Genet Resour Crop Evol (2020) 67:1853–1866
https://doi.org/10.1007/s10722-020-00945-0

RESEARCH ARTICLE

Survey and genetic diversity of wild *Brassica oleracea* L. germplasm on the Atlantic coast of France

Lorenzo Maggioni · Roland von Bothmer · Gert Poulsen · Karolina Härnström Aloisi

Received: 22 May 2019 / Accepted: 21 April 2020 / Published online: 21 May 2020

Abstract Wild populations of *Brassica oleracea* subsp. *oleracea* growing on European Atlantic coasts deserve attention since their diversity could contribute useful alleles to the *Brassica oleracea* cole crops (cabbage, broccoli, cauliflower, etc.). These populations have also been proposed as the source from which cole crops have been domesticated. Other authors have challenged their natural origin and suspected their derivation from vegetable brassicas escaped from fields and gardens. Our study surveyed northwestern French coastal areas and analysed with Amplified Fragment Length Polymorphism (AFLP) markers the genetic diversity and structure of nine wild populations, as well as of five accessions of locally grown *B. oleracea* crops. This study offered the highest level of detail ever presented about the distribution of wild *B. oleracea* populations along the French Atlantic coast. Populations analysed showed a low level of genetic differentiation, which might be explained by a relatively recent origin of all the populations from a common source, more likely than by insufficient physical or distance barriers to intercrossing. Traditional varieties commonly grown in the same area were not fully distinguishable from the wild populations on a molecular level. The level of genetic diversity of the wild populations was similar to, or lower than that of the cultivated crops. Therefore, the absence of a domestication bottleneck invited us to exclude the wild French populations as the likely source of original domestication events. Populations with higher levels of genetic diversity that could be targeted for conservation and breeding were the wild Penly and Petites Dalles and the Saint-Saëns cabbage.

Keywords *Brassica oleracea* · Genetic diversity · Crop wild relatives · AFLPs

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10722-020-00945-0) contains supplementary material, which is available to authorized users.

L. Maggioni (✉)
European Cooperative Programme for Plant Genetic Resources (ECPGR), c/o Bioversity International, Maccarese (Fiumicino), Rome, Italy
e-mail: l.maggioni@cgiar.org

R. von Bothmer
Swedish University of Agricultural Sciences, Alnarp, Sweden

G. Poulsen
Danish Seed Savers, Tjele, Denmark

K. Härnström Aloisi
Nordic Genetic Resource Centre, Alnarp, Sweden

Introduction

Since the 17th century, botanists have described a group of plants growing on the Atlantic cliffs of Great

Britain as having a close resemblance to the *Brassica oleracea* L. crops (kales, cabbages, cauliflowers, etc.). In his *Historia Plantarum*, Ray (1686) called these plants "wild coleworts". In *Species Plantarum*, Linnaeus assigned the name *Brassica oleracea* to this group of plants, also including the cultivated varieties under the same name (Linnaeus 1753). Currently recognized under the taxon name *Brassica oleracea* L. subsp. *oleracea* (Marhold 2011), populations growing on Atlantic cliffs have been recorded in Great Britain, northwest France, northern Spain, including the Cíes islands, the Portuguese islands of Berlengas, the German islands of Helgoland and Rügen, and one location in Denmark (Rødvig) (Rigueiro Rodríguez 1977; Snogerup et al. 1990; Tauleigne Gomes 2002; Gómez-Campo et al. 2005; Géhu 2009; Christensen et al. 2011; Drenckhahn 2017). The interpretation of the origin of this group of plants is controversial, considering that different authors have challenged their natural origin and even proposed that some or all of these populations could have derived from vegetable brassicas escaped from fields and gardens (Mitchell 1976; Christensen et al. 2011; Maggioni 2015). These populations currently occupy a specific niche on steep calcareous coastal cliffs, with discontinuous localized distribution.

Several studies have shown that the genetic background of these populations is very closely related to that of cultivated *B. oleracea* crops, with which they easily intercross with higher fertility rates than any of the Mediterranean wild relatives of the cole crops bearing the same chromosome number 2n = 18 (von Bothmer et al. 1995). This pattern has been interpreted by some authors as a valid indication of an Atlantic origin of the vegetable brassica crops (Song et al. 1990; Hodgkin 1995). Such interpretation has been challenged by other authors, based on historical, linguistic, literary and other considerations (Maggioni et al. 2014; Maggioni 2015). Independently from their origin, the Atlantic populations of *Brassica oleracea* growing in the wild deserve particular attention for their ability to intercross with their related crops. Since they belong to the primary genepool of the *Brassica oleracea* crops, their diversity should be explored, as it could contribute useful alleles that may be missing from the cultivated genepool. It is therefore important to accurately map the distribution of these populations, as a necessary step towards conserving valuable diversity and to investigate the nature and structure of such diversity. So far, knowledge about the distribution of the French populations was mainly limited to the map provided by Snogerup et al. (1990) and the phytosociological studies of some populations (Géhu 2009). Our study surveyed northwestern French coastal areas and enabled us to identify and collect samples from an unprecedented number of populations, some of them described for the first time.

In terms of genetic analysis, very few wild-growing *B. oleracea* populations from France have been previously studied. Lannér-Herrera et al. (1996) estimated by isozymes the diversity of four populations (from Mortagne, Granville, Fécamp and Mers-les-Bains) and by Random Amplified Polymorphic DNA (RAPD) markers the diversity of two populations (from Mortagne and Mers-les-Bains). They found a considerable intrapopulation diversity, with the exception of the Mortagne population, which was the smallest in size (ca. 550 individuals) and with lower diversity in the isozyme analysis. The four French populations also showed a considerable genetic diversity between populations (Gst = 0.39), with intermediate values compared to lower values between three Spanish populations (Gst = 0.11) and higher values between 11 British populations (Gst = 0.51). Lack of clustering of populations from the same or even adjacent regions suggested independence between genetic diversity and geographical distribution. No strong correlation was either detected between the level of gene diversity and population size (Lannér-Herrera et al. 1996). Other studies of French populations only include one or two accessions analysed for comparisons within multi-species sets (Lannér 1998; Lázaro and Aguinagalde 1998a, b). Eight wild *B. oleracea* populations from Britain were analysed with AFLPs by Watson-Jones et al. (2006) in a study aimed at detecting their level of genetic diversity as a baseline for monitoring and conservation purposes. They identified high levels of Nei's genetic diversity H (ranging from 0.18 to 0.33 in the different populations) and also significant differentiation between populations, with an Fst value of 0.2257 over 103 loci. On the other hand, no significant relationship was observed between genetic and geographic distance. The objectives and methodology of our study are comparable to this latter work, extending a similar baseline of knowledge to a subset of the French populations. We also included a few samples of cultivated accessions as a term of comparison.

Geneflow mechanisms that could influence the structure of the wild populations depend on pollen flow and seed dispersal. Pollinators of wild *B. oleracea*, studied by Raybould et al. (1999) and Warman (1999) in southern England, are mainly honeybees and bumblebees, foraging within distances of a few metres, with occasional longer flights of up to 500 m. The chance of seed dispersal of wild *Brassica* species from one site to another was considered by Snogerup et al. (1990) as being very small, based on the authors' observations and the seeds being heavy and round without any adaptation for dispersal by wind or animals. Occasional exceptions are seed-containing valves that can be transported by wind over longer distances. The role of birds, particularly seagulls, in the dispersal of *Brassica* seeds was hypothesized by Gillham (1970) but still requires further investigation.

The genetic diversity within *B. oleracea* crops appears to be generally lower than within wild taxa, as exemplified by van Hintum et al. (2007) on genebank accessions of white cabbages. Allender et al. (2007) found a wealth of diversity revealed by chloroplast simple sequence repeats (SSRs) among the Mediterranean wild species, contrasting with an apparent absence of diversity in *B. oleracea* crops and contemporary UK natural populations of *B. oleracea*. Another study showed that certain sets of landraces and commercial cultivars can show higher levels of diversity than wild populations (Maggioni et al. 2014). Our study intended to survey, with the greatest degree of accuracy, the geographic distribution and size of the French populations of *B. oleracea* growing in the wild. Additionally, we sought to analyse the genetic diversity and structure of wild populations, also in comparison with locally-grown traditional vegetable brassicas. Our data also aimed to identify target populations for conservation and use, as well as to offer insights regarding the origin of domestication of the brassica vegetable crops.

Materials and methods

Seed samples of 33 accessions of *B. oleracea* subsp. *oleracea*, listed in Online Resource 1, were collected during a mission to France in July 2012 (details on the collected populations in Online Resource 2). The mission was organized after obtaining prior consent from the French Ministry for Food and Agriculture and the Institute for Genetics, Environment and Plant Protection in Ploudaniel, France. Duplicate samples were sent to the French National Institute for Agricultural Research (INRA) in Ploudaniel and a Standard Material Transfer Agreement was signed by the Nordic Genetic Resource Centre (NordGen), where the original study material was deposited. Population sizes were determined by counting individuals of the small populations and estimating the large ones, also using binoculars. Seeds were collected keeping the offspring of each individual mother plant separate. Accessions from a subset of 12 of these populations were selected for molecular analysis, with the aim of representing the whole distribution range. In the end, only nine populations (Table 1) were included in the final analysis, since it was not possible to obtain sufficient samples from Paluel, Le Hode and Mortagne. Five additional accessions of traditional *B. oleracea* crops were provided by Dr Alain Label from INRA, Ploudaniel. This material comprised three cauliflower samples from the Barfleur area, one leafy kale from Montjoie-Saint-Martin area and the Chou de Saint-Saëns, which is a cabbage variety grown around Saint-Saëns at least since the 19th century (Table 1). The choice of the cultivated accessions included material that originated in, and is still grown as close as possible to, the distribution area of the wild populations. It is, however, unlikely that this cultivated material can naturally intercross with the wild populations since the area of cultivation is rather distant from the ocean and the flowering time is not overlapping (Label, personal communication 2014).

Seeds were germinated in soil trays under greenhouse conditions at NordGen, in Alnarp, Sweden, and leaf material was harvested at the 4–6 leaf stage. The leaf samples were powdered in a mixer mill (Merck Retsch mm 300) with a steel ball and DNA extraction carried out using the protocol described by Doyle and Doyle (1987), with the following modifications: after transferring the DNA-containing phase (approx. 450 µl) to clean tubes, 5 µl RNase (10 mg ml^{-1}) was added. Samples were then incubated at 37 °C for 30 min. DNA was precipitated with 0.8 volume cold isopropanol, gently mixed and centrifuged 10 min at 13,200 rpm. The supernatants were removed and the pellets were washed in 500 µl wash buffer (76% ethanol, 0.2 M sodium acetate), left at room temperature for 20 min and centrifuged for 5 min at

Table 1 Analysed populations and genetic diversity data (NI = number of individuals analysed)

Populations/Accessions	Accession code	NI	Mean Nei's genetic diversity (H)		Private bands	% polymorphic loci
			Mean	SE		
Wild Atlantic *B. oleracea*	Collecting number					
Mers-les-Bains	FRA12-01	22	0.229[b]	0.026	0	73
Le Tréport	FRA12-02	13	0.204[a]	0.028	0	61
Penly	FRA12-05	22	0.300[f]	0.024	0	91
Saint-Valery-en-Caux	FRA12-06	27	0.284[e]	0.024	0	93
Petites Dalles	FRA12-11	17	0.297[f]	0.024	0	93
Fécamp	FRA12-18	18	0.222[b]	0.025	0	75
Yport	FRA12-22	24	0.241[c]	0.026	0	79
Étretat	FRA12-28	17	0.218[ab]	0.027	0	66
Granville	FRA12-32	13	0.229[bc]	0.025	0	75
H_A (Mean Nei's genetic diversity)		173	0.247	0.009	–	78
H_T (Total Nei's genetic diversity)		173	0.297	0.022	0	100
Cultivated accessions	Accession number					
Barfleur 1 (cauliflower)	FLH FR50 0001	22	0.224[b]	0.027	0	70
Barfleur 2 (cauliflower)	FLH FR50 0002	24	0.265[d]	0.025	0	80
Barfleur 3 (cauliflower)	FLH FR50 0003	27	0.251[c]	0.027	0	75
Montjoie (leafy kale)	FO FR50 0001	22	0.244[c]	0.026	0	79
Saint-Saëns (cabbage)	PO FR76 0001	3	0.355[g]	0.019	2	93
H_A (Mean Nei's genetic diversity)		118	0.268	0.011	–	79
H_T (Total Nei's genetic diversity)		118	0.333	0.020	2	100

Means with different superscripts are significantly different at $P < .01$

13,200 rpm. They were then rinsed in a buffer (76% ethanol, 0.01 M ammonium acetate) and centrifuged at 13,200 rpm for 5 min. Discarding the supernatants, the pellets were left to dry at room temperature. Pellets were re-suspended in 50 µl double-distilled water. The amounts of DNA were determined using spectrophotometer analysis (BioSpectrometer, Eppendorf or QIAxpert, Qiagen) and AFLP reactions followed the procedure described in Maggioni et al. (2014), except for the last part where the polymerase chain reaction (PCR) products from the final PCR were visualized using capillary electrophoresis (QIAxcel Advanced, Qiagen). Specifically, a combination of two restriction enzymes (*Eco*RI and *Mse*I) was used. Preamplification was carried out with 45 ng of primers (EcoRI + C/MseI + A) having a single selective nucleotide. Two primer pairs with three selective nucleotides (EcoRI + CAG/MseI + AGG and EcoRI + CAC/MseI + AAC) were used for the selective amplification. Two standard samples and size standards (50–700 bp Sizing Standard, Licor) were included on all gels to ensure the quality of the individual gel runs and reproducibility of the AFLP markers. After alignment, only clear and reproducible bands between 50 and 600 bp were scored using QIAxcel ScreenGel Software v.4.0 (Qiagen) and transformed as either present (1) or absent (0) and recorded in a binary data matrix. The number of individuals analysed per population is listed in Table 1. Each individual analysed is the offspring of a different mother. Diversity analysis was performed comparing populations or groups of populations in different combinations. The groups were: all (nine) wild populations; three Barfleur cauliflower accessions; five cultivated accessions (three cauliflowers, one cabbage and one leafy kale); and all nine wild populations as a group, versus five cultivated accessions as a group. Four wild populations with minimal Nei genetic distance (lower than 0.03) were also included, to have an indication of the lowest value of

genetic differentiation among sub-groups of populations.

As an indicator of genetic diversity, we measured Nei's gene diversity (H = 2 * p * q) (Nei 1973). Nei's gene diversity was averaged over loci and individuals for each population to calculate the mean genetic diversity per population (H). It was averaged over loci for all individuals of two population groups (wild and cultivated) to calculate the total genetic diversity of the two groups (H_T). It was also averaged over loci and populations or accessions to calculate the mean genetic diversity of a group of populations or accessions (H_A). Student's *t* test was performed to determine the level of significance of the differences among the means. Presence or absence of private bands was determined within and across populations and accessions. Distance matrices were generated from Nei's genetic distances (Nei 1972) between pairs of populations and between pairs of individuals. Based on these distance matrices, principal coordinate analysis (PCoA) was used to visualize the differences between populations or between individuals. Total genetic variation was partitioned into two levels by an analysis of molecular variance (AMOVA) between and within populations/accessions. The significance was tested using 1000 permutations. PhiPT values are calculated as the proportion of the variance among populations, relative to the total variance. A Mantel test with 999 permutations was performed to test the relationship between linear genetic distance and geographic distance matrixes, based on 173 individuals belonging to the nine wild populations. All tests were computed in GenAlEx 6.2 (Peakall and Smouse 2006), calibrated for binary diploid data.

Results

Geographic survey

A list of locations visited and populations identified by geographic location, coordinates and size of the populations are provided in Table 2, while Fig. 1 shows their geographic distribution on a map of northwestern France. Our survey of the Atlantic coast of France allowed the identification of *B. oleracea* subsp. *oleracea* populations from 24 locations in the administrative regions of Somme (1), Seine-Maritime (20), Manche (1) and Charente-Maritime (2). Six additional locations with compatible habitat were visited without finding any populations. All the sites described by Snogerup et al. (1990) were visited and the presence of the taxon was reconfirmed, except in two cases, Ault and Cap de la Hève. At Ault, plants were sought without success, both northeast and southwest of the village. At Cap de la Hève, plants were sought without success at the beach, outside Sainte-Adresse, below the cliffs, as well as at the lighthouse above the cliffs. In both cases, previous observations made in 1985 had documented only two (Cap de la Hève) and 20 (Ault) plants, indicating that these two populations might either have become extinct or escaped our search.

Population sites not listed by Snogerup et al. (1990) were identified at the following locations (from NE to SW): Mesnil-Val, Biville, Penly, Paluel, Le Pont Rouge, Veulettes-sur-Mer, Les Petite Dalles, Grainval, Valleuse d'Étigues and Le Tilleul. Other locations with a favourable habitat where we did not find any populations were the following (from NE to SW): Sainte-Marguerite-sur-Mer, Quiberville-Plage, Veules-les-Roses and Cauville-sur-Mer.

Finally, two sites mentioned in existing databases or bibliographic references were also checked, i.e. Cayeux-sur-Mer, Somme, and Douvres-la-Délivrande, Calvados, but the habitat revealed itself to be unsuitable and therefore the original data should not be considered reliable. Cayeux-sur-Mer was purposefully checked since the European Cooperative Programme for Plant Genetic Resources (ECPGR) *Brassica* database (https://ecpgr.cgn.wur.nl/Brasedb/, accessed 10 October 2018) includes one accession (CGN0693) of wild *B. oleracea* that is registered as collected in 1981 at this location by the Foundation for Agricultural Plant Breeding, Wageningen. However, the habitat of this site (a flat beach with no cliffs) is not typical of wild *B. oleracea* and this taxon was not found there during our mission, although the beach is rich in *Crambe maritima* L. The latter was indeed the original botanical name used by the donor of the seed to the Dutch genebank, as indicated in the remarks to this database record in the ECPGR *Brassica* database. As for Douvres, the location of Douvres-la-Délivrande was visited since an herbarium specimen collected in 1834 and conserved at Kew is reported by Snogerup et al. (1990) as part of the French distribution of *B. oleracea*, in the following terms: 'Calvados: Douvres (North of Caen)'. Douvres-la-Délivrande (49.294502;

Table 2 Presence of *Brassica oleracea* in the locations visited during the collecting mission in northwest France in July 2012

	Locations visited	Department	Plants observed	Accessions collected	Population size
1	Cayeux-sur-Mer	Somme	No	–	–
2	Ault	Somme	No	–	–
3	Mers-les-Bains	Somme	Yes	Yes	> 5000
4	Le Tréport	Seine-Maritime	Yes	Yes	> 5000
5	Mesnil-Val	Seine-Maritime	Yes	Yes	< 50
6	Biville	Seine-Maritime	Yes	No	> 100
7	Penly	Seine-Maritime	Yes	Yes	> 1000
8	Sainte-Marguerite-sur-Mer	Seine-Maritime	No	–	–
9	Quiberville	Seine-Maritime	No	–	–
10	Veules-les-Roses	Seine-Maritime	No	–	–
11	Saint-Valery-en-Caux	Seine-Maritime	Yes	Yes	> 5000
12	Paluel	Seine-Maritime	Yes	Yes	> 200
13	Le Pont Rouge	Seine-Maritime	Yes	Yes	< 50
14	Veulettes-sur-Mer	Seine-Maritime	Yes	Yes	> 1000
15	Les Petites Dalles	Seine-Maritime	Yes	Yes	> 1000
16	Les Grandes Dalles	Seine-Maritime	Yes	Yes	> 300
17	Saint-Pierre-en-Port	Seine-Maritime	Yes	Yes	> 100
18	Fécamp	Seine-Maritime	Yes	Yes	> 1000
19	Grainval	Seine-Maritime	Yes	Yes	> 100
20	Yport	Seine-Maritime	Yes	Yes	> 5000
21	Vaucottes	Seine-Maritime	Yes	Yes	> 100
22	Valleuse d'Étigues	Seine-Maritime	Yes	Yes	> 100
23	Étretat	Seine-Maritime	Yes	Yes	> 5000
24	Le Tilleul	Seine-Maritime	Yes	Yes	> 1000
25	Cap d'Antifer	Seine-Maritime	Yes	Yes	> 100
26	Phare d'Antifer	Seine-Maritime	Yes	No	> 100
27	Cauville-sur-Mer	Seine-Maritime	No	–	–
28	Cap de la Hève	Seine-Maritime	No	–	–
29	Château d'Orcher	Seine-Maritime	Yes	No	< 50
30	Le Hode	Seine-Maritime	Yes	Yes	> 100
31	Luc-sur-Mer	Calvados	No	–	–
32	Douvres-la-Délivrande	Calvados	No	–	–
33	Granville	Manche	Yes	Yes	> 1000
34	L'Echailler	Charente-Maritime	Yes	No	< 100
35	Mortagne-sur-Gironde	Charente-Maritime	Yes	Yes	< 50

− 0.376493) is indeed an inland town north of Caen, 3 km from the ocean, but it is not a suitable habitat for wild brassicas. The closest location by the ocean, Luc-sur-Mer (49.317856; − 0.355909) was also visited, but it presents a flat beach with no cliffs. Only later it occurred to us that 'Douvres' is the French name for the English town of Dover, where indeed *B. oleracea* is growing on the cliffs and where this specimen was possibly collected.

Genetic diversity

A total of 56 polymorphic markers were scored, with sizes ranging from 53 bp to 584 bp. The mean Nei's

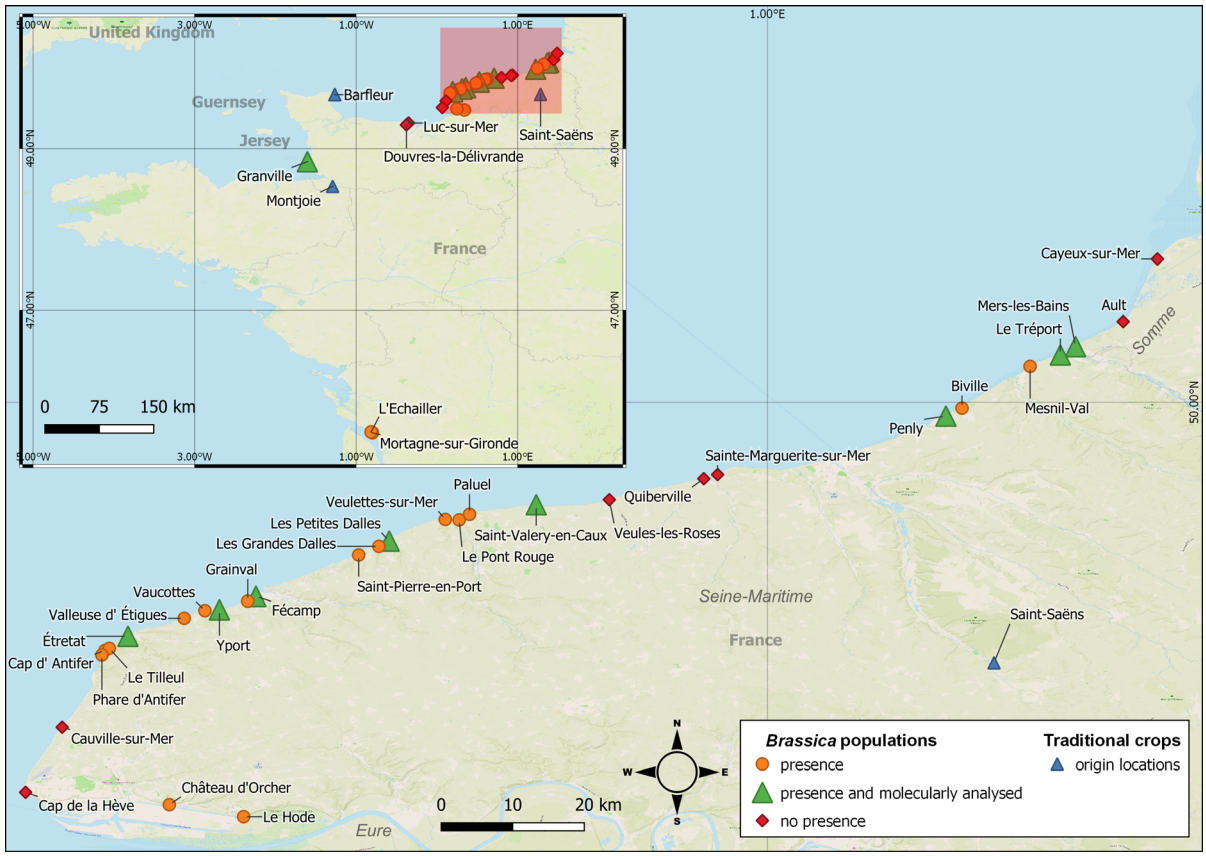

Fig. 1 Origin of the material collected and locations visited on the collecting mission to northwest France in 2012. Small triangles: origin locations of the cultivated crops. Circles and large triangles: locations with presence of *Brassica* populations (those corresponding to large triangles were molecularly analysed in the current paper). Squares: locations with no found presence of *Brassica* populations. Map created with QGIS Geographic Information System. Open Source Geospatial Foundation Project - http://qgis.osgeo.org

values of genetic diversity per population (H) are reported in Table 1, together with percentages of polymorphic loci and the number of private bands of each analysed population/accession within its respective group. No private bands could be detected to distinguish wild from cultivated populations, nor to characterise any of the wild populations from each other. Among the cultivated accessions, Saint-Saëns cabbage showed two private bands not present in any of the other tested cultivated accessions but present in the wild populations.

Mean genetic diversity of the wild populations was 0.247, ranging between 0.204 (Le Tréport) and 0.300 (Penly). This is comparable to the mean genetic diversity of the three Barfleur cauliflowers combined (0.247) and to the Montjoie leafy kale (0.244), while Saint-Saëns cabbage shows the highest value of H with 0.355.

Population structure

Principal coordinate analysis (PCoA) was used to visualize the differences between populations or between individuals. AMOVA results are reported in Table 3. The nine wild populations analysed did not show any distinguishable separation among them (Fig. 2). The AMOVA results show a variance of only 7% ($P < 0.001$) among the nine wild populations.

When populations are plotted with each symbol representing the average position of individuals of each population (Fig. 3), it is possible to identify a

Table 3 Analysis of molecular variance (AMOVA) (NP = number of populations, NI = number of individuals, P = probability)

Populations/Accessions	NP	NI	PhiPT	P	Variance within populations (%)	Variance among populations (%)
9 wild populations	9	173	0.066	0.001	93	7
4 wild populations (Étretat + Fécamp + St. Valery + Yport)	4	86	0.026	0.031	98	2
3 Barfleur cauliflowers	3	73	0.078	0.001	92	8
5 cauliflower, leafy kale and cabbage accessions	5	118	0.152	0.001	85	15
3 Barfleur cauliflower versus 1 cabbage versus 1 leafy kale	3	118	0.171	0.001	83	17
All wild versus all cultivated	2	291	0.059	0.001	94	6

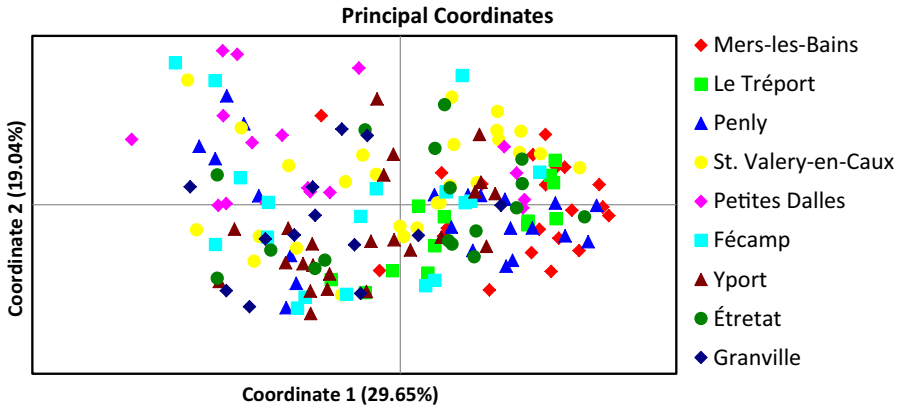

Fig. 2 PCoA of all wild populations. Symbols represent single individuals. Cumulated percentage of variation explained by the first two axes = 49%. For number of individuals in each population see Table 1

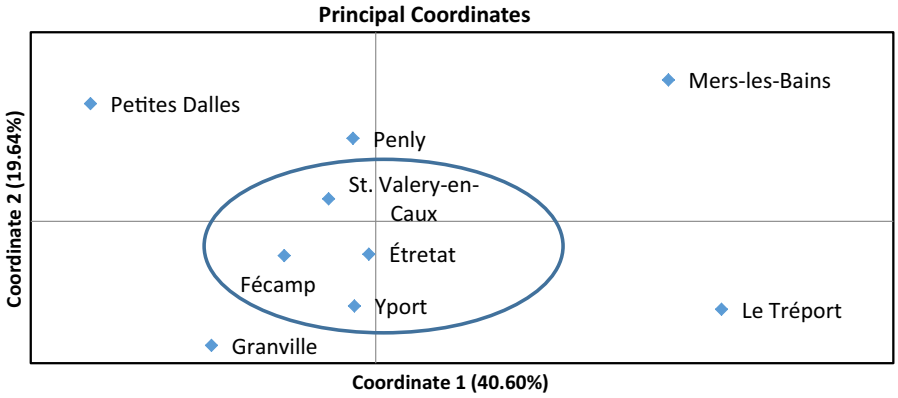

Fig. 3 PCoA of all wild populations. Symbols represent the average position of individuals of each population. Molecular variance among populations inside the ring is not significant ($P < 0.031$). Cumulated percentage of variation explained by the first two axes = 60%. For number of individuals in each population see Table 1

group of four populations (Étretat, Fécamp, Saint-Valery-en-Caux and Yport) with a non-significant 2% ($P < 0.031$) variance among them (Fig. 3 and Table 3).

A Mantel test was carried out on the nine wild populations, which indicated that there is no correlation between genetic and geographical distances ($R^2 = 0.0032$).

PCoA applied to the tested crop accessions confirmed that the three Barfleur cauliflower accessions are closely related (Fig. 4). AMOVA indicated 8% ($P < 0.001$) variation among these three accessions (Table 3). When all cultivated individuals were plotted on the PCoA (Fig. 5), it was evident that the kale, the cabbage and the three cauliflowers have ample areas of overlap, although Saint-Saëns cabbage is the

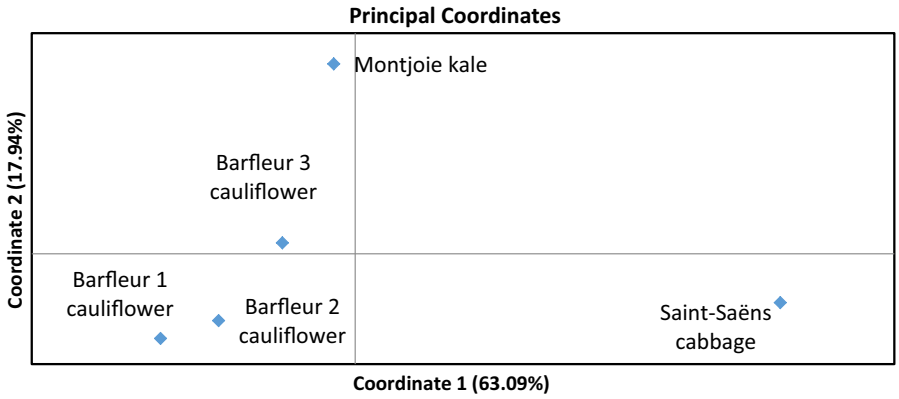

Fig. 4 PCoA of landraces. Symbols represent the average position of individuals of each accession. Cumulated percentage of variation explained by the first two axes = 81%. For number of individuals in each accession see Table 1

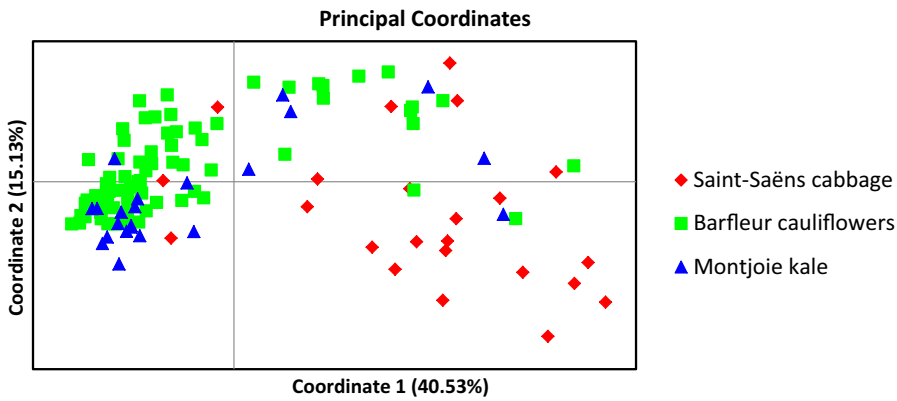

Fig. 5 PCoA of all landraces. Symbols represent single individuals. Cumulated percentage of variation explained by the first two axes = 56%. For number of individuals in each population/accession see Table 1

most genetically distinct of this group. AMOVA for these five accessions indicates 15% ($P < 0.001$) variance among them and 17% ($P < 0.001$) if the cauliflower accessions are combined (Table 3).

By plotting all samples (wild *B. oleracea* and landraces) together, the PCoA showed a large overlap of the various populations, indicating that it is not possible to clearly distinguish wild from cultivated accessions, based on the tested AFLPs (Fig. 6). The nine wild populations combined and compared to the crop accessions combined, showed a 6% ($P < 0.001$) variation among these two artificial populations.

When plotting the averaged positions of accessions (with all wild *B. oleracea* combined in one accession and the three Barfleur cauliflowers also combined) (Fig. 7), the cauliflowers are shown to be altogether genetically closer to the wild accessions than the Montjoie kale and the Saint-Saëns cabbage. The latter accession is confirmed to be genetically the most distant from all the other tested accessions.

Discussion

The present study has increased knowledge about the distribution and size of *B. oleracea* populations growing along the French Atlantic coast. As our collecting survey was rather quick and non-exhaustive, it is likely that other populations exist and that knowledge about the distribution area will be further refined by future investigations. The average level of genetic diversity of nine wild French populations ($H_A = 0.25$), ranging from 0.20 to 0.30, is comparable to the average value of eight wild British populations ($H_A = 0.26$), ranging from 0.18 to 0.33, previously detected by Watson-Jones et al. (2006), also with AFLPs. On the other hand, while a considerable differentiation among populations was measured in their study with an Fst value of 0.2257 over 103 loci, in our case variance among the nine populations was much lower (PhiPT = 0.07 – $P < 0.001$) and not significant (PhiPT = 0.02 – $P < 0.031$) if we consider four specific populations. The fact that these populations appear not to be genetically isolated from

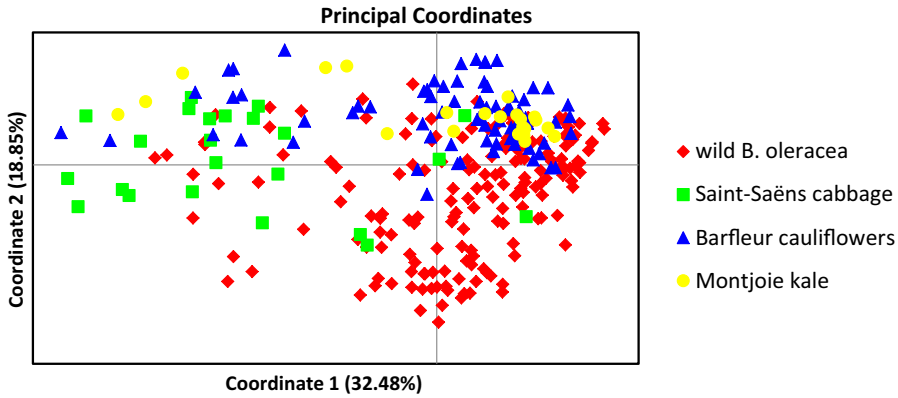

Fig. 6 PCoA of all wild *B. oleracea* compared to the landraces. Symbols represent single individuals. Cumulated percentage of variation explained by the first two axes = 51%. Number of individuals in each population/accession: wild *B. oleracea*: 173; Barfleur cauliflowers: 73; Saint-Saëns cabbage: 23; Montjoie kale: 22

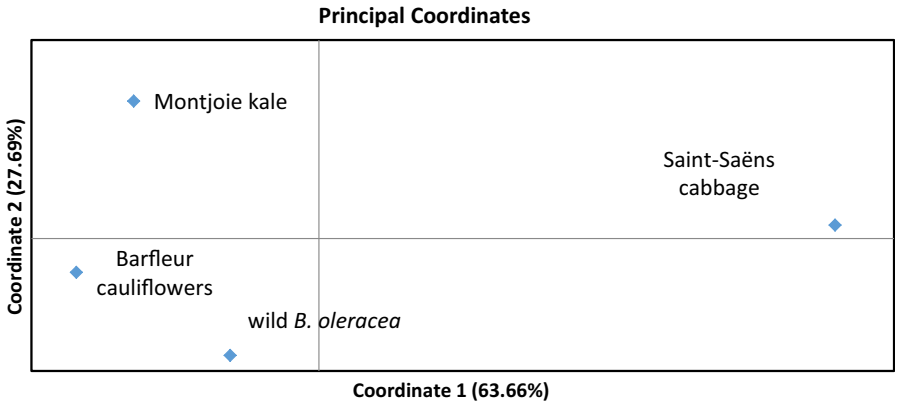

Fig. 7 PCoA of all wild *B. oleracea* compared to the landraces. Symbols represent the average position of individuals of each population. Cumulated percentage of variation explained by the first two axes = 91%. Number of individuals in each population/accession: wild *B. olera*cea: 173; Barfleur cauliflowers: 73; Saint-Saëns cabbage: 23; Montjoie kale: 22

each other is also supported by the lack of correlation between genetic diversity and geographic distances. A similar observation had already been obtained with isozymes by Lannér-Herrera et al. (1996) for four wild populations of *B. oleracea* from France. The observed pattern can be interpreted assuming the possibility of more frequent intercrossing among the French populations and/or that, after originating from a common source, these populations have not had sufficient time to differentiate since they settled in their current habitat. This seems especially the case of those four populations (Étretat, Yport, Fécamp and Saint-Valery), which are located at a maximum distance of 40 km from each other and show a negligible variation of 2% among themselves. Even for more distant populations, genetic drift is not occurring, or it is not sufficient to substantially differentiate any population from the others, but only to slightly reduce the genetic diversity of some of them (i.e. Le Tréport, H = 0.204).

The genetic differentiation (only 7% variance among nine populations) is remarkably low when compared to the results obtained by Watson-Jones et al. (2006) with the wild British populations. In this latter case, significant differentiation (F_{ST} = 0.22) was recorded among eight populations. This pattern is also different from what has been observed in related Mediterranean species such as *B. cretica* (Edh et al. 2007), *B. insularis* (Hurtrez-Boussès (1996), and *B. rupestris* (Maggioni et al. 2014), all showing the tendency of different populations to acquire much more distinct molecular patterns. Our results in terms of inter-population differentiation also differ from those of Lannér-Herrera et al. (1996) on similar French populations, but the sampling and type of markers are not fully comparable.

For conservation purposes, absence of private bands and of a population structure indicate that it is possible to capture most of the genetic diversity by sampling just one or two populations, for example Penly and Petites Dalles, which show the highest values of genetic diversity (H = 0.3).

The three cultivated crops, i.e. cabbage, cauliflower and kale, on average plot in different quadrants of the PCoA (Fig. 4). However, there is no clear-cut separation among the three crops when we look at the distribution of all individuals (Fig. 5). This pattern is in line with the difficulty to neatly separate all the morphotypes of *B. oleracea* with small number of molecular markers (< 200), either SSR (Tonguç and Griffiths 2004; Louarn et al. 2007) or AFLPs (Maggioni et al. 2017). More recently, the use of up to 6700 SNPs from selected loci allowed Cheng et al. (2016) to indicate a neat separation of seven morphotypes. However, the use of over 20,000 SNPs by other authors enabled good separation of various morphotypes, even though a few ambiguities in the separation between broccoli, cauliflower, cabbage and other kales still remained (Stansell et al. 2018; An et al. 2019). In our case, the possibility of intercrossing events across the three crops in the area of cultivation should also be considered. The Saint-Saëns cabbage is a traditional vegetable, appreciated for culinary and nutritional qualities. Its high genetic diversity is an additional important element that justifies attention for its conservation and possible use in plant breeding.

When wild *Brassica* populations are compared with landraces, it is not possible to genetically distinguish the two groups, which are confirmed to belong to the same genepool. Wild Atlantic populations have sometimes been considered the closer ancestors of the *B. oleracea* crops (Song et al. 1990; Hodgkin 1995; Gómez-Campo and Prakash 1999). Should this be the case, one would expect to find a reduction of diversity in the domesticated crops owing to a bottleneck effect (Gepts 2004). In our study, the crop accessions do not show an evident bottleneck effect in the PCoA (Fig. 6). Moreover, Saint-Saëns cabbage is even more diverse (H = 0.35) than the most diverse wild population (Penly, H = 0.30) (Table 1). Table 4 shows that the diversity of the French wild populations is in the same range as that of the British wild populations of *B. oleracea*. Based on the pattern in Table 4, genetic diversity of groups of accessions can show higher average values in cultivated crops than in the thriving wild populations so far analysed with the same methodology. This observation confirms the pattern already observed by Maggioni et al. (2014), which was interpreted by acknowledging that *Brassica* landraces and commercial varieties that are not maintained under strict selection tend to conserve (or to acquire) outstanding levels of genetic diversity, possibly as a result of open pollination. This pattern may seem counterintuitive, as domesticated crops present a domestication bottleneck if compared to their wild ancestors, but this is not necessarily true when the wild populations are not the true ancestors, or grow as isolated populations subject to genetic drift, or consist of feral populations derived from domesticated escapes. A domestication bottleneck between wild

Table 4 Mean Nei's genetic diversity (HA) of groups of accessions or populations, based on AFLPs, from previous literature

Species/Crop	Origin	Number of accessions or populations	Markers	H_A	References
White cabbages	The Netherlands	9	AFLP	0.12	van Hintum et al. (2007)
White cabbages	Worldwide	9	AFLP	0.15	van Hintum et al. (2007)
Leafy kales	Europewide	12	AFLP	0.20	Christensen et al. (2011)
Brassica rupestris	Calabria and Sicily	10	AFLP	0.21	Maggioni et al. (2014)
Leafy kales	Calabria and Sicily	22	AFLP	0.25	Maggioni et al. (2014)
Wild *Brassica oleracea*	France	9	AFLP	0.25	This paper
Wild *Brassica oleracea*	Britain	8	AFLP	0.26	Watson-Jones et al. (Watson-Jones et al. 2006)
Cauliflower, leafy kale and cabbage accessions	France	5	AFLP	0.27	This paper
Commercial broccoli and cauliflower	Italy	4	AFLP	0.28	Maggioni et al. (2014)

Atlantic *B. oleracea* and *B. oleracea* crops was excluded by Allender et al. (2007) and is also not visible from our data on the French populations. The low level of differentiation among the wild populations and between wild and cultivated crops seems to indicate that they are derived from a very similar genetic source. In fact, it seems unlikely that gene flow takes place between wild and cultivated crops, owing to their physical distance and different phenology. The occurrence of gene flow across most population sites tends to be excluded by the prevalent pollination carried out by insects and by the large majority of seeds falling at a very short distance from the mother plants (within 10 m from the base of the cliff) or on steep slopes up to about 200 m from the cliffs (Snogerup et al. 1990). The role of birds and wind in carrying seed over larger distances is not well known and might deserve more attention.

In a study on feral populations of *B. oleracea* in Britain and northern Spain, Mittel et al. (personal communication 2019) also found lower population structuring than might be expected. They suggested that populations could have been established by escapes from different cultivars. Further investigations comparing all the Atlantic populations and a wider range of old local landraces and improved cultivars, possibly using genotyping by sequencing methodologies, could help to clarify the source of the feral populations.

Conclusions

This study offers the highest detail ever presented about the distribution of wild *B. oleracea* populations along the Atlantic coast of France. These wild populations showed minimal differentiation from each other. Such a low level of genetic differentiation may be explained by a lack of sufficient physical or distance barriers to intercrossing, or more likely with a relatively recent origin of all the populations from a common source.

Traditional varieties representing three *B. oleracea* crops, commonly grown in the same area, are not fully distinguishable from the wild populations on a molecular level. Our results confirm that populations of *B. oleracea* growing in the wild belong to the same genepool as the cultivated *B. oleracea* crops. The level of genetic diversity of the wild populations is similar or lower than that of the cultivated crops. Therefore, the comparison between the French wild populations and the cultivated crops does not show any apparent sign of domestication bottleneck, which invites us to exclude these French populations as the likely source of original domestication events.

Populations with higher levels of genetic diversity have been identified and could be targeted for conservation measures (Penly and Petites Dalles). The Saint-Saëns cabbage shows the highest levels of genetic diversity of this study and can be singled out as the most interesting material for conservation and use for breeding.

Acknowledgements The authors wish to thank the Royal Physiographic Society of Lund, Sweden, for economic support to the molecular analysis, Sonja Ensslin for the creation of the maps and Olga Spellman (the Alliance of Bioversity International and CIAT Science Writing Services) for her critical proofreading and English language editing.

Funding The molecular analysis of this study was funded by the Royal Physiographic Society of Lund, Sweden.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Allender CJ, Allainguillaume J, Linn J, King GJ (2007) Simple sequence repeats reveal uneven distribution of genetic diversity in chloroplast genomes of *Brassica* L. and n = 9 wild relatives. Theor Appl Genet 114:609–618

An H, Qi X, Gaynor ML, Hao Y, Gebken SC, Mabry ME, McAlvay AC, Teakle GR, Conant GC, Barker MS, Fu T, Yi B, Pires JC (2019) Transcriptome and organellar sequencing highlights the complex origin and diversification of allotetraploid *Brassica napus*. Nat commun 10:2878. https://doi.org/10.1038/s41467-019-10757-1

Cheng F, Sun R, Hou X et al (2016) Subgenome parallel selection is associated with morphotype diversification and convergent crop domestication in *Brassica rapa* and *Brassica oleracea*. Nat Genet 48:1218–1224. https://doi.org/10.1038/ng.3634

Christensen S, von Bothmer R, Poulsen G, Maggioni L, Phillip M, Andersen BA, Bagger Jørgensen R (2011) AFLP analysis of genetic diversity in leafy kale (*Brassica oleracea* L. convar. *acephala* (DC.) Alef.) landraces, cultivars and wild populations in Europe. Genet Resour Crop Evol 58:657–666. https://doi.org/10.1007/s00122-010-9607-z

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Drenckhahn D (2017) Vorkommen des Atlantischen Wildkohls (*Brassica oleracea* L. subsp. *oleracea*) an den Kreidefelsen der Ostseeinsel Rügen. Deutschland. Forum geobotanicum 7:18–26. https://doi.org/10.3264/FG.2017.0116

Edh K, Widén B, Ceplitis A (2007) Nuclear and chloroplast microsatellites reveal extreme population differentiation and limited gene flow in the Aegean endemic *Brassica cretica* (Brassicaceae). Mol Ecol 16:4972–4983

Géhu JM (2009) Signification phytosociologique de *Brassica oleracea* ssp. *oleracea* sur les falaises atlantiques européennes. Fitosociologia 46:3–10

Gepts P (2004) Crop domestication as a long-term selection experiment. In: Janick J (ed) Plant breeding reviews, vol 24, part 2. Wiley, Hoboken, pp 1–44

Gillham ME (1970) Seed dispersal by birds. In: Perring FH (ed) The Flora of a changing Britain. Botanical Society of the British Isles, Hampton, pp 90–98

Gómez-Campo C, Prakash S (1999) Origin and domestication. In: Gómez-Campo C (ed) Biology of Brassica Coenospecies. Elsevier, Amsterdam, pp 33–52

Gómez-Campo C, Aguinagalde I, Ceresuela JL, Lázaro A, Martinez-Laborde JB, Parra-Quijano M, Simonetti E, Torres E, Tortosa ME (2005) An exploration of wild *Brassica oleracea* L. germplasm in Northern Spain. Genet Resour Crop Evol 52:7–13

Hodgkin T (1995) Cabbages, kales, etc. In: Smartt J, Simmonds NW (eds) Evolution of crop plants, 2nd edn. Longman, London, pp 76–82

Hurtrez-Boussès S (1996) Genetic differentiation among natural populations of the rare Corsican endemic *Brassica insularis*: implications for conservation guidelines. Biol Cons 76:25–30

Lannér C (1998) Relationship of wild *Brassica* species with chromosome number 2n = 18, based on comparison of the DNA sequence of the chloroplast intergenic region between *trnL* (UAA) and *trnF* (GAA). Can J Bot 76:228–237

Lannér-Herrera C, Gustafsson M, Fält AS, Bryngelsson T (1996) Diversity in natural populations of wild *Brassica oleracea* as estimated by isozyme and RAPD analysis. Genet Resour Crop Evol 43:13–23

Lázaro A, Aguinagalde I (1998a) Genetic diversity in *Brassica oleracea* L. (Cruciferae) and wild relatives (2n = 18) using isozymes. Ann Bot 82:821–828

Lázaro A, Aguinagalde I (1998b) Genetic diversity in *Brassica oleracea* L. (Cruciferae) and wild relatives (2n = 18) using RAPD markers. Ann Bot 82:829–833

Linnaeus C (1753) *Species Plantarum* 1. – Impensis Laurentii Salvii, Holmiae [L. Salvius, Stockholm]

Louarn S, Torp AM, Holme I, Andersen SB, Jensen BD (2007) Database derived microsatellite markers (SSRs) for cultivar differentiation in *Brassica oleracea*. Genet Resour Crop Evol 54:1717–1725. https://doi.org/10.1007/s10722-006-9181-6

Maggioni L (2015) Domestication of *Brassica oleracea* L. Dissertation, Swedish University of Agricultural Sciences. http://pub.epsilon.slu.se/12424/

Maggioni L, von Bothmer R, Poulsen G, Branca F, Bagger Jørgensen R (2014) Genetic diversity and population structure of leafy kale and *Brassica rupestris* Raf. in South Italy. Hereditas 151:145–158. https://doi.org/10.1111/hrd2.00058

Maggioni L, von Bothmer R, Poulsen G, Lipman E (2017) Domestication, diversity and use of *Brassica oleracea* L., based on ancient Greek and Latin texts. Genet Resour Crop Evol 65:1–23. https://doi.org/10.1007/s10722-017-0516-2

Marhold K (2011) Brassicaceae. In: Euro + Med Plantbase - the information resource for Euro-Mediterranean plant diversity. http://ww2.bgbm.org/EuroPlusMed/PTaxonDetail.asp?NameId=12940&PTRefFk=7200000. Accessed 10 Feb 2019

Mitchell ND (1976) The Status of *Brassica oleracea* L. subsp. *oleracea* (wild cabbage) in the British Isles. Watsonia 11:97–103

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Ray J (1686) *Historia Plantarum* 1. – Clark, London

Raybould AF, Mogg RJ, Clarke RT, Gliddon C, Gray A (1999) Variation and population structure at microsatellite and isozyme loci in wild cabbage (*Brassica oleracea* L.) in Dorset (UK). Genet Resour Crop Evol 46:351–360. https://doi.org/10.1023/A:1008658630440

Rigueiro Rodríguez A (1977) Trabajo Botánico sobre las Islas Cíes. ICONA, Madrid

Snogerup S, Gustafsson M, von Bothmer R (1990) *Brassica* sect. *Brassica* (Brassicaceae). I. Taxonomy and variation. Willdenowia 19:271–365

Song K, Osborn TC, Williams PH (1990) *Brassica* taxonomy based on nuclear restriction fragment length polymorphism (RFLPs). 3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*). Theor Appl Genet 79:497–506

Stansell Z, Hyma K, Fresnedo-Ramírez J, Sun Q, Mitchell S, Björkman T, Hua J (2018) Genotyping-by-sequencing of *Brassica oleracea* vegetables reveals unique phylogenetic patterns, population structure and domestication footprints. Hortic Res 5:38. https://doi.org/10.1038/s41438-018-0040-3

Tauleigne Gomes C (2002) Nouvelles citations pour la flore de l'archipel des Berlengas (Estremadura, Portugal). Acta Bot Malacitana 27:261–268

Tonguç M, Griffiths PD (2004) Genetic relationships of *Brassica* vegetables determined using database derived simple sequence repeats. Euphytica 137:193–201. https://doi.org/10.1023/B:EUPH.0000041577.84388.43

van Hintum ThJL, van de Wiel CCM, Visser DL, van Treuren R, Vosman B (2007) The distribution of genetic diversity in a *Brassica oleracea* gene bank collection related to the effects on diversity of regeneration, as measured with AFLPs. Theor Appl Genet 114:777–786. https://doi.org/10.1007/s00122-006-0456-2

von Bothmer R, Gustafsson M, Snogerup S (1995) *Brassica* sect. *Brassica* (Brassicaceae) II. Inter- and intraspecific crosses with cultivars of *B. oleracea*. Genet Resour Crop Evol 42:165–178

Warman EA (1999) Gene flow in wild cabbage. In: Gray AJ, Amijee F, Gliddon CJ (eds) Environmental impact of genetically modified crops. Department of the Environment, Transport and the Regions, London, pp 48–54

Watson-Jones SJ, Maxted N, Ford-Lloyd BV (2006) Population baseline data for monitoring genetic diversity loss for 2010: a case study for *Brassica* species in the UK. Biol Cons 132:490–499