plots of a 0.5 × 3 m area (n = 59–133 for a total of 380 ramets). Two plots (LP1 and LP2) were located in the LP zone (extractable [Zn] = 2792 and 5091 μg/g, respectively). They were 8 m apart. The other two (HP1 and HP2) were located 47 m apart in the HP zone (extractable [Zn] = 29 706 and 32 792 μg/g, respectively). Plots between zones were 295–318 m apart. Because we could detect the presence or absence of rosettes with the quadrat size (10 × 10 cm) of the grid, the number of ramets (= n) and genets (distinct multilocus genotypes, = G) can be considered as an estimate of plot ramet and genet density, respectively.

Microsatellite analysis

Leaves were dried at 55 °C for 24 h. DNA extractions were performed on dried leaf material (10–15 mg dry weight) using the extraction kit Dneasy® from Qiagen®. Polymorphism was assayed on each DNA sample at five microsatellite loci transferred from *A. thaliana* using the resources of the *Arabidopsis* community. Three loci were kindly provided by Thomas Mitchell-Olds (*LYR132*, *LYR133*, *LYR417*); *GC16* (= ATTS0392) was described in Clauss *et al*. (2002), and *ATH* (= ATHCTR1A) in Bell & Ecker (1994). In order for all loci polymerase chain reaction (PCR) products to be labelled with a single labelled primer, each forward primer contained a 5′-tail of 19 bp homologous to the universal consensus M13 forward primer sequence, followed by the specific locus sequence (Oetting *et al*. 1995).

PCR were carried out in a total volume of 15 μL containing 10 ng of template DNA, 3.5 mM $MgCl_2$, 200 μg/mL of BSA (for *LYR132*, *LYR133* and *GC16*), 100 μM dNTP, 0.2 μM each primer (the M13 forward primer was fluorescence-labelled with either IRD-700 or IRD-800), 20 mM Tris–HCl (pH 8.3), 50 mM KCl and 0.4 units of AmpliTaq DNA polymerase (Perkin–Elmer). PCR was carried out on a Perkin–Elmer Gene-Amp system 9700 (94 °C for 5 min, followed by locus-specific amplification: 94 °C for 30 s, annealing temperature for 35 s, 72 °C for 30 s, for eight cycles, followed by M13 labelling amplification: 94 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s, for 32 cycles, and a final extension 72 °C for 2 min). The annealing temperature of the locus-specific cycles varied between 50 and 60 °C, depending on the locus.

Amplification products were analysed on a Li-Cor automated DNA sequencer 4200 (Li-Cor-ScienceTec). PCR fragments were separated on 7% denaturing polyacrylamide gels (Long ranger, FMC) of 25 cm (for *LYR417*) and 33 cm, sizes were assessed with Li-Cor's BASE IMAGIR software v. 4.1 by comparison with an appropriate labelled molecular mass marker (Li-Cor-ScienceTec).

Data analysis

Extent of clonality. In order to assess the independence of the loci, a test for genotypic disequilibrium between pairs of loci and sequential Bonferroni-type correction (Rice 1989) was performed using FSTAT (Goudet 2001) for the transect (all sampled individuals are distinct genets, see Results).

To estimate the extent of clonality, we tested to see whether clustered individuals of the same multilocus genotype could be clones generated by asexual propagation or whether they could be identical genotypes (for the markers used) produced via sexual reproduction. The probability of obtaining a second encounter of the same multilocus genotype in the sampled individuals (p_{se}), under the assumption of linkage equilibrium, was calculated for the whole population (= transect + plots) according to Parks & Werth (1993). It may be assumed with over 95% confidence that the multiple occurrence of a multilocus genotype with $p_{se} < 0.05$ is likely to be generated by clonal propagation of a single genet. If $p_{se} > 0.05$, ramets are likely to be accounted for by sexual reproduction (Parks & Werth 1993). This analysis was performed for the 574 ramets showing no missing data.

Genetic variation within population. In order to analyse genets (sexually reproduced genotypes) only, distinct multilocus genotypes were identified for each plot (no clone for the transect), and one ramet per genotype was randomly chosen. The following measures of genetic variation were calculated for each locus using GEN-SURVEY (Vekemans & Lefèbvre 1997): mean number of alleles (A), observed heterozygosity (H_O) and expected heterozygosity (H_E) corrected for small sample size (Nei 1978) for the entire (transect + plots) 'genet' population (491 genets). The average measures were also computed for all loci. Wright's inbreeding coefficient (F_{IS}) corrected for small sample size (Kirby 1975) was estimated for the entire 'ramet' and 'genet' populations (612 ramets and 491 genets, respectively). The mean F_{IS} over all loci was also determined for the four plots with ramets and with genets, and the standard errors estimated using a jackknife procedure over the loci (Sokal & Rohlf 2000). The significance of F_{IS} for each locus, and over all loci for the entire population and for each plot was tested by randomization tests using FSTAT.

Spatial genetic structure within population: contributions of clonal growth and sexual reproduction. We assessed spatial genetic structure at a microgeographical scale using spatial autocorrelation analyses. These were performed with kinship coefficients (Loiselle *et al*. 1995) using SPAGEDI (Hardy & Vekemans 2002). In order to test for isolation-by-distance,

the multilocus kinship coefficient for each pair of individuals was plotted against the logarithm of the geographical distance separating them. The slope (b) of this linear regression provides a good estimator of the extent of gene dispersal at this scale, under the assumption that population genetic structure has reached equilibrium (Vekemans & Hardy 2004). For the graphical representation of kinship, average multilocus kinship coefficients per distance intervals (F_j) were computed for the following distance classes (upper bound distance in metres): $j = 22$ distance classes for the entire (transect + plots) 'ramet' and 'genet' population (0.2, 0.3, 0.4, 0.5, 0.75, 1, 2, 4, 8, 12, 16, 25, 50, 75, 100, 125, 150, 200, 250, 300, 350, 600), $j = 13$ for the four plots with ramets (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2, 3), and $j = 8$ for the four plots with genets (0.2, 0.4, 0.6, 0.8, 1, 1.5, 2, 3). Standard errors for the multilocus estimates of the kinship coefficients per distance class and the slope were estimated using a jackknife procedure over the loci. We tested the significance of the kinship coefficients and of the b estimates by comparing the observed values with those obtained after 2000 random permutations of the individuals among positions. We also quantified spatial genetic structure using the Sp statistics, calculated as $-b/(1 - F_{(1)})$ where $F_{(1)}$ is the mean F_j for the first distance interval. $F_{(1)}$ can be considered as an approximation of the kinship between pairs of neighbours, provided the first distance class contains enough pairs of individuals to get a reasonably precise $F_{(1)}$ value (Vekemans & Hardy 2004). $1/Sp$ can be an estimate of the neighbourhood size, $Nb = -(1 - F_{IS})/b$ (Vekemans & Hardy 2004).

In case of a significant structuring effect of clonal growth, a higher spatial genetic structure can be expected for the ramets than the genets (e.g. Heywood 1991). Therefore, pairwise Wilcoxon matched pairs tests between ramets and genets were performed for the entire population and for the four plots on the slopes (b) of the regression with the five polymorphic loci as replicates.

The relative contributions of seed and pollen dispersal to total gene flow were inferred by analysing, for the entire 'genet' population, the shape of the regression between the kinship coefficients and the logarithm of the distances obtained from the spatial autocorrelation analysis, using the methods described in Heuertz *et al*. (2003). The shape of the regression can be described by the k-value, calculated using the coefficients of the term of second and third power of a cubic regression between the residuals of the regression and the logarithm of the distance (for details, see Vekemans & Hardy 2004). A concave shape ($k > 0$) indicates leptokurtic gene flow, suggesting more restricted seed dispersal than pollen dispersal ($\sigma_s \ll \sigma_\rho$). For a convex curvature ($k < 0$), $\sigma_s \geq \sigma_\rho$.

Clonality and sexual reproduction in the context of heterogeneous spatial distribution of heavy metals. To compare clonal growth patterns between the LP and HP zones, the amount of multilocus genotypic variation was estimated for each of the four plots. The proportion of distinguishable multilocus genotypes was calculated as G/N (Ellstrand & Roose 1987), where G is the number of genets and N the number of ramets. Multilocus genotypic diversity (D_G) was calculated as a modification of the Simpson index (Pielou 1969). Ramets with missing data were excluded.

To test for the difference in genetic structure and sexual reproduction between the LP and HP zones, pairwise comparisons between plots (with genets) were performed on the slope (b) and on F_{IS} values using nonparametric Wilcoxon matched pairs tests (replicates = five polymorphic loci).

To detect genetic divergence between the LP and HP zones as a result of heavy metals heterogeneity, an analysis of molecular variance was performed on the four plots with ramets and with genets using ARLEQUIN (Schneider *et al*. 2000). The between-plot component of the total genetic diversity is given by F_{ST}. F-statistics were also used to partition the genetic diversity into its between-zone (LP and HP) component ($= F_{ZT}$) and its between-plot (within-zone) component ($= F_{SZ}$). Their relationship is described by $(1 - F_{ST}) = (1 - F_{SZ})(1 - F_{ZT})$ (Weir 1990). The significance of the F_{ST}, F_{SZ} and F_{ZT} values was tested using a nonparametric permutation approach (Excoffier *et al*. 1992).

Results

Extent of clonality

We did not find any linkage disequilibrium for the transect between the 10 pairs of loci after sequential Bonferroni correction ($P < 0.005$). In total, 459 (of 574 genotyped rosettes) distinct multilocus genotypes were identified for the entire population. The 232 ramets from the transect had distinct multilocus genotypes and can therefore be regarded as distinct genets. Five distinct multilocus genotypes occurred at least twice in different plots, four of which were located in LP and HP plots. At the plot scale, 38 distinct multilocus genotypes occurring at least twice within the same plot were identified; 30 of them (78.9%) were likely to be accounted for by clonal propagation ($p_{se} < 0.05$). These could, therefore, be considered as putative clones.

Genetic variation within population

At the 'genet' population level (transect + plots), A varied from 3 to 8, H_O varied from 0.428 to 0.710, H_E varied from 0.540 to 0.738 and F_{IS} varied from 0.013 to 0.208 (Table 1). Two of five loci had significant positive F_{IS} values (heterozygote deficiency). Similar F_{IS} values were found for the ramets.

Table 1 Estimates of within-population genetic variation (per locus and mean over all loci) for the entire population (transect + plots)

			Genets				Ramets	
Locus	Range in allele size	*A*	H_O	H_E	F_{IS}	*b*	F_{IS}	*b*
ATH	171–206	5	0.428	0.540	0.208**	−0.0021*	0.226**	−0.0027**
GC16	158–183	8	0.709	0.719	0.013 ns	−0.0027***	0.036 ns	−0.0039**
LYR132	229–238	3	0.562	0.610	0.079†	−0.0014***	0.040 ns	−0.0019**
LYR133	159–170	5	0.710	0.722	0.017 ns	−0.0042*	0.017 ns	−0.0063***
LYR417	185–244	6	0.683	0.738	0.074**	−0.0019***	0.090**	−0.0045***
Mean		5.4	0.618	0.666	0.071**	−0.0025***	0.075**	−0.0038***

For the genets: number of alleles (*A*), observed heterozygosity (H_O) and expected heterozygosity (H_E). For the genets and the ramets: Wright's inbreeding coefficient (F_{IS}) and slope (*b*) of the regression of pairwise kinship coefficients on the logarithm of geographical distance with significance level.
ns, not significant ($P > 0.10$); †$0.10 > P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

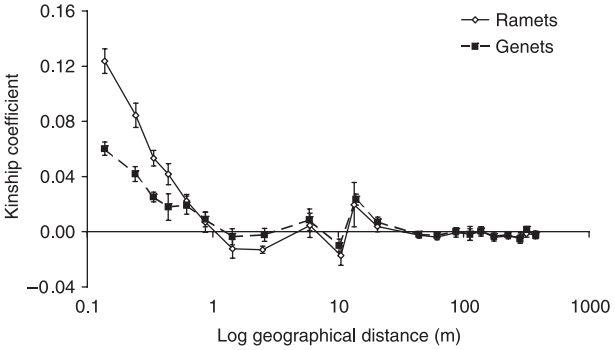

Fig. 1 Correlograms showing the spatial genetic structure for the entire (transect + plots) 'ramet' and 'genet' population, with mean Loiselle kinship coefficients over all loci (± SE) as a function of the geographical distance in metres (log-scale).

Spatial genetic structure within population: contributions of clonal growth and sexual reproduction

We found a significant linear relationship between decreasing pairwise kinship coefficients and the logarithm of increasing geographical distance in the entire (transect + plots) 'ramet' population (Fig. 1). The slopes (*b*) of the regression between pairwise kinship coefficients and spatial distances were significant ($P < 0.05$) for all five loci, with a mean *b*-value of –0.0038 (Table 1).

The entire 'genet' population (transect + plots) also showed a significant ($P < 0.05$) spatial genetic structure (mean *b*-value = –0.0025) (Fig. 1, Table 1). The slope of the 'ramet' population was, however, significantly steeper than that of the 'genet' population (Wilcoxon matched-pairs test $Z = 2.02$, $P < 0.05$). This indicated a more pronounced genetic structure for the 'ramet' population. For the genets, mean F_{IS} value (= 0.071) was similar to the average kinship coefficient between rosettes separated by the first distance interval ($F_{(1)} = 0.060$). This was not the case for the ramets ($F_{IS} = 0.075$ and $F_{(1)} = 0.124$) (Fig. 1). The estimate of neighbourhood size (*Nb*) for the genets was ~372; the *Sp* statistic was 0.0027. The shape of the linear regression between the kinship coefficients and the logarithm of the distance obtained for the spatial autocorrelation for the entire 'genet' population was found to be concave ($k = 2.19$; cubic regression $R^2 = 0.713$, $P < 0.001$), suggesting that seed dispersal was more restricted than pollen dispersal ($\sigma_s << \sigma_p$) according to Heuertz *et al.* (2003).

At the plot scale, the slopes (*b*) of the regression between pairwise kinship coefficients and spatial distances were significant ($P < 0.05$) for the four plots with genets and with ramets (Table 2). For LP1, HP1 and HP2, the slopes

Table 2 Mean Wright's inbreeding coefficient (F_{IS}) and mean slope over all loci (*b*) of the regression of pairwise kinship coefficients on the logarithm of geographical distance for the four plots with ramets and with genets, with significance level and standard error (SE), and for each plot, Wilcoxon matched pairs tests for *b* between ramets and genets

	Ramets				Genets				
Plot	F_{IS}	SE	*b*	SE	F_{IS}	SE	*b*	SE	Wilcoxon test *b*
LP1	−0.063 ns	0.082	−0.0890***	0.0181	−0.021 ns	0.078	−0.0482***	0.0072	*
LP2	−0.012 ns	0.099	−0.0763***	0.0184	−0.011 ns	0.064	−0.0393***	0.0050	†
HP1	0.149**	0.049	−0.0502***	0.0071	0.115**	0.041	−0.0333***	0.0087	*
HP2	0.075**	0.051	−0.0102***	0.0058	0.066*	0.059	−0.0063***	0.0056	*

ns, not significant ($P > 0.05$); †$P < 0.10$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Table 3 Estimates of multilocus genotypic variation for the four plots: number of distinct multilocus genotypes (*G*), proportion of distinct multilocus genotypes (*G*/*N*) and multilocus genotypic diversity (D_G)

Plot	*n*	*N*	*G*	*G*/*N*	D_G
LP1	75	74	26	0.351	0.906
LP2	59	59	32	0.542	0.958
HP1	133	124	103	0.831	0.996
HP2	112	104	98	0.942	0.999

n = number of ramets sampled, *N* = with no missing data.

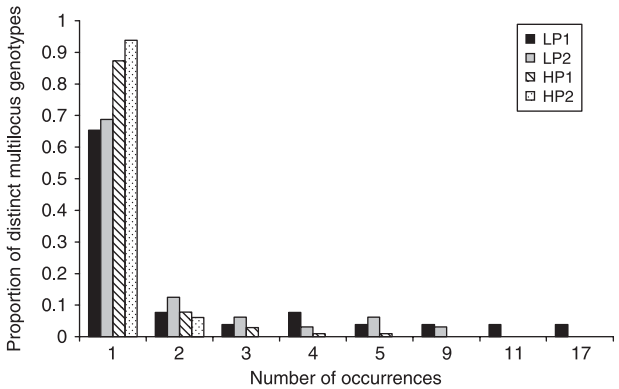

Fig. 2 Frequency-distribution diagram of the number of occurrences (= ramets) of the distinct multilocus genotypes for each of the four plots.

were significantly steeper ($P < 0.05$) for the ramets than for the genets, and marginally significant ($P = 0.08$) for LP2.

Clonality and sexual reproduction in the context of heterogeneous spatial distribution of heavy metals

Estimates of clonal diversity for each plot are given in Table 3. The proportion of distinct multilocus genotypes (*G*/*N*) varied from 0.351 to 0.942. Multilocus genotype diversity (D_G) was high, ranging from 0.906 to 0.999. Multiple occurrences of genotypes were mainly found in the LP1 and LP2 plots (Fig. 2). The distance between two identical multilocus genotypes did not exceed 1 m. These putative clones tended to aggregate in clumps, although there was some mixing of genets (Fig. 3). Ramet (*n*) and genet (*G*) density, *G*/*N* and D_G were lower for the LP1 and LP2 plots than for HP1 and HP2 (Table 3).

Pairwise comparison (Wilcoxon matched-pairs tests) between LP and HP plots (with genets) indicated that the slope was significantly ($Z = 2.02$, $P < 0.05$) lower for HP2 ($b = -0.0063$) than for LP1 and LP2 ($b = -0.0482$ and -0.0393, respectively), and not significant ($P > 0.10$) for the other pairwise comparisons. There was only a tendency for higher F_{IS} values in the HP plots than in the LP plots

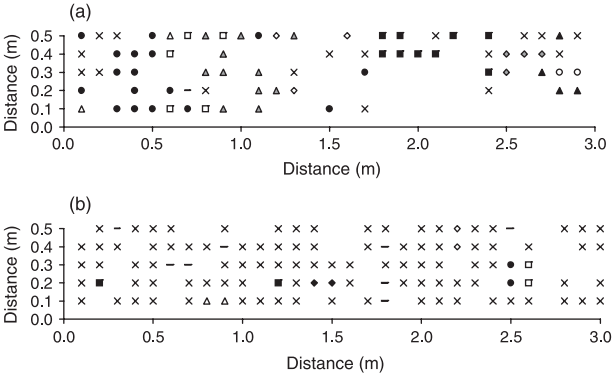

Fig. 3 Map of the distribution of ramets and their multilocus genotypes within (a) the LP1 plot ($G = 26$) and (b) the HP2 plot ($G = 98$). Symbols: X, multilocus genotypes with a single occurrence; –, incomplete genotype (missing data); others, multilocus genotypes with multiple occurrences (≥ 2). There is no correspondence between symbols in the two plots.

($Z = 1.75$, $P = 0.08$ between HP1 and LP1, HP1 and LP2, and HP2 and LP1) (Table 2). The other pairwise comparisons were not significant ($P > 0.10$).

Analysis of molecular variance on the plots with ramets and genets gave a value of –0.004 and –0.008 for the between-zone component of genetic diversity (F_{ZT}) and 0.031 and 0.019 for F_{SZ} (between-plots within LP and HP zones), respectively. The overall F_{ST} value was 0.027 for the ramets and 0.018 for the genets. The permutation test showed significance ($P < 0.001$) for the F_{SZ} and F_{ST} values, whereas F_{ZT} was nonsignificant ($P > 0.10$), indicating that the LP and HP zones were not genetically differentiated.

Discussion

Arabidopsis halleri *has a typical outcrosser behaviour on the metalliferous site of the Bois des Asturies*

Few studies have explored gene flow patterns and sexual reproduction in metallicolous plant populations. To our knowledge, this is the first study of genetic structure at a fine spatial scale conducted in a metallicolous population of an insect-pollinated plant species. The diversity of the *Arabidopsis halleri* population of 'Bois des Asturies' is high ($H_E = 0.666$; range: 0.540–0.738), and falls within the range of H_E values (0.11–0.75; mean 0.445) reported for 24 polymorphic loci by Clauss *et al*. (2002) in a metallicolous population of *A. halleri* in Germany. The slight, but significant, heterozygote deficiency ($F_{IS} = 0.07$, $P < 0.01$) is typical of an outcrossing species. However, a high F_{IS} value (0.208) was detected at one locus (*ATH*). A possible explanation might be the presence of null alleles, which can arise in microsatellite loci (Pemberton *et al*. 1995). Similar low values ($F_{IS} = 0.077$, $P = 0.10$) were reported in a German anthropogenic metallicolous population of *A. halleri* (M Clauss,

pers. commun.) and in the dioecious *Silene dioica* on serpentine (F_{IS} = 0.055, based on allozymes) (Westerbergh & Saura 1992). In contrast, high F_{IS} values were reported for metallicolous populations of the self-compatible *Thlaspi caerulescens* (F_{IS} = 0.107–0.510, based on allozymes) (Dubois *et al*. 2003). So despite the constrained ecological conditions in soils contaminated by heavy metals, which might affect genetic patterns, the breeding system appears to be a major determining factor of genetic structure.

Gene flow appears to be very extensive

We detected a significant pattern of fine-scale spatial genetic structure consistent with the model of isolation-by-distance: ramets separated by short distances are more likely to be genetically related than those that are farther apart. Sexual reproduction contributes significantly to the spatial genetic structure. Analysis of the relative contributions of seed and pollen dispersal to total gene flow indicates that seed dispersal is more restricted than pollen dispersal ($\sigma_s \ll \sigma_p$). Seeds of *A. halleri* have no particular mechanism for long-distance dispersal (Clapham & Akeroyd 1993). Hence, seed dispersal is expected to be restricted. By contrast, the flowering period is long (from April to September). Insect visitors (bumblebees, small bees, syrphids, butterflies), which may be potential pollinators, are very abundant in the 'Bois des Asturies' during the flowering season of *A. halleri* (F Van Rossum, pers. observ.). Bumblebees and butterflies can fly over long distances. As a result, long-distance pollen dispersal may be favoured and extensive pollen flow expected (Kwak *et al*. 1998). This is consistent with the particularly high estimate of neighbourhood size (*Nb* = 372) and low *Sp* value (0.0027) found for the genets. Such values are quite surprising, for an insect-pollinated herbaceous plant species like *A. halleri*, most *Nb* estimates do not exceed 200 (e.g. Richards 1997; Hardy *et al*. 2000), and mean *Sp* value is ~0.0171 (Vekemans & Hardy 2004). Such high *Nb* and low *Sp* values are generally found in wind-pollinated species (*Nb* > 200, mean *Sp* = 0.0064) (Richards 1997; Heuertz *et al*. 2003; Vekemans & Hardy 2004). Thus, our finding for *A. halleri* suggests extensive gene flow through pollen dispersal, revealing intense activity by insect pollinators. This also means that sites contaminated by heavy metals, although highly toxic, may be potentially important pollen and nectar providers for pollinators, especially in industrial, highly urban landscapes like in northern France, where suitable habitats for pollinators are few and far between.

Restricted seed dispersal possibly explains the steep slope of the spatial autocorrelation analysis at short distances (Fig. 1) and constitutes the main contribution of sexual reproduction to spatial genetic structure. The mean F_{IS} value is similar to the average kinship coefficient between genets of the first distance interval ($F_{(1)}$ = 0.060), indicating that it might partly reflect biparental inbreeding (Vekemans & Hardy 2004).

Our finding of extensive gene flow by pollen suggests that adjacent metallicolous and nonmetallicolous populations of *A. halleri* might exchange genes. In such a situation, Antonovics (1968) hypothesized that outcrossers on soils contaminated by heavy metals might evolve toward self-fertilization, so that gene flow from a nontolerant population can be prevented, and the breakdown of coadapted traits therefore avoided. Tolerance to Zn has been shown to be constitutive of the species, but variance for the tolerance has been observed among populations and individuals, metallicolous populations being more tolerant than nonmetallicolous populations (Bert *et al*. 2000). Therefore, it would prove very interesting to compare our results on genetic structure and gene flow patterns observed in an isolated metallicolous population with those found in metallicolous populations adjacent to nonmetallicolous populations of *A. halleri*.

Extent of clonality and sexual reproduction in the context of heterogeneous spatial distribution of heavy metals

A. halleri is reported to be a clonal species (e.g. Clapham & Akeroyd 1993). In the population we studied, most genets (80%) were represented by one ramet, and genotypic diversity (D_G > 0.90) was higher than the average (0.61) reported by Ellstrand & Roose (1987) in their review of clonally reproducing species. There is some spatial mixing of the genets (Fig. 3, LP plots). This can be explained by the development and rooting of vegetative rosettes on the creeping flowering stems, which can remain after the senescence of the stems and become separated from the 'mother' ramet. Clonal growth contributes significantly to within-population spatial genetic structure, as indicated by the significantly higher *b*-values for the ramets than for the genets, and by the higher F_{ST} value for the plots with ramets. However, this only happens at a small spatial scale (< 1 m) (Fig. 1). Although *A. halleri* is potentially clonal, we suggest that clonality remains a marginal mode of species' propagation.

Clonal ability can contribute to propagate- or perpetuate-adapted genotypes (e.g. Abrahamson 1980; Salemaa & Sievanen 2002). For instance, clonal spread was reported to be higher under harsh ecological conditions, e.g. at range margins (Eckert & Barrett 1993), increasing latitude (Stenström *et al*. 2001) and altitude (Young *et al*. 2002). A more intensive branching of the shoots was reported for *Arctostaphylos uva-ursi* and *Vaccinium uliginosum* plants growing on a copper–nickel polluted soil (Salemaa *et al*. 1999; Salemaa & Sievanen 2002). In our study, contrasted clonal growth patterns are observed between LP and HP zones. But as clonal spread occurs more readily in the LP

zone than in the HP zone (Table 3), it cannot be interpreted as a strategy for propagating or perpetuating adapted genotypes under the harshest ecological constraints (highest heavy metals concentrations). The contrasting clonal patterns observed for *A. halleri* seem to be a response to the variation in other environmental conditions in the site, such as humus forms (mull and mor in the LP and HP zones, respectively) or moisture levels (Gillet & Ponge 2002). This may result in locally different competition or nutrient availability (Lehman 1997).

The HP plots show significantly (or a trend for) weaker *b*-values than the LP plots. The LP and HP plots differ in clonal growth patterns but also in plant densities, as estimated by both genet and ramet numbers (Table 3). Plant density can affect the intensity of local genetic drift (Vekemans & Hardy 2004). With low plant density, genetic structure is expected to be more pronounced, as a result of a lower number of potential mates (Heywood 1991; Williams 1994). The higher plant density observed in the HP plots may reflect higher seed production and/or seedling recruitment and thus more marked sexual reproduction in this zone.

No evidence of genetic divergence due to heavy metal heterogeneity was found between LP and HP zones, as the F_{ZT} value (between zones) was not significant. Three hypotheses can be considered to explain this lack of divergence. First, within-population gene flow is too extensive to allow divergent selection processes. Second, the lowest concentrations of heavy metals (Zn, Pb, Cd) in the gradient may already correspond to very constraining conditions ([Zn] > 2792 μg/g) and therefore may not have resulted in differential selection. Third, our study is based on a small set of randomly chosen microsatellites (only five). These are generally expected to be neutral markers, but may also be in disequilibrium with selected genes. A larger array of loci should, therefore, be investigated. In a further study on the *A. halleri* population of 'Bois des Asturies', we will examine whether ecotypic divergence based on quantitative trait (survival, growth, metal accumulation) variation can be found between the LP and HP zones, using a reciprocal transplants experiment.

Acknowledgements

The authors thank Cécile Godé, Bertrand Spilliaert and Anne Petit for technical advice and support in genotyping and chemical analysis. They are very grateful to Myriam Heuertz for advice on analysis for relative contributions of seed vs. pollen dispersal, to Xavier Vekemans, Olivier Hardy, Jean-François Arnaud and three anonymous referees for helpful comments on the manuscript, and to Lycia Touzet for correction of the English. This study was supported by funding from the Contrat de Plan Etat/Région Nord-Pas de Calais (PRC), from the European FEDER, and from the 'Programme Environnement Vie et Société' (PEVS) from the CNRS (contract N°00 N55/0083).

References

Abrahamson WG (1980) Demography and vegetative reproduction. In: *Demography and Evolution in Plant Populations* (ed. Solbrig OT), pp. 89–106. Blackwell, Oxford.

Antonovics J (1968) Evolution in closely adjacent populations V. Evolution of self-fertility. *Heredity*, **23**, 219–238.

Antonovics J, Bradshaw AD (1970) Evolution in closely adjacent populations VIII. Clinal patterns at a mine boundary. *Heredity*, **25**, 349–362.

Bell CJ, Eckert JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics*, **19**, 137–144.

Bert V, Bonnin I, Saumitou-Laprade P, de Laguérie P, Petit D (2002) Do *Arabidopsis halleri* from nonmetallicolous populations accumulate zinc more effectively than those from metallicolous populations? *New Phytologist*, **155**, 47–57.

Bert V, Macnair MR, de Laguérie P, Saumitou-Laprade P, Petit D (2000) Zinc tolerance and accumulation in metallicolous and nonmetallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **146**, 225–233.

Berton A (1946) Présentation de plantes: *Arabidopsis halleri*, *Armeria elongata*, *Oenanthe fluviatilis*, *Galinsoga parviflora dicoidea*. *Bulletin de la Société Botanique Du Nord de la France*, **93**, 139–145.

Charpentier A (2002) Consequences of clonal growth for plant matings. *Evolutionary Ecology*, **15**, 521–530.

Clapham AR, Akeroyd JR (1993) *Cardaminopsis*. In: *Flora Europaea*, Vol. I (eds Tutin TG, Heywood VH, Burges NA, Valentine DH, Walters SM, Webb DA), p. 290. Cambridge University Press, Cambridge.

Clauss MJ, Cobban H, Mitchell-Olds T (2002) Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaceae). *Molecular Ecology*, **11**, 591–601.

Dubois S, Cheptou P-O, Petit C *et al.* (2003) Genetic structure and mating systems of metallicolous and nonmetallicolous populations of *Thlaspi caerulescens*. *New Phytologist*, **157**, 633–641.

Ducousso A, Petit D, Valero M, Vernet P (1990) Genetic variation between and within populations of a perennial grass: *Arrhenatherum elatius*. *Heredity*, **65**, 179–188.

Eckert CG, Barrett SCH (1993) Clonal reproduction and patterns of genotypic diversity in *Decodon verticillatus* (Lythraceae). *American Journal of Botany*, **80**, 1175–1182.

Ellstrand NC, Roose ML (1987) Patterns of genotypic diversity in clonal plant species. *American Journal of Botany*, **71**, 123–131.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Gillet S, Ponge JF (2002) Humus forms and metal pollution in soil. *European Journal of Soil Science*, **53**, 529–539.

Goudet J (2001) FSTAT, *a Program to Estimate and Test Gene Diversities and Fixation Indices, Version 2.9.3.2*. http://www.unil.ch/izea/softwares/fstat.html [updated from Goudet 1995].

Hardy OJ, Vanderhoeven S, Meerts P, Vekemans X (2000) Spatial autocorrelation of allozyme and quantitative traits within a natural population of *Centaurea jacea* (Asteraceae). *Journal of Evolutionary Biology*, **13**, 656–667.

Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Heuertz M, Vekemans X, Hausman JF, Palada M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology*, **12**, 2483–2495.

Heywood JS (1991) Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics*, **22**, 335–355.

Kevan PG (1999) Pollinators as bioindicators of the state of the environment: species, activity and diversity. *Agriculture, Ecosystems and Environment*, **74**, 373–393.

Kirby GC (1975) Heterozygote frequencies in small subpopulations. *Theoretical Population Biology*, **8**, 31–48.

Kwak MM, Velterop O, van Andel J (1998) Pollen and gene flow in fragmented habitats. *Applied Vegetation Science*, **1**, 37–54.

Lehmann C (1997) Clonal diversity of populations of *Calamagrostis epigejos* in relation to environmental stress and habitat heterogeneity. *Ecography*, **20**, 483–490.

Leita L, Muhlbachova G, Cesco S, Barbattini R, Mondini C (1996) Investigation of the use of honey bees and honey bee products to assess heavy metals contamination. *Environmental Monitoring and Assessment*, **43**, 1–9.

Li Y-C, Röder MS, Fahima T *et al.* (2000) Natural selection causing microsatellite divergence in wild emmer wheat at the ecologically variable microsite at Ammiad, Israel. *Theoretical and Applied Genetics*, **100**, 985–999.

Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, **27**, 237–277.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.

Lönn M, Prentice HC, Bengtsson K (1996) Genetic structure, allozyme–habitat associations and reproductive fitness in *Gypsophila fastigiata* (Caryophyllaceae). *Oecologia*, **106**, 308–316.

Macnair MR (1987) Heavy metal tolerance in plants: a model evolutionary system. *Trends in Ecology and Evolution*, **2**, 354–359.

McNeilly T (1967) Evolution in closely adjacent plant populations III. *Agrostis tenuis* on a small copper mine. *Heredity*, **23**, 99–108.

Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology and Evolution*, **16**, 693–700.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Oetting WS, Lee HK, Flanders DJ *et al.* (1995) Linkage analysis with multiplexed short tandem repeat polymorphism using infrared fluorescence and M13 tailed primers. *Genomics*, **30**, 450–458.

Parks JC, Werth CR (1993) A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, **80**, 537–544.

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4**, 249–252.

Pielou EC (1969) *An Introduction to Mathematical Ecology*. Wiley Interscience, New York.

Prentice HC, Lönn M, Lager H, Rosén E, van der Maarel E (2000) Changes in allozyme frequencies in *Festuca ovina* populations after a 9-year nutrient/water experiment. *Journal of Ecology*, **88**, 331–347.

Rice WR (1989) Analysing tables of statistical tests. *Evolution*, **43**, 223–225.

Richards AJ (1997) *Plant Breeding Systems*, 2nd edn. Chapman & Hall, Cambridge.

Salemaa M, Sievanen R (2002) The effects of apical dominance on the branching architecture of *Arctostaphylos uva-ursi* in four contrasting environments. *Flora*, **197**, 429–442.

Salemaa M, Vanha-Majamaa I, Gardner PJ (1999) Compensatory growth of two clonal dwarf shrubs, *Arctostaphylos uva-ursi* and *Vaccinium uligonosum* in a heavy-metal polluted environment. *Plant Ecology*, **141**, 79–91.

Salt DE, Smith RD, Raskin I (1998) Phytoremediation. *Annual Review of Plant Physiology and Plant Molecular Biology*, **49**, 643–668.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, *Version 2000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Sokal RR, Rohlf FJ (2000) *Biometry: The Principles and Practice of Statistics in Biology Research*, 3rd edn. W.H. Freeman, New York.

Stenström A, Jonsson BO, Jónsdóttir IS, Fagerstöm T, Augner M (2001) Genetic variation and clonal diversity in four clonal sedges (*Carex*) along the Arctic coast of Eurasia. *Molecular Ecology*, **10**, 497–513.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 912–935.

Vekemans X, Lefèbvre C (1997) On the evolution of heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. *Journal of Evolutionary Biology*, **10**, 175–191.

Weir BS (1990) *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.

Westerbergh A, Saura A (1992) The effect of serpentine on the population structure of *Silene dioica*. *Evolution*, **46**, 1537–1548.

Westerbergh A, Saura A (1994) Gene flow and pollinator behaviour in *Silene dioica* populations. *Oikos*, **71**, 215–224.

Williams CF (1994) Genetic consequences of seed dispersal in three sympatric forest herbs. II. Microspatial genetic structure within populations. *Evolution*, **48**, 1959–1972.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Young AG, Hill JH, Murray BG, Peakall R (2002) Breeding system, genetic diversity and clonal structure in the sub-alpine forb *Rutidosis leiolepis* F. Muell. (Asteraceae). *Biological Conservation*, **106**, 71–76.

The authors of the laboratory [Génétique et Evolution des Populations Végétales (GEPV)] are involved in population genetics of plant species with a special attention given to the genetics and evolution of zinc tolerance and zinc hyperaccumulation in *Arabidosis halleri*. F. Van Rossum's research interests are centred around a range of questions in plant evolutionary biology and conservation genetics related to fragmentation of natural plant populations and spatial environmental heterogeneity. I. Bonnin is population geneticist particularly interested in evolution of selected characters and relationships between migration, diversity and adaptation in experimental populations of wheat and *Arabidopsis*.
