Molecular Ecology (2004) **13**, 1357–1364 doi: 10.1111/j.1365-294X.2004.02150.x

Tracing back seed and pollen flow within the crop–wild *Beta vulgaris* complex: genetic distinctiveness vs. hot spots of hybridization over a regional scale

FRÉDÉRIQUE VIARD,* JEAN-FRANÇOIS ARNAUD,† MAXIME DELESCLUSE† and JOËL CUGUEN†
UMR CNRS 7127, Laboratoire EGPM, Université de Paris 6, Station Biologique, Place Georges Teissier, BP 74, 29682 Roscoff cedex, France, †UMR CNRS 8016, Laboratoire de Génétique et Évolution des Populations Végétales, Bâtiment SN2, Université de Lille 1, 59655 Villeneuve d'Ascq cedex, France

Abstract

Hybrids between transgenic crops and wild relatives have been documented successfully in a wide range of cultivated species, having implications on conservation and biosafety management. Nonetheless, the magnitude and frequency of hybridization in the wild is still an open question, in particular when considering several populations at the landscape level. The *Beta vulgaris* complex provides an excellent biological model to tackle this issue. Weed beets contaminating sugar beet fields are expected to act as a relay between wild populations and crops and from crops-to-crops. In one major European sugar beet production area, nine wild populations and 12 weed populations were genetically characterized using cytoplasmic markers specific to the cultivated lines and nuclear microsatellite loci. A tremendous overall genetic differentiation between neighbouring wild and weed populations was depicted. However, genetic admixture analyses at the individual level revealed clear evidence for gene flow between wild and weed populations. In particular, one wild population displayed a high magnitude of nuclear genetic admixture, reinforced by direct seed flow as evidenced by cytoplasmic markers. Altogether, weed beets were shown to act as relay for gene flow between crops to wild populations and crops to crops by pollen and seeds at a landscape level.

Keywords: *Beta vulgaris*, biosafety, cpDNA, crop–wild gene flow, hybridization, microsatellites, risk assessment

Received 27 October 2003; revision received 16 January 2004; accepted 16 January 2004

Introduction

Seeds and pollen are the vehicles of genes in plant species and their transport is the only way for sessile plants to overcome local genetic differentiation and adaptation. The relationships between geographical isolation, life-history traits, local adaptation and gene flow are critical issues when examining crop–weed complexes. Ellstrand *et al*. (1999) demonstrated that hybridization occurs between crop species and their wild relatives in 12 of 13 worldwide crops. Possible consequences of this gene flow include a decrease of the genetic diversity of wild populations and the evolution of more aggressive weeds. This has important implications for the management and conservation of genetic resources (Frankham *et al*. 2002), as well as for biosafety studies such as the likelihood of transgene escape from crop to wild relatives (Jarvis & Hodgkin 1999; Bartsch & Schuphan 2002). Weeds may be a potential way for the escape of transgenes through gene flow between cross-compatible taxa (Gray & Raybould 1998; Ellstrand 2003). However, the magnitude and the frequencies of hybridization and introgression in the field is still an open question, especially when considering several populations within a landscape (Hails 2000; Papa & Gepts 2003; Wilkinson *et al*. 2003). If wild individuals, crops and their weedy relatives are found over a narrow geographical area, overlap in flowering period, share pollinators and are cross-compatible, then hybridization events have a strong potential to continuously introduce transgenes into weedy and/or wild populations (Gray & Raybould 1998; Bartsch *et al*. 2003; Gepts & Papa 2003). Concerning the potential

Correspondence: Jean-François Arnaud.
E-mail: jean-francois.arnaud@univ-lille1.fr

effects of gene flow and introgression between domesticated plants and their wild relatives, the *Beta vulgaris* complex is of immediate concern as wild, crop and weedy forms of the same species are fully cross-compatible and can coexist in close parapatry in Europe (Letschert 1993; Bartsch *et al*. 1999; Desplanque *et al*. 1999).

In Western Europe, two subspecies of *B. vulgaris* are recognized: *B. vulgaris* ssp. *maritima* Arcang, the typical wild and coastal subspecies ('sea-beet') and *B. vulgaris* ssp. *vulgaris*, the cultivated beet (Letschert 1993). Sugar beets, extensively cultivated in northern France, are biennial and harvested before bolting and flowering. Within sugar beet fields, noncultivated bolting individuals are commonly found contaminating the fields and competing with cultivated individuals. These individuals are conspecific weeds that have been proved to arise from hybridization between cultivated lines and wild individuals in the seed production areas (southern France and eastern Italy, see Boudry *et al*. 1993; Mücher *et al*. 2000; Desplanque *et al*. 2002; Viard *et al*. 2002; Bartsch *et al*. 2003). Hence, gene flow from transgenic cultivars into surrounding weed individuals is likely to lead to more problematic weeds, in particular for traits such as herbicide resistance (Viard *et al*. 2002).

Weed beets might also play a major role (1) in the escape of genes by acting as a bridge between crops and wild relatives (Desplanque *et al*. 2002) and (2) in the spread of genes between sugar beet fields, a factor not yet well studied as stressed recently by Ellstrand (2003). In northern France, wild populations (i.e. sea beet *B. vulgaris maritima*) are coastal, located along estuaries, just at the upper level of the tide. In a previous study (Arnaud *et al*. 2003) we examined the likelihood for gene flow from weed to wild beets over a very restricted geographical scale, i.e. along a river bank characterized by a continuum between crop, weedy and wild individuals (Arnaud *et al*. 2003). Our results suggested that weed seeds dispersed and colonized the river-bank, creating a connection between cultivated and coastal areas. In contrast, we did not find strong evidence for pollen flow between weed and wild individuals. Non-overlapping flowering period and rapid extinction of the weed population were invoked to explain these results. However, the extent to which this particular situation is representative of the global dynamics of wild–weed gene exchanges in the sugar beet production area remains to be investigated. To address this issue, we examined the genetic differentiation between wild and weed beets at both populational and individual levels over a regional scale, thereby encompassing events on a longer time scale than previously studied. In order to trace back contemporary migration events and detect rare introgression events, we used a Bayesian admixture analysis to assign each individual into two clusters of wild and weedy lineages. This allowed us to analyse the extent of gene flow and introgression that occur over a landscape (i.e. at a regional scale) within the *B. vulgaris* complex in the light of the likelihood of transgene escape from genetically engineered sugar beet crops to wild beet populations.

Materials and methods

Species and sampling

Beta vulgaris ssp. *maritima* is a diploid short-lived perennial species ($2n = 18$) distributed widely around the Mediterranean Basin and along the coasts of Western Europe. In the North of Europe, sea beets colonize areas located along estuaries, just at the upper level of the tide and, more rarely, cliffs overhanging the sea. Northern France is a major area of sugar beet (*B. vulgaris* ssp. *vulgaris*) production and sugar beet fields are distributed in the vicinity of the coastal wild populations (*B. vulgaris* ssp. *maritima*). In many sugar beet fields, weed beets infest sugar beet fields (Boudry *et al*. 1993; Viard *et al*. 2002; Arnaud *et al*. 2003; Bartsch *et al*. 2003). Nine wild sea beet populations (309 individuals) were sampled together with 12 populations of weed beets (494 individuals sampled outside the sowing line, see Viard *et al*. 2002), located in sugar beet fields as close as possible from the coastline (Fig. 1, Table 1). The distance between wild and weed populations ranged between one and 10 km. This sampling illustrates the actual distribution of the cultivated and wild species in Northern France.

Genetic data collection

DNA was extracted and purified from dried leaves by using a DNeasy® 96 Plant Kit and following the standard protocol for isolation of DNA from plant leaf tissue outlined in the DNeasy 96 Plant protocol handbook (Qiagen Inc.).

The maternal cultivated origin was assessed by means of a chloroplastic polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) marker defined by Ran & Michaelis (1995). Weedy lineages carry a chloroplastic haplotype which is characteristic of cultivated lines, i.e. the Owen cytoplasmic male (CMS) sterility used universally in sugar beet breeding programmes (Owen 1945; Boudry *et al*. 1993). *Hin*dIII fragments digests of cpDNA from Owen CMS and nonOwen CMS lines are, respectively, characterized by one *Hin*dIII fragment of 563 base pairs (bp) and two *Hin*dIII fragment of 454 and 109 bp, respectively (Ran & Michaelis 1995). The nonOwen CMS group comprises several haplotypes that are nondistinguishable with this marker (Desplanque *et al*. 2000). Primers used, PCR conditions, DNA digestion and gel electrophoresis were applied as described by Ran & Michaelis (1995) and Viard *et al*. (2002).

The nuclear polymorphism was assessed using seven microsatellite loci (CT4, GTT1, GCC1, GAA1, BVM3, CAA1

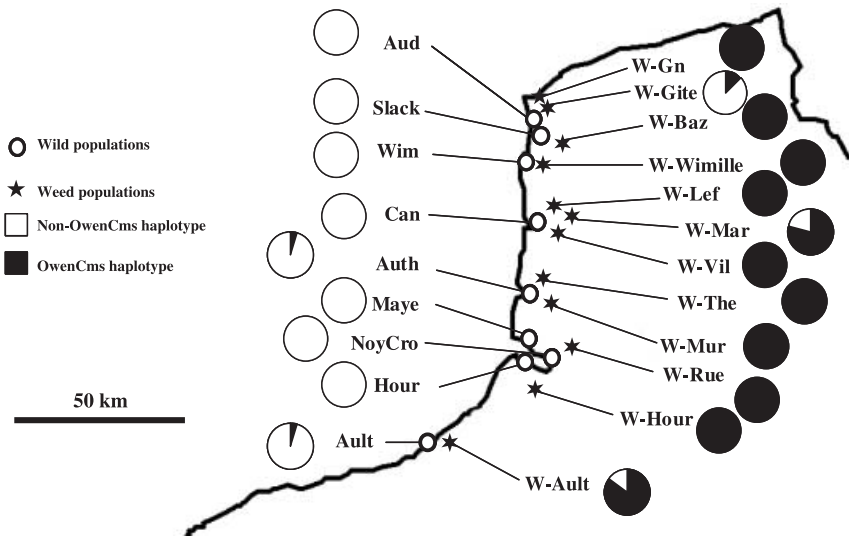

Fig. 1 Location of wild (coastal) and weed (within sugar beet fields) populations sampled along the coastline of Northern France and distribution of CMS vs. non-CMS haplotypes within each population.

Table 1 Geographical coordinates, occurence of CMS-Owen haplotype and summarizing statistics for nuclear genetic diversity across loci of study populations. *N*: sample size; CMS-Owen: percentage of CMS-Owen haplotype; A_r, allelic richness; H_O: observed heterozygosity; H_E: expected heterozygosity (gene diversity); F_{IS}: intrapopulation fixation index; *P*-value associated to the exact test for deviations to Hardy–Weinberg proportions. Mean values of these parameters are indicated for each of the two study population groups (i.e. weed vs. wild)

Population	Acronym	Latitude	Longitude	*N*	CMS-Owen	A_r	H_O	H_E	F_{IS}	*P*-value
Weed populations (sampled in sugar beet fields)										
Falaise Gris-nez	W-Gn	N 50°50.4′	E 1°35.2′	40	100%	3.84	0.47	0.53	0.126	0.019
Cap Gris-nez	W-Gite	N 50°51.0′	E 1°35.2′	39	23%	5.12	0.44	0.56	0.209	$< 10^{-6}$
Bazinghem	W-Baz	N 50°47.3′	E 1°38.6′	40	100%	4.94	0.53	0.58	0.084	$< 10^{-6}$
Wimille	W-Wim	N 50°45.9′	E 1°38.3′	40	100%	5.49	0.39	0.62	0.371	$< 10^{-6}$
Villiers	W-Vil	N 50°29.7′	E 1°39.0′	39	100%	6.12	0.47	0.62	0.234	$< 10^{-6}$
Maresville	W-Mar	N 50°31.3′	E 1°42.6′	42	81%	5.63	0.42	0.61	0.320	$< 10^{-6}$
Lefaux	W-Lef	N 50°34.3′	E 1°38.8′	39	100%	5.96	0.52	0.60	0.131	$< 10^{-4}$
Tigny	W-The	N 50°21.0′	E 1°41.7′	42	100%	4.51	0.42	0.54	0.216	$< 10^{-6}$
Le Muret	W-Mur	N 50°20.0′	E 1°39.0′	40	100%	6.22	0.48	0.65	0.276	$< 10^{-6}$
Le Champ Neuf	W-Rue	N 50°15.8′	E 1°37.0′	41	100%	4.62	0.41	0.52	0.225	$< 10^{-6}$
Le Hourdel	W-Hour	N 50°11.2′	E 1°33.2′	40	100%	6.09	0.49	0.58	0.170	0.003
Hable d'Ault	W-Ault	N 50°09.6′	E 1°29.6′	52	84.6%	4.89	0.41	0.55	0.263	$< 10^{-4}$
Mean: weed				494		4.93	0.45	0.58	0.22	
Wild populations (coastal populations — sea beet)										
Audresselles	Aud	N 50°49.4′	E 1°35.4′	30	0%	7.27	0.54	0.61	0.108	0.001
Ambleteuse	Slack	N 50°48.5′	E 1°36.4′	37	0%	6.65	0.57	0.63	0.095	0.003
Wimereux	Wim	N 50°46.1′	E 1°36.6′	38	0%	7.18	0.61	0.61	–0.006	0.107
Etaples	Can	N 50°30.9′	E 1°38.4′	39	0%	6.99	0.56	0.64	0.124	$< 10^{-3}$
La Madelon	Aut	N 50°22.3′	E 1°37.6′	33	6.3%	5.12	0.53	0.56	0.064	$< 10^{-6}$
Maye	Maye	N 50°15.7′	E 1°37.2′	36	0%	5.79	0.52	0.56	0.074	0.185
Noyelles Crotoy	NoyCro	N 50°12.8′	E 1°40.1′	34	0%	5.79	0.50	0.55	0.088	0.122
Le Hourdel	Hour	N 50°12.9′	E 1°33.9′	27	0%	4.86	0.50	0.48	–0.052	0.825
Ault	Ault	N 50°06.4′	E 1°27.2′	35	5.7%	5.93	0.54	0.64	0.164	$< 10^{-4}$
Mean: wild				309		6.18	0.55	0.59	0.079	

and CA2), as described by Mörchen *et al*. (1996), Viard *et al*. (2002) and Arnaud *et al*. (2003). Electrophoresis and genotyping were carried out on a LI-COR automated sequencer (LI-COR Inc., NB, USA).

Statistical analyses

For each population, genetic polymorphism was examined by calculating allele frequencies for each locus, allelic richness (A_r) (El Mousadik & Petit 1996), the observed heterozygosity (H_O), the genetic diversity H_E *sensu* Nei (1978) over all loci using GENEPOP version 3.3 (Raymond & Rousset 1995) and FSTAT version 2.9.3.2. (Goudet 1995). Population heterozygote deficiencies were quantified by calculating the Weir & Cockerham's (1984) $\hat{f}$, a multilocus estimator of the fixation index F_{IS}, and exact tests (Markov chain parameters set to 200 batches and 5000 iterations per batch) for deviations to Hardy–Weinberg proportions were carried out with GENEPOP. Mean values of genetic diversity parameters for each population group (i.e. weed vs. wild) was calculated and differences in H_E, H_O, F_{IS} and allelic richness were tested between weed and wild groups by using permutation procedures as implemented in FSTAT (tests carried out with 10 000 permutations). Genotypic linkage disequilibrium was estimated prior to other analyses using GENEPOP. Bonferroni adjustments for simultaneous statistical tests were applied (Rice 1989).

The overall genetic structure was investigated by calculating the $\hat{\theta}$ estimator of F_{ST} as described by Weir & Cockerham (1984). Significance of $\hat{\theta}$ values was assessed using the *G*-test that accounts for deviations to Hardy–Weinberg expectations (Goudet *et al*. 1996). A hierarchical analysis of the molecular variance was also conducted using ARLEQUIN version 2.0 (http://lgb.unige.ch/arlequin) to analyse the partition of the genetic variance within and between wild and weed groups. The overall genetic distances among populations were illustrated through a principal component analysis (PCA) carried out on gene frequency data with the PCA-GEN version 1.2.1 software (http://www2.unil.ch/izea/softwares/pcagen.html).

To test individual admixture proportions and the correspondence of genetic clusters with geographically labelled groups, we applied a model-based clustering algorithm described in Pritchard *et al*. (2000) by using STRUCTURE version 2.1. This software is based on a Bayesian method that enables the identification of clusters of genetically similar individuals from multilocus genotypes (e.g. Pierpaoli *et al*. 2003). We assumed here that each individual belonged either to a wild group or to a weedy group, given their sampling location (coastal populations vs. sugar beet fields). For each individual, the probability of belonging to the wild or the weed lineage was calculated as well as the probability to have ancestry in the other group, either in the sampled generation (immigrant from the alternative cluster) or in the first or second past generation (admixed individual). Each run consisted of a burn-in period of 200 000 steps followed by 10^6 MCMC (Monte Carlo Markov chain) replicates, with the number *K* of specified clusters being two and assuming that allele frequencies might be correlated.

Results

Chloroplastic diversity and seed flow

In the present study, 89.5% of weed individuals carried the Owen CMS type, demonstrating a maternal cultivated origin. However, three of the 12 weed populations showed nonOwen CMS haplotypes at high frequency (W-Gite: 87%, W-Mar: 19% and W-Ault: 15.4%, Fig. 1 and Table 1). Some rare sea beet individuals were found to have the Owen CMS cytoplasm (1.3%; Fig. 1) and only two wild populations contained individuals carrying the Owen CMS cytoplasm (6.3% and 5.7% in *Aut* and *Ault*, respectively) indicating the trace of gene flow through seeds from crop to wild populations.

Microsatellite diversity and within-population structure

Exact tests for genotypic linkage disequilibria between microsatellite loci within each population showed five (1.13%) significant *P*-values out of 441 comparisons, i.e. a proportion comparable to that expected by chance alone. A high level of nuclear diversity was observed over the whole study with a total of 96 alleles over seven microsatellite loci. Summary statistics describing the genetic diversity across loci within the 21 study populations are given in Table 1. From group comparisons (wild vs. weed), mean allelic richness (A_r) and observed heterozygosity (H_O) were significantly higher for wild beet populations ($P < 0.001$). The overall genetic diversity (H_E) was not statistically different between the two groups ($P = 0.56$). Weed beet populations exhibited a mean F_{IS} significantly higher ($P < 0.001$) than that observed for the wild group. Consistent heterozygote deficiencies were observed within all weed populations in contrast to wild sea beet populations that fitted Hardy–Weinberg expectations more closely (Table 1).

Gene flow and assignment analyses

Significant genetic differentiation ($\hat{\theta} = 0.115$, $P < 10^{-4}$) occurred among all the 21 populations. The AMOVA analysis showed a significant genetic structure between the two groups of wild and weed populations with an overall F_{CT} value (i.e. fixation index corresponding to the genetic variance among groups over total) equal to 0.11, whereas the mean genetic variance among populations within each

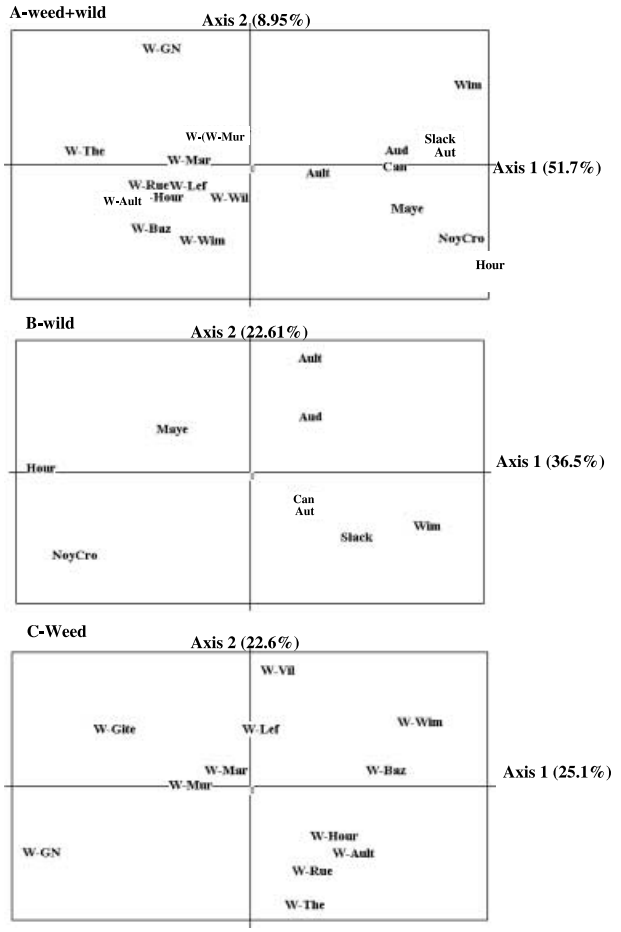

Fig. 2 Graphical ordination of sampled populations on the first frequency data. (a) Over the whole population; the first two axes represent 60.7% of the total inertia. (b, c) Analyses carried out over the weed and the wild populations only, respectively; the first two axes represent 59.1% and 47.8% of the total inertia, respectively.

group (F_{SC}) was equal to 0.065. This structure is illustrated by a graphical ordination of the samples following a PCA on gene frequency data (Fig. 2a). The first axis (51.71% of the total inertia) was found to be significant ($P < 10^{-4}$) and separated clearly the wild from the weed populations.

When analysing genetic structuring within each of the groups, the genetic variance among populations within wild and weed groups did not differ significantly ($\hat{\theta} = 0.063$, $P < 10^{-4}$ and 0.062, $P < 10^{-4}$, respectively). However, the spatial distribution of allelic diversity differed between wild and weedy lineages. As depicted in Fig. 2b,c, a geographical substructuring was apparent for the weed populations (e.g. W-GN and W-Gite; W-Wim and W-Baz; W-Vil, W-Lef and W-Mar) but less evident for the wild populations (axis 1 separates the three most southern populations NoyCro, Hour and Maye from the other populations). In agreement with the apparent step-by-step geographical structuring of weed populations, a significant isolation by distance pattern was found for weed populations (Mantel-like test, 5000 permutations, $P = 0.03$) but not for wild populations (Mantel-like test, 5000 permutations, $P = 0.08$).

Further insights about the individual clustering and possible immigrants came from the Bayesian analysis. By first carrying out a clustering analysis without prior information, we observed a high proportion of individual membership (0.913 and 0.916) to one of two clusters ($K = 2$). Using prior information on the status of collected individuals [i.e. sampled within coastal habitat ('wild') vs. sugar beet field ('weed')], the overall individual assignment increase up to 0.986 and 0.993 within each of the two assumed groups (Fig. 3). Of the 801 individuals analysed only 24 individuals (3%) showed probabilities ($P < 0.90$), indicating that they were not from the presumed cluster (Table 2), with 12 of 309 (3.9%) coastal individuals and 12 of 494 (2.4%) individuals sampled in sugar beet fields showing evidence for genetic admixture. In the coastal group, half of the admixed individuals belonged to the Ault population (17% of the individuals in this population). In general, the misclassified individuals sampled in the wild populations exhibited high probabilities of ancestry in the sampled generation, demonstrating that they were most probably (seed) immigrants. On the contrary, the misclassified individuals from the sugar beet fields generally showed high probabilities of resulting from hybridization one or two generations ago (Table 2). Interestingly, one of the two individuals carrying a CMS-cytoplasm from Ault (individual 289) exhibited a probability of 0.97 of being an immigrant, based on nuclear genes. This demonstrated that this individual arose from a seed migration at the present generation. The other individual with CMS in the wild did not show misclassification into the coastal wild group, suggesting that it resulted from several generations of cross-breeding by pollen from surrounding wild individuals.

Discussion

In spite of vast evidence for long-distance dispersal in wind-pollinated plants (reviewed in Hamrick & Godt 1996), our results show highly significant genetic differentiation between neighbouring wild and weedy populations (separated by 1–10 km) of the species *B. vulgaris*. This finding contrasts with the study of Bartsch *et al*. (1999), who found the presence of two sugar beet cultivar alleles at unusually high frequencies in natural sea beet populations from northeast Italy. Our results were unexpected as sea beets, weed beets and cultivated beets both belong to the same outcrossing species, are fully cross-compatible and mainly wind-pollinated. Additionally, they share a recent common coancestry, reinforced by the recurrent use of wild genetic resources in sugar beet breeding programmes and by the recent paternal ruderal origin of weed beets

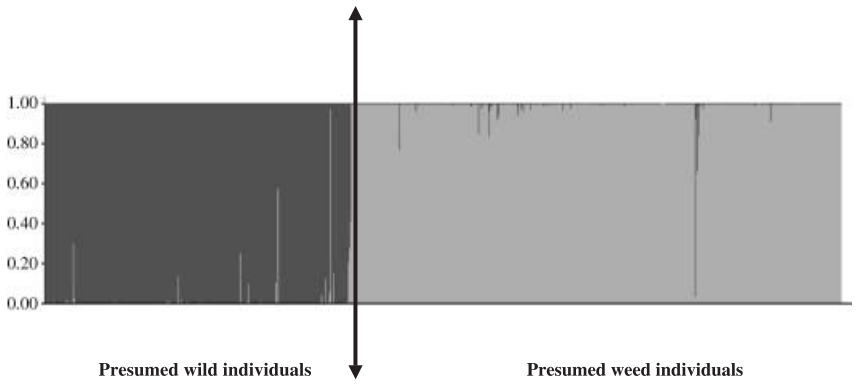

Fig. 3 Bayesian analysis of the nuclear genetic structure between wild and weed populations. Estimated individual membership assuming $K = 2$ using prior population information (geographical location: within or without a sugar beet field). Each individual is represented as a thin vertical line, which is partitioned into two coloured segments that indicate the individual's membership fractions in the two clusters (black and light grey picture wild and weed groups, respectively). The vertical black arrow separates individuals assumed to be either wild (on the left) or weedy (on the right).

Table 2 Population assignment and inferred ancestry of individuals. For each individuals with a probability (q) of being from its cluster of origin (clusters 1 and 2 are for wild and weed, respectively) inferior to 90%, the probability P to have an ancestry either in the sampled or first or second past generations (q-values computed with prior migration rate = 0.01) is given. The most significant values ($P < 0.75$) are indicated in bold characters

Population	Individual label	Presumed cluster	q-value	P alternative cluster/ancestry		
				Immigrant	Previous 1st	Previous 2nd
Wild-presumed (sampled along the seashore, coastal habitats)						
Aud	30	1	**0.577**	**0.216**	**0.142**	**0.065**
Can	135	1	**0.656**	**0.000**	**0.200**	**0.144**
Maye	198	1	**0.533**	**0.102**	**0.251**	**0.114**
Maye	206	1	0.865	0.077	0.026	0.033
NoyCro	234	1	0.816	0.063	0.048	0.073
NoyCro	236	1	**0.364**	**0.529**	**0.077**	**0.031**
Ault	284	1	0.800	0.066	0.087	0.048
Ault	289	1	**0.014**	**0.971**	**0.011**	**0.004**
Ault	292	1	**0.630**	**0.005**	**0.24**	**0.125**
Ault	307	1	**0.746**	**0.164**	**0.055**	**0.035**
Ault	308	1	**0.687**	**0.247**	**0.038**	**0.028**
Ault	309	1	**0.433**	**0.292**	**0.206**	**0.069**
Weed-presumed (sampled within sugar beet fields)						
W-Gite	375	2	0.853	0	0.015	0.133
W-Gite	358	2	**0.498**	**0.072**	**0.217**	**0.213**
W-Wim	448	2	**0.462**	**0.001**	**0.145**	**0.391**
W-Wim	438	2	**0.594**	**0.004**	**0.191**	**0.211**
W-Wim	457	2	0.858	0.039	0.057	0.045
W-Wim	458	2	0.859	0.021	0.053	0.067
W-Vil	483	2	0.872	0.001	0.035	0.092
W-Mur	656	2	**0.010**	**0.945**	**0.036**	**0.009**
W-Mur	657	2	0.787	0.005	0.057	0.152
W-Mur	658	2	**0.411**	**0.216**	**0.109**	**0.264**
W-Mur	660	2	**0.618**	**0**	**0.274**	**0.108**
W-Hour	732	2	**0.734**	**0.006**	**0.080**	**0.180**

(Boudry *et al*. 1993; Desplanque *et al*. 1999). Recent indirect estimates of pollen mediated gene flow in the seed production area of sugar beets suggested that gene dispersal greatly exceeds 1000 m (Saeglitz *et al*. 2000; Lavigne *et al*. 2002). However, our results demonstrate clearly that experimental cross-fertility, wind-mediated dispersal and close vicinity do not provide a basis for regular high gene flow from crops (through weed lineages) to wild populations in the study sugar beet production area.

The maintenance of such an overall genetic distinctiveness between weed and sea beets cannot be explained only by geographical isolation between weed and wild populations. Within each group, we found evidence for ample gene flow between populations on scales greater than 1–10

km. The hierarchical analyses of genetic variance showed a lower genetic differentiation ($F_{SC} = 0.06$) within each cluster than between them ($F_{CT} = 0.11$). In particular, a genetic isolation by geographical distance was observed within the weedy group. Weed pollen flow might therefore contribute to spread genes across sugar beet fields. The overall genetic differentiation in this wind-pollinated outcrossing complex might result from a time-lag between the peak of the flowering period of sea beets and wild beets: June and July, respectively. Pollen-mediated gene flow probably occurs between only a few individuals during a putative short overlapping period (Arnaud *et al*. 2003).

Interestingly, despite this general trend, rare immigration and hybridization/introgression events were found in several weed and wild populations. This was detected based not only on the frequency of CMS-Owen haplotypes found in the wild (demonstrating the occurrence of seed flow in several populations in the study landscape), but also on the nuclear admixture proportion in wild plants. Assignment analyses based on multilocus microsatellite genotypes proved to be a powerful tool for tracing contemporary cross-breeding events between domesticated/cultivated relatives in our case study: we detected 24 potentially admixed individuals of 803. Although rare (3%), this is the first evidence for nuclear gene flow between weed and wild individuals of *B. vulgaris* on a regional scale. In 'presumed' wild populations, half of the 12 admixed individuals were found in the Ault population, in which they represented a large fraction (17%) of the total number of individuals. It is noteworthy that Ault is surrounded by crops infested heavily by weed beets, a situation hardly encountered elsewhere. Regardless of whether GMOs are involved, consequences of crop–wild gene flow may also involve genetic assimilation or demographic swamping of wild populations (Gepts & Papa 2003; Haygood *et al*. 2003). Such intraspecific hybridizations are a main concern, as they can lead to loss of local adaptation of native/wild populations, which can be of primary importance during periodic disturbances or extreme environmental conditions (Allendorf *et al*. 2001). The main difficulty now is to determine if any proportion of admixture can be acceptable, a question difficult to tackle in the case of genetically engineered crops for which the level of acceptable hybridization (if any) will rely on the fitness conferred by the transgenes to the hybrids in the wild (Hails 2000; Meagher *et al*. 2003; Wilkinson *et al*. 2003).

Another critical feature in the *B. vulgaris* complex is the occurrence of individuals characterized by a cytoplasm typical of cultivated lines in several wild populations (i.e. CMS Owen). The results reported here extend, over several populations at a regional scale, our preliminary findings carried out in one population only and on a very fine scale (Arnaud *et al*. 2003). Seed flow appears to be a major vector for transportation of cultivated genes through weeds (proved to arise from hybridization between wild and cultivated lines) in the wild in the *B. vulgaris* complex. Furthermore, the presence of individuals carrying the nonOwen CMS cytoplasm in three weed populations could be the result of seed migration from natural populations to the sugar beet fields, or might be the relict of an ancient contamination before the worldwide use of Owen CMS in sugar beet breeding. However, additional studies are needed to identify which nonOwen CMS were present in these weed populations (Desplanque *et al*. 2000). The population of Ault is worth considering carefully, as displaying both seed flow and the highest frequency of nuclear admixture. In spite of our flowering time-lag hypothesis, the situation in Ault suggested that seed migration may have created a local hot-spot of hybridization and enhanced further introgression between wild and weed individuals due to simultaneous flowering within the population.

This study demonstrated that both seed and pollen flow contribute to the transfer of genes into wild populations of sea beet. However, seed flow plays a prominent role in the settlement of weed individuals, which act as pollen donors and enhance the probability of overlap between the flowering periods of wild and weedy subspecies. As pointed out recently by Gepts & Papa (2003), the spread of crop genes in wild populations must be monitored more closely. Our finding suggests that a reduction of gene transfers from weed to wild populations of sea beets would require setting up a large zone (at least several kilometres) free of crops along the coastline. Pollen flow between weed populations also has implications for biosafety study, as such pollen dispersal may promote the spread of transgenes across crops (Ellstrand 2003). However, risk assessment analyses based primarily on gene flow by pollen/seed transfers is not in itself a sufficient information and, as suggested by Meagher *et al*. (2003) and Wilkinson *et al*. (2003), the focus of attention should be on the relative fitness of plants that have acquired any transgenes.

Acknowledgements

The authors are grateful to J. Bernard for technical assistance and advices and to C. R. Engel for stimulating discussions and improvement of the manuscript. They also thank two anonymous referees and J. M. Bergelson for critical reading and helpful comments. This work was funded by a grant from the 'Bureau des Ressources Génétiques', the ACI 'Impact des OGM' from the French Ministry of Research, the 'Contrat de Plan Etat/Région Nord-Pas-de-Calais'. MD was supported by an INRA/region Nord-Pas de Calais fellowship. F. Viard thanks the CNRS for supporting, through an ATIP grant, her research projects in the Station Biologique de Roscoff.

References

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution*, **16**, 613–622.

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*, **270**, 1565–1571.

Bartsch D, Cuguen J, Biancardi E, Sweet J (2003) Environmental implications of gene flow from sugar beet to wild beet — current status and future research needs. *Environmental Biosafety Research*, **2**, 105–115.

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.

Bartsch D, Schuphan I (2002) Lessons we can learn from ecological biosafety research. *Journal of Biotechnology*, **98**, 71–77.

Boudry P, Mörchen M, Saumitou-Laprade P, Vernet P, Van Dijk H (1993) The origin and evolution of weed beets: consequences for the breeding and release of herbicide resistant transgenic sugar beets. *Theoretical and Applied Genetics*, **87**, 471–478.

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.

Desplanque B, Viard F, Bernard J *et al.* (2000) The linkage disequilibrium between chloroplast DNA and mitochondrial DNA haplotypes in *Beta vulgaris* ssp. *maritima* (L.): the usefulness of both genomes for population genetic studies. *Molecular Ecology*, **9**, 141–154.

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

Ellstrand NC, Prentice HC, Hancock JF (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics*, **30**, 539–563.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK.

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.

Gray AJ, Raybould AF (1998) Reducing transgene escape routes. *Nature*, **392**, 653–654.

Hails RS (2000) Genetically modified plants — the debate continues. *Trends in Ecology and Evolution*, **15**, 14–18.

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London, Series B*, **351**, 1291–1298.

Haygood R, Ives AR, Andow DA (2003) Consequences of recurrent gene flow from crops to wild relatives. *Proceedings of the Royal Society of London B*, **270**, 1879–1886.

Jarvis DI, Hodgkin T (1999) Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agrosystems. *Molecular Ecology*, **8**, S159–S173.

Lavigne C, Klein EK, Couvet D (2002) Using seed purity data to estimate an average pollen mediated gene flow from crops to wild relatives. *Theoretical and Applied Genetics*, **104**, 139–145.

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

Meagher TR, Belanger FC, Day PR (2003) Using empirical data to model transgene dispersal. *Philosophical Transactions of the Royal Society of London B*, **358**, 1157–1162.

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

Mücher T, Hesse P, Pohl-Orf M, Ellstrand NC, Bartsch D (2000) Characterization of weed beet in Germany and Italy. *Journal of Sugar Beet Research*, **37**, 19–38.

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

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

Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. *Theoretical and Applied Genetics*, **106**, 239–250.

Pierpaoli M, Birò ZS, Herrmann M *et al.* (2003) Genetic distinction of wildcat (*Felis silvestris*) populations in Europe, and hybridization with domestic cats in Hungary. *Molecular Ecology*, **12**, 2585–2598.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

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.

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

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

Saeglitz C, Pohl M, Bartsch D (2000) Monitoring gene flow from transgenic sugar beet using cytoplasmic male-sterile bait plants. *Molecular Ecology*, **9**, 2035–2040.

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.

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

Wilkinson MJ, Elliott LJ, Allainguillaume J *et al.* (2003) Hybridization between *Brassica napus* and *B. rapa* on a national scale in the United Kingdom. *Science*, **302**, 457–459.

The authors of the 'Laboratoire de Génétique et Évolution des Populations Végétales' (GEPV) are involved in population genetics of plant species, with special attention given to the interactions between wild and cultivated relatives and the evolution of breeding reproductive systems. More information about the activities of GEPV can be found at the following web site: http://www.univ-lille1.fr/gepv/. Frédérique Viard is a population geneticist particularly interested in the study of weedy and invasive plants and animals in terrestrial and marine environments (see http://www.sb-roscoff.fr).
