Plant Science 253 (2016) 194–205

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Spatial genetic structure in wild cardoon, the ancestor of cultivated globe artichoke: Limited gene flow, fragmentation and population history

D. Rau[a,*], M. Rodriguez[a,b], E. Rapposelli[c], M.L. Murgia[a], R. Papa[d], A.H.D. Brown[e], G. Attene[a,b]

[a] Dipartimento di Agraria, Università degli Studi di Sassari, Via E. De Nicola, 07100 Sassari, Italy
[b] Centro per la Conservazione e Valorizzazione della Biodiversità Vegetale, Università degli Studi di Sassari, Surigheddu, 07040 Alghero, Italy
[c] AGRIS Sardegna, Servizio per la Ricerca in Arboricoltura, Via Mameli 126/D, Cagliari, Italy
[d] Dipartimento di Scienze Agrarie, Alimentari ed Ambientali (D3A), Università Politecnica delle Marche (UNIVPM), Via Brecce Bianche, 60131 Ancona, Italy
[e] Bioversity International (Honorary Research Fellow), Via dei Tre Denari, 472/a, 00054 Maccarese, Fiumicino, Italy

ARTICLE INFO

Article history:
Received 13 July 2016
Received in revised form
23 September 2016
Accepted 26 September 2016
Available online 29 September 2016

Keywords:
Cynara cardunculs var. *sylvestris*
Crop wild relatives
Domestication
Spatial genetic structure
Fragmentation
Colonisation history

ABSTRACT

Nuclear and chloroplast markers and phenotypic characters were integrated to analyse the population genetic structure of wild cardoon, *Cynara cardunculus var. sylvestris*, the ancestor of cultivated globe artichoke, *Cynara cardunculus* var. *scolymus* on the island of Sardinia, Italy. The spatial scale ranged from a few metres to ~200 km. Wild cardoon appears to be genetically fragmented, with significant genetic divergence at various scales, indicating that gene flow is insufficient to counterbalance the effects of genetic drift or founder effects. Divergence between populations was higher for chloroplast (40%) than for nuclear markers (15%), suggesting that gene flow via seed was lower than via pollen. Two main genetic groups were detected; these correlated with differences in flowering time, capitula size, glossiness, and anthocyanin pigmentation. A complex population structure of wild cardoon emerged over small spatial scales, likely resulting from the interplay between gene dispersal, colonisation history and selective forces. Indeed, Sardinia appears to be a 'hybrid zone' of different gene pools. The island has unique diverse germplasm that has originated from hybridisation among different gene pools. The sampling of seeds from a few plants but from many sites is suggested as the best strategy to harvest the genetic diversity of wild cardoon.

1. Introduction

Cynara cardunculus L. is a diploid species (2n = 2 x = 34) that belongs to the family Asteraceae, which is native to the Mediterranean basin. The species comprises two cultivated subspecies, the globe artichoke (var. *sativa moris*, var *scolymus* [L.] Fiori, ssp. *scolymus* [L.] Hegi) and the leafy cardoon (var. *altilis* DC) [1], along with their ancestor, the perennial wild cardoon (var. *sylvestris* [Lamk] Fiori). The outcrossing rates of all three of these *taxa* are high, with pollination mediated by bees [2]. The globe artichoke is a horticultural crop that contributes in particular to the agricultural economy of southern Europe (i.e., mainly Spain, France, Italy), and more recently those of some African countries that border the Mediterranean Sea [3–5]. The marketable product is the large immature inflorescence, known as the capitulum or head, although other parts of the plants are also edible (i.e., stem, part of the leaf) and are used in traditional cuisine [3,6]. Several studies have also underlined the pharmaceutical relevance of this crop; e.g., [7–12]. Its progenitor, wild cardoon, is also appreciated for the preparation of traditional foods [13–15], and has been studied in terms of its pharmaceutical properties and as a crop for energy purposes in the Mediterranean environment [16–18].

The domestication of globe artichoke from wild cardoon is not yet fully understood [19–21]. It has been hypothesised that globe artichoke was domesticated in either southern Italy or northern Africa or both, while leafy cardoon was probably cultivated in Spain [19,20]. Selection for large, non-spiny heads probably led to globe artichoke, while selection for non-spiny, large stalked, tender leaves led to leafy cardoon [2]. Studies based on molecular markers have indicated that both of these crops evolved from the wild cardoon gene pool, which is therefore the progenitor of both of them [19,22–25]. Wild cardoon and these two cultivated subspecies are completely interfertile, and therefore they belong to the same gene pool [1,2,26].

* Corresponding author.
E-mail address: dmrau@uniss.it (D. Rau).

http://dx.doi.org/10.1016/j.plantsci.2016.09.013
0168-9452/

Nowadays, it is widely recognised that knowledge of the population structure of the wild relatives of a crop is key to defining appropriate optimal strategies for sampling genetic variation. Furthermore, this knowledge sheds light on aspects of the biology of crop plants, serves in the study of the domestication process [27,28], and allows better exploitation of the potential of genetic resources for plant breeding [29,30]. However, few studies are available on the level and structure of population variation in wild cardoon. Foury and Wiklund [31,32] studied the morphological variation of herbarium accessions of wild cardoon from the *circum*-Mediterranean area. Their study of bracteal traits suggested that there were two phylogenetically meaningful gene pools: the eastern and the western. This was confirmed by Gatto et al. [19], who determined the population structure of the *C. cardunculus* complex around the Mediterranean basin, using microsatellite (SSR) markers. Indeed, they described a large-scale geographical pattern of genetic variation that mainly distinguished wild cardoon into the eastern (i.e., mainly Italy, Greece) and the western (i.e., Spain, Portugal) gene pools, with the Tunisian gene pool as 'intermediate' between these two. At the regional scale, Khaldi et al. [33,34] reported high variation in six Tunisian wild cardoon populations for 19 morphological traits and SSR markers, without any evident genetic dichotomy within the sample analysed. Portis et al. [35] compared Sardinian and Sicilian provenances using amplified fragment length polymorphism (AFLP) and SSR markers, and showed that the wild cardoon gene pools from these Islands were clearly distinct. However, data on the spatial organisation of the genetic variation of *C. cardunculus* at lower scales are lacking.

In terms of the history of *C. cardunculus* as a crop, the island of Sardinia occupies an interesting geographic position as a 'bridge' in the centre of the Mediterranean Sea. Indeed, it lies between southern Italy and Spain, as the two areas where domestications of globe artichoke (Italy) and leafy cardoon (Spain) appear to have taken place, with wild cardoon as the common ancestor.

In this study, we characterised a hierarchical collection of wild cardoon over spatial scales from a few metres up to ~200 km, on the island of Sardinia. We integrated AFLP, chloroplast (cp)SSRs and phenotypic analyses to determine genetic variation of wild cardoon over various geographical scales. In particular we aimed to compare the extent of molecular divergence among populations for chloroplast with nuclear markers in order to detect the footprint of gene flow via pollen versus seed migration. A better understanding of population genetic structure allows the determination of the correct strategies for sampling genetic diversity, to improve our understanding of the role of intrinsic (e.g., gene flow, mating system, drift) and extrinsic (e.g., selection) factors, colonisation history, and inter-gene-pool hybridisation in wild cardoon evolution.

2. Materials and methods

2.1. The collection of Sardinian wild cardoon

The accessions of wild cardoon were taken from a hierarchical collection from Sardinia that was sampled in July and August 2003. Based on ecology and pedology, seven main zones of the geographic distribution of these wild cardoon were defined. From each area (except for the Sulcis zone) two populations were chosen (one from Sulcis), for a total of 13 populations. Plants from the sampled populations were distributed without any (evident) discontinuity. For each population, three transects were sampled, with each represented by 15 plants; thus there were 45 plants collected per site, for a total of 585 plants.

For each plant collected, the mature inflorescence was cut and threshed separately. The progenies (i.e., single-head progenies) were kept in separate bags, labelled, and stored at −20 °C. The geographic coordinates were recorded (i.e., latitude, longitude, altitude). The minimum and maximum distances between populations were 14 km and 207 km, respectively. The distances between populations within the areas ranged from 14 km to 62 km. On average, the transects within each population were 630 m apart, and the distances between the plants within the transects varied from a minimum of 2 m–3 m to a maximum of 150 m, with an average of 30 m. Details for each population are reported in Supplementary Table S1.

Molecular characterisation was carried out on 118 individuals, considering three plants per transect, three transects per population, and a total 13 populations (Supplementary Fig. S1). Each of the populations was represented by nine individuals except for one population (S. Sabina), which comprised 10 individuals. Each individual originated from a seed that was randomly extracted from one single-head progeny. Thus, the 118 individuals analysed represented 118 different mother plants. Details of the accessions analysed are given in Supplementary Table S2. For 33 of the 118 mother plants, three further seeds were used to grow plants under field conditions for the phenotypic characterisation. These 33 mother plants were representative of all of the localities, as two to three plants per population were included, except for the Posada site, which was represented by one mother plant (Supplementary Table S2).

2.2. Molecular analysis

The seeds were grown in a growth chamber at 20 °C. The DNA was extracted from young leaves of one-month-old plants, according to Doyle and Doyle [36].

Amplified fragment length polymorphism analysis was conducted using micro-organism AFLP kits (Gibco BRL, Grand Island, NY, USA), with four primer combinations and following the manufacturer instructions. Digested and ligated DNA fragments were first pre-amplified with primers that were complementary to the adapters and included an additional selective 30 nucleotides (i.e., EcoRI + A, MseI + C primers). Selective amplification was subsequently carried out using primers with three selective nucleotides and considering four primer combinations: E-AAC/MseI + CTG, E-AAC/MseI + CAT, E-AGC/MseI + CTA and E-AGC/MseI + CAT. The AFLPs were resolved on 6% denaturing polyacrylamide gels run at 50 W, using Tris-taurine-EDTA (TTE) buffer (10.8 g Trizma base, 3.6 g taurine, 0.2 g $Na_2EDTA{\cdot}2H_20$), and fixed (10% [v/v] glacial acetic acid, and 20% [v/v] methanol). The bands on the gels were visualised using silver staining [37].

Chloroplast simple sequence repeat variation was screened for 35 cpSSR loci using nine of the primer pairs designed by Weising and Gardner [38], 20 designed by Chung and Staub [39], and six designed by Xu et al. [40].

2.3. Hierarchical analysis of molecular variance

The genetic diversity estimated by AFLPs was partitioned using the analysis of molecular variance framework (AMOVA; [41]), as implemented in Arlequin version 3.5.1 [42]. Total genetic variance was partitioned into three hierarchical levels: among areas ($\sigma^2{}_a$), among populations within areas ($\sigma^2{}_b$), and among individuals within populations ($\sigma^2{}_c$). The genetic divergence among areas (F_{CT}) was quantified according to Eq. (1), the divergence among populations (within areas; F_{SC}) according to Eq. (2), and the total genetic divergence among populations according to Eq. (3):

$$F_{CT} = \sigma^2{}_a/\sigma^2{}_T \quad (1)$$

$$F_{SC} = \sigma^2{}_b/(\sigma^2{}_b + \sigma^2{}_c) \quad (2)$$