Chloroplastic and nuclear diversity of wild beets at a large geographical scale: Insights into the evolutionary history of the Beta section

Pascal Touzet | Sarah Villain | Laetitia Buret | Hélène Martin | Anne-Catherine Holl | Céline Poux | Joël Cuguen

Univ. Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, Lille, France

Correspondence
Pascal Touzet, Evolution Ecologie Paléontologie, UMR CNRS 8198, Université de Lille1, Villeneuve d’Ascq Cedex, France
Email: pascal.touzet@univ-lille1.fr

Funding information
Région Nord-Pas de Calais; European Regional Development Fund

Abstract

Historical demographic processes and mating systems are believed to be major factors in the shaping of the intraspecies genetic diversity of plants. Among Caryophyllales, the Beta section of the genus Beta, within the Amaranthaceae/Chenopodiaceae alliance, is an interesting study model with species and subspecies (Beta macrocarpa, Beta patula, Beta vulgaris maritima and B.v. adanensis) differing in geographical distribution and mating system. In addition, one of the species, B. macrocarpa, mainly diploid, varies in its level of ploidy with a tetraploid cytotype described in the Canary Islands and in Portugal. In this study, we analyzed the nucleotide diversity of chloroplastic and nuclear sequences on a representative sampling of species and subspecies of the Beta section (except B. patula). Our objectives were (1) to assess their genetic relationships through phylogenetic and multivariate analyses, (2) relate their genetic diversity to their mating system, and (3) reconsider the ploidy status and the origin of the Canarian Beta macrocarpa.

KEYWORDS
allo-polyploidy, Beta genus, mating systems, phylogeny

1 | INTRODUCTION

The nature of forces that shape genetic diversity of species is a long-standing question in evolutionary biology (Leffler et al., 2012). Both historical demographic process that occurred during glaciation periods and life history traits are generally admitted to be the major factors influencing the present intraspecies genetic diversity. In plants, mating systems are believed to be of main importance, in particular the frequent transition to self-fertility that is expected to affect both neutral diversity and the efficacy of selection (Glémín, 2007; Glémín, Bazin, & Charlesworth, 2006). Empirical studies in a set of species have partially confirmed these theoretical expectations (reviewed in Glémín & Galtier, 2012 and in Castric, Billiard, & Vekemans, 2013). Another evolutionary mechanism influencing plant species diversification is polyploidization. This can occur after interspecies hybridization (allo-polyploidy) or intraspecific genome duplication (autopolyploidy) (reviewed by Soltis, Marchant, Van de Peer, & Soltis, 2015). It has been generally believed that allopolyploids are more frequent than autopolyploids thanks to the expected gain in fitness of hybrids combining two diverged genomes and thus enlarging their ability of conquering new environments (Abbott et al., 2013). However, autopolyploid occurrence seems to have been underestimated as it appears to be as frequent as allopolyploids, partly due to the difficulty in phenotypically distinguishing them from their diploid counterparts (Barker, Arrigo, Baniaga, Li, & Levin, 2016).
The Betoideae constitute a small subfamily of the Amaranthaceae/Chenopodiaceae alliance that is characterized by a unique fruit type, a capsule that normally opens with a circumscissile lid (Kadereit, Hohmann, & Kadereit, 2006). Within this subfamily, two groups have been defined: Hablitzieae and Beteae which is composed by a single genus, Beta. This genus is partitioned in two sections: sect. Corollinae (including the previous section Nanae) and sect. Beta (see Biancardi, Panella, & Lewellen, 2012 for the recent evolution of Beta taxonomy). This last section is composed of B. macrocarpa, B. patula and the species complex B. vulgaris, within which can be found wild forms (B. v. maritima, B. v. adanensis), cultivars (B. v. vulgaris) and weeds, a hybrid between B. v. maritima and B. v. vulgaris (Desplanque et al., 1999).

Species of the Beta section differ in their respective geographical distribution. Beta patula is endemic to two islets of the Madeira Island and one islet at Desertas Islands (Romeiras et al., 2016). Beta macrocarpa has been described as two different cytotypes: one diploid cytotype distributed from Portugal to Turkey, along the Mediterranean Basin, and a tetraploid one found in the Canary Islands (Buttler, 1977) and in Portugal (Castro, Romeiras, Castro, Duarte, & Loureiro, 2013). Within Beta vulgaris, while B.v. maritima populations are found on a large geographical area, along both the Atlantic coasts of Western Europe and most of the Mediterranean coast, B. v. adanensis is restricted in the eastern part of the Mediterranean Basin (Aegean islands, Turkey and Syria). In addition, subspecies of the Beta section differ in their mating system: B.v. maritima is allogamous and self-incompatible, while B. macrocarpa and B. v. adanensis have been described as self-compatible (Bruun et al., 1995; Letschert, 1993).

This section exhibiting variation in breeding systems, ploidy but also life history traits (Hautekèete, Piquot, & Van Dijk, 2001; Letschert, 1993) is therefore an interesting group to infer their impact on genetic diversity. Former genetic studies have focused on the B. vulgaris species complex (Desplanque et al., 1999, 2000; Letschert, 1993; Nishizawa, Kubo, & Mikami, 2000; Nishizawa, Mikami, & Kubo, 2007), or more specifically on B.v. maritima as the main genetic resource of cultivated beet (Andrello, Henry, Devaux, Desprez, & Manel, 2016; Andrello et al., 2017; Cuguen et al., 1994; Fénart, Touzet, Arnaud, & Cuguen, 2006; Fievet, Touzet, Arnaud, & Cuguen, 2007; Leys et al., 2014; Raybould, Mogg, & Clarke, 1996; Raybould, Mogg, Gliddon, Thorpe, & Clarke, 1998; Richards, Reeves, Fenwick, & Panella, 2014), while some information is available at the section level (Jung et al., 1993; Shen, Newbury, & Ford-Loyd, 1996; Letschert, 1993; Hohmann, Kadereit, & Kadereit, 2006; Kadereit et al., 2006; Andrello et al., 2016, 2017; Romeiras et al., 2016). Therefore, in the present study, we analyze the nucleotide diversity of a representative sampling of species and subspecies of the Beta section (except B. patula) at chloroplastic and nuclear loci in order to: (1) assess their genetic relationships through phylogenetic and multivariate analyses, (2) relate species/subspecies diversity of the section to their mating system, and (3) reconsider the ploidy status and the origin of the Canarian Beta macrocarpa.

2 | MATERIAL AND METHODS

2.1 | Plant species and sampling

Seeds from the Beta section were obtained from the Federal Centre for Breeding Research on Cultivated Plants of Braunschweig, from the University of Birmingham and from our lab’s collection. Details on sampling are given in Table 1. For the study of chloroplastic and nuclear nucleotide diversity, a total of 33 individuals of Beta v. maritima, 12 Beta v. adanensis and 12 Beta macrocarpa were analyzed (Figure 1). These accessions were chosen on the basis of their geographical location. For each location, DNA was extracted from a single individual.

Additional samples from the Beta genus, belonging to the Corollinae section, were sequenced in order to root the phylogenetic trees: Beta lomatogona (PI198401), Beta macrorhiza (BETA 545) (kindly provided by Lothar Freese, Julius Kühn-Institut, Quedlinburg, Germany) and Beta nana (kindly provided by Lee Panella, USDA, Fort Collins, USA).

2.2 | DNA amplification and sequencing

The DNA extraction from dried leaf tissue was carried out with a Nucleospin® 96Plant kit (Macherey-Nagel) on a MicroLab® 8 Star robot (Hamilton).

2.2.1 | cpDNA sequences

Four cpDNA regions were selected for sequencing: the trnK intron (K1K2) including the matK gene, the trnD-trnT intergenic spacer (DT), the trnL-trnF intergenic spacer (LF), and the 5′ part of the intergenic spacer HK ranging between trnH and psbA. On account of its size (about 1,900 base pairs [bp]), the K1K2 region was amplified in two overlapping fragments.

The set of primers (forward/reverse) used was 5′-GTGTCGCGGGATTGCAA-3′/5′-ATTAGGGCATCCATTAGTA-3′ for the first part of K1K2 (annealing temperature [T_a] = 54°C for the Beta section/58°C for the Corollinae section) (modified from Grivet & Petit, 2003) and 5′-CTAGCAAAAGAACTGGAAG-3′/5′-GGATTTCTAACCATCTGGTTT-3′ for the second part of K1K2 (T_a = 50°C/58°C); 5′-ACCAATTGAACTAACTCCCC-3′/5′-CTACCCACAGATGATAAGG-3′ for DT (T_a = 56.5°C/58°C) (Grivet & Petit, 2003); 5′-GGTTCAAGTCCCTCTATCC-3′/5′-ATTGAACCTGTGACACAGG-3′ for LF (T_a = 57.5°C) (Taborlet et al., 1991); 5′-CGACAAAAATACCATGACC-3′/5′-GCTATGCATGTCCTGTG-3′ for HK (T_a = 57°C). This last fragment could not be amplified for the 3 Corollinae species.

PCR amplification was performed in a 25 μl mix containing 25 ng of DNA template, 3 mmol/L of MgCl_2, 1.5 μmol/L of Buffer 10X (Perkin-Elmer, Norwalk, CT, USA), 0.2 μmol/L of each primer, 200 μmol/L of each dNTP, and 0.625 U/μl of hot start Taq polymerase (AmpliTaq Gold, Perkin-Elmer, Norwalk, CT, USA). PCR mixture underwent the following conditions on a 9700 thermal cycler (Perkin-Elmer, Norwalk, CT, USA): 12-min denaturing at 94°C, 40 cycles of 30″ denaturing at 94°C, 45″ annealing at T_a (see above) and from 1 to 2 min extension (depending on the fragment length) at 72°C and a final extension step at 72°C for 10 min,
**Table 1** Localities of samples. The species, the sample numbers, the site of origin (country and location), the IDDBNR accession number (unique identification number assigned to an accession by the Beta International Database) are given, as well as the donor institution: BGRC: Braunschweig Genetic Resources, Birm.: University of Birmingham, Lille: our lab collection.

| Species          | Sample number | Country | Location         | IDDBNR | Donor       |
|------------------|---------------|---------|------------------|--------|-------------|
| *B. v. maritima* | 1             | Ireland | Sligo            | 5905   | BGRC        |
|                  | 2             | Great Britain | Scarborough   | 5915   | BGRC        |
|                  | 3             | Great Britain | Ramsgate      |        | Lille       |
|                  | 4             | Great Britain | Land's end     |        | Lille       |
|                  | 5             | Netherlands | Zwin            |        | Lille       |
|                  | 6             | France    | Roscoff         |        | Lille       |
|                  | 7             | France    | Sables d'Olonne |        | Lille       |
|                  | 8             | France    | Erromardie      |        | Lille       |
|                  | 9             | Spain     | Foz             |        | Lille       |
|                  | 10            | Spain     | Punta Fouxeria  |        | Lille       |
|                  | 11            | Spain     | Playa de la Lanzada |    | Lille       |
|                  | 12            | Portugal | Obidos          | 7069   | BGRC        |
|                  | 13            | Morocco   | Casablanca      | 8550   | BGRC        |
|                  | 14            | Morocco   | Essaouira       | 8560   | BGRC        |
|                  | 15            | Morocco   | Safi            | 8556   | BGRC        |
|                  | 16            | Portugal | Madeira         | 6069   | BGRC        |
|                  | 17            | Portugal | Ponto do Parvo  |        | Lille       |
|                  | 18            | Spain     | Los Arenetes    |        | Lille       |
|                  | 19            | France    | Bages           |        | Lille       |
|                  | 20            | Italy     | Fosso d'Arno, Toscana | 9452 | BGRC        |
|                  | 21            | Italy     | Lazio           | 9461   | BGRC        |
|                  | 22            | Italy     | Sicily          | 2205   | BGRC        |
|                  | 23            | Malta     |                 | 8615   | BGRC        |
|                  | 24            | Tunisia   | Sfax            | 3542   | BGRC        |
|                  | 25            | Tunisia   | Bor. Djilidj    | 415    | BGRC        |
|                  | 26            | Italy     | Veneto          | 9481   | BGRC        |
|                  | 27            | Croatia   | Istria          | 6952   | BGRC        |
|                  | 28            | Greece    | Levkas          | 139    | BGRC        |
|                  | 29            | Greece    | Khalkidhiki     | 208    | BGRC        |
|                  | 30            | Greece    | Kissamos, Crete |        | Lille       |
|                  | 31            | Greece    | Lesbos          |        | Lille       |
|                  | 32            | Egypt     | Matruh          | 9742   | BGRC        |
|                  | 33            | Turkey    | Hatay           | 8440   | BGRC        |
| *B. v. adanensis* | a1         | Greece    | Samos           |        | Lille       |
|                  | a2           | Turkey    | Canakkale       | 3010   | BGRC        |
|                  | a3           | Greece    | Lesbos          |        | Lille       |
|                  | a4           | Turkey    | Izmir           | 3016   | BGRC        |
|                  | a5           | Greece    | Chios           |        | Lille       |
|                  | a6           | Greece    | Kos             |        | Lille       |
|                  | a7           | Greece    | Kokinos         |        | Lille       |
|                  | a8           | Cyprus    | Paphos          | 7119   | BGRC        |
|                  | a9           | Turkey    | Aydin           | 8462   | BGRC        |
|                  | a10          | Israel    | Zomet Lakhish   | 3798   | BGRC        |
|                  | a11          | Iran      | Sorkan, Khuzestan | 8623 | BGRC        |
|                  | a12          | Iran      | Minab, Hormozgan | 8622  | BGRC        |

(Continued)
after 40 cycles. The PCR products were then purified using a QIAquick PCR Purification Kit (Qiagen, Inc., Valencia, CA, USA) and directly sequenced with an ABI Prism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Norwalk, CT, USA). Sequence data were obtained on a 3100-Avant Genetic Analyser (Applied Biosystems).

2.2.2 Nuclear DNA sequences

For nuclear analysis, three genes, largely used in phylogenetic studies, were partially sequenced: the alcohol dehydrogenase (adhl) with primers 5′-TGTCCTGCCTGTTTTCACTG-3′/5′-TACTGCTCCTAGGCCGAAA-3′ (Ta = 61°C/53°C) anchored in exons 1 and 2, the chlorophyll a/b-binding protein cab11 with primers 5′-CTTCATTAGCTGAGGAAC-3′/5′-GCTCTGACATTGGAAACCC-3′ (Ta = 55°C) anchored in exons 1 and 2, and the ITS region (internal transcribed spacers ITS1 and ITS2 of nuclear ribosomal DNA and the 5.8S rRNA gene) with primers 5′-GGAAGTAAAAGTCGTAACAAGG-3′/5′-TCCTCCGCTATATGATGC-3′ (Ta = 53°C) anchored in ITS1 and ITS2 (White et al., 1990). Both PCR and sequencing were done as described in the cpDNA section. PCR products were directly sequenced for the autogamous diploid species B. v. adanensis and B. macrocarpa and for the Corollinae species. For the outcrossers B. v. maritima and tetraploid B. macrocarpa, PCR products were cloned into pCR2.1-TOPO using TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) before sequencing. A minimum of six clones was sequenced to reliably identify both haplotypes and examine PCR-generated errors due to nucleotide misincorporation and/or recombination.

All sequences generated in the present study have been registered in Genbank (KP747713–KP748171).

2.3 Data analyses

DNA sequences were assembled with SEAVIEW (Gouy, Guindon, & Gascuel, 2010), aligned with MAFFT v.7 (Katoh & Standley, 2013) and manually checked and cleaned using Gblocks (Castresana, 2000) when necessary (Cab11). The cpDNA alignment with outgroups (for the phylogenetic reconstruction) or without outgroups (for diversity analyses) displayed a total size of 3742 bp and 3752 bp, respectively (K1K2: 1892 bp/1905 bp, DT: 914 bp/911 bp, LF: 301 bp and HK: 635 bp). For the nuclear alignment, the discrepancy between both alignments (with and without outgroups) is mainly due to Cab11 for which the intron could not be aligned between Beta and Corollinae sections. The Adh alignment was 349 bp, the Cab11 displayed 797 bp without outgroup and 1140 bp when outgroups were included (692 bp after removing the poorly aligned sites), and the ITS region was 668 bp/674 bp long (without/with outgroup).

2.3.1 Phylogenetic and haplotype network reconstructions

The alignment resulted in a dataset of 3,742 bp for the chloroplastic dataset (K1K2, LF, DT and KH) and of 1715 bp (adh, cab11 and ITS) for the nuclear alignment. Phylogenetic reconstructions based on both chloroplastic and nuclear concatenated datasets and on each nuclear gene separately were performed by maximum likelihood (ML) with PHYML v.3.0 (Guindon et al., 2010) and by Bayesian analyses with MrBayes, version 3.2.2 (Ronquist et al., 2012). For the individual nuclear genes analyses, heterozygotes samples were represented by both alleles. For the concatenated analysis, each individual was represented by only one sequence per gene because alleles from different nuclear loci cannot be phased; heterozygous sites were therefore encoded according to the DNA ambiguity code. However, all alleles from the Beta macrocarpa 4× individuals were kept in the concatenated analysis as they obviously were from different origins.

The best fitting model of sequences evolution was selected from the BIC (Bayesian Information Criterion) output of jMODELTEST, version 2.1.3 (Darriba, Taboada, Doallo, & Posada, 2012) for each data partition. For the Bayesian analysis of the concatenated chloroplastic dataset, four partitions corresponding to the four genes were defined. Similarly for the concatenated nuclear dataset, five partitions were
considered: they correspond to the intronic and exonic regions of the adh and cab11 genes and to ITS. For the ML analyses, datasets, concatenated or not, were considered as one partition.

Analyses with MrBAYES were done as follows: two runs of four Markov chains were calculated simultaneously for 1,000,000 to 5,000,000 generations depending on the dataset, with initial equal probabilities for all trees and a random starting tree. Trees were sampled each 100 generations, and the consensus tree with posterior probabilities (PP) was calculated after removal of the first 25% to 50% (according to the analysis) of the total number of generated trees (according to the analysis). The average standard deviation of split frequencies between the two independent runs was lower than 0.01.

PopART v1.7 (Leigh & Bryant, 2015) was used to construct the chloroplastic haplotype TCS network.

### 2.3.2 Principal component analysis

In order to assess the existence of genetic clusters within the Beta section, we conducted a principal component analysis (PCA) on the concatenated nuclear sequences of all individuals except for the samples B. v. maritima 6 and B. v. adanensis a10 (adh sequence was missing for 6, and cab11 sequence for a10) using adegenet R package (Jombart, 2008; R Core Team Development 2014).

### 2.3.3 Statistical analyses—nucleotide diversity parameters

For each species/subspecies of the Beta section, we estimated the nucleotide diversity both as \( \pi \), the average number of nucleotide differences per site between a pair of randomly chosen sequences (Nei,
(a) GBR
   a7 TUR
   a8 GRC
   a9 CYP
   a11 IRN
   a12 GRC
   a13 GRC
   a14 GRC
   a15 GRC
   a16 HRV
   a18 GRC
   a19 ESP
   a20 ESP
   a21 ESP
   a22 ESP
   a23 ESP
   a24 ESP
   a25 ESP
   a26 ESP
   a27 ESP
   a28 ESP
   a29 ESP
   a30 ESP
   a31 ESP
   a32 ESP
   a33 ESP
   a34 ESP
   a35 ESP
   a36 ESP
   a37 ESP
   a38 ESP
   a39 ESP
   a40 ESP
   a41 ESP
   a42 ESP
   a43 ESP
   a44 ESP
   a45 ESP
   a46 ESP
   a47 ESP
   a48 ESP
   a49 ESP
   a50 ESP
   a51 ESP
   a52 ESP
   a53 ESP
   a54 ESP
   a55 ESP
   a56 ESP
   a57 ESP
   a58 ESP
   a59 ESP
   a60 ESP
   a61 ESP
   a62 ESP
   a63 ESP
   a64 ESP
   a65 ESP
   a66 ESP
   a67 ESP
   a68 ESP
   a69 ESP
   a70 ESP
   a71 ESP
   a72 ESP
   a73 ESP
   a74 ESP
   a75 ESP
   a76 ESP
   a77 ESP
   a78 ESP
   a79 ESP
   a80 ESP
   a81 ESP
   a82 ESP
   a83 ESP
   a84 ESP
   a85 ESP
   a86 ESP
   a87 ESP
   a88 ESP
   a89 ESP
   a90 ESP
   a91 ESP
   a92 ESP
   a93 ESP
   a94 ESP
   a95 ESP
   a96 ESP
   a97 ESP
   a98 ESP
   a99 ESP
   a100 ESP

(b) IRL
   b1 GBR
   b2 GBR
   b3 GRC
   b4 GRC
   b5 NLD
   b6 FRA
   b7 FRA
   b8 FRA
   b9 ESP
   b10 ESP
   b11 ESP
   b12 PRT
   b13 ESP
   b14 PRT
   b15 MAR
   b16 MAR
   b17 PRT
   b18 ESP
   b19 FRA
   b20 ESP
   b21 ESP
   b22 ESP
   b23 ESP
   b24 ESP
   b25 TUR
   b26 TUR
   b27 HRV
   b28 GRC
   b29 GRC
   b30 GRC
   b31 GRC
   b32 GRC
   b33 ESP
   b34 ESP
   b35 ESP
   b36 ESP
   b37 ESP
   b38 ESP
   b39 ESP
   b40 ESP
   b41 ESP
   b42 ESP
   b43 ESP
   b44 ESP
   b45 ESP
   b46 ESP
   b47 ESP
   b48 ESP
   b49 ESP
   b50 ESP
   b51 ESP
   b52 ESP
   b53 ESP
   b54 ESP
   b55 ESP
   b56 ESP
   b57 ESP
   b58 ESP
   b59 ESP
   b60 ESP
   b61 ESP
   b62 ESP
   b63 ESP
   b64 ESP
   b65 ESP
   b66 ESP
   b67 ESP
   b68 ESP
   b69 ESP
   b70 ESP
   b71 ESP
   b72 ESP
   b73 ESP
   b74 ESP
   b75 ESP
   b76 ESP
   b77 ESP
   b78 ESP
   b79 ESP
   b80 ESP
   b81 ESP
   b82 ESP
   b83 ESP
   b84 ESP
   b85 ESP
   b86 ESP
   b87 ESP
   b88 ESP
   b89 ESP
   b90 ESP
   b91 ESP
   b92 ESP
   b93 ESP
   b94 ESP
   b95 ESP
   b96 ESP
   b97 ESP
   b98 ESP
   b99 ESP
   b100 ESP

B. v. maritima
B. v. adanensis
B. macrocarpa
1987), and as Watterson’s \( \theta_w \) (Watterson, 1975). Among species/subspecies of the Beta section, we calculated shared and fixed polymorphisms and the nucleotide divergence (Dxy), using DnaSP version 5 (Librado & Rozas, 2009).

### 3 | RESULTS

#### 3.1 | Phylogenetic analyses

The concatenated chloroplastic sequences from the 57 samples of the Beta section and 3 samples from the Corollinae section enabled us to generate a rooted phylogenetic tree that revealed (Figure 2a) several clades however with low bootstrap (BP) and posterior probabilities (PP): (1) a clade composed of all Beta macrocarpa samples except two samples from the Canary Islands (samples from islands Tenerife and Gran Canaria—m3 and m4), (2) a large clade within which we found most of the B. v. adanensis samples and some Eastern B. v. maritima samples as well as the two B. macrocarpa samples from the Canary Islands, and (iii) a large clade composed mainly of Western B. v. maritima samples. The remaining samples were not assigned to a particular clade (see also the haplotype network, Figure S1).

The low bootstrap values reflect the extremely low diversity level displayed by the sequences.

The three nuclear loci sequenced on the same samples (partial sequences of adh, cab11 and the ITS region) led to a nuclear consensus phylogenetic tree (Figure 2b) on which we found this time with a good phylogenetic support: (1) a clear distinct B. macrocarpa clade (BP = 93, PP = 1.00) and (2) a large clade composed of B. v. maritima and B. v. adanensis (BP = 93, PP = 1.00). All B. v. adanensis sequences were regrouped in a clade (BP = 67, PP = 0.99) displaying as well one B. v. maritima individual.

Notably, only one allele was found for the B. v. adanensis and B. macrocarpa samples, as expected for autogamous species, whereas two alleles could be found for the allogamous B. v. maritima samples. Two alleles per individual were also found for the two Canary B. macrocarpa individuals (m3 and m4) with one allele belonging to the B. macrocarpa clade and the other to the B. v. maritima clade. These two individuals are most likely tetraploid, resulting from the hybridization between B. v. maritima and B. macrocarpa. One-locus trees can be found in the supplementary data (Figures S2–S4).

Overall, chloroplastic and nuclear phylogenetic trees showed that (1) within the B. vulgaris species, B. v. maritima exhibited the largest diversity, while B. v. adanensis represented a sublineage within the B. v. maritima clade, (2) the B. macrocarpa samples formed a distinct monophyletic lineage from the B. vulgaris subspecies (except for the distinctive m3 and m4 samples), and (3) the two B. macrocarpa samples from the Canary Islands (m3 and m4) displayed a B. v. maritima chloroplastic haplotype and exhibit a hybrid pattern with two nuclear alleles, one maritima-like and one macrocarpa-like (Figure 2a,b).

#### 3.2 | Principal component analysis

The principal component analysis (PCA) was consistent with the phylogenies (Figure 3). The first axis of the PCA, representing 26.6% of the variance, separated B. macrocarpa from B. vulgaris subspecies. As expected by the phylogenetic analysis, the m3 and m4 samples were at an intermediate position between the B. macrocarpa cluster and the B. v. maritima one, confirming their hybrid status. The diversity of B. v. maritima was spread along the second axis that explained 15.6% of the variance, with no clear geographical pattern, while B. v. adanensis accessions remained aggregated.

#### 3.3 | Nucleotide diversity of the Beta section

The representative distribution of the sampling enabled us to measure the overall nucleotide diversity of the members of the Beta section, at both chloroplastic and nuclear levels (Table 2). At the species/inastraspecies level, for both genomes, B. v. maritima exhibited the highest level of diversity, followed by B. v. adanensis and last B. macrocarpa displaying the more conserved sequences. This result was obtained whatever parameter was considered (except for \( \pi \) calculated on adh) (Table 2). Note that we did not include in the analyses the two B. macrocarpa from Tenerife and Gran Canaria islands since they are most likely allo-tetraploid. We will call them 4× in the rest of the article for the purpose of clarity.

At the chloroplastic level, B. v. maritima exhibited 3 times as many haplotypes as B. v. adanensis and 5 times as many haplotypes as B. macrocarpa. The same pattern was also observed when estimating nucleotide diversity, which differed almost in an order of magnitude between B. v. maritima and B. macrocarpa.

At the nuclear level, once again, B. v. maritima was the most polymorphic species/subspecies at any analysed locus. As previously mentioned, B. v. adanensis and B. macrocarpa were homozygous for every analysed locus. Note that it was also the case for the two 4× B.
macrocarpa individuals as we found only one allele for the locus coming from the B. v. maritima genome, and one allele for the locus coming from the B. macrocarpa genome, suggesting that 4× B. macrocarpa may preferentially reproduce by selfing.

For B. v. maritima, cab11 was the most polymorphic locus with 25 segregating alleles, then adh with 9 alleles, and then ITS with only 2 alleles. For B. v. adanensis, polymorphism was reduced with two alleles on adh and ITS, and only one allele on cab11. Strikingly, B. macrocarpa was fixed on each analysed locus despite the large geographical distribution of the species sample.

In parallel of the phylogenetic and PCA analyses, the level of divergence between the members of the Beta section can be described by assessing the number of private and shared polymorphisms among members, as well as the number of fixed differences (Table 3). Accordingly, B. macrocarpa represents a distinct genetic pool from Beta vulgaris, as it exhibits fixed differences at both genomic compartments with B. v. maritima and B. v. adanensis, while B. v. maritima and B. v. adanensis exhibit none.

The same pattern is less obvious when considering the nucleotide divergence among Beta section members (Dxy, Table 3). B. macrocarpa divergence with B. v. maritima or B. v. adanensis at the chloroplastic level is comparable with the divergence among subspecies of Beta vulgaris at the chloroplastic level, but is higher at the nuclear loci, especially at the ITS locus with a level of nucleotide divergence that is 5 time as

**TABLE 2** Species diversity of the Beta section. At each locus, chloroplastic (cp) and nuclear loci (Adh, Cab11, and ITS) and for each species/subspecies are given: the number of populations per species (Pop) and sequences (Seq), number of haplotypes, number of segregating sites, diversity per site estimated from the total number of mutations (Θw), diversity as the average number of nucleotide differences per site between a pair of randomly chosen sequences (π) with standard deviation (SD).

| Locus | Species        | Pop/Seq | Length (bp) | Number of haplotypes | Segregating sites | Θw ± SD (×10⁻³) | π ± SD (×10⁻³) |
|-------|----------------|---------|-------------|----------------------|-------------------|-----------------|----------------|
| Cp    | B. v. maritima | 33/33   | 3,752       | 16                   | 16                | 1.05 ± 0.40     | 0.97 ± 0.08     |
|       | B. v. adanensis| 12/12   | 3,752       | 5                    | 5                 | 0.44 ± 0.25     | 0.26 ± 0.09     |
|       | B. macrocarpa  | 10/10   | 3,752       | 3                    | 2                 | 0.19 ± 0.14     | 0.11 ± 0.05     |
| Adh   | B. v. maritima | 31/62   | 349         | 9                    | 8                 | 4.88 ± 2.10     | 1.26 ± 0.28     |
|       | B. v. adanensis| 12/24   | 349         | 2                    | 1                 | 0.77 ± 0.77     | 1.49 ± 0.09     |
|       | B. macrocarpa  | 6/12    | 349         | 1                    | 0                 | 0               | 0              |
| Cab11 | B. v. maritima | 32/64   | 797         | 25                   | 37                | 10.53 ± 3.21    | 9.61 ± 0.53     |
|       | B. v. adanensis| 11/22   | 797         | 1                    | 0                 | 0               | 0              |
|       | B. macrocarpa  | 7/14    | 797         | 1                    | 0                 | 0               | 0              |
| ITS   | B. v. maritima | 32/64   | 669         | 2                    | 3                 | 0.95 ± 0.58     | 2.28 ± 0.05     |
|       | B. v. adanensis| 12/24   | 669         | 2                    | 1                 | 0.40 ± 0.40     | 0.58 ± 0.14     |
|       | B. macrocarpa  | 7/14    | 669         | 1                    | 0                 | 0               | 0              |

**TABLE 3** The number of species-specific polymorphisms, shared polymorphisms, fixed differences, and nucleotide divergence (Dxy) (Jukes-Cantor) between Beta species.

| Locus | Species comparison (species1/species2) | Species1 only | Species2 only | Shared | Fixed | Dxy ± SD (×10⁻³) |
|-------|----------------------------------------|---------------|---------------|--------|-------|-----------------|
| Cp    | Maritima/adanensis                     | 13            | 2             | 3      | 0     | 1.04 ± 0.22     |
|       | Maritima/macrocarpa                    | 15            | 1             | 1      | 1     | 1.21 ± 0.25     |
|       | Adanensis/macrocarpa                   | 4             | 1             | 1      | 2     | 1.07 ± 0.33     |
| Adh   | Maritima/adanensis                     | 8             | 1             | 0      | 0     | 4.68 ± 0.91     |
|       | Maritima/macrocarpa                    | 8             | 0             | 0      | 2     | 8.88 ± 2.23     |
|       | Adanensis/macrocarpa                   | 1             | 0             | 0      | 2     | 7.24 ± 2.09     |
| Cab11 | Maritima/adanensis                     | 37            | 0             | 0      | 0     | 7.15 ± 1.18     |
|       | Maritima/macrocarpa                    | 37            | 0             | 0      | 0     | 9.74 ± 1.48     |
|       | Adanensis/macrocarpa                   | 0             | 0             | 0      | 6     | 7.84 ± 2.61     |
| ITS   | Maritima/adanensis                     | 3             | 1             | 0      | 0     | 2.70 ± 0.69     |
|       | Maritima/macrocarpa                    | 3             | 0             | 0      | 7     | 12.86 ± 2.88    |
|       | Adanensis/macrocarpa                   | 1             | 0             | 0      | 8     | 12.45 ± 3.53    |
high between B. macrocarpa and B. vulgaris subspecies than the divergence among Beta vulgaris subspecies.

4 | DISCUSSION

The present study aimed to survey the chloroplastic and nuclear genetic diversities of Beta species (Beta section) and explore the phylogenetic relationships among them.

Accordingly with former studies (Andrello et al., 2016, 2017; Kaderiret et al., 2006; Letschet, 1993; Romeiras et al., 2016), Beta macrocarpa appeared to be a distinct monophyletic lineage from Beta vulgaris that comprised the two subspecies B.v. maritima and B.v. adanensis. The divergence date between B. macrocarpa and B. vulgaris has recently been estimated to be 1.4 Mya (Romeiras et al., 2016).

Within Beta vulgaris, the two subspecies B.v. maritima and B.v. adanensis were analysed on a representative geographical sampling. It must be noted that the two subspecies differ in their distribution and in their mating system. While B.v. maritima populations are found on a large geographical area, along the Atlantic coasts of Western Europe and the coasts of most Mediterranean countries, B.v. adanensis is restricted in the eastern part of the Mediterranean Basin (Aegean islands, Turkey, Syria and Iran). Therefore, the observation of a lower genetic diversity of B.v. adanensis when compared with B.v. maritima was expected. The low divergence between B. vulgaris subspecies can be explained by a recent differentiation of B.v. adanensis (indeed the B. v. adanensis lineage is not clearly defined and is moreover nested within the B. v. maritima clade) and/or contemporary gene flow between the subspecies, as populations of both subspecies can be found in close proximity. Controlled crosses are possible between the subspecies, confirming that reproductive barriers are limited (Hautekèete, 2001). In addition, differentiation of B.v. adanensis populations in situation of sympatry with B.v. maritima is most likely due to a transition from self-incompatibility to self-compatibility. Indeed, the present study suggests that B. v. adanensis reproduces mainly by selfing, as all analysed individuals were found homozygous at nuclear loci. This lack of heterozygosity could also be explained by the low level of diversity in the subspecies. However, the hypothesis of selfing conforms to the cytological pattern observed on self-pollon germination in B. v. adanensis (Bruun et al., 1995). Further studies on a larger sampling and including a population level, in particular by contrasting parapatric versus allopatric situations, are necessary to estimate current gene flow that could occur between the subspecies, and the level of self-fertilization in B.v. adanensis. The development of population genomic approaches thanks to next-generation sequencing methodology would be worthwhile to propose a demographic scenario of B.v. adanensis differentiation, measure the level introgression between both subspecies, the direction of gene flow, as well as the impact of the transition toward selfing on its genomic diversity (synonymous and nonsynonymous) as exemplified in Capsella or Mimulus (Brandvain, Kenney, Flagel, Coop, & Sweigart, 2014; Foxe et al., 2009).

In previous studies, Beta macrocarpa has been described as two cytotypes: one diploid cytotype widely distributed from Portugal to Turkey, along the Mediterranean Basin, and a tetraploid cytotype first found in the Canary Islands (Buttler, 1977). Earlier studies on this tetraploid cytotype have suggested a hybrid origin of the taxon between B. v. maritima and B. macrocarpa: (1) cytological observations revealed a complete diploidised meiosis as expected for an alloploid (Lange & de Bock, 1989), (2) genetic analyses on nuclear allozyme loci showed B.v. maritima and B. macrocarpa alleles-like (Abe & Tsuda, 1987; Letschet, 1993), and (3) a maritima-like chloroplastic haplotype was found in a Canarian individual (Kishima, Mikami, Hirai, Sigauri, & Kinoshita, 1987). Nevertheless, the occurrence of tetraploid individuals does not seem to be restricted to the Canary Islands as formerly believed: recent studies localized 4× individuals on another Macaronesian island, Santo Porto (Madeira Archipelago) (Leys et al., 2014) but also in continental populations from Southern Portugal (Castro et al., 2013). The present study confirms the hybrid origin of 4× B. macrocarpa from two Canary Islands (Gran Canaria and Tenerife): at the nuclear level each individual bears a maritima-like allele and a macrocarpa-like allele with the exception of ITS where only one allele, belonging to the B.v. maritima clade, was found. This is most likely due to concerted evolution as observed in allopolyploid Gossypium species (Wendel, Schnabel, & Seelanan, 1995), rice (Bao, Wendel, & Ge, 2010), or tobacco (in Bao et al., 2010).

At the chloroplastic level, both 4× Canarian individuals shared the same haplotype with B.v. maritima individuals. This suggests that the initial maternal parent of the hybrid was B.v. maritima, and thus B. macrocarpa was the pollen donor. The hybridization between self-incompatible B.v. maritima and self-compatible B. macrocarpa led to an allopolyploid species, described as self-compatible in early studies (Buttler, 1977). Our results suggest that 4× individuals mainly reproduce by selfing, as we did not find any heterozygosity at the homeologous loci.

It must be noted that if the present study confirms the allopolyploid nature of 4× B. macrocarpa found in Canary Islands, it also shows for the first time that B. macrocarpa individuals found in the Canary Islands are not all tetraploid. Indeed, the individual from Fuerteventura (m2) exhibits all the genetic features of 2× continental B. macrocarpa at both chloroplastic and nuclear levels. This result raises the question of the occurrence of 2× B. macrocarpa populations in the Canary Islands where they were until now considered as absent. It remains to know the relative occurrence of the two forms in the Canary Islands as well as the geographical origin of 4× macrocarpa populations: whether the hybridization occurred in the islands or in the continent followed by long-distance dispersal (Linder & Barker, 2014). Further studies are needed to describe the phenotypic characteristics and the ecological preferences of the different macrocarpa cytotypes in order to better distinguish them taxonomically but also to understand how the two types coexist in the Macaronesian archipelago and the adjacent regions.

ACKNOWLEDGMENTS

This work was supported by grants from the Région Nord-Pas de Calais and the European Community (European Regional Development Fund). We wish to warmly thank Lothar Frese, Brian
Ford-Lloyd, and Lee Panella for providing seeds, leaves, and DNA samples.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

PT and JC conceived and designed the study. SV, LB, and A-CH carried out the laboratory experiments. PT, SV, and CP carried out the analyses. PT and SV wrote the draft manuscript that was edited by CP and JC.

ORCID

Pascal Touzet http://orcid.org/0000-0001-7166-0952

REFERENCES

Abbott, R. J., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., ... Zinner, D. (2013). Hybridization and speciation. Journal of Evolutionary Biology, 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Abe, J., & Tsuda, C. (1997). Genetic analysis for isozyme variation in the section Vulgares, genus Beta. Japanese Journal of Breeding, 37, 253–261. https://doi.org/10.1270/jsjbs1951.37.253

Andrello, M., Henry, K., Devaux, P., Desprez, B., & Manel, S. (2016). Taxonomic, spatial and adaptive genetic variation of Beta section Beta. Theoretical and Applied Genetics, 129, 257–271. https://doi.org/10.1007/s00122-015-2625-7

Andrello, M., Henry, K., Devaux, P., Verdelet, D., Desprez, B., & Manel, S. (2017). Insights into the genetic relationships among plants of Beta section Beta using SNP markers. Theoretical and Applied Genetics, 130, 1857–1866. https://doi.org/10.1007/s00122-017-2929-x

Bao, Y., Wendel, J. F., & Ge, S. (2010). Multiple patterns of rDNA evolution following polyploidy in Oryza. Molecular Phylogenetics and Evolution, 55, 136–142. https://doi.org/10.1016/j.ympev.2009.10.023

Barker, M. S., Arrigo, N., Baniaga, A. E., Li, Z., & Levin, D. A. (2016). On multiple patterns of rDNA evolution and chromosome number in Oryza. Phytologist, 210, 55, 136–142. https://doi.org/10.1016/j.ympev.2009.10.023

Bao, Y., Wendel, J. F., & Vernet, P. (1994). Gynodioecy and mitochondrial DNA polymorphism in natural populations of Beta vulgaris ssp. maritima. Genetics Selection Evolution, 26, 87–101. https://doi.org/10.1186/1297-9686-26-S1-587

Castro, S., Romeiras, M. M., Castro, M., Duarte, M. C., & Loureiro, J. (2013). Hidden diversity in wild Beta taxa from Portugal: Insights from genome size and ploidy level estimations using flow cytometry. Plant Science, 207, 72–78. https://doi.org/10.1016/j.plantsci.2013.02.012

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. Genetics Selection Evolution, 26, 87–101. https://doi.org/10.1186/1297-9686-26-S1-587

Cuguen, J., Boudry, P., Brookerm, B., 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 microsatellite markers. Theoretical and Applied Genetics, 98, 1194–1201. https://doi.org/10.1007/s001220051184

Desplanque, B., Viard, F., Forcioli, D., Bernard, J., Saumitou-laprade, J., Cuguen, J., & Van Dijk, H. (2000). The linkage disequilibrium between cpDNA and mtDNA haplotypes in Beta vulgaris subsp maritima (L): The usefulness of both genomes for population genetic studies. Molecular Ecology, 9, 141–154. https://doi.org/10.1046/j.1365-294x.2000.00843.x

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. Proceedings of the Royal Society of London B, 273, 1391–1398. https://doi.org/10.1098/rspb.2005.3464

Fievet, V., Touzet, P., Arnaud, J.-F., & Cuguen, J. (2007). Spatial analysis of nuclear and cytoplasmic DNA diversity in wild sea beet (Beta vulgaris ssp. maritima) populations: Do marine currents shape the genetic structure? Molecular Ecology, 16, 1847–1864. https://doi.org/10.1111/j.1365-294x.2006.03208.x

Foxe, J. P., Slotte, T., Stahl, E. A., Neuffer, B., Hurka, H., & Wright, S. (2009). Recent speciation associated with the evolution of selfing in Capsella. Proceedings of the National Academy of Sciences of the United States of America, 106, 5241–5245. https://doi.org/10.1073/pnas.0807679106

Glémis, S. (2007). Mating systems and the efficacy of selection at the molecular level. Genetics, 177, 905–916. https://doi.org/10.1534/genetics.107.073601

Glémis, S., Bazin, E., & Charlesworth, D. (2006). Impact of mating systems on patterns of sequence polymorphism in flowering plants. Proceedings of the Royal Society of London B, 273, 3011–3019. https://doi.org/10.1098/rspb.2006.3657

Glémis, S., & Galtier, N. (2012). Genome evolution in outcrossing versus selfing versus asexual species. In M. Anisimova (Ed.), Evolutionary genetics: Statistical and computational methods (pp. 311–335). New York, NY: Springer. https://doi.org/10.1007/978-1-61779-582-4

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution, 27, 221–224. https://doi.org/10.1093/molbev/msp259

Grivet, D., & Petit, R. J. (2003). Chloroplast DNA phylogeography of the hornbeam in Europe: Evidence for a bottleneck at the outset of postglacial colonization. Conservation Genetics, 4, 47–56. https://doi.org/10.1023/A:1021804009832

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0.
Katoh, K., & Standley, D. (2013). MAFFT multiple sequence alignment software.

Kadereit, G., Hohmann, S., & Kadereit, J. W. (2006). A synopsis of Mediterranean-Californian disjunctions: Molecular evidence from Chenopodiaceae-Betoideae. Taxon, 55, 67–78. https://doi.org/10.2307/25065529

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jung, C., Pillen, K., Frese, L., Fähr, S., & Melchinger, A. E. (1993). Phylogenetic relationships between cultivated and wild species of the genus Beta revealed by DNA "fingerprinting". Theoretical and Applied Genetics, 86, 449–457. https://doi.org/10.1007/BF00838560

Kaderireit, G., Hohmann, S., & Kadereit, J. W. (2006). A synopsis of Chenopodiaceae subfam. Betoideae and notes on the taxonomy of Beta. Willedenowia, 36, 9–19. https://doi.org/10.3372/wi.36.36101

Katoh, K., & Standley, D. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. (Outlines version 7). Molecular Biology and Evolution, 30, 772–780. https://doi.org/10.1093/molbev/mso010

Kishima, Y., Mikami, T., Hirai, A., Sugiura, M., & Kinoshita, T. (1987). Beta vulgaris ssp. maritima (sea beet) populations: RFLPs and isozymes show different patterns of gene flow. Heredity, 77, 245–250. https://doi.org/10.1038/hdy.1996.138

Raybould, A. F., Mogg, R. J., Gliddon, C., Thorpe, R., & Clarke, R. T. (1998). The genetic structure of sea beet (Beta vulgaris ssp maritima) populations. III. Detection of isolation by distance at microsatellite loci. Heredity, 80, 127–132. https://doi.org/10.1046/j.1365-2540.1998.00265.x

Richards, C. M., Reeves, P. A., Fenwick, A. L., & Panella, L. (2014). Genetic structure and gene flow in Beta vulgaris subspecies maritima along the Atlantic coast of France. Genetic Resources and Crop Evolution, 61, 651–662. https://doi.org/10.1007/s10722-013-0066-1

Romeiras, M. M., Vieira, A., Silva, D. N., Moura, M., Santos-Guerra, A., Batista, D., ... Paulo, O. S. (2016). Evolutionary and biogeographic insights on the Macaronesian Beta-Patellifolia species (Amaranthaceae) from a time-scaled molecular phylogeny. PLoS ONE, 11(3), e0152456. https://doi.org/10.1371/journal.pone.0152456

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D., Darling, A., Höhna, S., ... Hulsenbeck, J. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Shen, Y., Newbury, H. J., & Ford-Lloyd, B. V. (1996). The taxonomic characterization of annual Beta germplasm in a genetic resources collection usingRAPD markers. Euphytica, 91, 205–212.

Soltis, P. S., Marchant, D. B., Van de Peer, Y., & Soltis, D. E. (2015). Polyplody and genome evolution in plants. Current Opinion in Genetics & Development, 35, 119–125. https://doi.org/10.1016/j.gde.2015.11.003

Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology, 17, 1105–1109. https://doi.org/10.1007/BF0037152

Watterson, G. A. (1975). On the number of segregating sites in genetic models without recombination. Theoretical Population Biology, 7, 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Wendel, J. F., Schnabel, A., & Seelanan, T. (1995). Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (Gossypium). Proceedings of the National Academy of Sciences of the United States of America, 92, 280–284. https://doi.org/10.1073/pnas.92.1.280

White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. Innis, D. Gelfand, J. Sninsky & T. White (Eds.), PCR protocols: A guide to methods and applications (pp. 315–322). Orlando, FL: Academic Press. Chapter 38.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.