Theor Appl Genet (2008) 116:1063–1077
DOI 10.1007/s00122-008-0735-1

ORIGINAL PAPER

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

Stéphane Fénart · Jean-François Arnaud · Isabelle De Cauwer · Joël Cuguen

Received: 15 June 2007 / Accepted: 18 February 2008 / Published online: 12 March 2008

Abstract Hybridization between cultivated species and their wild relatives is now widely considered to be common. In the *Beta vulgaris* complex, the sugar beet seed multiplication areas have been the scene of inadvertent pollination of sugar beet seed bearers by wild ruderal pollen donors, generating a weedy form of beet which infests sugar beet fields in European countries. Up to now, investigations of evolutionary dynamics of genetic diversity within the *B. vulgaris* complex were addressed using few genetical markers and few accessions. In this study, we tackled this issue using a panel of complementary markers: five nuclear microsatellite loci, four mitochondrial minisatellite loci and one chloroplastic PCR-RFLP marker. We sampled 1,640 individuals that illustrate the actual distribution of inland ruderal beets of South Western France, weed beets and wild sea beets of northern France as well as the diversity of 35 contemporary European diploid cultivars. Nuclear genetic diversity in weed beets appeared to be as high as those of ruderal beets and sea beets, whereas the narrowness of cultivar accessions was confirmed. This genetic bottleneck in cultivars is even more important in the cytoplasmic genome as only one haplotype was found among all sugar beet cultivars. The large majority of weed beet populations also presented this unique cytoplasmic haplotype, as expected owing to their maternal cultivated origin. Nonetheless, various cytoplasmic haplotypes were found within three populations of weed beets, implying wild-to-weed seed flows. Finally, our findings gave new insights into the genetical relationships between the components of the *B. vulgaris* complex: (1) we found a very strong genetic divergence between wild sea beet and other relatives, which was unexpected given the recent evolutionary history and the full cross-compatibility of all taxa and (2) we definitely confirmed that the classification into cultivated, wild, ruderal and weed forms according to their geographical location, phenotype or their domesticated status is clearly in accordance with genetic clustering despite the very recent domestication process of sugar beet.

Communicated by A. Schulman.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0735-1) contains supplementary material, which is available to authorized users.

S. Fénart · J.-F. Arnaud · I. De Cauwer · J. Cuguen (✉)
Laboratoire de Génétique et Évolution des Populations Végétales, UMR CNRS 8016, Université des Sciences et Technologies de Lille, Lille 1, Bâtiment SN2, 59655 Villeneuve d'Ascq Cedex, France
e-mail: joel.cuguen@univ-lille1.fr

Introduction

Species complexes consist of clusters of closely related species or subspecies that are able to exchange genetic material in natural conditions (e.g. Pernès 1984; Coyne and Orr 2004). Species complexes are thus of great interest to understand hybridization, introgression or speciation processes related to different environmental features, spatial isolation or variations in life-history traits (e.g. Bowen et al. 2001; Jørgensen et al. 2002; Shaffer et al. 2004). The knowledge of the extent of genetic diversity and relationships within and among crop species and their wild relatives is also essential for the efficient use of plant genetic resource collections, in order to prevent wild populations from the introgression of characters from cultivated accessions, or to improve crop quality (Soleimani et al. 2002; Fernie et al. 2006; Mariac et al. 2006; Tani et al. 2006).

The major crop species have been generated by the assortment of characters selected from their wild relatives

during domestication processes (Olsen and Schaal 1999; Zohary and Hopf 2000; Goodrich and Wiener 2005; Zeder et al. 2006; Ross-Ibarra et al. 2007). Gene movements in crop–wild species complexes are not limited to the movements of genes from wild to crops through human selection, but introgression of cultivated traits into their wild relatives has clearly been documented in 12 of the 13 most important food crops in some parts of their agricultural distribution (Ellstrand et al. 1999; see also Stewart et al. 2003). Moreover, in seven cases out of 13, introgression of domesticated traits has had consequences on weed species by increasing their competitiveness (Ellstrand et al. 1999; Ellstrand and Schierenbeck 2000).

A recent and extensive essay on crop–wild interactions was devoted for a large part to the special case of the *Beta vulgaris* complex (Ellstrand 2003). This species complex is of particular interest since crop, wild and weed forms of *B. vulgaris* can be found in sympatric situations in Europe and are all interfertile (Santoni and Berville 1992; Boudry et al. 1993; Bartsch et al. 1999). According to their habitat, four types of beets can be distinguished within the *B. vulgaris* complex: (1) sea beets [*B. vulgaris* subsp. *maritima* (L.) Arcangeli] mostly found along the western European Mediterranean coastlines (Letschert 1993) and colonising areas located along estuaries, just at the upper level of the tide and, more rarely, cliffs overhanging the sea (Raybould et al. 1996; Laporte et al. 2001; Fievet et al. 2007); (2) cultivated beets which have been known for more than 2,000 years in the eastern Mediterranean regions (Ford-Lloyd and Williams 1975) and now including sugar beet, fodder beet, table beet and leaf beet. While sugar beet production fields are principally found in northern European countries, sugar beet seeds are mainly produced in South-Western France or Northern Italy (Bartsch et al. 1999). In seed production fields, two rows of pollen donors (classically tetraploid but increasingly diploid) frame four rows of diploid male-sterile seed bearers that are characterised by a particular cytoplasmic male sterility (CMS) only found in cultivars (the "OwenCMS", see Owen 1945), (3) weed beets, infesting some sugar-beet fields and leading to severe agronomic problems since the seventies (Horsney and Arnold 1979); (4) inland ruderal beets that are associated with disturbed man-made habitats and whose status is not clear, being either viewed as feral beets escaped from private gardens or cultivated fields and running wild, or typical wild beets originating from Mediterranean coast (Desplanque et al. 1999).

Gene flow between cultivated and wild relatives is likely to occur and has already been demonstrated by local studies either from wild-to-crop (Boudry et al. 1993; Desplanque et al. 1999) or from crop-to-wild (Bartsch et al. 1999), mainly in seed production areas. In the main European seed production area in south-western France, sugar beet seed bearers can be accidentally pollinated by ruderal pollen donors (Boudry et al. 1993; Desplanque et al. 1999). The ensuing crop–wild F1 hybrids are mixed with sugar beet seeds that are sown in growing fields in Northern France as well as in other European countries. While sugar beets are biennial, crop–wild F1 hybrids, considered as weeds, have inherited from their wild parents their early bolting ability and as a consequence can flower and reproduce during the crop season (Boudry et al. 1993, 1994; Van Dijk 2004). This unwanted recurrent hybridization is thus at the origin of weed beets, identified as the cause of serious agronomic problems for the last 30 years in Europe, with significant effects on sugar beet yield and quality (Longden 1989; Brants and Hermann 1998; Bartsch et al. 1999; Desplanque et al. 2002; Bartsch et al. 2003). This hybrid origin of weed beet has been firstly attested indirectly (i.e. not genetically) by their capacity to bolt without vernalization, revealing the presence of the dominant *B* allele, inherited from the ruderal parent and cancelling any cold requirement to bolt (Santoni and Berville 1992; Boudry et al. 1993). It has been shown that weed beets could act as an escape route for cultivated traits to wild beet populations (Arnaud et al. 2003; Cuguen et al. 2004; Viard et al. 2004) and that pollen flow between weed beet populations can occur over large distances (Fénart et al. 2007).

Despite the documented knowledge of each form of the *B. vulgaris* complex, only fragmented information is available about their genetical relationships and, as far as we are aware, only one study tackled this problem but with only few accessions and few genetic PCR–RFLP markers (Desplanque et al. 1999). In fact, comparison of the genetic diversity and its evolutionary dynamics among the different forms has never been performed over a large sample set with a sufficient number of highly polymorphic loci. In this respect, this study aimed at filling this lack and was based on the assessment of the nuclear and cytoplasmic genetic diversity found within the four forms of beet, each represented by a very large sample dataset. We thus investigated the genetic polymorphism of a large sample of (1) weed beet populations from different locations in the French and Belgian sugar beet production areas, (2) major accessions of the contemporary sugar beet cultivars, (3) representative populations of inland ruderal beets found within the French seed production area, in south-western France and (4) wild sea beet populations collected along 1,000 km coasts, from south Brittany to the North of France. All these individuals were characterized both at the cytoplasmic and nuclear level using highly polymorphic markers. The nuclear genetic diversity of each accession was assessed using five highly polymorphic microsatellites and we traced back the maternal genome using both maternally inherited mitochondrial minisatellites and one chloroplastic PCR–RFLP marker specific of the cytoplasmic male sterility used in cultivars (i.e. "OwenCMS", see Owen 1945).

The large sample set as well as the diversity and complementarity of the molecular tools used in this study allowed us to precisely investigate and compare the amount of genetic diversity found within each forms of *B. vulgaris* and to assess their genetical relationships and, in particular, to clarify the taxonomic position of weed and inland ruderal beets.

Materials and methods

Sampling

Beta vulgaris is a diploid short-lived perennial species ($2n = 18$) widely distributed along the European coastline and around the Mediterranean Basin (Letschert 1993), where the domestication process of beets occurred more than 2000 years ago (Ford-Lloyd and Williams 1975; Zohary and Hopf 2000). To investigate the genetic diversity and relationships within the *B. vulgaris* complex, we relied on a large sample set of representative populations of wild, cultivated and weed beets:

A total of 11 populations of wild sea beets (38 ± 11 individuals per population for a total of 416 individuals) were sampled along the Channel French coastline and labelled with acronyms S_{01}–S_{12} (Table 1, Fig. 1).

In the French seed production area, 12 populations of inland ruderal beets (50 ± 2 individuals per populations for a total of 596 individuals) were collected and labelled from R_{01} to R_{12} (Table 1, Fig. 1). These inland ruderal populations were found to be associated with past and present road works (roadsides, car parks), rubble deposits or garden edges.

Weed beets' genetic diversity has been assessed on a total of 481 weed beet individuals (40 ± 14 individuals per population), collected among 12 cultivated fields from Northern France and Southern Belgium. These 12 populations of weed beets were labelled from W_{01} to W_{12} (Table 1, Fig. 1).

Sugar beet genetic diversity was studied from a panel of 35 cultivars kindly provided by the 'Institut Technique de la Betterave industrielle' (ITB, the French Institute that studies and promotes specific and new agronomic qualities of beet cultivars). These 35 diploid cultivars, launched on the European market between 1999 and 2003, have been released by 13 of the major European seed companies. As only few individuals were available for each cultivar (4–5 individuals per cultivar, for a total number of 147 individuals) and for the sake of statistical robustness, cultivars were grouped into 13 "populations" according to seed companies and are labelled from C_{01} to C_{13} (Table 1).

This large sample set of 1,640 individuals allowed us (1) to precisely characterise the genetic diversity of the contemporary cultivars as well as those of inland ruderal beets of south western France, weed beets and wild sea beets of northern France and (2) to assess the genetic relationships within the *B. vulgaris* complex.

Genetic data collection

DNA extraction

Extraction and purification of total DNA were performed using a DNeasy 96 Plant Kit (Qiagen Inc.) according to manufacturer's protocol.

Cytoplasmic diversity

We checked for the occurrence of OwenCMS cytotype using a diagnostic chloroplastic PCR–RFLP marker related to a polymorphic *Hin*dIII site mapped in the *petG-psbE* chloroplast fragment (Ran and Michaelis 1995). Primers used, PCR conditions and DNA digestion of this PCR–RFLP method were applied as described by Ran and Michaelis (1995). This polymorphism allows to distinguish OwenCMS from non-OwenCMS lines (Desplanque et al. 2000; Viard et al. 2002, 2004; Arnaud et al. 2003). In order to obtain more precise information about the cytoplasmic diversity, we characterised the mitochondrial polymorphism by genotyping individuals at four mitochondrial minisatellite loci named Tr1, Tr2, Tr3 and Tr4 (Nishizawa et al. 2000). Multiplex PCR amplifications were carried in a 10.5 µl volume containing 3 mM $MgCl_2$, 200 µM of each dNTP, 0.2 µg/µL of BSA, 0.2 µM of each forward and reverse primer, 0.625 U of *Taq* polymerase (Applied Biosystems) and ≈50 ng of template DNA. Cycling condition included an initial denaturation step of 5 min at 94°C followed by 30 cycles of 30 s at 94°C, 60 s annealing at 62°C, 30 s at 72°C. Final extension was conducted 10 min at 72°C, using a 9700 thermal cycler (Applied Biosystems).

Among wild and cultivated accessions, *B. vulgaris* carries about 20 different mitotypes, previously described in Forcioli et al. (1998) and Desplanque et al. (2000). In order to assign a mitotype name to each wild, weed and ruderal individual genotyped in the present study, we identified the association between mitotypes and minisatellite haplotypes within a reference database recently used to examine the genealogical relationship between the different mitotypes depicted in Fénart et al. (2006).

Nuclear diversity

Individuals were genotyped at five nuclear microsatellite loci named GAA1, GTT1, GCC1, BVM3, CAA1 (Mörchen et al. 1996; Viard et al. 2002). Loci GTT1, GCC1 and

Table 1 Population acronym, sample location (name of the nearest town), sample size (*N*), GPS coordinates (WGS84 system) and taxon (wild sea beets, wild ruderal beets, weed beets or cultivar) of each population sampled for this study

Population acronym	Sample location	*N*	GPS coordinates	Taxon
S_{01}	Nieuwpoort	37	N 51°09.000′ –E 2°43.000’	Wild sea beets
S_{02}	Grand-Fort Philippe	41	N 51°00.334′–E 2°05.802′	Wild sea beets
S_{03}	Hourdel	38	N 50°12.900′–E 1°33.900′	Wild sea beets
S_{04}	Pourville-sur-mer	16	N 49°55.083′–E 1°01.944′	Wild sea beets
S_{05}	Cabourg	46	N 49°17.509′–E −0°07.651′	Wild sea beets
S_{06}	Cap Lévi	47	N 49°41.303′–E −1°28.397′	Wild sea beets
S_{07}	Pointe de l’Agon	18	N 49°00.093′–E −1°34.497′	Wild sea beets
S_{08}	Trébeurden	50	N 48°46.442′–E −3°34.950′	Wild sea beets
S_{09}	Argenton	48	N 48°31.391′–E −4°45.610′	Wild sea beets
S_{10}	Guidel	35	N 47°46.092′–E −3°31.528′	Wild sea beets
S_{11}	Lihou (Gernsey)	40	N 49°27.298′–E −2°39.196′	Wild sea beets
R_{01}	Saint Clar	49	N 43°53.508′–E 0°46.381′	Wild ruderal beets
R_{02}	Laplume	51	N 44°06.613′–E 0°31.867′	Wild ruderal beets
R_{03}	Pargan-Taillac	51	N 44°03.459′–E 0°35.361′	Wild ruderal beets
R_{04}	Cuq	50	N 44°04.979′–E 0°41.953′	Wild ruderal beets
R_{05}	Flamarens	50	N 44°01.055′–E 0°47.653′	Wild ruderal beets
R_{06}	Nérac	50	N 44°08.104′–E 0°20.520′	Wild ruderal beets
R_{07}	Vic-Fézensac	50	N 43°45.292′–E 0°17.730′	Wild ruderal beets
R_{08}	Plieux	50	N 43°57.019′–E 0°43.925′	Wild ruderal beets
R_{09}	Port-sainte Marie	45	N 44°14.999′–E 0°23.567′	Wild ruderal beets
R_{10}	Galapian	50	N 44°17.981′–E 0°24.831′	Wild ruderal beets
R_{11}	Montpézat	50	N 44°20.759′–E 0°31.493′	Wild ruderal beets
R_{12}	Prayssas	50	N 44°17.218′–E 0°30.494′	Wild ruderal beets
W_{01}	Illies	39	N 50°32.603′–E 2°49.365′	Weed beets
W_{02}	Illies	39	N 50°32.506′–E 2°49.823′	Weed beets
W_{03}	Fournes	40	N 50°34.972′–E 2°54.372′	Weed beets
W_{04}	Beaucamps-Ligny	24	N 50°35.822′–E 2°55.906′	Weed beets
W_{05}	Herlies	75	N 50°34.799′–E 2°52.631′	Weed beets
W_{06}	Ault	45	N 50°09.627′–E 1°29.631′	Weed beets
W_{07}	Vieille-Chapelle	45	N 50°35.945′–E 2°44.032′	Weed beets
W_{08}	Wissant	20	N 50°53.937′–E 1°41.891′	Weed beets
W_{09}	Wissant	20	N 50°53.057′–E 1°41.523′	Weed beets
W_{10}	La Goudinière	45	N 50°40.453′–E 3°23.782′	Weed beets
W_{11}	Merkem	45	N 50°57.574′–E 2°52.864′	Weed beets
W_{12}	Ingelmunster	44	N 50°56.323′–E 3°16.684′	Weed beets
C_{01}	2 Cultivars	6	–	Cultivars
C_{02}	1 Cultivar	4	–	Cultivars
C_{03}	2 Cultivars	9	–	Cultivars
C_{04}	2 Cultivars	9	–	Cultivars
C_{05}	2 Cultivars	10	–	Cultivars
C_{06}	3 Cultivars	14	–	Cultivars
C_{07}	2 Cultivars	6	–	Cultivars
C_{08}	1 Cultivar	5	–	Cultivars
C_{09}	9 Cultivars	36	–	Cultivars
C_{10}	1 Cultivar	5	–	Cultivars
C_{11}	3 Cultivars	14	–	Cultivars
C_{12}	5 Cultivars	21	–	Cultivars
C_{13}	2 Cultivars	8	–	Cultivars

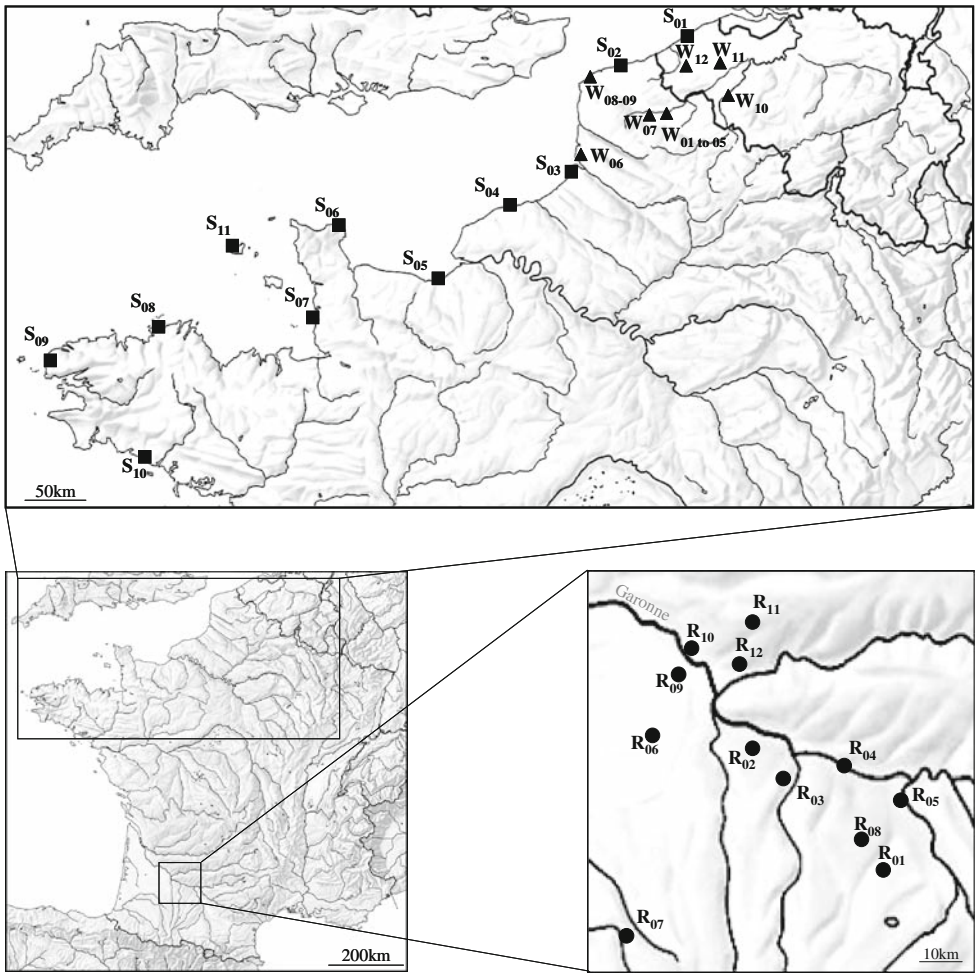

Fig. 1 Spatial location of sampled populations of wild and weed populations of *Beta vulgaris*. Wild inland ruderal beet populations (labelled with an "R") are visualised by *black circles*; wild sea beet populations ("S") are visualised by *black triangles* and weed beet populations ("W") are visualised by *black squares*

BVM3 have been recently mapped on chromosomes VI, II and IX, respectively (Laurent et al. 2007). These five microsatellite loci were amplified into a multiplex polymerase chain reactions (PCR) performed in a 10.5 µL volume mix as follows: 2.5 µL of DNA template (corresponding to a quantity of 25 ng), 1 µL of PCR Buffer 10× (Applied Biosystems), 2.9 mM $MgCl_2$, 0.2 µg/µL of BSA, 2% of DMSO (Dimethyl sulfoxid), 0.1 µM of each forward and reverse primer for loci GTT1, BVM3, CAA1 and GAA1 and 0.05 µM of each primer for loci GCC1, 290 µM of each dNTP and 0.9 U/µL of hot start *Taq* polymerase (*AmpliTaq* Gold, Applied Biosystems). PCR was carried out on a 9700 thermal cycler (Applied Biosystems) under the following conditions: 10 min denaturing at 94°C followed by 40 cycles of 45 s denaturing at 94°C, 45 s annealing at 54°C and 45 s extension at 72°C and a final extension step at 72°C for 10 min.

Detection and analysis of PCR products

To check for the presence of OwenCMS, *petG–psbE* cpDNA *Hind*III-digested products were separated using 0.8% agarose gel electrophoresis and visualized after ethidium bromide staining under UV light. Individuals carrying the OwenCMS are visualized by a two-bands pattern (454 base pairs (bp) and 109 bp), while non-OwenCMS individuals are characterised by an undigested 563 bp fragment (Ran and Michaelis 1995).

Detection of both minisatellite and microsatellite fragments was performed with an ABI Prism® 3100 Genetic Analyzer 16-capillary array system (Applied Biosystems) following manufacturer's protocols. For each individual, 2.5 µL of PCR product were mixed with 9.6 µL of Hi-Di™ formamide (Applied Biosystems) and 0.4 µL of GeneScan™–1000ROX™ size standard (Applied Biosystems) for the mitochondrial DNA minisatellites or 0.4 µL of GeneScan™-500LIZ size standard (Applied Biosystems) for microsatellites loci. The ABI Prism® 3100 Genetic Analyzer was set with the D matrix filter to detect the four dyes VIC™ (green), PET™ (red), NED™ (yellow) and 6-FAM™ (blue) used to label the forward primers of the four minisatellite markers Tr1, Tr2, Tr3 and Tr4 respectively. For nuclear microsatellites, the G5 matrix filter was used to detect alleles of GCC1, GTT1, BVM3, CAA1 and GAA1,

forward primers of which were labelled with dyes PET™ (red), NED™ (yellow), VIC™ (green), 6-FAM™ (blue) and NED™ (yellow), respectively. Raw data of electrophoresis obtained were read using GENEMAPPER v3.7 (Applied Biosystems). Individuals with doubtful genotypes (i.e. with missing data or presenting new alleles) were genotyped a second time at all loci.

Nuclear genetic data analysis

For each population, nuclear genetic diversity was examined by calculating allelic frequencies, allelic richness (A_r) following the rarefaction procedure of El Mousadik and Petit (1996), the genetic diversity (*He*) sensu Nei (1978) and the unbiased intra-population fixation index (F_{IS}) for each microsatellite locus and over all loci using GENEPOP version 3.3 (Raymond and Rousset 1995). Genotypic linkage disequilibrium was estimated prior to other analyses using GENEPOP version 3.3 (Raymond and Rousset 1995). Heterozygote deficiencies and significance of deviations from Hardy–Weinberg equilibrium within each population of weed, wild and inland ruderal beets were tested using a score test (Raymond and Rousset 1995). Permutation tests, implemented in the software FSTAT version 2.9.3.2. (Goudet 1995), were used to compare allelic richness, genetic diversity and F_{IS} between wild, ruderal and weed groups and also between French and Belgian weed beets populations (10,000 permutations).

Assuming that it should be useless and erroneous to compute genetic diversity (*He*) and intra-population fixation index (F_{IS}) in the 13 composite cultivar samples, we only compared allelic richness of sugar beet cultivars with either weed, wild, or inland ruderal beet groups using permutation tests as described above.

Genetic relationships among the *B. vulgaris* complex

Genetic relationships among the *B. vulgaris* complex were described in an unrooted neighbour-joining tree based on Cavalli-Sforza and Edwards' (1967) chord distance (D_{CE}) and using the software POPULATIONS v1.1.24 (Olivier Langella, available at http://www.pge.cnrs-gif.fr/bioinfo/). Bootstrap values were obtained based on 10,000 replications over populations. The ensuing populations' tree was visualised using TREEVIEW (Rod Page, available at http://www.taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Genetic differentiation between populations was assessed by pairwise F_{ST} estimates between populations following the ANOVA procedure of Weir and Cockerham (1984) using FSTAT version 2.9.3.2. Significance of pairwise F_{ST} was tested by randomly permuting multilocus genotypes among population samples (log-likelihood statistics *G*, 10,000 permutations) as suggested by Goudet et al. (1996). We also performed a hierarchical analysis of molecular variance (AMOVA) using Arlequin v3.1 (Excoffier et al. 2005, available at http://www.cmpg.unibe.ch/software/arlequin3/) to analyse the partition of the genetic variance within (F_{SC}) and between (F_{CT}) wilds, weed and cultivar groups. Bonferroni adjustments for simultaneous statistical tests were applied following Rice (1989).

We further tested the correspondence of taxonomic origin with genetic clusters by applying a model-based clustering algorithm, using the STRUCTURE software (Pritchard et al. 2000). This Bayesian method identifies clusters of genetically related individuals from multilocus genotypes, using or not prior knowledge of their group affiliation. This approach assumes that there are *K* groups contributing to the gene pool of the sampled populations. Individuals can have membership in multiple clusters, membership coefficients summing to 1 across clusters. In our case study, the membership of each individual was tested for a range of genetic clusters from $K = 2$ to $K = 5$, without prior information on their taxonomical affiliation. Each run consisted of a burn-in period of 1,000 steps followed by 10^6 MCMC (Monte Carlo Markov Chain) replicates, assuming that allele frequencies are uncorrelated across clusters. Repeated runs of STRUCTURE produced identical results to those shown.

Results

Cytoplasmic diversity

The presence of the *Hin*dIII restriction site in the *petG–psbE* chloroplast fragment revealed the occurrence of OwenCMS cytotype in all sugar beet individuals as well as in all individuals sampled within nine out of the 12 weed beet populations (Table 2). The association between mitotypes and mitochondrial minisatellite haplotypes obtained from the analysis of the collection data set of Fénart et al. (2006) are presented into brackets in the legend of Fig. 2 and a total of 10 mitotypes were represented according to their association with minisatellite haplotypes. A clear dichotomy appeared between, on the first hand, cultivars and weed beet populations where the OwenCMS is majoritary and, on the other hand, wild ruderal and sea beet populations characterized respectively by eight and nine mitotypes and a quasi-absence of the OwenCMS. All individuals carrying the OwenCMS cytotype also exhibited a unique combination of alleles (haplotype) from the four mitochondrial minisatellites loci: 500, 404, 420 and 438 bp, for Tr1, Tr2, Tr3 and Tr4, respectively (see Fig. 2). These results confirmed the maternal cultivated origin of the large majority of weed beet individuals and highlighted the lack of cytoplasmic diversity among cultivars. Nonetheless,

Table 2 Cytoplasmic and nuclear diversity within each sampled population and mean values over all samples for each taxon

	Cytoplasmic diversity		Nuclear diversity					
	Rate of OwenCMS cytotype	Number of minisatellite haplotypes	BVM3	CAA1	GCC1	GTT1	GAA1	All
Wild sea beet populations: (mean nuclear F_{ST}: 0.147, $P < 10^{-3}$; mean cytoplasmic F_{ST}: 0.359, $P < 10^{-3}$)								
S_{01}	0%	1	A_r: 4.33/H_e: 0.769 F_{IS}: 0.016*	2.57/0.443 0.025^{NS}	1.99/0.485 0.276^{NS}	2.06/0.372 0.345*	2.72/0.520 0.117**	2.73/0.518 0.134^{NS}
S_{02}	0%	5	4.60/0.773 −0.072^{NS}	3.97/0.695 0.439***	2.03/0.357 0.043^{NS}	2.66/0.548 −0.113^{NS}	2.35/0.460 −0.06^{NS}	3.12/0.567 0.062***
S_{03}	0%	2	3.35/0.567 −0.161^{NS}	3.86/0.742 0.078^{NS}	2.28/0.534 0.064^{NS}	1.97/0.454 −0.274^{NS}	1.61/0.191 −0.104^{NS}	2.61/0.498 −0.058^{NS}
S_{04}	0%	1	3.91/0.746 0.330**	3.58/0.694 0.279^{NS}	1.25/0.063 0.000^{NS}	2.00/0.517 0.153^{NS}	1.00/0.000 NC[a]	2.35/0.404 0.257*
S_{05}	2.38%	6	4.60/0.818 −0.009^{NS}	4.28/0.791 −0.017^{NS}	1.41/0.105 −0.032^{NS}	3.08/0.649 −0.224^{NS}	3.05/0.589 −0.246^{NS}	3.28/0.590 −0.106**
S_{06}	0%	3	4.37/0.777 0.124^{NS}	4.33/0.779 0.178*	2.57/0.511 0.101^{NS}	1.99/0.491 −0.140^{NS}	1.70/0.217 0.446**	2.99/0.555 0.113**
S_{07}	5.56%	4	3.69/0.722 0.154^{NS}	4.26/0.785 0.026^{NS}	1.40/0.108 −0.030^{NS}	1.98/0.458 0.029^{NS}	2.42/0.384 −0.157^{NS}	2.75/0.491 0.033^{NS}
S_{08}	0%	6	4.54/0.812 −0.010^{NS}	4.51/0.784 0.031^{NS}	1.35/0.096 −0.043^{NS}	1.98/0.479 −0.294^{NS}	1.88/0.264 −0.134^{NS}	2.85/0.487 −0.067^{NS}
S_{09}	0%	5	4.61/0.822 0.138**	3.76/0.687 0.040^{NS}	2.08/0.420 0.057^{NS}	2.60/0.549 −0.441^{NS}	1.64/0.179 0.166^{NS}	2.94/0.531 −0.018**
S_{10}	0%	5	5.41/0.880 0.091^{NS}	4.36/0.738 0.071*	1.22/0.056 −0.015^{NS}	2.47/0.479 −0.014^{NS}	2.22/0.395 −0.014^{NS}	3.14/0.510 0.047^{NS}
S_{11}	0%	6	3.88/0.725 0.000^{NS}	5.31/0.872 −0.007^{NS}	1.20/0.051 −0.013^{NS}	1.95/0.418 −0.166^{NS}	1.68/0.183 −0.068^{NS}	2.80/0.450 −0.040^{NS}
Mean	0.72%	4	4.30/0.765 0.061^{NS}	4.07/0.728 0.092*	1.71/0.253 0.009^{NS}	2.25/0.492 −0.102^{NS}	2.03/0.307 −0.016^{NS}	2.87/0.509 0.025*
Wild ruderal beet populations: (mean nuclear F_{ST}: 0.067, $P < 10^{-3}$; mean cytoplasmic F_{ST}: 0.449, $P < 10^{-3}$)								
R_{01}	0%	4	4.94/0.843 −0.065^{NS}	4.23/0.760 0.113^{NS}	2.41/0.466 0.124^{NS}	3.03/0.680 0.159^{NS}	2.08/0.376 0.132^{NS}	3.34/0.625 0.079^{NS}
R_{02}	0%	3	5.05/0.838 0.181^{NS}	4.22/0.786 0.053^{NS}	2.56/0.580 0.324*	3.27/0.707 −0.137^{NS}	2.26/0.428 0.038^{NS}	3.47/0.668 0.090***
R_{03}	0%	2	5.06/0.857 0.405***	4.10/0.757 0.120^{NS}	3.03/0.659 0.227*	3.38/0.714 0.011^{NS}	1.99/0.266 0.043^{NS}	3.51/0.651 0.187***
R_{04}	0%	3	4.01/0.730 0.096^{NS}	3.66/0.681 0.148^{NS}	1.94/0.416 −0.107^{NS}	2.90/0.636 0.150^{NS}	1.89/0.252 0.049^{NS}	2.88/0.543 0.086^{NS}
R_{05}	0%	4	5.18/0.868 0.055^{NS}	2.81/0.570 0.264^{NS}	3.05/0.598 0.096^{NS}	3.33/0.705 0.064^{NS}	1.95/0.361 −0.273^{NS}	3.27/0.620 0.065^{NS}
R_{06}	0%	1	3.68/0.744 0.194*	3.98/0.759 0.157^{NS}	1.91/0.380 0.210^{NS}	3.60/0.750 −0.067^{NS}	2.37/0.547 −0.024^{NS}	3.11/0.636 0.088^{NS}
R_{07}	0%	3	4.49/0.810 −0.061^{NS}	3.21/0.580 −0.103^{NS}	2.80/0.642 0.222*	2.75/0.617 −0.168^{NS}	2.08/0.314 −0.018^{NS}	3.07/0.593 −0.026**
R_{08}	0%	2	4.73/0.829 0.083^{NS}	3.40/0.668 0.132*	2.79/0.589 −0.053^{NS}	3.20/0.679 0.204*	1.50/0.149 −0.077^{NS}	3.12/0.583 0.087^{NS}
R_{09}	0%	2	3.17/0.655 0.017^{NS}	3.78/0.758 0.237*	2.21/0.492 0.052^{NS}	3.13/0.663 −0.039^{NS}	1.78/0.281 0.131^{NS}	2.81/0.570 0.080^{NS}
R_{10}	0%	3	4.48/0.785 0.006^{NS}	2.63/0.437 0.359***	1.97/0.332 0.398**	2.83/0.551 0.165*	2.25/0.414 0.227^{NS}	2.83/0.504 0.190***
R_{11}	0%	4	5.25/0.871 0.104^{NS}	3.21/0.651 −0.044**	2.61/0.578 0.308***	3.06/0.653 −0.010^{NS}	1.69/0.187 0.252*	3.16/0.588 0.095***
R_{12}	0%	5	4.24/0.775 0.278**	3.82/0.720 0.348**	2.39/0.541 0.056^{NS}	3.03/0.659 0.083^{NS}	1.99/0.394 0.086^{NS}	3.09/0.618 0.189***
Mean	0%	3	4.52/0.800 0.108**	3.59//0.677 0.149***	2.47/0.523 0.155***	3.13/0.668 0.035^{NS}	1.99/0.331 0.047^{NS}	3.14/0.600 0.101***

Table 2 continued

	Cytoplasmic diversity		Nuclear diversity					
	Rate of OwenCMS cytotype	Number of minisatellite haplotypes	BVM3	CAA1	GCC1	GTT1	GAA1	All
Weed beet populations: (mean nuclear F_{ST}: 0.056, $P < 10^{-3}$; mean cytoplasmic F_{ST}: 0.380, $P < 10^{-3}$)								
W_{01}	100%	1	4.46/0.799 0.165**	3.72/0.675 −0.025^{NS}	2.10/0.514 −0.198^{NS}	2.79/0.580 −0.106^{NS}	1.53/0.145 −0.058^{NS}	2.92/0.543 −0.021^{NS}
W_{02}	100%	1	4.56/0.787 0.023^{NS}	3.65/0.671 0.083^{NS}	2.43/0.537 0.045^{NS}	2.40/0.420 −0.161^{NS}	1.38/0.100 −0.031^{NS}	2.88/0.503 0.011^{NS}
W_{03}	100%	1	4.36/0.789 0.050^{NS}	3.56/0.694 0.243***	2.47/0.564 −0.152*	2.82/0.572 −0.135^{NS}	1.57/0.145 0.310*	2.96/0.553 0.032**
W_{04}	100%	1	3.97/0.699 0.047^{NS}	4.57/0.814 0.079*	2.04/0.371 0.215^{NS}	2.76/0.620 0.260^{NS}	2.16/0.330 0.115^{NS}	3.10/0.567 0.133***
W_{05}	100%	1	3.63/0.688 0.147**	2.83/0.601 −0.220^{NS}	2.23/0.519 0.050^{NS}	2.86/0.527 −0.087^{NS}	1.34/0.090 0.261^{NS}	2.58/0.485 −0.011^{NS}
W_{06}	75.56%	2	4.70/0.810 0.204**	3.64/0.694 0.199**	2.16/0.523 −0.105^{NS}	2.74/0.599 0.295*	1.63/0.200 0.111^{NS}	2.97/0.565 0.158**
W_{07}	100%	1	5.40/0.880 0.102^{NS}	3.22/0.672 0.308***	2.72/0.607 −0.035^{NS}	2.49/0.563 0.050^{NS}	1.00/0.000 NCa	2.97/0.544 0.111***
W_{08}	60%	5	6.04/0.922 0.187*	2.58/0.559 0.016^{NS}	2.18/0.497 −0.106^{NS}	2.42/0.491 0.185*	1.69/0.226 0.337***	2.98/0.539 0.110***
W_{09}	100%	1	5.63/0.889 −0.066^{NS}	3.45/0.693 0.089^{NS}	2.17/0.461 0.086^{NS}	2.74/0.607 0.306***	1.21/0.053 NC[a]	3.04/0.541 0.084*
W_{10}	100%	1	4.80/0.828 0.034^{NS}	3.72/0.738 0.308**	2.07/0.489 −0.045^{NS}	2.76/0.564 −0.065^{NS}	1.49/0.128 −0.043^{NS}	2.97/0.549 0.070*
W_{11}	100%	1	4.06/0.780 0.202**	3.63/0.752 0.498***	1.99/0.489 −0.044^{NS}	2.07/0.489 0.046^{NS}	1.89/0.258 0.053^{NS}	2.73/0.554 0.197***
W_{12}	33.33%	3	4.76/0.821 0.004*	2.92/0.619 −0.137^{NS}	1.99/0.503 −0.039^{NS}	2.81/0.641 0.113^{NS}	1.55/0.148 0.249^{NS}	2.81/0.546 0.003^{NS}
Mean	89.07%	1.58	4.70/0.808 0.100***	3.46/0.682 0.120***	2.21/0.506 −0.027^{NS}	2.64/0.556 0.058**	1.54/0.152 0.130***	2.91/0.541 0.073**
Cultivar accessions: (mean nuclear F_{ST}: 0.082, $P < 10^{-3}$; mean cytoplasmic F_{ST}: not calculable (monomorphic))								
C_{01}	100%	1	2.58[b]	1.98	2.00	1.98	1.00	1.91
C_{02}	100%	1	3.00	3.00	2.00	2.00	1.00	2.20
C_{03}	100%	1	2.40	2.00	1.71	1.99	1.00	1.82
C_{04}	100%	1	3.34	1.71	2.42	2.00	1.00	2.09
C_{05}	100%	1	2.05	2.00	1.90	2.21	1.65	1.96
C_{06}	100%	1	1.94	2.31	1.92	2.00	1.00	1.83
C_{07}	100%	1	2.67	2.00	2.00	2.65	1.00	2.06
C_{08}	100%	1	1.98	2.00	2.00	2.00	1.00	1.80
C_{09}	100%	1	2.88	2.76	1.99	1.99	1.21	2.17
C_{10}	100%	1	3.60	1.98	1.98	2.00	1.00	2.11
C_{11}	100%	1	2.47	1.50	1.98	2.00	1.00	1.79
C_{12}	100%	1	2.06	1.99	1.96	1.94	1.00	1.79
C_{13}	100%	1	2.80	2.00	2.00	1.99	1.00	1.96
Mean	100%	1	2.60	2.09	1.99	2.06	1.07	1.96

For cytoplasmic diversity, we estimated the rate of OwenCMS cytotype and the number of mitochondrial minisatellite haplotypes found within each population. For each nuclear microsatellite loci (BVM3, CAA1, GCC1, GTT1 and GAA1), we estimated the allelic richness (A_r), the expected heterozygosity (H_e) and the intrapopulation fixation index (F_{IS}). Mean genetic differentiation (F_{ST}) values within each taxon are also presented for nuclear and cytoplasmic diversity

NS non significant

[a] NC: not calculable. locus GAA1 showed only one allele in S_{04} and W_{09} or a second allele present in only one individual in W_{12}, making it impossible to estimate F_{IS} value

[b] Cultivar samples presented in this study are made artificially of different cultivars pooled together according to their producers and then computing either H_e or F_{IS} for these composite populations is not relevant

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, for significance of heterozygotes deficiency

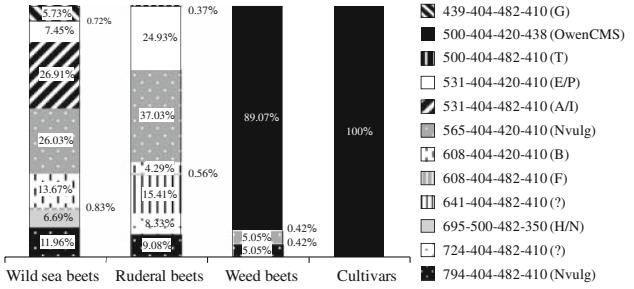

Fig. 2 Proportions of each minisatellite haplotype found within wild sea beets, inland ruderal beets, weed beets and cultivars. Each haplotype is based on the association of alleles of Tr1, Tr2, Tr3 and Tr4 loci, respectively, indicated by their size in base pairs. Associations of minisatellites haplotypes and mitotypes defined from the collection data set used in Fénart et al. (2006) are presented into brackets. The question mark indicates that the concerned minisatellite haplotype was not represented in the core collection

non-OwenCMS individuals were also found within the weed beet populations at a rate of 10.94% (Fig. 2).

The occurrence of non-OwenCMS individuals concerns the weed populations W_{06}, W_{08} and W_{12} (see Table 2) where four different minisatellite haplotypes also found in ruderal and wild populations (Fig. 2) were detected. Two of these haplotypes (565-404-420-410 and 794-404-482-410) are commonly found in association with a non-OwenCMS mitotype called *Nvulg* widely distributed in wild populations (Cuguen et al. 1994; Desplanque et al. 2000; Fénart et al. 2006). The two others (531-404-420-410 and 695-500-420-350) were, respectively found in association with mitotypes H/N and E/P following the reference database nomenclature defined in Desplanque et al. (2000) and Fénart et al. (2006) (see Fig. 2).

Nuclear diversity

Exact tests for genotypic linkage disequilibria between microsatellite loci within each population (except the composite samples of cultivars) showed five significant P values out of 360 comparisons (1.39%), 18 being expected from type I error. Multiple tests across all populations yielded no significant adjusted P values, also attesting that the five detected linkage disequilibria may be artifactual, either imputed to a small number of alleles and/or to population substructuring.

Statistics of population genetic diversity (A_r, H_e and F_{IS}) are presented in Table 2 for each locus and overall loci in each sample as well as their mean values overall sample for each form of *B. vulgaris*. Allelic richness (A_r) ranged from 1.00 (locus GAA1; population W_{07} and S_{04}) to 6.04 (Locus BVM3; population W_{08}) and amount of expected heterozygosity (H_e) were relatively high across all loci and samples, except for locus GAA1, which showed the lowest values for allelic richness and expected heterozygosity. A_r, H_e and F_{IS} did not significantly differ between wild ruderal, weed and wild sea beet populations. Nonetheless, a general trend toward higher F_{IS} values for weed and inland ruderal beet populations can be visualised in Table 2 (mean F_{IS} of 0.073 and 0.101, respectively) compared to wild sea beet populations (0.025). Finally, permutation tests confirmed that allelic richness was significantly lower ($P < 10^{-3}$) for cultivated beet (mean A_r of 1.960) compared to ruderal, weed and sea beets groups (mean A_r of 3.14, 2.91 and 2.87, respectively, see Table 2).

Genetic relationships within the *B. vulgaris* complex

Within each taxon, a very clear genetic differentiation between populations was found at the cytoplasmic level (F_{ST} of 0.359, 0.449 and 0.380; all at $P < 10^{-3}$ for wild sea beet, inland ruderal beet and weed beet populations, respectively), except between the cultivar accessions exhibiting only the OwenCMS mitotype (Table 2). Results of the AMOVA analysis indicates that the level of genetic differentiation is high either between or within weed beet, wild sea beet and inland ruderal beet groups ($F_{SC} = 0.416$ and $F_{CT} = 0.479$; both at $P < 10^{-3}$).

Mean nuclear population differentiation estimated over all studied populations—including cultivars—highlighted a significant genetic differentiation ($F_{ST} = 0.285$; $P < 10^{-3}$). Within each form of beets, mean population differentiation was low: mean F_{ST} values were of 0.056, 0.082 and 0.067; all at $P < 10^{-3}$, for weed, cultivated and ruderal beets respectively but were more pronounced between the wild sea beet populations ($F_{ST} = 0.147$; $P < 10^{-3}$). Pairwise nuclear F_{ST} estimates over all populations can be visualised in the Electronic supplementary material S1. Highly significant pairwise differentiation between all ruderal and wild populations and all other populations was detected. In contrast, only few differentiation estimates appeared to be significant between cultivars and weed beet populations, suggesting evidence for close genetical affinities between these two forms. It should also be emphasised that no significant genetic differentiation was observed among cultivar samples, except for 3 pairwise comparisons (see Electronic supplementary material S1).

The neighbour-joining tree of populations presented in Fig. 3 showed a clear genetic distinctiveness between the four forms, each corresponding to a single cluster of wild sea beet populations, wild ruderal beet populations, weed beet populations and cultivar accessions. Despite this clear genetic clustering by taxons, we can nonetheless depict a relative proximity of weed beet and cultivar samples populations with ruderal population, compared to wild sea beet populations. This highlighted the evolutionary divergence between wild sea beet and wild ruderal beet populations. In addition, the congruence between taxonomic and genetic dis-

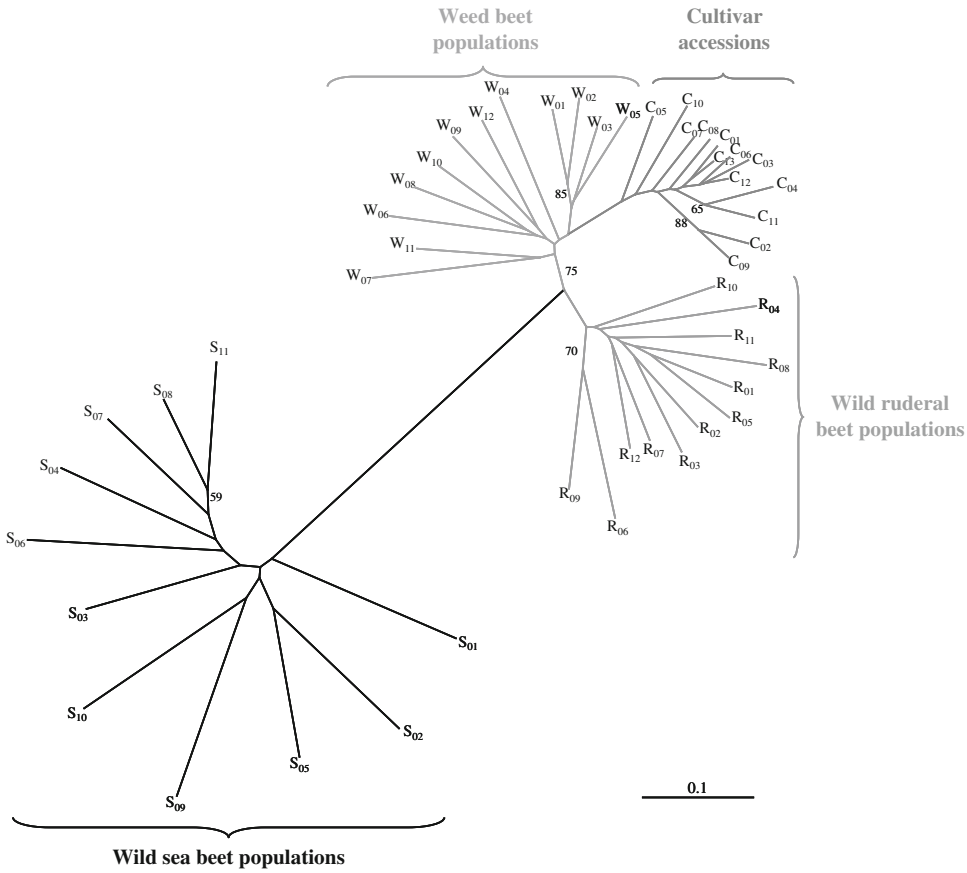

Fig. 3 Neighbour-joining tree describing genetic relationships among the *Beta vulgaris* complex based on the Cavalli-Sforza and Edwards' (1967) chord distance (D_{CE}). Wild inland ruderal beet populations are labelled with an "R"; wild sea beet populations with an "S"; weed beet populations with a "W" and cultivar accessions with a "C". Significance of each node was tested with 10,000 permutations over populations and only bootstrap values >50% are reported on the figure

tinctiveness was also well supported by results of the Bayesian analysis. Indeed, the most likelihood number of genetic clusters was $K = 4$, each corresponding to a taxon. Figure 4 showed that wild sea beets, inland ruderal beets and cultivars clearly constituted different genetic clusters. For weed beets, most individuals belonged to a single and distinct genetic cluster. However, a large part of individuals (23%) presented high values of individual memberships ($q > 0.75$) assigning them into the genetic cluster of cultivars. Moreover, but to a lesser extent, weed beets have a large proportion of their nuclear genome that come from the inland ruderal beet group, result suggestive of a genetic admixture between cultivated and inland ruderal beets. These differences of membership of weed beet individuals into the cultivar cluster may reflect that the weed beet populations are continuously replenished by new crop–wild F1 hybrids or cultivar bolters.

In the same way, results of the AMOVA analysis were in complete agreement with the topology of the neighbour-joining tree and showed a significant genetic differentiation between sea beet, ruderal beet, weed beet and cultivar groups with an overall F_{CT} value (i.e. fixation index corresponding to the genetic variance among groups over total) of 0.266, whereas the mean genetic variance among populations within each group (F_{SC}) was equal to 0.084, meaning that genetic differences were significantly more pronounced between forms than within each form. Pairwise estimates of F_{CT} were also computed. Comparisons of "cultivar-weed" and "weed-ruderal" produced the weakest F_{CT} values (0.054 and 0.057; all at $P < 10^{-4}$, respectively) whereas medium values were found for comparisons of "sea-ruderal" and "cultivar-ruderal" (0.104 and 0.146; all at $P < 10^{-4}$, respectively) and comparisons of "sea-cultivar" and "sea-weed" presented the highest F_{CT} values (0.434 and 0.386; all at $P < 10^{-4}$, respectively).

Discussion

Bottleneck in sugar beet cultivars

In this study we compared genetic diversity at both nuclear and mitochondrial level of a large set of accessions from the four forms of the *B. vulgaris* complex. Results of nuclear polymorphism analysis using five microsatellite loci clearly show a deficit in allelic richness in sugar beet cultivars, compared to wild, ruderal and weeds beet populations. This low level of genetic diversity is likely to result from the bottleneck associated to domestication and the ensuing breeding process and has been widely documented in both plant and animal domesticated species (e.g. Eyre-Walker et al. 1998; Clark et al. 2004; Otero-Arnaiz et al. 2005; Vasemagi et al. 2005; Zhu et al. 2007).

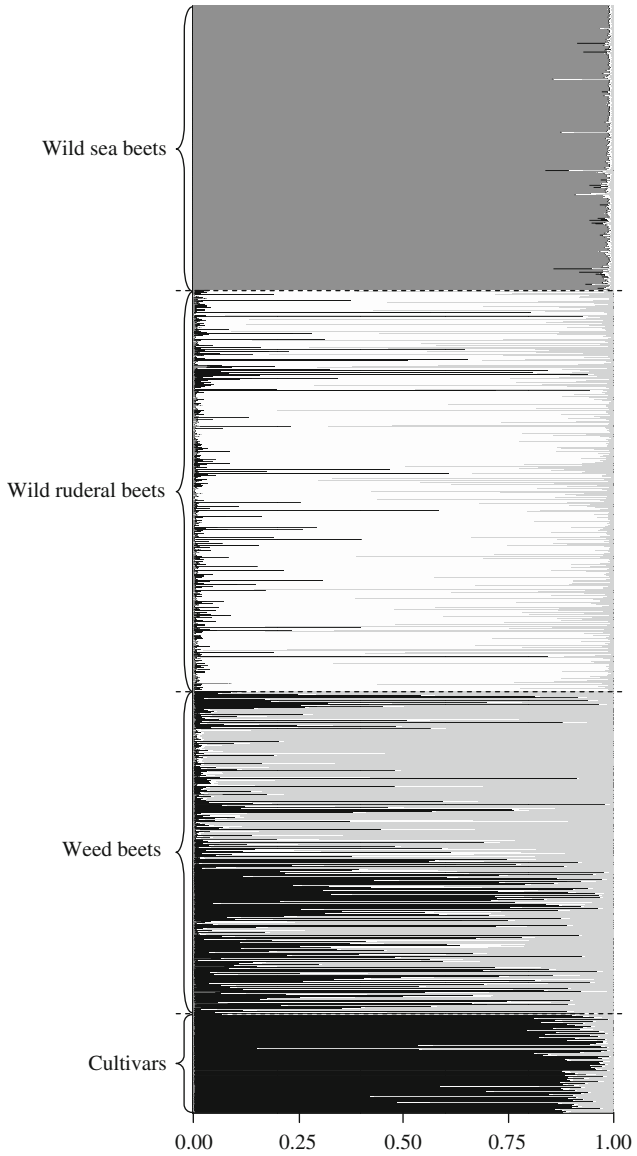

Fig. 4 Bayesian analysis of the nuclear genetic structure within the *Beta vulgaris* species complex. Individual membership was estimated assuming a number of $K = 4$ clusters and using no prior information on taxon membership (see text for explanations). Each individual was represented by a thin vertical line, which was partitioned into four coloured segments that indicated the individual's membership fractions into the four clusters (*dark grey* wild sea beets, *white* wild ruderal beets, *light grey* weed beets and *black* cultivars)

The genetic bottleneck related to the breeding process is expected to be higher in cytoplasmic genomes, especially when a particular character such as cytoplasmic male sterility is used (Provan et al. 1999). This is currently the case in sugar beet since, using four mitochondrial minisatellites, we only found a single haplotype over the 35 analysed cultivars coming from 13 European seed companies, compared to a total of ten mitochondrial haplotypes found within the wild sea beet and ruderal beet populations (Fig. 2). This unique haplotype found in cultivars corresponds to the combination of alleles 500, 404, 420 and 438, for minisatellite loci Tr1, Tr2, Tr3 and Tr4, respectively and is strictly associated with the OwenCMS cytoplasm (see also Fénart et al. 2006). While important efforts have been made to improve sugar beet using wild relatives (reviewed in Panella and Lewellen 2007), only one CMS germplasm has been used in sugar beet cultivars for decades (Owen 1945). Several different CMS types have been described in *B. vulgaris* ssp. *maritima* (Cuguen et al. 1994; Desplanque et al. 2000) and for exemple Touzet et al. (2004) proposed that *G* CMS could be a valuable alternative in sugar beet breeding.

Interestingly, a very low genetic differentiation (based on pairwise F_{ST} estimates, see Electronic supplementary material S1) was observed among the 13 cultivar samples each corresponding to a single sugar beet seed company. Together, these findings confirm the wide use of a restricted set of related maternal lines in cultivars (Owen 1945) as well as the use of more differentiated paternal lines which may result from a choice of the breeders to maintain a sufficient level of diversity as a baseline to future selection programs (McGrath et al. 1999).

Weed beet genetic diversity

Nuclear genetic diversity appeared to be high in weed beet populations, compared to sugar beet cultivars. Allelic richness and gene diversity for weed beets are of the same magnitude as in wild sea beet and wild ruderal beet populations and are significantly higher than in the cultivars. This is consistent with a wild paternal origin of weed beets and, as a consequence, confirms the hybrid origin of weed beets, characterised by a high level of genetic diversity owing to mixing of different gene pools. These high values of A_r and H_e observed in all weed beet populations highlight the large polymorphism introduced by the ruderal pollen donors as well as the recurrent introduction of new hybrids within these populations (see also Viard et al. 2002).

Furthermore, global F_{ST} value over weed beet populations (0.056; $P < 10^{-3}$) as well as pairwise genetic differentiation among weed beet populations (see Electronic supplementary material S1) underlined the genetic isolation of weed beet populations even when populations are in the vicinity of each other (e.g. W_{08} and W_{09}), implying independent founding events and distinct demographic histories, even if pollen of weed beet can disperse over large distances (Fénart et al. 2007). Additionally, no significant correlation between pairwise F_{ST} estimates and geographical distances between weed beet populations was observed ($r^2 = 0.0318$; $P = 0.78$, Mantel test after 10,000 permutations), that is no genetic isolation by distance was detected.

In addition, strong significant heterozygote deficiencies were found within eight populations of weed beets out of 12 with a mean F_{IS} estimate of 0.073 ($P < 10^{-3}$). This find-

ing, almost unexpected in a self incompatible species such as *B. vulgaris* (Larsen 1977; Bruun et al. 1995), is nevertheless in accordance with previous studies that already revealed such departures from Hardy–Weinberg equilibrium within weed beet populations (Viard et al. 2002; Arnaud et al. 2003; Cuguen et al. 2004; Viard et al. 2004) compared to wild sea beet populations (Fievet et al. 2007). These significant heterozygote deficiencies can be due either to population spatial substructuring or to a mixture of cohorts, related to recurrent infestation events and to differential recruitments of the soil seed bank; both events producing either a spatial or temporal Wahlund effect (Wahlund 1928; Hattemer 1982). Moreover, besides the intrinsic self-incompatibility of *B. vulgaris* taxa, a dominant mendelian self-fertility factor has been identified by Owen (1942). It was then widely introduced in the cultivated germplasm to produce inbred lines (Mackay et al. 1999). Some weed beets could have inherited this self-fertility factor from their cultivated maternal parent within the seed production area and, therefore introduced it into the weed populations present in sugar beet fields, with in an associated way, significant heterozygote deficiencies.

As weed beets result from accidental pollination events of seed bearers by ruderal beets, they are expected to carry the same cytoplasm as their cultivated maternal parent, i.e. OwenCMS, and thus to present a uniformity of mitochondrial DNA (Boudry et al. 1993). The signature of a cultivated maternal origin of weed beets has been revealed in 9 weed beets populations out of 12 (89.07% of the sampled individuals). However, three different cases of discrepancy between sugar beet and weed beet cytoplasm were revealed by our study and concerned three populations (W_{06}, W_{08} and W_{12}, see Table 2). In those populations, besides the OwenCMS mitotype, the most frequent minisatellite haplotypes are associated with the mitotype *Nvulg* (Fig. 2), the most widely found mitotype in the wild (Cuguen et al. 1994; Forcioli et al. 1998; Desplanque et al. 2000) and considered to be the mitochondrial ancestral state in *B. vulgaris* (Fénart et al. 2006). The presence of this cytotype could be related to the cultivation of ancient cultivars that did not carried the OwenCMS cytoplasm and that could have been conserved in the seed bank. The two remaining non-OwenCMS minisatellite haplotypes concerned two individuals of the coastal population of Wissant A (W_{08}). The first one, 531-404-420-410, was described in a large panel of wild sea beet populations of the French Channel Coasts where it was found in a strict association with mitotype *E*, a wild CMS mitotype (Fievet et al. 2007). Its occurrence in weed beets can be related to the presence of wild populations in the neighbourhood of this field (J.-F. Arnaud and S. Fénart, personal observations), highlighting a possible wild-to-weed seed movement. Arnaud et al. (2003) illustrated the possibility of human mediated seed movement from sugar beet fields to a wild population 1.5 km away. Our results show that the opposite movement can also be detected, underlining locally a possible influence of wild sea beet populations on the genetic diversity of weed beet populations. The last minisatellite haplotype, 695-500-482-350, has been found in association with two rare mitotypes *H* and *N* (Fig. 2, see also Fénart et al. 2006), both observed only in wild ruderal populations within the cultivar seed multiplication area (Desplanque et al. 2000). Their presence in weed beet populations suggests the possibility of accidental contamination by wild ruderal plants during the seed multiplication process.

Relationship within the *B. vulgaris* complex

The status of weed beets as the result of accidental and recurrent hybridisation between ruderal beets and cultivated seed bearers in the seed production area has been widely discussed (Boudry et al. 1993; Cuguen et al. 1994). However, since the preliminary study of Desplanque et al. (1999) based on few polymorphic markers, no further genetical investigations have been performed to finely assess genetic relationships within the *B. vulgaris* complex, i.e. cultivated, weed, ruderal and sea beets.

To address this issue, we investigated the genetic divergence within the *B. vulgaris* complex by focusing on a large panel of populations including the 35 sugar beet cultivars, 12 weed beet populations, 12 populations of ruderal beets sampled in the sugar beet seed production area in south-western France and 11 populations of sea beet from the French Channel coasts. The unrooted neighbour-joining tree, presented in Fig. 2, revealed a striking monophyletic clustering of accessions as a function of their origin, showing that the classification into cultivated, wild, ruderal and weed forms according to their geographical location, phenotype or their domesticated status is clearly supported by genetic data. Results of the analyses of molecular variance are also in complete agreement with a strict clustering of wild, weed and cultivated forms of beet. To gain further insights, pairwise F_{CT} were computed for each possible pair of groups and clearly documented: (1) the close intermediate taxonomical position of weed beets between cultivars and wild ruderal beets, reinforcing their hybrid status between these two forms; (2) closer genetic affinities of wild ruderal beets with weed beets lineages and cultivars compared to wild sea beet populations sampled along the Channel coastline (3) the clear genetic divergence of wild sea beets compared either to cultivars and weed beets. In particular, nuclear genetic diversity was strongly pronounced when wild sea beet populations were compared with cultivated and weed populations (F_{CT} of 0.434 and 0.386; both at $P < 10^{-4}$, respectively).

This striking genetic distinctiveness was not completely expected given (1) the recent evolutionary divergence of

the different groups through domestication process and (2) the full cross-compatibility within the *B. vulgaris* complex (Letschert 1993). The topology of the tree is also in agreement with the one presented in Desplanque et al. (1999) but highlights the intermediate position of weed beets between cultivars and wild populations. Furthermore, the results of the Bayesian analysis of individual clustering also indicated a clear genetic distinctiveness between wild sea beet, wild ruderal beets and cultivars. A genetic proximity between weed beets and cultivars was also revealed as a significant part of weed beet individuals clustered into the group of cultivars. Such individuals may be cultivar bolters with a low level of vernalization requirement.

Cytoplasmic data revealed an even more important genetic distinctiveness, not only between taxa but also within populations of the same group ($F_{CT} = 0.479$ and $F_{SC} = 0.416$ and see Table 2 for F_{ST} values within each taxa). Indeed, cytoplasmic effective population size is at least two fold lower than nuclear one in a gynodioecious species (Laporte et al. 2000), and seeds of *B. vulgaris* have no particular mechanism for long distance dispersal and, except accidental seed movements previously invoked, seed dispersal is expected to be restricted, leading to a stronger cytoplasmic differentiation compared to nuclear markers (see Fievet et al. 2007, for similar results on wild sea beet populations).

The origin of inland ruderal beet has been subject to debate, being presented either as natural forms of non-coastal wild beets (De Bock 1986) or as originated from feral forms and as a consequence related to cultivated beets (Bartsch et al. 1993, 1999). Our finding of high cytoplasmic and nuclear genetic diversities is consistent with the results of Desplanque et al. (1999) considering ruderal form as a wild relative of *B. vulgaris*. However, even if ruderal beets belong to the wild compartment, as revealed by their genetic diversity both at the nuclear and cytoplasmic level, their genetic distinctiveness from wild sea beet populations sampled on the Channel coast suggests distinct evolutionary histories and may support a common ancestry of French South Western ruderal beets with Mediterranean sea beet populations. Nevertheless, the topology of the neighbour-joining tree as well as the results of the AMOVA also suggests a genetic proximity between ruderal populations and sugar beet cultivars. It should also be noted that few individuals of inland ruderal beet (2.3%) exhibited high values of individuals clustering (>0.75) to the cultivar cluster (Fig. 4). This may reflect a possible gene flow from sugar beet cultivars to wild ruderal beet populations, due to the spatial proximity of ruderal populations with seed multiplication fields that produce high amounts of pollen. As a consequence, these populations have to be taken into account with extreme care for the risk assessment of (trans)gene escape in the wild within the seed production areas (Bartsch et al. 1999).

Acknowledgments We thank M. Richard-Molard, from the 'Institut Technique de la Betterave industrielle' (ITB), for providing us the 35 sugar beet cultivars included in this study. We are also very grateful to Marion Ballenghien and Adeline Courseaux for their technical assistance in molecular analysis. We are very grateful to Alan Schulman and two anonymous referees for providing helpful comments and suggestions on previous versions of the manuscript. This work was funded by the "Contrat de Plan État/Région Nord-Pas-de-Calais". Stéphane Fénart was supported by an INRA/Région Nord-Pas-de-Calais PhD fellowship and Isabelle De Cauwer is supported by a CNRS/Région Nord-Pas-de-Calais PhD fellowship.

References

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. Proc R Soc Lond 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. Environ Biosafety Res 2:105–115

Bartsch D, Lehnen M, Clegg J, Pohl-Orf M, Schuphan I, Ellstrand NC (1999) Impact of gene flow from cultivated beet on diversity of wild sea beet populations. Mol Ecol 8:1733–1741

Bartsch D, Sukopp H, Sukopp U (1993) Introduction of plants with special regard to cultigens running wild. In: Wöhrmann K, Tomiuk J (eds) Transgenic organisms. Birkhaüser Verlag, Basel, pp 135–151

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. Theor Appl Genet 87:471–478

Boudry P, Wieber R, Saumitou-Laprade P, Pillen K, Van Dijk H, Jung C (1994) Identification of RFLP markers closely linked to the bolting gene B and their significance for the study of the annual habit in beets (*Beta vulgaris* L.) Theor Appl Genet 88:852–858

Bowen BW, Bass AL, Rocha LA, Grant WS, Robertson DR (2001) Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale. Evolution 55:1029–1039

Brants I, Hermann H (1998) Herbicide tolerant sugar beet. In: Proceedings of the 61st IIRB congress. IIRB, Brussels, pp 195–204

Bruun L, Haldrup A, Petersen SG, Frese L, De Bock TSM, Lange W (1995) Self-incompatibility reactions in wild species of the genus *Beta* and their relation to taxonomical classification and geographical origin. Genet Resour Crop Evol 42:293–301

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Am J Hum Genet 19:233–257

Clark RM, Linton E, Messing J, Doebley JF (2004) Pattern of diversity in the genomic region near the maize domestication gene *tb1*. Proc Natl Acad Sci USA 101:700–707

Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Sunderland

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: Den Nijs HCM, Bartsch D, Sweet J (eds) Introgression from genetically modified plants into wild relatives. CABI Publishers, Inc., Oxfordshire, pp 183–201

Cuguen J, Wattier R, Saumitou-Laprade P, Forcioli D, Mörchen M, Van Dijk H, Vernet P (1994) Gynodioecy and mitochondrial DNA polymorphism in natural populations of *Beta vulgaris* ssp *maritima*. Gen Sel Evol 26:S87–S101

De Bock TSM (1986) The genus *Beta*: domestication, taxonomy and interspecific hybridization for plant breeding. Acta Hort 182:335–343

Desplanque B, Boudry P, Broomberg K, Saumitou-Laprade P, Cuguen J, Van Dijk H (1999) Genetic diversity and gene flow between wild, cultivated and weedy forms of *Beta vulgaris* L (Chenopodiaceae), assessed by RFLP and PCR-based methods. Theor Appl Genet 98:1194–1201

Desplanque B, Hautekèete N-C, Van Dijk H (2002) Transgenic weed beets: possible, probable, avoidable? J Appl Ecol 39:561–571

Desplanque B, Viard F, Bernard J, Forcioli D, Saumitou-Laprade P, Cuguen J, Van Dijk H (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. Mol Ecol 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. Theor Appl Genet 92:832–839

Ellstrand NC (2003) Dangerous liaisons? When cultivated plants mate with their wild relatives. The Johns Hopkins University Press, Baltimore, London

Ellstrand NC, Prentice HC, Hancock JF (1999) Gene flow and introgression from domesticated plants into their wild relatives. Annu Rev Ecol Syst 30:539–563

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? Proc Natl Acad Sci USA 97:7043–7050

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

Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS (1998) Investigation of the bottleneck leading to the domestication of maize. Proc Natl Acad Sci USA 95:4441–4446

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. Mol Ecol 16:3801–3813

Fénart S, Touzet P, Arnaud J-F, Cuguen J (2006) Emergence of gynodioecy in wild beet (*Beta vulgaris* ssp. *maritima* L.): a genealogical approach using chloroplastic nucleotide sequences. Proc R Soc Lond B 273:1391–1398

Fernie AR, Tadmor Y, Zamir D (2006) Natural genetic variation for improving crop quality. Curr Opin Plant Biol 9:196–202

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? Mol Ecol 16:1847–1864

Forcioli D, Saumitou-Laprade P, Valero M, Vernet P, Cuguen J (1998) Distribution of chloroplast DNA diversity within and among populations of the gynodioecious species *Beta vulgaris* ssp *maritima* (Chenopodiaceae). Mol Ecol 7:1193–1204

Ford-Lloyd BV, Williams JT (1975) A revision of *Beta* section *Vulgares* (*Chenopodiaceae*), with new light on the origin of cultivated beets. Bot J Linean Soc 71:89–102

Goodrich J, Wiener P (2005) A walk from the wild side: the genetics of domestication of livestock and crops. BioEssays 27:574–576

Goudet J (1995) FSTAT (Version 1.2). A computer program to calculate *F*-statistics. J Hered 86:485–486

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

Hattemer HH (1982) Genetic distance between populations—Part 3: Wahlund's principle as related to genetic distance and an application. Theor Appl Genet 62:219–223

Horsney KG, Arnold MM (1979) The origins of weed beet. Ann Appl Biol 92:279–285

Jørgensen S, Hamrick JL, Wells PV (2002) Regional patterns of genetic diversity in *Pinus flexilis* (Pinaceae) reveal complex species history. Am J Bot 89:792–800

Laporte V, Cuguen J, Couvet D (2000) Effective population sizes for cytoplasmic and nuclear genes in a *Gynodioecious* species: the role of the sex determination system. Genetics 154:447–458

Laporte V, Viard F, Bena G, Valero M, Cuguen J (2001) The spatial structure of sexual and cytonuclear polymorphism in the gynodioecious *Beta vulgaris* ssp. *Maritime*: I/at a local scale. Genetics 157:1699–1710

Larsen K (1977) Self incompatibility in *Beta vulgaris* L. 1. Four gametophytic, complementary S-loci in sugar beet. Hereditas 85:227–248

Laurent V, Devaux P, Thiel T, Viard F, Mielordt S, Touzet P, Quillet M-C (2007) Comparative effectiveness of sugar beet microsatellite markers isolated from genomic libraries and GenBank ESTs to map the sugar beet genome. Theor Appl Genet 115:793–805

Letschert JPW (1993) *Beta* section *Beta*: biogeographical patterns of variation and taxonomy. Agricultural University of Wageningen

Longden PC (1989) Effects of increasing weed-beet density on sugarbeet yield and quality. Ann Appl Biol 114:527–532

Mackay IJ, Gibson JP, Caligari PDS (1999) The genetics of selfing with concurrent backcrossing in breeding hybrid sugar beet (*Beta vulgaris altissima* L.). Theor Appl Genet 98:1156–1162

Mariac C, Luong V, Kapran I, Mamadou A, Sagnard F, Deu M, Ndjeunga J, Bezançon G, Pham J-L, Vigouroux Y (2006) Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. Theor Appl Genet 114:49–58

McGrath JM, Derrico CA, Yu Y (1999) Genetic diversity in selected, historical US sugarbeet germplasm and *Beta vulgaris* ssp. *maritima*. Theor Appl Genet 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. Theor Appl Genet 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. Curr Genet 37:34–38

Olsen KM, Schaal BA (1999) Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. Proc Natl Acad Sci USA 96:5586–5591

Otero-Arnaiz A, Casas A, Hamrick JL, Cruse-Sanders J (2005) Genetic variation and evolution of Polaskia chichipe (*Cactaceae*) under domestication in the Tehucan Valley, central Mexico. Mol Ecol 14:1603–1611

Owen FV (1942) Inheritance of cross- and self-sterility and self-fertility in *Beta vulgaris*. J Agric Res 64:679–698

Owen FV (1945) Cytoplasmically inherited male-sterility in sugar beets. J Agric Res 71:423–440

Panella L, Lewellen RT (2007) Broadening the genetic base of sugar beet: introgression from wild relatives. Euphytica 154:383–400

Pernès J (1984) Gestion des Ressources Génétiques des Plantes, Tome 2. Lavoisier, Paris, 346 pp

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

Provan J, Powell EA, Dewar H, Bryan G, Machray GC, Waugh R (1999) An extreme cytoplasmic bottleneck in the modern European cultivated potato (*Slanum tuberosum*) is not reflected in decreased levels of nuclear diversity. Proc R Soc Lond B 266:633–639

Ran Z, Michaelis G (1995) Mapping of a chloroplast RFLP marker associated with the CMS cytoplasm of sugar beet (*Beta vulgaris*). Theor Appl Genet 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. J Hered 86:248–249

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

Ross-Ibarra J, Morell PL, Gaut BS (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. Proc Natl Acad Sci USA 104(suppl 1):8641–8648

Santoni S, Berville A (1992) Evidences for gene exchanges between sugar beet (*Beta vulgaris* L.) and wild beets: consequence for transgenic sugar beets. Plant Mol Biol 20:575–580

Shaffer HB, Fellers GM, Randal Voss S, Oliver JC, Pauly GB (2004) Species boundaries, phylogeography and conservation genetics of the red-legged frog (*Rana aurora/draytonii*) complex. Mol Ecol 13:2667–2677

Soleimani VD, Baum BR, Johnson DA (2002) AFLP and pedigree-based genetic diversity estimates in modern cultivars of durum wheat [*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.]. Theor Appl Genet 104:350–357

Stewart Jr CN, Halfhill MD, Warwick SI (2003) Transgene introgression from genetically modified crops to their wild relatives. Nat Rev Genet 4:804–816

Tani N, Yoshimaru H, Kawahara T, Hoshi Y, Nobushima F, Yasui T (2006) Determination of the genetic structure of remnant *Morus boninensis* Koidz. trees to establish a conservation program on the Bonin Islands, Japan. BMC Ecol 6:14

Touzet P, Hueber N, Bürkholz A, Barnes S, Cuguen J (2004) Genetic analysis of male fertility restoration in wild cytoplasmic male sterility *G* of beet. Theor Appl Genet 109:240–247

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: Den Nijs HCM, Bartsch D, Sweet J (eds) Introgression from genetically modified plants into wild relatives and its consequences. CABI publishers, Inc., Oxfordshire, pp 53–69

Vasemagi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in atlantic salmon (*Salmo salar* L.). Mol Biol Evol 22:1067–1076

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 vs. hot spots of hybridization over a regional scale. Mol Ecol 13:1357–1364

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. Theor Appl Genet 104:688–697

Wahlund S (1928) Zusammensetzung von populationen und korrelationsercheinungen von standpunkt der vererbungslehre aus betrachtet. Hereditas 11:65–106

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

Zeder MA, Emshwiller E, Smith BD, Bradley DG (2006) Documenting domestication: the intersection of genetics and archaeology. Trends Genet 22:139–155

Zhu Q, Zheng X, Luo J, Gaut BS, Ge S (2007) Multilocus analysis of nucleotide variation of *Oryza sativa* and iIts wild relatives: severe bottleneck during domestication of rice. Mol Biol Evol 24:875–888

Zohary D, Hopf M (2000) Domestication of plants in the Old World. Oxford University Press, New York