MOLECULAR ECOLOGY

Molecular Ecology (2009) **18**, 3201–3215 doi: 10.1111/j.1365-294X.2009.04279.x

Fine-scale geographical structure of genetic diversity in inland wild beet populations

JEAN-FRANÇOIS ARNAUD, STÉPHANE FÉNART, CÉCILE GODÉ, SYLVIE DELEDICQUE, PASCAL TOUZET and JOËL CUGUEN

Laboratoire de Génétique et Évolution des Populations Végétales, UMR CNRS 8016, Bâtiment SN2, Université des Sciences et Technologies de Lille – Lille 1, F-59655 Villeneuve d'Ascq Cedex, France

Abstract

Introgression arising from crop-to-wild gene flow provides novel sources of genetic variation in plant species complexes. Hybridization within the *Beta vulgaris* species complex is of immediate concern; crop lineages (*B. vulgaris* ssp. *vulgaris*) hybridize easily with their wild relatives (*B. vulgaris* ssp. *maritima*) thereby threatening wild beet gene diversity with genetic swamping. Hybridization 'hotspots' occur in European seed production areas because inland ruderal wild beets occur and reproduce in sympatry with cultivated beets. We studied gene flow occurring between seed-producing cultivars and ruderal wild *B. vulgaris* in southwestern France to determine whether feral beets, arising from unharvested cultivated seed, represent an opportunity for crop-to-wild gene flow. We surveyed 42 inland ruderal beet populations located near seed production fields for nucleo-cytoplasmic variation and used a cytoplasmic marker diagnostic of cultivated lines. Occurrence of cultivated-type cytoplasm within ruderal populations clearly reflected events of crop seed escape. However, we found no genetic signatures of nuclear cultivated gene introgression, which suggests past introgression of cultivated cytoplasm into a wild nuclear background through seed escape rather than recent direct pollen flow. Overall, patterns of genetic structure suggested that inland ruderal wild beet populations act as a metapopulation, with founding events involving a few sib groups, followed by low rates of seed or pollen gene flow after populations are established. Altogether, our results indicate that a long-lived seed bank plays a key role in maintaining cultivated-type cytoplasm in the wild and highlight the need for careful management of seed production areas where wild and cultivated relatives co-occur.

Keywords: *Beta vulgaris*, conservation genetics, crop-to-wild gene flow, genetic pollution, nuclear and cytoplasmic differentiation, seed dispersal

Received 10 November 2008; revision received 9 May 2009; accepted 14 May 2009

Introduction

Domesticated plants and their wild relatives often form large species complexes composed of crops, accompanying feral weeds and related wild species. These species complexes constitute evolving systems connected through both shared ancestry and multiple pathways of gene exchange that have important ecological and evolutionary consequences (Ellstrand 2003a; Gepts & Papa 2003; Stewart *et al.* 2003; Reagon & Snow 2006). Gene flow is thus a crucial parameter in the cultivation of crops because it can affect the genetic diversity of related landraces and wild relatives. Crop-to-wild gene flow has been implicated in the increased likelihood of extinction of wild relatives by displacing native wild allelic diversity (Ellstrand & Elam 1993; Bartsch *et al.* 1999; Gepts & Papa 2003; Chapman & Burke 2006; Bitocchi *et al.* 2009). Crop-to-wild gene flow leading to successful hybridization also appears to have played a significant role in the evolution of weediness including new invasive species (Ellstrand & Schierenbeck 2000;

Correspondence: Jean-François Arnaud, Fax: +33 3 20 43 69 79; E-mail: jean-francois.arnaud@univ-lille1.fr

Campbell *et al.* 2006; Fénart *et al.* 2007; Sørensen *et al.* 2007). As pointed out by Stewart *et al.* (2003), introgression of domesticated traits generally creates maladapted crop–wild hybrids with low fitness, where local adaptation and intrinsic co-adapted gene complexes are disrupted. Thus, assessment of the consequences of gene flow leading to crop–wild hybrids often dismisses the potential for hybrid lineages to persist or even thrive in natural or semi-natural environments (Warwick *et al.* 2008; Zapiola *et al.* 2008).

Several conditions must be fulfilled for gene flow to take place among crops and their wild relatives and for the long-term establishment of domesticated genes in the wild gene pool: first, wild populations must grow within the range of pollen and seed dispersal from cultivated crops; second, crop–wild crosses must yield viable and fertile progeny so that backcrossing between hybrids and wild progenitors can occur and therefore maintain crop genes in the wild gene pool; and, finally, ongoing gene flow by pollen or seed requires also synchrony, or at least partial overlap, in flowering periods (Ellstrand 2003a; Gepts & Papa 2003; Chapman & Burke 2006).

The *Beta vulgaris* species complex is of immediate concern as the crop lineages (*Beta vulgaris* ssp. *vulgaris* including sugar, fodder, red table or leaf beets) have a high potential for hybridization with their wild relatives, *Beta vulgaris* ssp. *maritima*, also known as sea beet. Wild beet populations mostly occur along the Northern Atlantic coastline, with a large geographical distribution from Cape Verde and Canary Islands in the west, northwards along continental Europe's coast to the North and Baltic Seas, and also extend eastwards through the Mediterranean area (Letschert 1993). Typically, wild beets are not found in inland situations but rather grow along the shoreline at the upper high tide level (Raybould *et al.* 1996; Viard *et al.* 2004; Fievet *et al.* 2007). However, along the Mediterranean Basin and especially within the French seed production area (southwestern France), atypical, noncoastal wild beet populations have colonized ruderal habitats, i.e. disturbed, human-influenced habitats.

In this context, the main goal of this study was to determine whether cultivated seed escape represents an opportunity for gene flow from cultivars grown for seed production to wild relatives. To this end, 42 putative inland wild populations, either situated close to or far from seed production fields, were sampled and examined for nuclear and cytoplasmic variation within the French seed production area. The amount and spatial distribution of genetic diversity will be assessed and discussed with respect to management of crop gene escape. The questions addressed in this study are the following:

1 Using a cytoplasmic marker specific to cultivated lines to trace seed dispersal, can we find evidence of introgression from cultivated crops to wild, ruderal populations? If so, can this introgression be correlated with the proximity of seed production fields? We expect to find cultivated-type cytoplasm in suspected feral beet populations located along field margins within the seed production area.
2 Are there any differences in the amount of genetic diversity and population genetic structure between ruderal wild populations found in proximity to seed production fields compared with others located far from agricultural fields? For wild populations located in the close vicinity of seed production fields, the expectations are twofold: (i) an increase in allelic richness and/or displacement of native genetic diversity because of potential pollen and seed flow from genetically differentiated crop lines, as observed by Fénart *et al.* (2008) in weed beets infesting sugar beet fields in Northern France (see also Bartsch *et al.* 1999; for coastal wild sea beets in Italy) and (ii) a departure from panmixia related to the introgression of crop genes involving a Mendelian self-fertility factor widely used in cultivated germplasm and leading to the breakdown of the self-incompatibility system (Owen 1942; Cuguen *et al.* 2004). Similar predictions can be made for wild beet populations characterized by the occurrence of individuals carrying a cultivated-type cytoplasm.
3 Is the genetic diversity geographically structured following a classical isolation-by-distance model that supposes that gene dispersal is spatially restricted? Assuming that wild inland populations are in migration–drift genetic equilibrium without any external pollen or seed flow events from cultivars, we expect a progressive decline of genetic similarity with geographical distance between populations, as it is classically observed in wild beet populations located along the shoreline (e.g. Viard *et al.* 2004; Fievet *et al.* 2007).

Materials and methods

The species and the study site

The wild beet (*Beta vulgaris* ssp. *maritima*) is diploid ($2n = 18$), wind-pollinated, perennial, self-incompatible, gynodioecious and is the wild ancestor of all cultivated beets (*Beta vulgaris* ssp. *vulgaris*). Gene flow has been found to be efficient between cultivars and wild relatives, illustrated by both overlapping flowering periods in the areas of sympatry and successful controlled cross-pollinations (Bartsch *et al.* 1999; Desplanque *et al.* 2002). There is no vegetative reproduction and dispersal

can only occur through seeds and/or pollen movement. Seeds are aggregated in a seedball that has no particular dispersal mechanism and is primarily dispersed by gravity. Therefore, except for accidental long-distance dispersal events, seed movements usually show a short-range pattern of dispersal, in contrast to pollen which is the most efficient source of gene flow through wind dispersal (Fénart *et al.* 2007; Fievet *et al.* 2007). Cultivated beets and F1 crop–wild hybrids are morphologically very different from wild beets (thick roots, large leaves of different shape and colour, Arnaud *et al.* 2003; Van Dijk 2004) and crop and wild lineages have been shown to be genetically distinct in both their nuclear and cytoplasmic compartments (Viard *et al.* 2004; Fénart *et al.* 2008).

Although, in northern Europe, wild beets are tightly restricted to the shoreline, they have a looser distribution in the Mediterranean area (Desplanque *et al.* 1999; Ellstrand 2003b; Van Dijk 2004). As a result, wild beets occur both on French coasts and in human-disturbed (ruderal) habitats in the inland areas and valleys of southwestern France. Ruderal wild beet populations form a continuum from Narbonne (Mediterranean region) up to Agen (halfway between Bordeaux and Toulouse), with no connection to the Atlantic coast near Bordeaux (Fig. 1). Feral plants are defined as plants

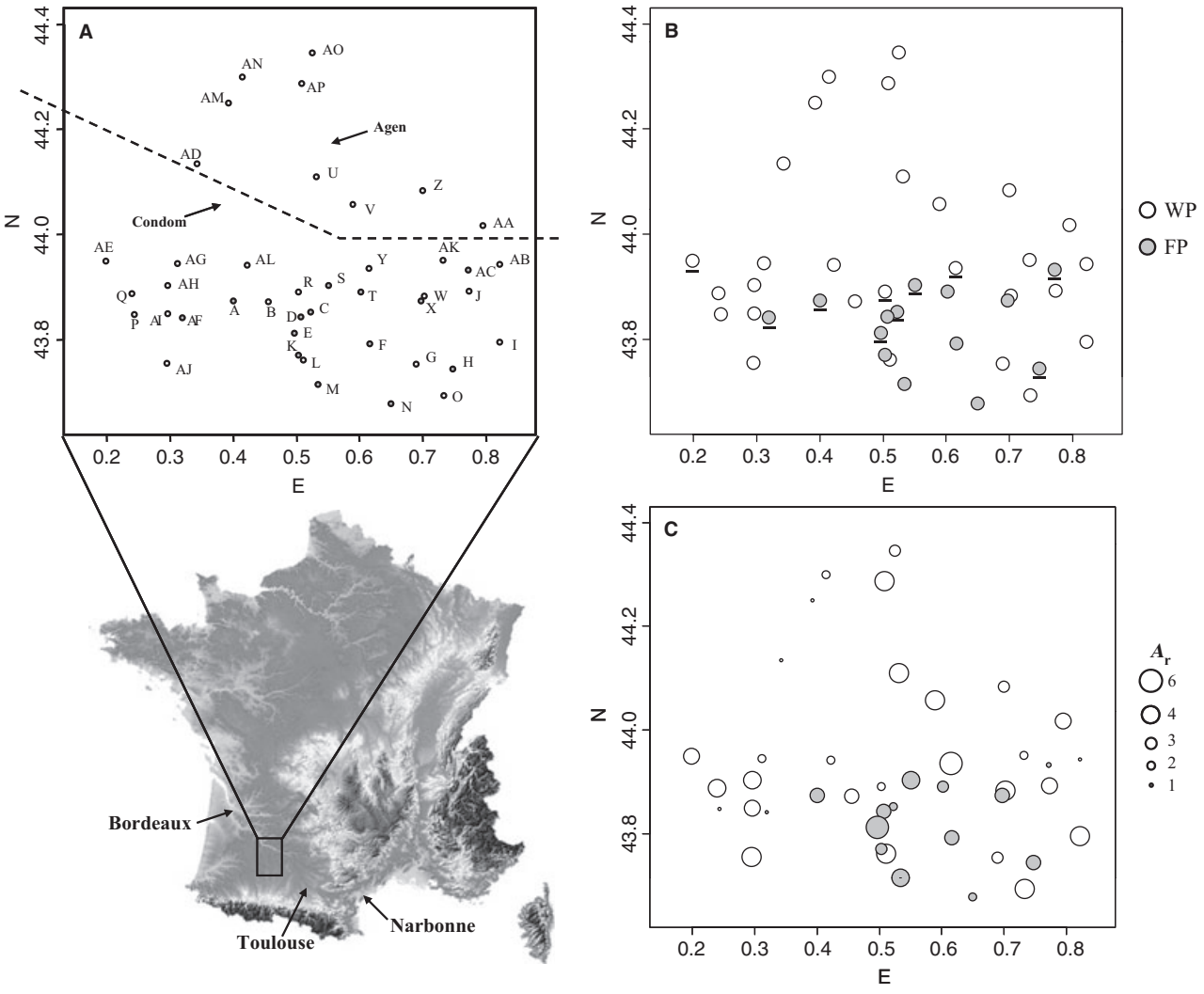

Fig. 1 Map of the study area and location of sampled populations of *Beta vulgaris* within the French seed production area. Typical wild populations (WP) and feral populations (FP) are represented by open circles and grey filled circles, respectively. (A) Geographical location and acronyms of populations; the dashed line separates the former seed production area (southern part) from the current seed production area (northern part). (B) Populations characterized by the occurrence of CMS Owen cytoplasm (specific to cultivated lines) are underlined. (C) Graphical representation of mean cytoplasmic allelic richness, as revealed by the mitochondrial minisatellite TR1.

that have escaped from domestication and returned to the wild state. Up to now, it has not been clear whether some of these ruderal beets are feral beets, because many crops are assumed to form transient feral populations, or if they were already present before seed production developed in southwestern France. In addition, small farmers in this region frequently grow red beet or Swiss chard in kitchen gardens for seed production and consumption, which may be an additional source of cultivar gene escape if they bolt and flower before being harvested (Desplanque *et al.* 2002).

French seed crop production is situated in southwestern France in the vicinity of Agen. Cultivated beet seed has been produced in this region since the 1960s and the major proportion of beet seeds produced in Western Europe come from this region (Van Dijk 2004). In seed production fields, four lines of male-sterile seed bearers alternate with two lines of pollen donors. In the European Community, the recommended minimum distance between beet seed production fields and pollen sources from neighbouring wild relatives is at least 1000 m to avoid contamination of seed lots by wild beets as the pollen parent. As the sugar beet seed producers have implemented isolation measures by removing all flowering ruderal beets 1000 m around the fields, contamination of seed lots by crop–wild hybrids has fallen to 0.04% in France (Alibert *et al.* 2005). However, over the last 10 years, the area of beet seed production has shifted northwards to the Lot-et-Garonne department within the French seed production area (Fig. 1A). The former areas of seed production have been abandoned because of the high incidence of ruderal beets, which makes it very difficult to check and remove them. For this reason, we carefully checked for the presence of populations along field margins, both in the former area of beet seed production situated in the south and the present area of beet seed production situated northwards between Condom and Agen (see Fig. 1A).

Sampling design

After carefully investigating the whole seed production area in the summer of 2001, we collected a total of 42 populations, with 33 populations in the former beet seed production area and nine populations in the present beet seed production area (Fig. 1A, B). Collected populations can be partitioned into 28 typical ruderal populations found within villages and 14 populations sampled within or along field margins. The 28 ruderal populations were associated with past and present construction (roadsides, parking lots), rubble deposits or garden edges within villages. These populations were all located more than 1000 m from crop fields or kitchen gardens and will a priori be classified as 'wild populations' (WP) throughout this study. The second category of populations was composed of beets located within or in the direct vicinity (i.e. along field margins) of crop seed production fields. These 14 sampled populations were therefore suspected to be feral populations originating from cultivated seed escape and will a priori be referred as 'feral populations' (FP) hereafter.

The salient feature of this sampling was that no FP populations were found in the present beet seed production area, as depicted in Fig. 1B. In contrast, within the former seed production area, 14 populations of 33 were found to be within or in the vicinity of crop fields and consequently classified as FP populations (Fig. 1B). It must be emphasized that none of the collected individuals displayed cultivated-like morphological traits. Therefore, our classification into WP and FP populations is based solely on geographical considerations: WP populations occurred at least 1000 m away from former and present beet seed production fields, whereas FP populations were all located along or within former beet seed production fields.

Leaves from single individuals were collected to minimize kinship and biparental inbreeding. For each population, the sampling area did not exceed 30 m^2. The number of individuals sampled ranged from 38 to 57 per population with a mean sample size of 49.78 ± 2.40. Sample sizes and a priori population classification are detailed in Table 1 for each population. For molecular studies, leaves from 2091 individuals were collected and dried in silica gel prior to DNA extraction.

Molecular typing

Extraction and purification of total DNA was performed using a DNeasy® 96 Plant Kit following the standard protocol for isolation of DNA from plant leaf tissue outlined in the DNeasy 96 Plant protocol handbook (QIAGEN Inc.) and as described in Arnaud *et al.* (2003).

Sugar and fodder beets typically have a combination of mitochondrial and chloroplast genotypes designated 'CMS' to signify their inheritance from a male-sterile seed parent used for seed production. This CMS has been coined 'Owen CMS' (Owen 1945) and is carried by 100% of cultivated lineages because of its worldwide use in sugar beet breeding programmes since the 1960s. This CMS-associated cytoplasm is known to be absent in wild beet populations found along the shoreline, or to occur at very marginal frequencies as a result of crop seed escape (Cuguen *et al.* 1994; Viard *et al.* 2004; Fénart *et al.* 2008; Dufay *et al.* 2009). In this study, we checked all populations for the occurrence of Owen CMS cytoplasm using a diagnostic chloroplastic polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) marker related to a polymorphic

Table 1 Locations, codes, geographical coordinates (GPS), population classification and sample sizes (*N*, number of individuals) of collected wild beet populations. Mean estimates of genetic diversity per population are also provided: number of alleles (A_n), allelic richness (A_r), observed heterozygosity (H_O), gene diversity (H_E) over the seven microsatellites used for nuclear data and for the mitochondrial minisatellite locus TR1 used for cytoplasmic genetic diversity. Mean F_{IS}-estimates and significance of departures from Hardy–Weinberg equilibrium are reported for nuclear data. Mean percentage of individuals carrying a CMS Owen cytoplasm (% CMS Owen), specific to cultivated lines, is also indicated

Population location	Code	Class	*N*	Latitude N	Longitude E	Nuclear diversity					Cytoplasmic diversity		
						A_n	A_r	H_O	H_E	F_{IS}	% CMS Owen	A_n	A_r
St Puy	B	WP	48	43.873	0.456	6.14	5.79	0.57	0.62	0.07^{NS}	0.00	4	3.91
Tourrenquets	G	WP	50	43.753	0.690	6.00	5.79	0.52	0.60	0.14**	0.00	3	3.00
Montfort	I	WP	51	43.795	0.821	7.00	6.69	0.56	0.64	0.12^{NS}	0.00	5	5.00
Saint Clar	J	WP	49	43.892	0.773	6.71	6.48	0.55	0.63	0.14***	0.00	4	4.00
Lavardens village	L	WP	51	43.761	0.511	6.57	6.34	0.60	0.65	0.07^{NS}	0.00	5	5.00
Nougaroulet	O	WP	50	43.693	0.733	6.29	6.08	0.54	0.64	0.15*	0.00	5	5.00
Courrensan	P	WP	38	43.848	0.244	5.00	5.00	0.47	0.56	0.16^{NS}	0.00	1	1.00
Gondrin	Q	WP	49	43.888	0.240	6.57	6.18	0.51	0.60	0.16***	0.00	5	4.69
Mas d'Auvignon	R	WP	53	43.890	0.503	6.00	5.81	0.55	0.59	0.07^{NS}	1.89	2	1.89
Laplume	U	WP	51	44.110	0.531	6.71	6.46	0.56	0.66	0.15***	0.00	5	4.86
Pargan-Taillac	V	WP	51	44.058	0.589	6.57	6.38	0.51	0.64	0.21***	0.00	5	4.88
Castelnau d'Arbieu	W	WP	50	43.884	0.703	6.00	5.72	0.57	0.62	0.08^{NS}	0.00	5	4.98
Lectoure	Y	WP	50	43.936	0.615	7.00	6.74	0.61	0.67	0.09^{NS}	24.00	6	5.95
Cuq	Z	WP	50	44.083	0.699	4.86	4.76	0.49	0.53	0.07^{NS}	0.00	3	2.89
Flamarens	AA	WP	50	44.018	0.794	5.57	5.37	0.58	0.64	0.09^{NS}	0.00	4	4.00
Marsac	AB	WP	50	43.944	0.822	4.71	4.49	0.55	0.59	0.07^{NS}	0.00	1	1.00
Nérac	AD	WP	50	44.135	0.342	5.00	4.77	0.58	0.61	0.06^{NS}	0.00	1	1.00
Montréal sur Gers	AE	WP	50	43.949	0.199	6.43	6.25	0.65	0.66	0.03^{NS}	44.00	4	4.00
Larressingle	AG	WP	50	43.945	0.311	6.57	6.22	0.54	0.59	0.09^{NS}	0.00	2	2.00
Mouchan	AH	WP	50	43.903	0.297	6.00	5.82	0.56	0.66	0.16***	0.00	5	4.55
Roques	AI	WP	50	43.849	0.297	5.57	5.44	0.59	0.66	0.12***	0.00	4	4.00
Vic-Fézensac	AJ	WP	50	43.755	0.296	5.29	5.12	0.58	0.61	0.04^{NS}	0.00	5	4.89
Plieux	AK	WP	50	43.950	0.732	6.29	5.99	0.56	0.61	0.08^{NS}	0.00	2	2.00
Caussens	AL	WP	50	43.942	0.422	4.14	3.95	0.41	0.46	0.11^{NS}	0.00	2	2.00
port-sainte Marie	AM	WP	45	44.250	0.393	4.71	4.55	0.57	0.59	0.02^{NS}	0.00	1	1.00
Galapian	AN	WP	50	44.300	0.414	5.86	5.58	0.46	0.56	0.18***	0.00	2	2.00
Montpézat	AO	WP	50	44.346	0.525	6.29	6.10	0.52	0.62	0.17***	0.00	3	3.00
Prayssas	AP	WP	50	44.287	0.508	5.29	5.13	0.52	0.64	0.19***	0.00	5	4.91
Maignaut-Tauziat	A	FP	50	43.875	0.400	6.71	6.35	0.50	0.55	0.09^{NS}	22.00	5	5.00
La sauvetat (friche)	C	FP	50	43.852	0.523	5.57	5.41	0.59	0.63	0.07^{NS}	6.00	3	3.00
La sauvetat (chêne)	D	FP	50	43.843	0.507	6.14	5.91	0.53	0.62	0.15^{NS}	0.00	5	4.89
Cézan	E	FP	50	43.812	0.497	6.71	6.44	0.58	0.60	0.04^{NS}	2.00	8	7.91
Montestruc-sur-Gers	F	FP	50	43.792	0.616	6.86	6.58	0.51	0.61	0.15*	0.00	5	5.00
Puycasquier	H	FP	57	43.744	0.747	6.00	5.78	0.58	0.68	0.14***	17.54	5	4.84
Lavardens	K	FP	50	43.771	0.503	6.43	6.17	0.60	0.65	0.08^{NS}	0.00	4	3.99
Larrama	M	FP	50	43.716	0.533	5.29	5.17	0.57	0.62	0.08^{NS}	0.00	6	5.91
Leboulin	N	FP	49	43.678	0.650	6.29	6.16	0.58	0.63	0.09^{NS}	0.00	3	3.00
Terraube	S	FP	50	43.903	0.551	7.43	7.05	0.58	0.64	0.10^{NS}	10.00	6	5.98
La Haille	T	FP	50	43.891	0.602	5.43	5.26	0.57	0.64	0.10^{NS}	0.00	4	3.99
Arnaud Lanne	X	FP	50	43.874	0.697	5.43	5.26	0.44	0.56	0.20***	0.00	5	5.00
Gramont	AC	FP	50	43.933	0.772	6.57	6.27	0.60	0.66	0.10^{NS}	2.00	2	1.91
Lagardère	AF	FP	49	43.842	0.320	5.00	4.83	0.49	0.59	0.18***	100.00	1	1.00

WP, wild populations found in association with past and present road construction (roadsides, parking lots), rubble deposits or garden edges within villages; FP, feral populations found within or in close vicinity of ancient fields devoted to seed production.
^{NS}nonsignificant; *$P < 0.05$; **$P < 0.01$; **$P < 0.001$.

*Hin*dIII site mapped in the *petG-psbE* chloroplast fragment. *Hin*dIII digests of cpDNA from non-Owen CMS and Owen CMS lines are characterized by one single *Hin*dIII fragment of 563 bp and by two *Hin*dIII fragments of 454 and 109 bp respectively (Ran & Michaelis 1995). Primers used, PCR conditions and DNA

digestion of this PCR-RFLP method were applied as described by Ran & Michaelis (1995).

This polymorphism makes it possible to distinguish Owen CMS (cultivated) from non-Owen CMS (wild) lines, but wild cytoplasmic diversity cannot be detected with this PCR-RFLP assay. Therefore, to obtain further information on cytoplasmic diversity, we used an additional mitochondrial minisatellite locus named TR1, first described by Nishizawa *et al.* (2000) and found to be highly polymorphic in wild beet populations (see Fievet *et al.* 2007; Fénart *et al.* 2008). PCR amplifications and cycling conditions were performed as described in Fievet *et al.* (2007). Polymorphism was revealed by electrophoresis on 2% agarose gels and visualized after ethidium bromide staining under UV-light using known individuals as internal size standards.

With respect to nuclear genetic variation, individuals were all genotyped at seven microsatellite loci (CT4, GTT1, GCC1, GAA1, BVM3, CAA1 and CA2) according to PCR protocols previously described in Mörchen *et al.* (1996), Viard *et al.* (2002) and Arnaud *et al.* (2003). These loci were selected for their polymorphism and their unambiguous amplification patterns for allelic identification (Cuguen *et al.* 2004). Electrophoresis and genotyping were performed on a LI-COR automated DNA sequencer model 4200s (LI-COR Inc.) as described by Viard *et al.* (2002). To minimize the rate of genotyping errors, a second round of PCR and electrophoresis were performed for individuals with dubious multilocus genotypes (i.e. with missing data or displaying rare alleles).

Statistical analyses of cytoplasmic and nuclear genetic variation

Basic and descriptive statistics such as the mean number of alleles (A_n), the allelic richness (A_r), the level of observed (H_O) and expected (H_E) heterozygosities, and the intra-population fixation index F_{IS} (a measure of departure from panmixia within populations) per locus and per population were calculated using the software FSTAT 2.9.3 (Goudet 1995). H_E refers to the gene diversity *sensu* Nei (1978), and A_r is an unbiased estimate of the expected number of alleles independent of the sample size, following the rarefaction procedure described in El Mousadik & Petit (1996). Differences in mean A_n, A_r, H_O, H_E and F_{IS} between WP and FP populations were tested using the Mann–Whitney *U*-test. With regard to nuclear data, genotypic disequilibrium and deviations from Hardy–Weinberg equilibrium were estimated with exact probability tests across loci and populations using GENEPOP version 3.4 (Raymond & Rousset 1995). The Markov chain method provided unbiased estimates of the Fisher's exact test probability using the following parameters: 3000 dememorizations, 500 batches, 200 000 iterations per batch. The overall critical significance of multiple tests (over populations or loci) was estimated by Fisher's combined probability test, applying sequential Bonferroni adjustments for simultaneous statistical tests (Rice 1989).

We first assessed population subdivision with estimates of genetic differentiation derived from Wright's *F*-statistics and using the analysis of variance procedure of Weir & Cockerham (1984). Mean *F*-estimates per locus and averaged *F*-indices over loci were calculated by jackknifing among the samples and over loci respectively, except for the mitochondrial minisatellite locus TR1. For nuclear data, 95% confidence intervals for averaged *F*-estimates were obtained by bootstrapping over loci. To determine whether estimates of F_{IS} and F_{IT} were significantly different from zero, we used 10 000 randomizations of alleles within and overall population samples respectively. When testing for significance of population differentiation (F_{ST}), we used the log-likelihood *G*-test (10 000 runs of multilocus genotypes randomization among population samples) as suggested when it is suspected that some populations do not perfectly meet Hardy–Weinberg requirements (Goudet *et al.* 1996). All *F*-calculations were performed using FSTAT version 2.9.3 (Goudet 1995). Bonferroni adjustments for simultaneous multiple statistical tests were applied as described by Rice (1989). To test for a difference in the apportionment of genetic variation between the two a priori groups of *B. vulgaris* populations, we further performed a hierarchical analysis of molecular variance (AMOVA) using Arlequin v3.1 (Excoffier *et al.* 2005) to analyse the partition of the genetic variance within (F_{SC}) and between (F_{CT}) samples classified as WP and FP populations.

To depict the genetic relationships of populations through neighbour-joining trees, pairwise genetic divergence between populations of *B. vulgaris* was quantified from cytoplasmic and nuclear microsatellite data using the Cavalli-Sforza & Edwards (1967) chord distance (D_{CE}) based on allelic frequencies. This genetic distance appears to be one of the most efficient in depicting the correct tree topology among closely related populations under various mutation model assumptions (Takezaki & Nei 2008).

In an attempt to test whether the genetic differentiation between populations followed a classical isolation-by-distance process because of geographically restricted gene flow, one-dimensional Mantel correlograms were designed following the approach described in Oden & Sokal (1986) and Sokal *et al.* (1987) and using the normalized Mantel statistics *rz* (Smouse *et al.* 1986). Tests of association between genetic divergence and geographical distances among populations were run by calculating a series of normalized Mantel correlations (*rz*)

where a set of weight matrices are analysed against a genetic divergence matrix. The set of weight matrices describes the geographical relationships between pairwise populations in terms of geographical distance classes to be tested (e.g. Torres *et al.* 2003; Fievet *et al.* 2007). Pairwise multilocus F_{ST}-values were used as a measure of genetic divergence among populations. Statistical significance of each *rz* value was tested by randomly permuting rows and columns of one matrix while keeping the other constant (10 000 permutations).

As an effect of spatially restricted gene flow between populations can be constrained in a nonisotropic way, we further performed two-dimensional Mantel correlogram analyses using the method of Oden & Sokal (1986). A directional spatial correlogram not only incorporates the geographical distance between pairwise localities but also the angular direction between them, therefore providing the scale as well as the direction over which a spatial process occurs (e.g. Sokal *et al.* 1987; Torres *et al.* 2003). Two-dimensional analyses were only applied to the complete data set of 42 populations because analysing the WP and FP populations independently would have resulted in an insufficient number of pairwise localities in each distance-class direction for statistically relevant analysis of normalized Mantel *rz* calculations. Overall significance of each one and two-dimensional correlograms was assessed using a Bonferroni test as suggested by Oden (1984). All correlogram calculations were performed using the software PASSAGE version 1.1.2.3 (Rosenberg 2001).

Results

Occurrence of CMS Owen cytoplasm within populations

The first survey of cytoplasmic diversity was designed to assess the within-population occurrence of the Owen CMS cytoplasm, which is universally used in sugar beet breeding programmes, but absent in wild beet populations. Of 42 sampled populations, 10 were characterized by the presence of Owen CMS cytoplasm (Table 1), suggesting a maternal cultivated origin for a total of 115 individuals of 2091. This cultivated cytoplasm occurred only in populations situated in the former seed production area, as can be shown in Fig. 1B (southern part of the seed production area). Frequencies of Owen CMS ranged from 0% to 44% in WP populations (mean: 2.5%) and from 0% to 100% in FP populations (mean 11.4%) (see Table 1). In terms of number of individuals carrying this cultivated cytoplasm, FP populations were characterized by a significantly higher incidence of the Owen CMS cytoplasm ($\chi^2_1 = 69.98$; $P < 0.0001$) compared with WP populations.

Genotypic linkage and Hardy–Weinberg disequilibrium

Among 21 possible pairs of microsatellite loci, three were characterized by an overall significant genotypic disequilibrium after Bonferroni correction, involving CAA1/BVM3, CAA1/CT4 and BVM3/CT4 (all at $P < 0.001$, Fisher's exact tests). Within populations, exact tests for genotypic linkage disequilibrium indicated only one unbiased, significant *P*-value of 882 comparisons after Bonferroni corrections (44 expected from type I error at $\alpha = 0.05$), involving the CAA1/CT4 pair within the population AJ. Altogether, these results suggested that there was no genotypic disequilibrium in the studied populations.

Mean F_{IS}-estimates per locus ranged from 0.043 to 0.166, all being significant at $P < 0.01$, with a mean multilocus estimate of 0.111 (±0.023) as shown in Table 2. Using multilocus probability tests, 27 of 43 populations were found to behave as expected under Hardy–Weinberg equilibrium (Table 1). As single-locus F_{IS}-values corresponding to significant heterozygote deficiency were not specific to any one locus or any one population (data not shown), observed heterozygote deficiencies cannot be attributed to a null allele. No significant differences in mean F_{IS}-values were found between WP and FP populations (Mann–Whitney *U*-test, $P = 0.86$, see Table 3). The same result held when comparing populations carrying the Owen CMS cytoplasm with populations without this cultivated cytoplasm ($P = 0.13$, Table 3). Overall, these results suggest either substructuring or inbreeding effects within WP and FP populations, irrespective of the occurrence of Owen CMS cytoplasm.

Nuclear and cytoplasmic diversity

Nuclear microsatellite loci displayed moderate to high numbers of alleles, ranging from 4 to 31 for the GCC1 and CAA1 loci respectively (Table 2). Mean observed heterozygosity (H_O) varied from 0.310 (locus GAA1) to 0.749 (locus BVM3) with an associated gene diversity H_E of 0.338 to 0.819 respectively. No significant differences in mean estimates of A_n, A_r, H_O and H_E were detected between WP and FP populations (Mann–Whitney *U*-test, all at $P > 0.05$, Table 3). As for mitochondrial genetic variation, strong differences in the distribution of allelic richness were detected between populations (Fig. 1C), with the minisatellite TR1 locus exhibiting from one to eight alleles per population for a total number of 10 alleles (Tables 1 and 2). However, we observed no significant differences between WP and FP populations in terms of mean number of alleles or allelic richness ($P = 0.14$ and 0.12 respectively; Table 3). Likewise, for both cytoplasmic and nuclear data, we

Table 2 Distribution of nuclear and cytoplasmic genetic diversity among 42 *Beta vulgaris* populations sampled within the seed production area in southwestern France, as measured by total number of alleles (A_n), allelic richness (A_r), observed heterozygosity (H_O), gene diversity (H_E) and *F*-statistics following Weir & Cockerham (1984)

	Allelic size range (bp)	A_n	A_r	H_O	H_E	F_{IS}	F_{IT}	F_{ST}
Nuclear loci								
GTT1	114–126	5	4.01	0.625	0.652	0.043 (0.016)**	0.107 (0.020)***	0.067 (0.013)***
GCC1	97–106	4	3.96	0.409	0.490	0.166 (0.025)***	0.218 (0.027)***	0.063 (0.013)***
GAA1	189–216	9	3.20	0.310	0.338	0.122 (0.025)***	0.178 (0.024)***	0.063 (0.011)***
CAA1	142–184	31	11.28	0.602	0.688	0.130 (0.017)***	0.191 (0.027)***	0.070 (0.008)***
BVM3	97–130	30	13.44	0.749	0.819	0.102 (0.015)***	0.163 (0.017)***	0.068 (0.007)***
CT4	141–161	15	10.63	0.612	0.759	0.175 (0.022)***	0.238 (0.020)***	0.076 (0.011)***
CA2	225–255	7	3.22	0.523	0.554	0.067 (0.019)***	0.114 (0.020)***	0.050 (0.012)***
Mean over all loci						0.111 (0.023)***	0.166 (0.024)***	0.061 (0.004)***
C.I. (95%)						0.071–0.153	0.125–0.209	0.055–0.067
Cytoplasmic locus								
TR1	439–794	10	7.56	—	—	—	—	0.359 (0.053)***

Mean *F*-estimates per locus (standard errors) and average *F*-indices (standard errors) are calculated by jackknifing among the samples and over loci respectively. C.I.: 95% confidence intervals for nuclear data are obtained by bootstrapping over loci. Allelic size ranges in base pairs (bp) are also indicated for each locus.
**$P < 0.01$, ***$P < 0.001$; —, not applicable (haploid data).

Table 3 Summary of results from tests for significant differences (Mann–Whitney *U*-tests) in mean genetic parameters in *Beta vulgaris* (i) when comparing wild populations (WP) vs. feral populations (FP) and (ii) when comparing populations with presence vs. absence of CMS Owen cytoplasm

	WP vs. FP populations Mean WP/mean FP/significance	Presence vs. absence of CMS Owen cytoplasm within populations Mean Owen CMS/mean non-Owen CMS/significance
Nuclear data		
H_O	0.54/0.55/NS	0.57/0.54/NS
H_E	0.61/0.62/NS	0.62/0.61/NS
A_n	5.89/6.13/NS	6.34/5.86/NS
A_r	5.67/5.90/NS	6.09/5.65/NS
F_{IS}	0.11/0.11/NS	0.09/0.11/NS
Mean nuclear pairwise F_{ST}	0.067/0.049/***	0.043/0.066/***
Cytoplasmic data		
A_n	3.53/4.42/ NS	4.2/3.72/ NS
A_r	3.48/4.38/ NS	4.15/3.66/ NS
Mean cytoplasmic pairwise F_{ST}	0.340/0.281/**	0.399/0.298/*

Results are indicated for the seven nuclear microsatellite loci (nuclear data) and for the TR1 minisatellite locus (cytoplasmic data).
NS, not significant; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

failed to detect significant differences in mean parameters of gene diversity between the 10 populations exhibiting the Owen CMS cytoplasm and those without this cultivated cytoplasm.

Genetic relationships of populations

Neighbour-joining (NJ) trees (not shown) showed no clear geographical pattern. All bootstrap values were below 50%, indicating that the NJ trees had low resolution. Furthermore, there was no concordance between the spatial location of populations and their genetic clustering, using either nuclear data or cytoplasmic polymorphism. Moreover, populations did not cluster according to population classification (WP or FP), and most FP populations were scattered throughout the NJ trees within clusters of WP populations. Likewise, populations did not cluster according to the occurrence of CMS Owen cytoplasm within populations.

Levels of genetic differentiation

For nuclear data, a moderate, but significant level of population genetic structure was found among populations, with a mean F_{ST}-estimate of 0.061 (±0.004; $P < 0.001$, Table 2). All single-locus F_{ST}-values were significantly different from zero and very similar, ranging from 0.050 (locus CA2) to 0.076 (locus CT4). As a result of this significant genetic differentiation for all nuclear loci, single-locus F_{IT}-estimates were all higher than 0.1, with a multilocus mean of 0.166 (±0.024; $P < 0.001$, Table 2). Using a hierarchical analysis of the partition of the genetic variance within (F_{SC}) and between (F_{CT}) groups of populations classified as WP or FP, we found significant genetic differentiation within groups ($F_{SC} = 0.06145$; $P < 10^{-6}$) but not between groups of WP and FP populations ($F_{CT} = -0.00063$; $P = 0.673$). This result indicates that the genetic structure was only

present within groups and corroborates the NJ tree results. All pairwise F_{ST}-values between the 42 populations were significant at $P < 0.05$. By further analysing the WP and FP population data sets independently, we found that mean nuclear pairwise F_{ST}-estimates among populations were significantly higher for WP populations compared with FP populations, mean estimates being equal to 0.067 and 0.049 respectively (Mann–Whitney *U*-test, $P < 0.0001$, Table 3).

Compared with nuclear data, the amount of genetic differentiation was considerably higher for cytoplasmic genetic data, with an overall significant F_{ST}-value of 0.359 (±0.053) for the mitochondrial TR1 locus (Table 2). Like nuclear data, we found significant cytoplasmic genetic differentiation within groups ($F_{SC} = 0.34678$; $P < 10^{-6}$), but not among groups of WP and FP populations ($F_{CT} = 0.01486$; $P = 0.111$). Likewise, mean cytoplasmic pairwise F_{ST}-estimates were significantly higher for WP populations compared with FP populations (Mann–Whitney *U*-test, $P = 0.0051$, Table 3).

In addition, we tested whether there was an effect of the occurrence of Owen CMS cytoplasm on genetic differentiation between pairwise populations, regardless of WP and FP classification. Mean nuclear pairwise F_{ST}-values were significantly lower for populations carrying the Owen CMS cytoplasm compared with populations without this cultivated cytoplasm (mean pairwise F_{ST}-values of 0.043 and 0.066 respectively, $P < 0.0001$, Table 3). In contrast, the opposite result was obtained for cytoplasmic data: populations carrying the Owen CMS cytoplasm were significantly more differentiated at the TR1 locus compared with populations where this cultivated cytoplasm did not occur (mean pairwise F_{ST}-values of 0.399 and 0.298 respectively, $P = 0.012$). This result can be explained by the large variance in frequencies of Owen CMS cytoplasm.

Spatial genetic structure

Despite significant genetic differentiation between populations, Mantel tests and one-dimensional Mantel correlograms failed to depict a continuous decrease in genetic similarity with geographical isolation of populations as would be expected under an isolation-by-distance process at migration–drift equilibrium (Table S1). Only significant short-distance spatial structuring was detected for nuclear data (Fig. 2A). The outline of the Mantel correlogram for cytoplasmic data did not indicate statistically significant spatial structuring for any distance class, except one significantly negative *rz* value for populations separated by around 43 km (Fig. 2C). A cutoff can be observed beyond 43 km for both nuclear and cytoplasmic data; beyond this distance, population allelic frequencies were clearly independent of each other and reflected the effect of random genetic drift rather than gene flow. Likewise, two-dimensional Mantel correlograms failed to detect obvious spatial trends at both nuclear and mitochondrial loci (Fig. 2B, D). Except a slight short-distance structuring in a northern–southern direction, windrose correlograms were characterized by nonsignificant, positive or negative, *rz*-values that were independent of angular direction or geographical distance. Only 6 of 80 (7.5%) *rz*-values were significant at $P < 0.05$. After Bonferroni correction, none of the Mantel correlograms yielded an overall significance at $P < 0.05$. When analysing the WP and FP populations separately, the one-dimensional Mantel correlograms did not illustrate any substantial spatial patterns (Table S1). However, a strong isolation-by-distance pattern was found for cytoplasmic variation after removing the populations that exhibited the Owen CMS cytoplasm. This result again suggests that the occurrence of the Owen CMS cytoplasm varies greatly among populations (Table S1).

Discussion

Seed escape from crop to inland wild beet populations: evidence from cytoplasmic diversity

Seeds have the potential to travel over long distances and remain viable for longer periods than pollen. This gives rise to a long-lived seed bank, which can act as an efficient pathway for crop-to-wild gene flow (Reagon & Snow 2006; Hills *et al*. 2007). In this study, we showed that 115 plants of 2091 carried the Owen CMS cytoplasm, i.e. a cultivated germplasm, and can therefore be considered as having a cultivated maternal parent originating from seeds lost during harvest. Beyond this evidence of crop-to-wild seed escape, two main conclusions can be drawn from the distribution of Owen CMS cytoplasm among the surveyed populations: (i) this cytoplasm did not occur in the present seed production area, and feral populations were absent from the vicinity of seed production fields. This implies that careful harvesting and efficient removal of wild beet individuals within 1000 m around the beet seed production fields, as practised by breeders, are very effective and reduce the opportunities for seed escape and the occurrence of undesirable feral populations; (ii) However, in the former beet seed production area, where 14 field margin (FP) populations were found, Owen CMS cytoplasm occurred at higher rates than in wild (WP) populations. The occurrence of Owen CMS cytoplasm in FP populations clearly reflected the establishment of cultivated beet seeds that escaped during harvesting; the remaining individuals characterized by a wild-type cytoplasm were the result of the long-lived

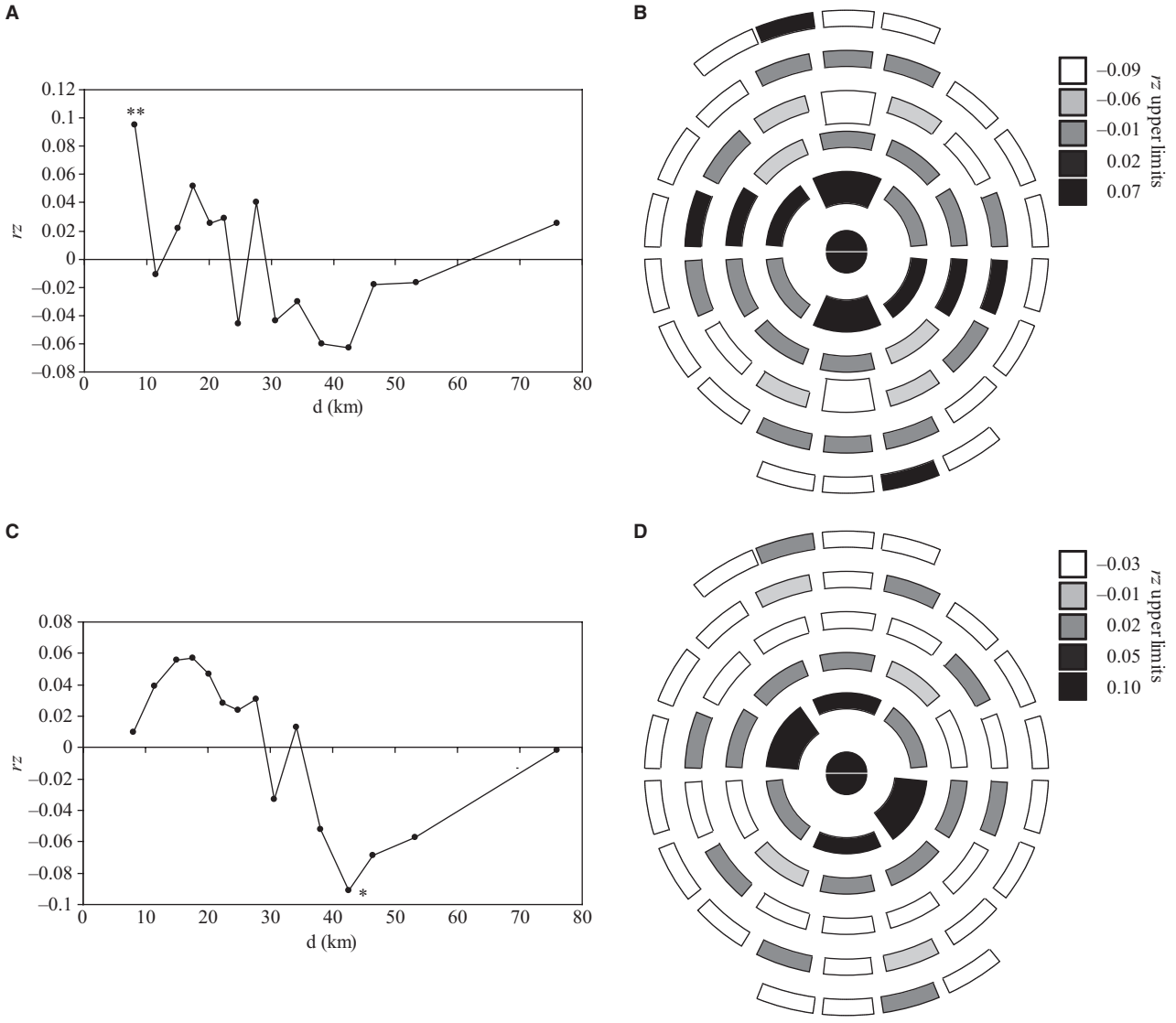

Fig. 2 Spatial genetic variation for the 42 populations of *Beta vulgaris* sampled within the French seed production area. One-dimensional Mantel correlograms for patterns of variation of F_{ST}-estimates against geographical distances (km) between populations using the information from seven nuclear microsatellite loci (A) and from the mitochondrial minisatellite locus (C). A minimum number of 57 pairs of populations occurred in each distance class for one-dimensional Mantel correlograms. *rz*: normalized Mantel statistic; $*P < 0.05$, $**P < 0.01$. Two-dimensional (windrose) Mantel correlograms for pattern of variation of F_{ST}-estimates against geographical distances (km) between populations using the information from seven nuclear microsatellite loci (B) and from the mitochondrial minisatellite locus (D). The lower half of the correlogram is symmetrical to the upper half. Sectors of windrose correlograms represent distance–direction classes. The six circular and successive annuli represent six interval distance classes with upper limits of 11, 20, 31, 44, 59 and 76 km respectively. Full-width boxes correspond to significant *rz*-values at $P < 0.05$, half-width boxes represent nonsignificant *rz*-values. Boxes with dashed outlines are distance–direction classes based on an insufficient number of population pairs (<15) and are not considered for spatial analyses. Shading represents approximate quintiles of the range of values taken by the normalized Mantel statistic *rz*.

seed bank of wild populations, which had been destroyed when the seed production was set up in the area. With regard to wild beet populations found within villages, the presence of Owen CMS cytoplasm could be partly explained by the fact that landraces, such as red beet or Swiss chard, are frequently grown in kitchen gardens for consumption and/or seed production (Desplanque *et al.* 2002). If, nonetheless, this Owen CMS cytoplasm reflected cultivated lines originating from the former seed production fields, it may

involve long-distance seed dispersal events because of human activities. Such human-mediated, long-distance dispersal can be frequent in cropping systems through seed loss by agricultural machines during harvesting and by spillage during postharvest transport (von der Lippe & Kowarik 2007).

Level of nuclear and cytoplasmic genetic diversity in inland wild beets

In the beet seed production area, of primary interest is the likelihood of genetic swamping due to crop-to-wild gene flow, which can lead to displacement of native genetic diversity when hybridization is asymmetric owing to a difference in population size (Ellstrand & Elam 1993). In *Beta vulgaris*, regional or very local studies suggest that introgression of cultivated beet genes in weed or wild populations appear to have no more impact than enhancing local genetic diversity in the weed or wild gene pool (Bartsch *et al.* 1999; Viard *et al.* 2002). Similarly, in *Zea mays*, despite several decades of gene flow from modern cultivars into maize Italian landraces, no genetic erosion has been found to date (Bitocchi *et al.* 2009). In this study on inland wild beet populations, levels of genetic diversity at nuclear loci were very similar to those of wild beet populations located along the coastline far from crop fields (see Viard *et al.* 2004; Fievet *et al.* 2007; Fénart *et al.* 2008).

Moreover, beet populations along field margins (FP), which were presumably exposed to pollen flow from cultivars, did not significantly differ from wild populations (WP) in their level of nucleo-cytoplasmic genetic diversity, even when accounting for the occurrence of Owen CMS cytoplasm. Furthermore, no substantial genetic divergence was observed between FP and WP populations, nor between populations carrying or not carrying the Owen CMS cytoplasm. The salient result from this study lies in the absence of genetic signatures of introgression of nuclear crop genes within the surveyed populations. Using Bayesian analysis of population memberships based on multilocus microsatellite genotypes (see Beaumont & Rannala 2004) and after extensive runs of Monte Carlo Markov Chain replicates, we failed to detect any clearly admixed individuals or hybridizing populations, even when accounting for the Owen CMS occurrence as a prior for individual membership probability (data not shown). Our study on inland wild beets therefore contrasts with previous results that depicted a clear genetic distinctiveness between weedy crop–wild hybrid lineages and coastal wild beet populations (Arnaud *et al.* 2003; Cuguen *et al.* 2004; Viard *et al.* 2004).

Several explanations can be offered for these findings. First, our study lacks past and present cultivars as referents of crop genetic diversity, which could be used to quantify introgression events through Bayesian and admixture analyses (Breton *et al.* 2008; Bitocchi *et al.* 2009). The limited number of microsatellite loci used, the absence of any crop-specific alleles (see Arnaud *et al.* 2003; Viard *et al.* 2004) and the use of genetically diverse paternal lines selected by breeders for seed production (McGrath *et al.* 1999) may also limit identification of nuclear introgression from crops to wild populations *via* feral beets. Second, changes in the genetic diversity of wild populations would also, however, mainly depend on whether the crop and wild populations were initially genetically distinct (Ellstrand 2003b). The evolutionary origin of ruderal beets, still unclear and unresolved, may explain the lack of significant detection of introgressants during the screening. Ruderal beets are either viewed as domesticated beets that escaped from private gardens in the Middle Ages, and/or that have escaped from cultivated fields since the 1960s, and associated with wild beets originating from Mediterranean coast (Desplanque *et al.* 1999). As a result, compared with Atlantic wild beet relatives, which have a distinct evolutionary trajectory, genetic distinctiveness between ruderal and cultivar taxa is very low (Fénart *et al.* 2008). Third, even with such low genetic divergence, the weed populations found within sugar beet fields and resulting from hybridization between crops and ruderal wild pollen donors are, nonetheless, clearly genetically differentiated from coastal and inland wild beet populations (Arnaud *et al.* 2003; Fénart *et al.* 2008). Therefore, at least pure cultivars or first-year crop–wild hybrids would have been detected, even at a low rate of introgression and even without cultivar referents. Furthermore, cultivars and crop–wild hybrids are morphologically very different from pure wild plants (Van Dijk 2004). However, we did not detect any cultivated traits (such as thick roots or large leaves) in the collected individuals during the sampling.

Taken together, these observations suggest old introgressions of cultivated cytoplasm in a nuclear wild background, involving crop-to wild gene flow in the past through crop seed escape rather than through direct, recent pollen flow. This conclusion was also reached by Desplanque *et al.* (2002) and Van Dijk (2004) using field, morphological and flow cytometry measurements. Assuming a crop seed escape dating from at least more than 10 years ago in the former seed production area, the dilution of nuclear crop genes could result from several rounds of recombination through backcrossing via pollen from surrounding inland wild individuals. Likewise, Alibert *et al.* (2005) showed that crop and wild ruderal pollen effectively compete, thereby impeding the emergence of numerous crop-hybrid

progeny (see also Fénart *et al.* 2007, for density-level effect of pollen saturation in beets).

Contrasted patterns of differentiation but lack of spatial structuring for nuclear and cytoplasmic genetic diversity

A significant, but moderate overall genetic differentiation (F_{ST} = 0.061) was depicted among the 42 populations at nuclear loci. This low level of genetic differentiation may result from recent evolutionary history and a common ancestry of inland wild populations during colonization and/or efficient historical gene flow via pollen dispersal. However, in contrast to the spatial genetic structure commonly observed in coastal wild beet populations (e.g. Raybould *et al.* 1996; Viard *et al.* 2004; Fievet *et al.* 2007), we found no clear isolation-by-distance pattern among inland beet populations. Fievet *et al.* (2007) demonstrated that genetic connectedness among wild beet populations was primarily because of pollen flow strictly confined along the open habitat of the shoreline because of inland barriers to pollen dispersal. In our study area, pollen flow may be contained by topography including villages, forests, hills and valleys. This containment may therefore explain why gene flow did not continuously decrease with geographical distance.

Stronger spatial genetic structure is expected for maternally inherited mitochondrial genes that are dispersed through seed flow compared with nuclear genes that are dispersed by both pollen and seed flow (Ennos 2001). In *B. vulgaris*, seeds are agglomerated in a corky seedball that can contain up to eight seeds primarily dispersed by gravity. Therefore, the unit of dispersal is a multiseeded fruit, in which significant proportions of the seeds may have strong level of correlated paternity as well as sharing the same mother plant. Such kin-structured seed dispersal provides further opportunities for strong genetic differentiation when the group dispersal is composed of related seeds within fruits (Torimaru *et al.* 2007). Hence, inland wild populations were strongly genetically differentiated for cytoplasmic variation (F_{ST} = 0.359) and followed an isolation-by-distance pattern when populations carrying the Owen CMS cytoplasm were removed from the analyses. These findings provide further evidence that accidental crop seed escape can lead to a departure from migration–drift equilibrium and make it difficult to identify spatially sensitive genetic structure in wild populations. Overall, this assemblage of wild beet populations resembles that of a metapopulation, with local founding events involving a small number of sib groups, followed by low genetic connectedness via seed flow after population establishment.

Implications for crop management and for the evolution of life-history traits inherited from the crop and the wild

Our results highlight the need to carefully survey and manage wild populations with respect to feral populations. No field-margin feral populations were observed in the present beet seed production area, and we found no evidence of crop seed escape in wild populations. This reflects the extreme precautions that breeders take to avoid seed escape and to eradicate all wild or weedy beets in the vicinity of seed production fields. However, because of laxer practices in previous years, the persistence of maternally inherited crop genes is expected within the former beet seed production fields, thereby increasing exposure to wild populations. This may have at least two consequences. First, *B. vulgaris* is a gynodioecious species, referring to a breeding system in which female and hermaphrodites co-occur within populations. It has recently been shown that variation in the sex-ratio is likely to be driven by frequency-dependent selection in wild beet populations (Dufay *et al.* 2009). Therefore, the introduction of a new cultivated male-sterilizing cytoplasm (Owen CMS) in the wild is likely to modify the evolutionary dynamics of gynodioecy in this species. Second, transgenic beet varieties resistant to total herbicide are not expected to be selected for outside agrosystems, especially in more natural coastal populations. However, ruderal wild beets are mostly encountered in human-disturbed habitats (canal banks, irrigation and drainage ditches, garden edges, roadsides or on the edges of parking lots) where herbicides may be employed. As such, escape of crop seed containing any transgenes for herbicide resistance would be of immediate concern in this area. As a case in point, Warwick *et al.* (2008) demonstrated the persistence and spread of a herbicide-resistance trait within a weedy population of *Brassica rapa* over a 6-year period, even in the absence of herbicide selection pressure and in spite of a fitness cost related to hybridization (see also Zapiola *et al.* 2008).

Crop–wild gene flow also provides opportunities for studying the potential role of hybridization in generating novel genetic variation that may facilitate adaptation of plants to edge-of-the-range habitats or in agrosystems. From an evolutionary point of view, fitness effects of crop genes following their transfer into a wild genetic background may not be straightforward (Sørensen *et al.* 2007; Campbell & Snow 2009). For instance, it has recently been shown in sunflower and radish that under competitive conditions, the relative fitness of crop–wild F1s or more advanced-generation hybrids was greater relative to wild plants (Campbell & Snow 2007). These findings highlight the need to study the fate of weedy crop–wild hybrid beet lineages,

whose traits of seed dormancy and earliness of flowering inherited from the wild gene pool pre-adapt them for invasive success in disturbed man-made and agricultural habitats (Ellstrand & Schierenbeck 2000; Hautekèete *et al.* 2002). On the other hand, farming practices also can have profound effects on the evolution of life-history traits related to the mating system (e.g. Guillemin *et al.* 2008; Campbell & Snow 2009). In this regard, we found that 16 of 43 wild populations surveyed in this study departed from genotypic equilibrium expected under Hardy–Weinberg assumptions, whereas coastal wild beet populations rarely show departures from equilibrium (Raybould *et al.* 1996; Fievet *et al.* 2007). Beyond inbreeding effects because of population structure, one explanation could be related to the introgression of crop genes involving a Mendelian self-fertility factor widely used in cultivated germplasm to produce inbred lines, the gene *Sf* (Cuguen *et al.* 2004). This hypothesis remains speculative in ruderal wild beets, as we did not depict any significant differences in mean F_{IS}-values between WP populations and FP populations, nor between populations characterized by the presence of the Owen CMS cytoplasm and populations lacking this cultivated cytoplasm. However, the combination of domesticated (breakdown of self-incompatibility) and wild (earliness of flowering) life-history traits may facilitate adaptive evolution in weedy lineages found within sugar beet fields (Ellstrand 2003b; Van Dijk 2004). We are currently investigating whether these two key life-history traits are positively selected for in weed beets living in agrosystem-like conditions (S. Fénart & J.-F. Arnaud, unpublished data).

Acknowledgements

We wish to express our gratitude to M. Delescluse and the FNAMS for advice and help in sampling populations in the field and to A. Petit and L. Despinoy for technical assistance in molecular analyses. We are also grateful to C.R. Engel, H. Van Dijk, X. Vekemans, A. Widmer and five anonymous referees for critical reading and providing very helpful comments and suggestions on previous version of the manuscript. J.-F. Arnaud is grateful to the CNRS for supporting him as a full-time research scientist during the 2008–2009 academic year. S. Fénart was supported by an 'INRA/Région Nord-Pas-de-Calais' fellowship. This work was funded by the ACI 'Impact des OGM', the 'Contrat de Plan État/Région Nord-Pas-de-Calais' and by a CNRS grant 'Appel à proposition, Impact des Biotechnologies dans les Agro-écosystèmes'.

References

Alibert B, Sellier H, Souvré A (2005) A combined method to study gene flow from cultivated sugar beet to ruderal beets in the glasshouse and open field. *European Journal of Agronomy*, **23**, 195–208.

Arnaud J-F, Viard F, Delescluse M, Cuguen J (2003) Evidence for gene flow via seed dispersal from crop to wild relatives in *Beta vulgaris* (Chenopodiaceae): consequences for the release of genetically modified crop species with weedy lineages. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **270**, 1565–1571.

Bartsch D, Lehnen M, Clegg J *et al.* (1999) Impact of gene flow from cultivated beet on genetic diversity of wild sea beet populations. *Molecular Ecology*, **8**, 1733–1741.

Beaumont AR, Rannala B (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics*, **5**, 251–261.

Bitocchi E, Nanni L, Rossi M *et al.* (2009) Introgression from modern hybrid varieties into landrace populations of maize (*Zea mays* ssp. *mays* L.) in central Italy. *Molecular Ecology*, **18**, 603–621.

Breton C, Pinatel C, Médail F, Bonhomme F, Bervillé A (2008) Comparison between classical and Bayesian methods to investigate the history of olive cultivars using SSR-polymorphisms. *Plant Science*, **175**, 524–532.

Campbell LG, Snow AA (2007) Competition alters life history and increases the relative fecundity of crop–wild radish hybrids (*Raphanus* spp.). *New Phytologist*, **173**, 648–660.

Campbell LG, Snow AA (2009) Can feral weeds evolve from cultivated radish (*Raphanus sativus*, Brassicaceae)? *American Journal of Botany*, **96**, 498–506.

Campbell LG, Snow AA, Ridley CE (2006) Weed evolution after crop gene introgression: greater survival and fecundity of hybrids in a new environment. *Ecology Letters*, **9**, 1198–1209.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics*, **19**, 233–257.

Chapman MA, Burke JM (2006) Letting the gene out of the bottle: the population genetics of genetically modified crops. *New Phytologist*, **170**, 429–443.

Cuguen J, Wattier R, Saumitou-Laprade P *et al.* (1994) Gynodioecy and mitochondrial DNA polymorphism in natural populations of *Beta vulgaris* ssp *maritima*. *Genetics, Selection & Evolution*, **26**, S87–S101.

Cuguen J, Arnaud J-F, Delescluse M, Viard F (2004) Crop-wild interaction within the *Beta vulgaris* complex: a comparative analysis of genetic diversity between sea beet and weed beet populations within the French sugarbeet production area. In: *Introgression from Genetically Modified Plants into Wild Relative* (eds Den Nijs HCM, Bartsch D, Sweet J), pp. 183–201. CABI Publishers Inc., Oxfordshire, UK.

Desplanque B, Boudry P, Broomberg K *et al.* (1999) Genetic diversity and gene flow between wild, cultivated and weedy forms of *Beta vulgaris* L. (Chenopodiaceae), assessed by RFLP and microsatellite markers. *Theoretical and Applied Genetics*, **98**, 1194–1201.

Desplanque B, Hautekèete N-C, Van Dijk H (2002) Transgenic weed beets: possible, probable, avoidable? *Journal of Applied Ecology*, **39**, 561–571.

Dufay M, Cuguen J, Arnaud J-F, Touzet P (2009) Sex ratio variation among gynodioecious populations of sea beet: can it be explained by negative frequency-dependent selection? *Evolution*, **63**, 1483–1497.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among population of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, **92**, 832–839.

Ellstrand NC (2003a) Current knowledge of gene flow in plants: implications for transgene flow. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, **358**, 1163–1170.

Ellstrand NC (2003b) *Dangerous Liaisons? When Cultivated Plants Mate with Their Wild Relatives*. The Johns Hopkins University Press, Baltimore, MA.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology & Systematics*, **24**, 217–242.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences, USA*, **97**, 7043–7050.

Ennos RA (2001) Inferences about spatial processes in plant populations from the analysis of molecular markers. In: *Integrating Ecology and Evolution in a Spatial Contex* (eds Silvertown J, Antonovics J), pp. 45–71. Blackwell Science Ltd, Oxford, UK.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Fénart S, Austerlitz F, Cuguen J, Arnaud J-F (2007) Long distance pollen-mediated gene flow at a landscape level: the weed beet as a case study. *Molecular Ecology*, **16**, 3801–3813.

Fénart S, Arnaud J-F, De Cauwer I, Cuguen J (2008) Nuclear and cytoplasmic genetic diversity in weed beet and sugar beet accessions compared to wild relatives: new insights into the genetic relationships within the *Beta vulgaris* complex species. *Theoretical and Applied Genetics*, **116**, 1063–1077.

Fievet V, Touzet P, Arnaud J-F, Cuguen J (2007) Spatial analysis of nuclear and cytoplasmic DNA diversity in wild sea beet (*Beta vulgaris* ssp. *maritima*) populations: do marine currents shape the genetic structure? *Molecular Ecology*, **16**, 1847–1864.

Gepts P, Papa R (2003) Possible effects of (trans)gene flow from crops on the genetic diversity from landraces and wild relatives. *Environmental Biosafety Research*, **2**, 89–103.

Goudet J (1995) FSTAT (Version 1.2). a computer program to calculate F-Statistics. *Journal of Heredity*, **86**, 485–486.

Goudet J, Raymond M, De Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.

Guillemin M-L, Faugeron S, Destombe C *et al.* (2008) Genetic variation in wild and cultivated populations of the haploid-diploid red alga *Gracilaria chilensis*: how farming practices favor asexual reproduction and heterozygosity. *Evolution*, **62**, 1500–1519.

Hautekèete N-C, Piquot Y, Van Dijk H (2002) Life span in *Beta vulgaris* ssp. *maritima*: the effects of age at first reproduction and disturbance. *Journal of Ecology*, **90**, 508–516.

Hills MJ, Hall L, Arnison PG, Good AG (2007) Genetic use restriction technologies (GURTs): strategies to impede transgene movement. *Trends in Plant Science*, **12**, 177–183.

Letschert JPW (1993) *Beta* section *Beta*: biogeographical patterns of variation and taxonomy. *Wageningen Agricultural University Papers*, **93**, 1–137.

von der Lippe M, Kowarik I (2007) Crop seed spillage along roads: a factor of uncertainty in the containment of GMO. *Ecography*, **30**, 483–490.

McGrath JM, Derrico CA, Yu Y (1999) Genetic diversity in selected, historical US sugarbeet germplasm and *Beta vulgaris* ssp. *maritima*. *Theoretical and Applied Genetics*, **98**, 968–976.

Mörchen M, Cuguen J, Michaelis G, Hanni C, Saumitou-Laprade P (1996) Abundance and lenght polymorphism of microsatellite repeats in *Beta vulgaris* L. *Theoretical and Applied Genetics*, **92**, 326–333.

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

Nishizawa S, Kubo T, Mikami T (2000) Variable number of tandem repeat loci in the mitochondrial genomes of beets. *Current Genetics*, **37**, 34–38.

Oden NL (1984) Assessing the significance of a spatial correlogram. *Geographical Analysis*, **16**, 1–16.

Oden NL, Sokal RR (1986) Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Zoology*, **35**, 608–617.

Owen FV (1942) Inheritance of cross- and self-sterility and self-fertility in *Beta vulgaris*. *Journal of Agricultural Research*, **64**, 679–698.

Owen FV (1945) Cytoplasmically inherited male-sterility in sugar beets. *Journal of Agricultural Research*, **71**, 423–440.

Ran Z, Michaelis G (1995) Mapping of a chloroplast RFLP marker associated with the CMS cytoplasm of sugar beet (*Beta vulgaris*). *Theoretical and Applied Genetics*, **91**, 836–840.

Raybould AF, Goudet J, Mogg RJ, Gliddon CJ, Gray AJ (1996) Genetic structure of a linear population of *Beta vulgaris* ssp. *maritima* (sea beet) revealed by isozyme and RFLP analysis. *Heredity*, **76**, 111–117.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): a population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reagon M, Snow AA (2006) Cultivated *Helianthus annuus* (Asteraceae) volunteers as a genetic ''bridge'' to weedy sunflower populations in North America. *American Journal of Botany*, **93**, 127–133.

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

Rosenberg MS (2001) *Pattern Analysis, Spatial Statistics, and Geographic Exegesis*. Department of Biology, Arizona State University, Tempe, AZ.

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627–632.

Sokal RR, Oden NL, Barker JSF (1987) Spatial structure in *Drosophila buzzatii* populations: simple and directional spatial autocorrelation. *The American Naturalist*, **129**, 122–142.

Sørensen BS, Kiaer LP, Jørgensen RB, Hauser TP (2007) The temporal development in a hybridizing population of wild and cultivated chicory (*Cichorium intybus* L.). *Molecular Ecology*, **16**, 3292–3298.

Stewart CN Jr, Halfhill MD, Warwick SI (2003) Transgene introgression from genetically modified crops to their wild relatives. *Nature Reviews Genetics*, **4**, 806–817.

Takezaki N, Nei M (2008) Empirical tests of the reliability of phylogenetic trees constructed with microsatellite DNA. *Genetics*, **178**, 385–392.

Torimaru T, Tani N, Tsumura Y, Nishimura N, Tomaru N (2007) Effects of kin-structured seed dispersal on the genetic

structure of the clonal dioecious shrub *Ilex leucoclada*. *Evolution*, **61**, 1289–1300.

Torres E, Iriondo JM, Escudero A, Pérez C (2003) Analysis of within-population spatial genetic structure in *Antirrhinum microphyllum* (Scrophulariaceae). *American Journal of Botany*, **90**, 1688–1695.

Van Dijk H (2004) Gene exchange between wild and crop in *Beta vulgaris*: how easy is hybridization and what will happen in later generations? In: *Introgression from Genetically Modified Plants into Wild Relatives and its Consequence* (eds Den Nijs HCM, Bartsch D, Sweet J), pp. 53–69. CABI publishers, Inc., Oxfordshire, UK.

Viard F, Bernard J, Desplanque B (2002) Crop-weed interactions in the *Beta vulgaris* complex at a local scale: allelic diversity and gene flow within sugar beet fields. *Theoretical and Applied Genetics*, **104**, 688–697.

Viard F, Arnaud J-F, Delescluse M, Cuguen J (2004) Tracing back seed and pollen flow within the crop-wild *Beta vulgaris* complex: genetic distinctiveness versus hot spots of hybridization over a regional scale. *Molecular Ecology*, **13**, 1357–1364.

Warwick SI, Légère A, Simard M-J, James T (2008) Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy Brassica rapa population. *Molecular Ecology*, **17**, 1387–1395.

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

Zapiola ML, Campbell CK, Butler MD, Mallory-Smith CA (2008) Escape and establishment of transgenic glyphosate-resistant creeping bentgrass *Agrostis stolonifera* in Oregon, USA: a 4-year study. *Journal of Applied Ecology*, **45**, 486–494.

All the authors are involved in population genetic studies of plant species, with a special interest in the evolution of mating system in plants and in its consequences on patterns of spatial genetic structure. More information about the activities of the 'Laboratoire de Génétique et Evolution des Populations Végétales' can be found in the following web site: http://www.univ-lille1.fr/gepv/.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1 Overall correlations (*rz*) of genetic and geographical distances and Mantel correlograms among populations of *Beta vulgaris*

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