

# Comparison of different methodologies in order to perform a representative *Cynara cardunculus* L. core collection

Cravero, V., Crippa, I., Martin, E. and Cointry, E.

## SUMMARY

This study was carried out in order to compare different methodologies to select a representative core collection from an initial collection of *Cynara cardunculus* L. species, using both morphological and molecular markers. The combination of two stratification criteria with three sampling strategies allowed the establishment of six different core collections. The Maximization strategy was applied in order to obtain the seventh one. All the obtained collections were validated through phenotypic and molecular parameters, establishing as an initial criterion that the core collection should include, at most, 35-40 % of the accessions belonging to the original collection. All collections passed molecular validation; nevertheless morphological validation determined that the Proportional sampling strategy is the best to keep the initial variability while retaining the least number of accessions, especially when combined with the first stratification criterion. Although Maximization strategy allowed to preserve the original variability, it retained the largest number of accessions. In conclusion, the combination PrS1 (proportional sampling and first stratification criterion) is the best strategy to perform a representative core collection from a *Cynara cardunculus* L. initial collection, using both morphological and molecular data.

**Keywords:** Cardoon, globe artichoke, genetic variability, genetic resources.

Cravero, V., Crippa, I., Martin, E. y Cointry, E., 2019. Comparación de diferentes metodologías a fin de conformar una colección núcleo representativa de *Cynara cardunculus* L. Agriscientia 36: 25-38

## RESUMEN

El objetivo del trabajo fue comparar diferentes metodologías e identificar la más adecuada para conformar una colección núcleo representativa a partir de una colección inicial de *Cynara cardunculus* L., utilizando marcadores morfológicos y moleculares. Combinando dos criterios de estratificación y tres técnicas

de muestreo fue posible establecer seis colecciones núcleo. La aplicación de la estrategia M (maximización) permitió establecer la séptima colección. Las siete colecciones fueron validadas a través de parámetros fenotípicos y moleculares, estableciendo como criterio inicial que la colección núcleo debía incluir, como máximo, el 35-40 % de las accesiones pertenecientes a la colección inicial. Todas las colecciones superaron la validación molecular; sin embargo, la validación morfológica determinó que la estrategia de muestreo proporcional fue superior al resto respecto a su capacidad de conservar la variabilidad inicial reteniendo el menor número de accesiones, especialmente, cuando fue combinada con el primer criterio de estratificación. La estrategia M, si bien permitió preservar la variabilidad original, retuvo el mayor número de accesiones. En conclusión, la combinación PrS1 (muestreo proporcional y primer criterio de estratificación) es la mejor estrategia para formar una colección núcleo representativa a partir de una colección inicial de *Cynara cardunculus* L., utilizando datos morfológicos y moleculares en conjunto.

**Palabras clave:** Cardo, alcaucil, variabilidad genética, recursos genéticos..

*Cravero, V., Martin, E. and Cointry, E.: Instituto de Investigaciones en Ciencias Agrarias de Rosario (IICAR, CONICET-UNR). Campo Experimental Villarino (S2125ZAA). Zavalla, Santa Fe, Argentina. Crippa, I.: INDEAR, Ocampo 210 bis (2000) Rosario, Santa Fe, Argentina. Correspondence to: cravero@iicar-conicet.gov.ar.*

## INTRODUCTION

The species *Cynara cardunculus* L. includes three botanical varieties: globe artichoke (*C. cardunculus* L. var. *scolymus* (L.) Fiori), cultivated cardoon (*C. cardunculus* L. var. *altilis* DC) and wild cardoon (*C. cardunculus* L. var. *sylvestris* (Lamk) Fiori). They are all cross-pollinated and hybridization between them is possible.

DNA fingerprinting confirmed that both cultivated forms (globe artichoke and cultivated cardoon) evolved independently from the wild cardoon, through anthropogenic selection (Portis, Barchi, Acquadro, Macua and Lanteri, 2005). Sonnante, Carluccio, De Paolis and Pignone (2008) proposed that globe artichoke might originate from the eastern wild cardoon, while the western genotypes might be the ancestors of the cultivated cardoon. Therefore, the domestication of both cultivated varieties occurred, probably, at different times and different places.

The species has great potential not only as human food but also for several purposes such as forage for livestock, extraction of inulin and other pharmacological and cosmetic active compounds, exploitation of its lignocellulosic biomass for energy

and paper pulp, and production of biodiesel from the oil (Fernández, Curt and Aguado, 2006).

Genetic diversity is the main requirement for species evolution and provides the necessary adaptation to the prevailing biotic and abiotic environmental conditions. Genetic diversity allows changes in the crops genetic structure as a response to environmental changes. Understanding the magnitude and patterns of the available genetic diversity has important implications for breeding programs and for the conservation of genetic resources.

In the genus *Cynara*, many diversity studies have been performed using morphological traits. The first attempts to determine the genetic variation among globe artichoke cultivars, for qualitative and quantitative traits, goes back to the studies by Dellacecca, Magnifico, Marzi, Porceddu and Mugnozza (1976) who generated a list of descriptors based on 20 morphological and productive characters. Studies to detect genetic diversity at molecular level have also been conducted. The first evaluations were performed by Tivang, Skroch, Nienhuis and De Vos (1996) and Lanteri, Di Leo, Ledda, Mammeli and Portis (2001), who assessed diversity within globe artichoke accessions using RAPD and determined the multiclonal origin of

most of these accessions. After that, several studies were carried out in order to quantify and determine genetic diversity patterns between and within *C. cardunculus* botanical varieties; most of them only included some European accessions (Raccuia, Mainolfi, Mandolino and Melilli, 2004; Sonnante *et al.*, 2008; Mauro *et al.*, 2009; Gatto, De Paola, Bagnoli, Vendramin and Sonnante, 2013; De Felice *et al.*, 2015; Pagnotta, Fernández, Sonnante and Egea-Gilabert, 2017).

Different *C. cardunculus* accessions, including local and foreign ones, belonging to the three botanical varieties, were collected over time. It is known that many *C. cardunculus* cultivars or ecotypes, growing in different geographical areas, are usually named according to the locality of cultivation (Bianco, 1990). Thus, a unique cultivar could be called differently depending on the cultivation region, while, different cultivars could take the same name, producing misunderstanding.

The management of many accessions hinders the activities related to their conservation, such as their multiplication and regeneration, as well as their optimal utilization. For those accessions that must be sexually reproduced, spatial or temporal isolation is required during multiplication and regeneration in order to preserve their genetic integrity; while, for clonal accessions, only field or tissue culture conservation is possible. In both situations, to decrease the number of retained accessions is a critical point.

In this way, Frankel (1984) proposed the “core collections” (CC) creation as a way to preserve most of the available genetic variability with the least possible repetition. A CC was defined as a limited set of accessions which, using a minimum of repetitions, represents most of the genetic diversity of a crop species and its wild relatives. It is considered a cheap alternative for improving the management of germplasm banks while promoting the evaluation and use of genetic resources (Frankel, 1984; Brown, 1995).

At the time of creating a CC, two main problems are faced. The first one is the optimal number of accessions needed to keep most of the initial collection (IC) genetic diversity and, the second, which accessions will be included.

On the basis of the neutral allele model, Brown (1989) determined that a sample of 10 % of the IC, with a maximum of 3000 accessions per species, allowed to keep 70 % of the alleles. Nevertheless, this ratio should be modified according to the amount of accessions included in the IC (Li *et al.*, 2003). It should be increased when the initial

number of accessions is low and should be reduced when this number is large.

Regarding which accessions will be included in the CC, several strategies have been proposed. The most common approaches are the M (maximization) strategy and stratified sampling. M strategy tends to maximize the total allele diversity of the CC whereas the stratified sampling objective is to include accessions that optimize the representativeness of the genetic diversity in the CC (Ruiz Valcárcel, Giraldo Carbajo, Royo and Carrillo Becerril, 2013). Some authors suggested performing a hierarchical clustering (stratification), based on taxonomy, origin or knowledge about domestication patterns or improvement history, previous to sampling (Brown, 1989, 1995). Others proposed the application of multivariate analysis (cluster, discriminant analysis or principal components) from molecular, phenotypic or agronomic data (Crossa, De Lacy and Taba, 1995). Both approaches could be applied independently or in combination (Zhang *et al.*, 2000).

Once established the CC, its quality and representativeness with regard to the IC should be tested, process known as validation.

In previous studies (Cravero, Martin and Cointry, 2007; Casadevall, Martin and Cravero, 2011), the variability available between some accessions, belonging to the local *C. cardunculus* collection, was evaluated using Sequence Related Amplified Polymorphism (SRAP) (Li and Quiros, 2001). Those studies proved that SRAP is also a helpful tool to detect genetic diversity in this species and to classify accessions into groups based on their genetic distance values.

The aim of this study was to compare different methodologies to select a representative CC from an IC using morphological and molecular markers.

## MATERIAL AND METHODS

### Plant material

The research plant material was the *Cynara cardunculus* L. living collection belonging to the National University of Rosario (Argentina) which includes 52 accessions: 6 of *C. cardunculus* L. var. *sylvestris*, 18 of *C. cardunculus* L. var. *altilis*, and 28 of *C. cardunculus* L. var. *scolymus*. This collection includes natural populations, landraces, and commercial varieties (OP, clones and hybrids) with different geographical origins (Table 1).

At least 30 plants from each accession were implanted at the University Experimental Field

**Table 1.** Accessions belonging to the living initial collection, country of origin and genetic status. OP: Open-Pollination variety, NP: natural population

<i>Globe artichoke</i>			<i>Cultivated cardoon</i>		
Code	Origin	Genetic status	Code	Origin	Genetic status
A01	Argentina	Clone	CC02	unknown	Unknown
A02	Argentina	OP	CC03	Argentina	Landrace
A03	Argentina	Clone	CC04	Argentina	Landrace
A04	Argentina	Clone	CC05	Argentina	Landrace
A05	Argentina	Clone	CC06	Brazil	OP
A06	Argentina	Clone	CC07	Canada	OP
A07	Canada	OP	CC08	USA	OP
A08	USA	OP	CC09	Spain	OP
A09	USA	OP	CC10	Spain	OP
A10	USA	OP	CC13	Spain	OP
A11	USA	OP	CC14	Spain	OP
A12	USA	OP	CC16	Spain	OP
A13	USA	OP	CC17	Spain	OP
A14	USA	OP	CC18	Spain	OP
A15	USA	OP	CC19	Spain	OP
A16	Italy	OP	CC20	Spain	OP
A17	Spain	Hybrid	CC21	Spain	OP
A18	Spain	Hybrid	CC26	Italy	OP
A19	Spain	Hybrid			
A20	Spain	Hybrid			
A21	Spain	OP	CS01	Argentina	NP
A22	Spain	Hybrid	CS02	Argentina	NP
A23	Spain	OP	CS03	Argentina	NP
A24	Spain	Hybrid	CS04	Argentina	NP
A25	Spain	Hybrid	CS05	Uruguay	NP
A26	Spain	OP	CS06	Uruguay	NP
A27	Spain	Hybrid			
A28	France	Clone			

Station (33° 1' S, 60° 53' W) in a completely randomized design with three replications.

### Morphological evaluation

Morphological evaluation was carried out during 2012 over each individual plant considering thirteen quantitative traits: plant height, plant diameter, length and width of the leaves, days from planting to harvest, harvest length, weight, length, diameter, and height/diameter ratio of the main head, number of heads per plant, total yield, and market yield. Market yield was calculated as total yield adjusted by a quality factor. Mean values and standard deviation was calculated for each accession.

### Molecular evaluation

Genomic DNA was extracted from 0.10 g of fresh leaf tissue of five plants, randomly selected

from each accession, using the DNeasy Plant mini Kit (Qiagen Inc., Valencia, CA, USA). Two PCR (Polymerase Chain Reaction) based marker systems were employed: Sequence Related Amplified Polymorphism (SRAP) and Simple Sequence Repeats (SSR).

### Sequence Related Amplified Polymorphism (SRAP)

For SRAP, all combinations between five primers forward and five reverse (Table 2) were tested in five accessions. The six most polymorphic combinations (**C1**: MS1 – Em4, **C2**: MS1 – Em5, **C3**: MS2 – Em 4, **C4**: MS2 – Em3, **C5**: Me4 – Em3, and **C6**: Me4 – Em4) were chosen for all accessions analysis. The PCR reaction mixture (20 mL total volume) contained 30 ng genomic DNA, 0.20 mM of each dNTPs, 50 mM MgCl<sub>2</sub>, 0.50 mM of each primer, 1X GO Taq buffer solution, and 1.50 units of GO Taq polymerase (Promega). All samples were

**Table 2.** SRAP primers sequences.

Forward	Reverse
<b>MS1:</b> TGAGTCCAAACCGGATA	<b>Em1:</b> GACTGCGTACGAATTAAT
<b>MS2:</b> TGAGTCCAAACCGGAGC	<b>Em2:</b> GACTGCGTACGAATTTGC
<b>Me3:</b> TGAGTCCAAACCGGAAT	<b>Em3:</b> GACTGCGTACGAATTGAC
<b>Me4:</b> TGAGTCCAAACCGGACC	<b>Em4:</b> GACTGCGTACGAATTTGA
<b>Me5:</b> TGAGTCCAAACCGGAAG	<b>Em5:</b> GACTGCGTACGAATTAAC

subjected to the following thermal profile: 5 min of denaturing at 94 °C and five cycles of three steps including 1 min of denaturing at 94 °C, 1 min of annealing a 35 °C, and 1 min of elongation at 72 °C. For the next 35 cycles, annealing temperature was elevated to 50 °C ending with an elongation step of 10 min at 72 °C.

### Simple Sequence Repeat (SSR)

Sixteen SSR primers previously developed for the species by Acquadro *et al.* (2009) were tested (CELMs 01, 06, 13, 15, 20, 21, 26, 27, 29, 33, 36, 37, 39, 41, 53 and 59). The PCR reactions were carried out following the protocol reported by these authors. Those SSR that showed inconsistent amplification were discarded. Therefore, the analysis of the whole population was performed with seven SSR (CELMs 01, 06, 26, 27, 33, 41 and 59).

Separation of the amplified fragments was performed, in both techniques (SRAP and SSR), on 6 % (w/v) polyacrylamide gels. Gels were 34 cm x 50 cm x 0.15 mm in size. Electrophoresis conditions were held at 100 W for 2 hours at room temperature for SSR and 2 hr 30 min for SRAP. Gels were stained with AgNO<sub>3</sub>. Each amplified fragment was considered as a single locus and was scored as band presence (1) or band absence (0).

### Stratification and cluster analysis

Several CC were constructed combining molecular and phenotypic data in order to preserve most of the genetic variation with minimum redundancy. Two stratification criteria were applied for cluster analysis: in the first one (S1), all the accessions were considered at the same level (Crossa *et al.*, 1995) whereas in the second one (S2), accessions were stratified by botanical varieties before cluster analysis (van Hintum, 1995). In both cases, Gower's similarity coefficients (Gower, 1975) between accessions were calculated. This kind of distances is usually applied for mixed data (i.e. databases with continuous, ordinal, categorical, and/or

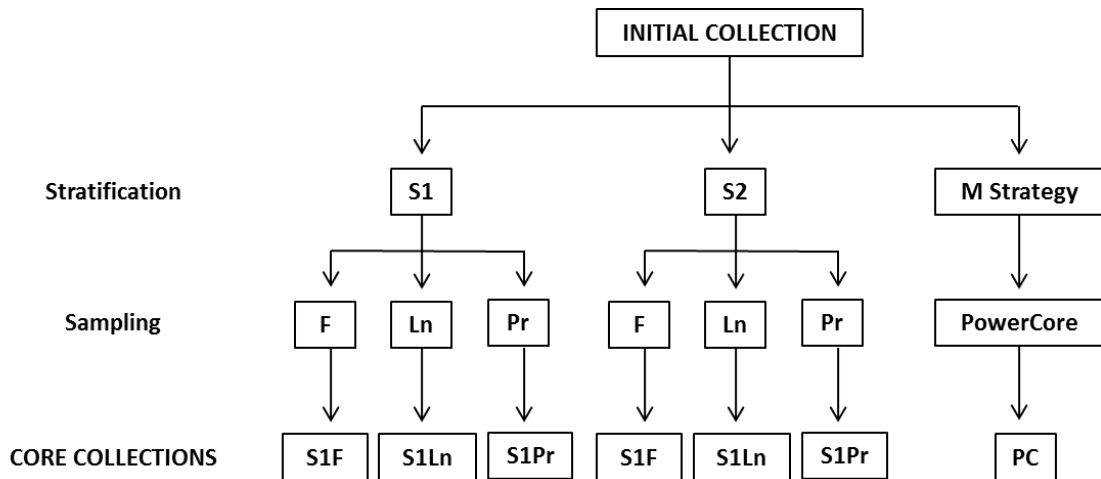
binary data). Following, multivariate hierarchical clustering analyses were carried out throughout the algorithm average linkage: the distance between two clusters is the average of the distances of all pairs of observations, one observation in the pair taken from the first cluster and the other from the second cluster. For each dendrogram, the cophenetic correlation coefficient was calculated in order to determine the goodness of fit between the dendrogram and the original distance matrix (Sokal and Rohlf, 1962). The Mantel's test was applied in order to test the null hypothesis of correlation between the dendrograms and the distance matrix.

Trees were cut at the threshold value (cut-off point) 0.60 and the resulting groups were analyzed.

### Core collections establishment

A predetermined sampling intensity of 35-40 % was considered optimal to capture most of the available genetic diversity included in the IC (Li *et al.*, 2010). Three sampling strategies proposed by Brown (1989) were applied to conform the CC: **Fixed sampling (F)**, where an equal number of accessions were selected from each group; **Logarithmic sampling (Ln)**, where the number of accessions selected from each cluster was proportional to the natural logarithm of the number of accessions in the cluster, with a minimum of one; and **Proportional sampling (Pr)**, where 35 % of the accessions included in each cluster was selected, with a minimum of one. Also, the **maximization strategy (M)** (Brown and Schoen, 1994) was applied. The M strategy is a marker-based method which examines all possible core collections that can be formed and singles out this one that maximizes the number of alleles observed in the marker *loci* (Gouesnard *et al.*, 2001). This strategy does not require a prior stratification and uses a deterministic (not statistical) criterion to define not only the number of accessions, but also identifies which accessions should be included in the CC. M strategy was applied using the PowerCore software (Kim, Jung and Cho, 2007). Mean values and phenotypic variance for each quantitative trait was calculated for each accession and each cluster. Those accessions that showed means values similar to their own cluster and low phenotypic variance were selected to reach the number determined by each strategy.

Combination between both stratification criteria (S1 – S2) and the three sampling strategies (F, Ln and PR) gave rise to six different CC. The seventh CC was formed by the M strategy (Figure 1).



**Figure 1.** Core collections resulting from two stratification (S1, S2) and three sampling (F, Ln, Pr) strategies combination plus core collection obtained by M strategy.

### Core collections validation

All CC were compared in their ability to capture the initial diversity of the IC, using morphological and molecular data. For quantitative traits, parameters proposed by Hu, Zhu and Xu (2000) were calculated: mean difference percentage (MD), variance difference percentage (VD), coincidence rate of range (CR), and coefficient of variation (CV). A CC was considered representative if MD and VD  $\leq 20\%$ , CR  $\geq 80\%$ , and CV  $\geq 100\%$  (Hu *et al.*, 2000). Validation at a molecular level was performed by the following parameters: average number of alleles per marker (AN), total alleles (TA), effective number of alleles (Ne), Shannon-Weaver index (SW), polymorphic information content (PIC), and heterozygosity (H) (Nagy *et al.*, 2012). All analyses were performed by InfoGen statistical software (Balzarini and Di Rienzo). Values for each parameter in the CC were compared with the IC by a t-test ( $P < 0.05$ ). A CC was considered representative when these parameters remained unchanged between the IC and the CC (Wang, Hu, Zhang and Zhang, 2007; Zhang *et al.*, 2010).

## RESULTS AND DISCUSSION

### Stratification and cluster analysis

A core collection can effectively represent the genetic diversity of an initial collection if the accessions are first separated into groups with some significance. Following this concept, Brown (1989) affirmed that stratification is important

to establish efficiently a good core collection. While there are some stratification criteria already established, which vary in complexity, they can also be freely developed by curators to maximize the diversity between groups. Among the simplest criteria, Knüpffer and van Hintum (1995) proposed the stratification of a barley collection segregating spring from winter types. Zhang *et al.* (2000) stratified a sesame (*Sesamum indicum*) collection according to the different regions where this crop was produced in China. Tohme, Jones, Beebe and Iwanaga (1995) applied a stratification method with high level of complexity. They divided a *Phaseolus vulgaris* collection into three groups: "grown in primary centers of diversity", "grown in non-primary centers of diversity" and "wild". In the first group, they made a subdivision according to the country of origin, generating 11 subgroups. Each group was then subdivided according to the accessions adaptation ability, separating areas with a long production history from those where the species was recently introduced. These areas were further divided in agroecological classes, based on the combination of soil type, altitude, drought and photoperiod in each one, allowing the formation of 54 possible settings. Diwan, McIntosh and Bauchan (1995) compared 11 methods to perform core collections from the national collection of Medicago US, finding that collections formed from morphological stratification and subsequent cluster analysis represented the original collection better than those formed by random selection.

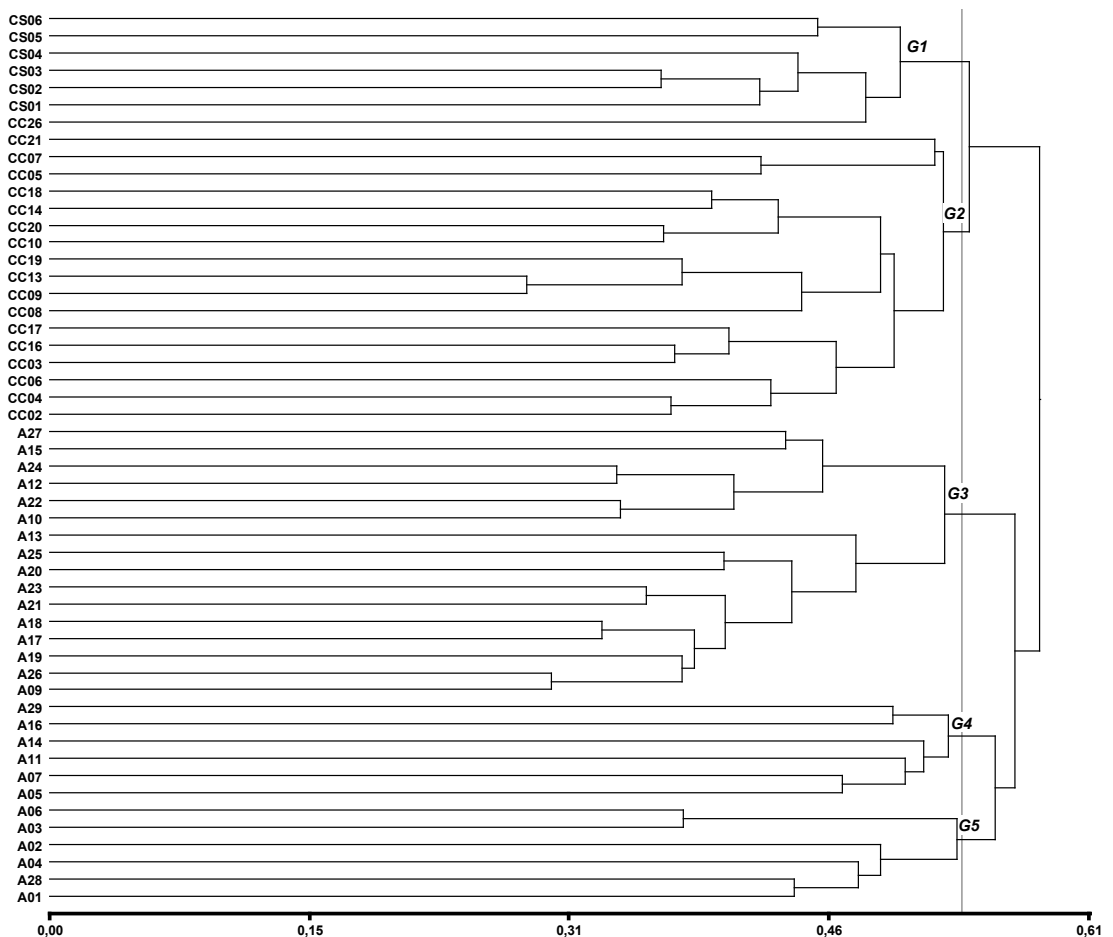
In our research, intermediate complexity methods of stratification were applied. The first one (S1) consisted of a cluster analysis (Cossa *et al.*, 1995) based on morphological and molecular

data; genetic distances were calculated and a dendrogram was constructed considering all the accessions together (Hu *et al.*, 2000; Zhang *et al.*, 2010). In the second methodology (S2), accessions were separated into botanical varieties before the cluster analysis was carried out.

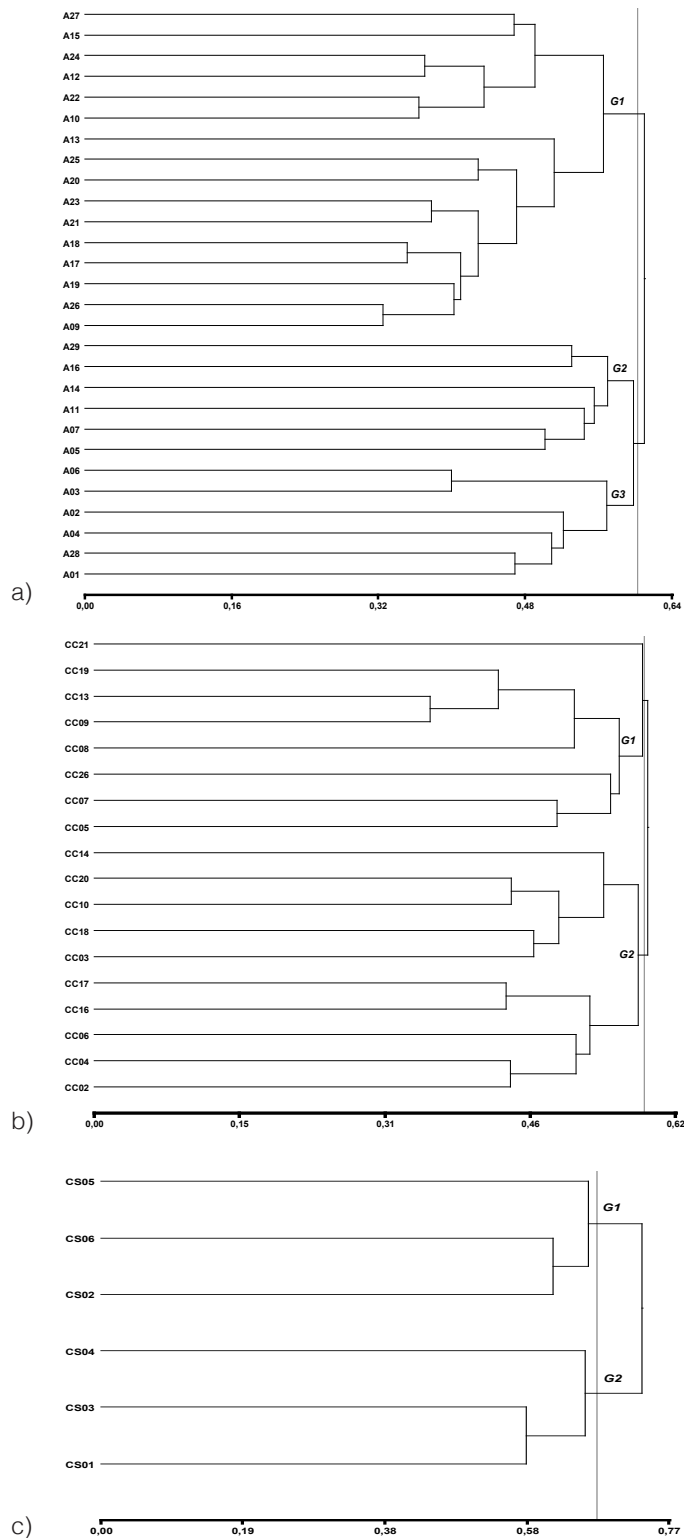
Cluster analysis performed with the stratification criterion S1 showed five main groups or clusters (Figure 2) at a distance of 0.55. Cluster 1 (G1) included all the wild accessions and one cultivated cardoon (CC26). Cluster 2 (G2) grouped most cultivated cardoons except for CC26. The globe artichoke accessions were distributed in three clusters: cluster 3 (G3), which included 16 accessions, 11 of them were Spanish, including five (A20, A22, A24, A25 and A27) belonging to the same breeding program and three (A17, A18 and A19) to another one; the remaining five accessions belonging to G3 are from the United States; cluster 4 (G4) included 6 accessions of different genetic

origins; while cluster 5 (G5) was composed of accessions from the local breeding program and one French accession.

With the second stratification criterion (S2) three dendrograms, one per each botanical variety, were constructed. In the first one (globe artichoke) (Figure 3a) three main clusters could be distinguished at a distance of 0.60, which had identical components to those of G3, G4 and G5 of Figure 2. The second dendrogram (cultivated cardoon) showed two main groups at the same cut-off point (0.60) (Figure 3b). The first one included eight accessions and the second one, 10; which generates a difference with the dendrogram shown in Figure 2 where all cultivated cardoon accessions were included in a single group. Finally, the dendrogram obtained for wild cardoons (Figure 3c) included two clusters at the cut-off point 0.60; each one conformed by three accessions. Again, differences were observed with Figure 2 where, at the same distance,



**Figure 2.** Hierarchical cluster analysis performed by S1. Gower distances. Algorithm average linkage. Cophenetic correlation coefficient: 0.70.



**Figure 3.** Hierarchical cluster analysis performed by S2. Gower distances. Algorithm average linkage. a) *Cynara cardunculus* var. *scolymus*. Cophenetic correlation coefficient: 0.84. b) *Cynara cardunculus* var. *altilis*. Cophenetic correlation coefficient: 0.63. c) *Cynara cardunculus* var. *sylvestris*. Cophenetic correlation coefficient: 0.69.



all wild cardoon accessions define a single group. Regardless of the stratification criterion, cophenetic correlation varied between 0.63 and 0.84 for all dendrograms, showing an acceptable to high adjustment between the dendrograms and the original distance matrix. The lowest value was obtained for *C. cardunculus* var. *altilis* with the stratification criteria S2; nevertheless, all correlation values were significant ( $p < 0.001$ ) by Mantel's test.

### Core collections establishment

Another important step in a core collection set up is the selection of the accessions that will be included. At this point, the initial collection has been divided into many clusters containing accessions with certain similarity degree. Selected accessions should be the best representatives of each group. Selection can be made at random (Hu *et al.*, 2000; Zhang *et al.*, 2010) or based on a formal analytical procedure (van Hintum, Brown, Spillane and Hodgkin, 2003) or practical considerations like their importance and prior knowledge (Igartua *et al.*, 1998).

In our research, it was established as a prior criterion that the CC should include no more than 35 - 40 % (18-21 accessions) of the accessions in the IC (52). To determine the number of accessions to be selected within each cluster, we applied different procedures proposed by Brown (1989) and known as fixed (F), proportional (Pr), and log (Ln) strategies. The efficiency of each strategy has been compared by several researchers (Brown, 1989; van Hintum, 1995) which concluded that the most appropriate strategies would be Ln or Pr, as the case. F and Ln strategies favor small groups, i.e., a higher percentage of the entries will be selected for small groups, compared with large groups (Brown, 1989), while the Pr strategy favors larger groups, retaining a greater number of accessions belonging to these groups, which achieve greater representation within the core collection. Chandra, Huaman, Hari Krishna and R. Ortiz (2002) compared the three strategies using isoenzymatic data from different potatoes samples and concluded that the Pr strategy is superior to others in terms of the CC representativeness. Although most authors agree on the higher efficiency of the Pr strategy, Li, Shi, Cao and Wang (2005) reported that the optimum sampling strategy depends on the structure of germplasm as well as the groups defined before sampling.

Tables 3 and 4 show the number of accessions included in each group performed by both stratification criteria (S1 and S2), the number of

**Table 3.** Number of accessions included in each cluster (n) performed by the stratification criteria S1, number of accessions of each group included in the CCs constructed by the different sampling criteria (F, Ln, Pr) and percentage of CI accessions included in each CC (%).

Group	n	F	Ln	Pr
<b>G1</b>	7	4	3	2
<b>G2</b>	17	4	6	7
<b>G3</b>	16	4	5	6
<b>G4</b>	6	4	3	2
<b>G5</b>	6	4	3	2
<b>Total</b>	52	20	20	19
<b>Percentage of CI accessions included in each CC</b>	100%	38%	38%	36%

**Table 4.** Number of accessions included in each cluster (n) performed by the stratification criteria S2, number of accessions of each group (G) included in the CCs constructed by the different sampling criteria (F, Ln, Pr) and percentage of CI accessions included in each CC (%).

<i>C. cardunculus</i>	n	F	Ln	Pr
<i>var. scolymus</i> "globe artichoke"				
G1	16	3	4	6
G2	6	3	3	2
G3	6	3	3	2
Total	28	9	10	10
<i>var. altilis</i> "cultivated cardoon"				
G1	8	3	3	3
G2	10	3	3	4
Total	18	6	6	7
<i>var. sylvestris</i> "wild cardoon"				
G1	3	3	2	1
G2	3	3	2	1
Total	6	6	4	2
Grand Total	52	21	20	19
Percentage of CI accessions included in each CC (%)	100%	40%	38%	36%

accessions of each group which will be included in each CC, as well as the percentage of the IC accessions included in each CC.

Core collections constructed by S1 include 20 accessions (38 % of IC) when F and Ln criteria were applied and 19 (36 %) with Pr. As was previously said, the sampling criterion F allows that smaller groups are, proportionally, more represented than the larger ones. In this way, 66 % of wild cardoons, 22 % of cultivated cardoons, and 43 % of globe artichoke accessions were included in the CC. Ln and Pr criteria take into account the number of accessions included in each group making a more balanced representation of each one. In this case, the CC S1Ln included 50 % of the wild cardoon, 33 % of the cultivated cardoon and 39 % of the

globe artichoke accessions; whereas by S1Pr 19 accessions were included in the CC, reducing in one the wild cardoon and globe artichoke accessions and adding one cultivated cardoon.

The CC set up from S2 included 21 (40 %), 20 (38 %) and 19 accessions (36 %), respectively, when F, Ln and Pr strategies were applied (Table 4). When the F strategy was applied, all wild cardoon accessions present in the IC were included in the CC while cultivated cardoon and globe artichoke were underrepresented, with six and nine accessions, respectively. The Ln strategy threw out two wild cardoons, kept the same number of cultivate cardoon and added one globe artichoke accession to the CC, reducing in one the total accessions with respect to the previous strategy. The Pr strategy assigned to the CC the same number of accessions of each botanical variety than those included by S1Pr (10 accessions of globe artichoke, 7 of cultivated cardoon and 2 of wild cardoon). Therefore, the number of accessions included in the CC through the Pr strategy is independent of the applied stratification method.

Finally, the CC obtained by M strategy included 28 accessions. It seems that the great allele richness in the CC is achieved keeping 53 % of the IC accessions. This percentage is higher than that fixed as initial criterion to construct the CC (35 - 40 %).

In order to select the accessions of each cluster that would integrate the CC, mean values of both the accessions and clusters for each quantitative trait as well as phenotypic variances of the accessions were taken into account. Those accessions that showed mean values close to their cluster's mean values and exhibited low phenotypic variances were selected one by one until reaching

the number pre-established by each sampling criterion. This procedure was successfully applied by Noirot, Hamon and Anthony (1996) to form a coffee (*Coffea arabica*) core collection; and also by Zewdie, Tong and Bosland (2004) to construct a pepper (*Capsicum annuum*) CC.

The accessions of each botanical variety included in the CCs are showed in Table 5. Most CC, except PC, kept constant some accessions (CS04, CS05, CC05, CC08, CC10, CC18, A04, A05, A06, A10, A12, A14, and A15) which formed the "base set". These accessions, which belong to different clusters, are those that best represent the IC variability. The other accessions included in each CC varied with each strategy. On the other hand, although a greater number of accessions were included in the CC set out by PC, the "base set" was not maintained.

### Core collections validation

Once the core collections were formed, we had to evaluate to what extent these collections achieved the original objectives to represent the IC diversity avoiding repetitions. This validation process usually includes a comparison of each CC with the initial collection, using any suitable method. The criteria used for CC validation are often based on phenotypic data (Ortiz, Ruiz-Tapia and Mujica-Sanchez, 1998; Balakrishnan, Nair and Sreenivasan, 2000). Nevertheless, many authors have used molecular parameters (Bonierbale, Maya, Claros and Iglesias, 1995; Hatz, Jahoor and Fischbeck, 1996). This approach is relevant especially when phenotypes are evaluated in a limited range of environments, which can result in

**Table 5.** Accessions included in each core collection (CC). In bold, accessions belonging to the "base set".

CC	Globe artichoke	Cultivated cardoon	Wild cardoon	Total accessions
S1F	A02, <b>A04, A05, A06, A10, A12, A14, A15</b> , A16, A26, A28, A29	<b>CC05, CC08, CC10, CC18</b>	CS01, CS02, <b>CS04, CS05</b>	20
S1Ln	<b>A04, A05, A06, A10, A12, A14, A15</b> , A16, A20, A26, A28.	<b>CC05, CC08, CC10, CC18</b> , CC17, CC21	CS02, <b>CS04, CS05</b>	20
S1Pr	<b>A04, A05, A06, A10, A12, A14, A15</b> , A20, A25, A26.	<b>CC05, CC08, CC10, CC18</b> , CC19, CC21	<b>CS04, CS05</b>	19
S2F	<b>A04, A05, A06, A10, A12, A14, A15</b> , A16, A28,	<b>CC05, CC08, CC10, CC17</b> , CC18, CC21	CS01, CS02, CS03, <b>CS04, CS05</b> , CS06.	21
S2Ln	<b>A04, A05, A06, A10, A12, A14, A15</b> , A16, A26, A28.	<b>CC05, CC08, CC10, CC17</b> , CC18, CC21	CS01, CS02, <b>CS04, CS05</b>	20
S2Pr	<b>A04, A05, A06, A10, A12, A14, A15</b> , A20, A25, A26.	CC03, <b>CC05, CC08, CC10, CC17</b> , CC18, CC21	<b>CS04, CS05</b>	19
PC	A01, A02, <b>A04, A06</b> , A07, A13, <b>A14, A15</b> , A19, A21, A24, A26, A27, A28	CC03, CC04, <b>CC05</b> , CC06, CC07, <b>CC08, CC10</b> , CC13, CC14, CC19, CC20, CC21	CS02, <b>CS04</b>	28

an underestimation of the genetic variability due to genotype x environment interaction effects. In this work, a combination of both approaches (phenotypic and molecular) was used to validate the CCs. This strategy has already been successfully used by Wang *et al.* (2007) and Agrama *et al.* (2009) in the development of CCs of rice and apricot, respectively.

Validation of the CCs through quantitative traits (Table 6) showed that in all CCs CR %  $\geq$  80 % and MD % = 0. This last value indicates that the mean values observed for all variables in IC remain constant in the CCs. For VD %, the CCs formed by S1 strategy, as well as S2Pr and PC, remained within the acceptable range ( $\leq$  20 %) whereas in S2F and S2Ln this value increased. While values close to 100 % for CV % were observed in all the CCs, this value was only exceeded by S1Pr and PC. Therefore, through the CCs validation by quantitative parameters it was determined that only two CCs (S1Pr and PC) achieved the predetermined requirements (MD  $\leq$  20 %, VD  $\leq$  20 %, CR  $\geq$  80 %, and CV  $\geq$  100 %).

According to molecular validation, all CCs kept the IC diversity since no significant differences were observed between the CCs and the IC, for any of the estimated parameters, either from SRAP and SSR markers. The average number of alleles for both types of molecular markers remained equal. Some variation was observed for the total number of alleles, with loss of one or two alleles according to CC, but differences were not significant. In all

other parameters there was a tendency to increase in most CCs and decrease in others, but the differences were not significant either.

Several studies comparing different sampling strategies have been published with different results. Erksine and Muehlbauer (1991) did not find significant differences in the number of isozyme alleles captured when three sampling methods (including at random and stratified strategies) were compared. In our study, while similar results were obtained with the three evaluated sampling strategies (F, Ln and PR) F was less satisfactory because values out of the pre-established ones were obtained for some of the morphological parameters. On the other hand, Pr was the best strategy to preserve the IC variability, especially when the stratification method S1 was applied. In S1Pr, all validation parameters (morphological and molecular) fitted the preset ranges. Moreover, with the Pr strategy, it was possible to retain the lowest number of accessions (36 % of the IC).

The M strategy was proposed by Schoen and Brown (1993) as the strategy that allows to set the number of accessions that should form the CC as well as to identify which accessions should be included. In our work, CC performed by the M strategy showed all phenotypic parameters within acceptable ranges, however, it included too many accessions (52 % of the IC), widely exceeding the pre-established percentage of retention (35 - 40%). Bataillon, David and Schoen (1996), stated that the M strategy performs well when the accessions

**Table 6.** Parameters for CC validation. **In bold:** acceptable values for a CC

	IC	S1F	S1Ln	S1Pr	S2F	S2Ln	S2Pr	PC
<b>MD%</b>		<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
<b>VD%</b>		<b>15.92</b>	<b>13.61</b>	<b>14.15</b>	32.08	28.67	<b>16.40</b>	<b>17.45</b>
<b>CR%</b>		<b>94.52</b>	<b>90.51</b>	<b>90.27</b>	<b>87.83</b>	<b>87.80</b>	<b>90.23</b>	<b>99.09</b>
<b>CV%</b>		99.76	98.33	<b>100.28</b>	96.81	97.30	99.23	<b>108.50</b>
<b>AN SSR</b>	4.00	3.85 <sup>ns</sup>	4.00 <sup>ns</sup>	3.85 <sup>ns</sup>	4.00 <sup>ns</sup>	4.00 <sup>ns</sup>	3.85 <sup>ns</sup>	4.00 <sup>ns</sup>
<b>AN SRAP</b>	14.00	14.00 <sup>ns</sup>	14.00 <sup>ns</sup>	14.00 <sup>ns</sup>	14.00 <sup>ns</sup>	14.00 <sup>ns</sup>	14.00 <sup>ns</sup>	14.00 <sup>ns</sup>
<b>TA</b>	114	112 <sup>ns</sup>	113 <sup>ns</sup>	112 <sup>ns</sup>	113 <sup>ns</sup>	113 <sup>ns</sup>	112 <sup>ns</sup>	113 <sup>ns</sup>
<b>Ne</b>	2.75	2.72 <sup>ns</sup>	2.86 <sup>ns</sup>	2.88 <sup>ns</sup>	2.80 <sup>ns</sup>	2.84 <sup>ns</sup>	2.84 <sup>ns</sup>	2.72 <sup>ns</sup>
<b>S-W</b>	4.49	4.50 <sup>ns</sup>	4.49 <sup>ns</sup>	4.49 <sup>ns</sup>	4.48 <sup>ns</sup>	4.49 <sup>ns</sup>	4.49 <sup>ns</sup>	4.51 <sup>ns</sup>
<b>PIC SSR</b>	0.81	0.83 <sup>ns</sup>	0.83 <sup>ns</sup>	0.83 <sup>ns</sup>	0.83 <sup>ns</sup>	0.83 <sup>ns</sup>	0.83 <sup>ns</sup>	0.82 <sup>ns</sup>
<b>PIC SRAP</b>	0.35	0.36 <sup>ns</sup>	0.35 <sup>ns</sup>	0.35 <sup>ns</sup>	0.36 <sup>ns</sup>	0.36 <sup>ns</sup>	0.35 <sup>ns</sup>	0.34 <sup>ns</sup>
<b>H SSR</b>	0.84	0.85 <sup>ns</sup>	0.85 <sup>ns</sup>	0.85 <sup>ns</sup>	0.84 <sup>ns</sup>	0.84 <sup>ns</sup>	0.85 <sup>ns</sup>	0.84 <sup>ns</sup>
<b>H SRAP</b>	0.45	0.47 <sup>ns</sup>	0.47 <sup>ns</sup>	0.47 <sup>ns</sup>	0.47 <sup>ns</sup>	0.47 <sup>ns</sup>	0.46 <sup>ns</sup>	0.45 <sup>ns</sup>
Percentage of CI accessions retained in the CC (%)	100%	38%	38%	36%	40%	38%	36%	53%

CI: initial collection, MD%: mean differences percentage, VD%: variance differences percentage, CR%: coincidence rate of range, CV%: coefficient of variation; AN: average number of alleles per marker, TA: total alleles, Ne: effective number of alleles, SW: Shannon-Weaver index, PIC: polymorphic information content, H: heterozygosity, %CI: percentage of CI accessions retained in the CC. ns: not significant, \*: p<0.05.

come from populations with restricted gene flow or when the accessions are predominantly selfing, which does not occur in *Cynara cardunculus*.

## CONCLUSIONS

Although in all core collections the molecular parameters calculated in the initial collection remain unchanged, the sampling strategy Pr is the most efficient to keep the initial morphological variability, especially if it is combined with the stratification criterion S1. On the other hand, this last strategy achieves the requirement to maintain as few accessions as possible, which is essential for core collections, especially in species like *C. cardunculus*, a cross-pollinated species in which, two reproductive systems may coexist. For those accessions that must be sexually reproduced, spatial or temporal isolation is required during multiplication and regeneration in order to preserve their genetic integrity; meanwhile, for clonal accessions, only field or tissue culture conservation is possible. In both situations reducing the number of retained accessions is a critical point. In this sense, we conclude that the combination PrS1 is the best strategy to perform a representative core collection from a *Cynara cardunculus* L. initial collection, using both morphological and molecular data.

## ACKNOWLEDGMENTS

This work was supported by Agencia Nacional de Promoción Científica y Tecnológica, Ministerio de Ciencia Tecnología e Innovación Productiva de la Nación Argentina (Grant: PICT-1018/2012).

## REFERENCES

- Acquadro, A., Lanteri, S., Scaglione, D., Arens, P., Vosman, B. and Portis, E. (2009). Genetic mapping and annotation of genomic microsatellites isolated from globe artichoke. *Theoretical and Applied Genetics*, 118 (8), 1573–1587. doi: 10.1007/s00122-009-1005-6.
- Agrama, H. A., Yan, W. G., Lee, F., Fjellstrom, R., Chen, M. H., Jia, M. and McClung, A. (2009). Genetic assessment of a mini-core subset developed from the USDA Rice Genebank. *Crop Science*, 49 (4):1336–1346. doi: 10.2135/cropsci2008.06.0551.
- Balakrishnan, R., Nair, N. V. and Sreenivasan, T. V. (2000). A method for establishing a core collection of *Saccharum officinarum* L. germplasm based on quantitative morphological data. *Genetic Resources and Crop Evolution*, 47 (1):1–9. doi: 10.1023/A:1008780526154
- Balzarini, M. and Di Rienzo, J. Info-Gen: Software para análisis estadístico de datos genéticos (versión 2011) (Computer software). Córdoba, Argentina: Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba.
- Bataillon, T. M., David, J. L. and Schoen, D. J. (1996). Neutral genetic markers and conservation genetics: simulated germplasm collections. *Genetics*, 144:409–417.
- Bianco, V. V. (1990). Carciofo (*Cynara scolymus* L.). In: V. V. Bianco and F. Pimpini (Eds.); *Orticoltura* (209-251). Bologna, Italy: Patron Editore.
- Bonierbale, M., Maya, M., Claros, J. and Iglesias, C. (1995). Application of molecular markers to describing the genetic structure of cassava gene pool. In: *Proceedings of the Second International Scientific Meeting of the Cassava Biotechnology Network*. Working document no. 150 (pp. 106-122). Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.
- Brown, A. H. D. (1989). The case for core collection. In: A. H. D. Brown, O. H. Frankel, D. R. Marshall and J. T. Williams (Eds.), *The use of plant genetic resources* (136-156). Cambridge, U. K. Cambridge University Press.
- Brown, A. H. D. (1995). The Core Collection at the Crossroads. In: T. Hodgkin, A. H. D. Brown, T. J. L van Hintum and E. A. V. Morales (Eds.), *Core Collections of Plant Genetic Resources* (3-19). New York, USA. John Wiley and Sons.
- Brown, A. H. D. and Schoen, D. J. (1994). Optimal sampling strategies for core collections of plant genetic resources. In: Loeschcke, V., Jain, S. K. and Tomiuk, J. (Eds.), *Conservation Genetics*. vol 68. Basel, Switzerland. Birkhäuser, EXS.
- Casadevall, R., Martin, E. and Cravero, V. (2011). Simple Sequence Repeat (SSR) vs. Sequence-Related Amplified Polymorphism (SRAP) markers for *Cynara cardunculus* characterization. *Spanish Journal of Agricultural Research*, 9 (2), 453-459. doi: 10.5424/sjar/20110902-161-10
- Chandra, S., Huaman, Z., Hari Krishna, S. and Ortiz, R. (2002). Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data: A simulation study. *Theoretical and Applied Genetics*, 104(8), 1325–1334. doi: 10.1007/s00122-001-0854-4
- Cravero, V. P., Martin, E. A. and Cointry, E. L. (2007). Genetic diversity in *Cynara cardunculus* determined by sequence-related amplified polymorphism markers. *Journal of the American Society for*

- Horticultural Science*, 132 (2), 208-212. Doi: 10.21273/JASHS.132.2.208.
- Crossa, J., DeLacy, I. H. and Taba, S. (1995). The use of multivariate methods in developing a Core Collection. In: T. Hodgkin, A. H. D. Brown, T. J. L. van Hintum and E. A. V. Morales (Eds.), *Core Collections of Plant Genetic Resources* (77-89). New York, USA. John Wiley and Sons.
- De Felice, B., Borra, M., Manfellotto, F., Anna, S., Biffali, E. and Guida, M. (2015). Assessment of genetic diversity between wild and cultivated artichokes using SSR markers. *Genetic Resources and Crop Evolution*, 63 (8), 1363–1369. doi: 10.1007/s10722-015-0323-6.
- Dellacecca, V., Magnifico, V., Marzi, V., Porceddu, E. and Mugnozza, G.T. (1976). Contributo alla conoscenza delle varietà di carciofo coltivate nel mondo. In: Minerva Médica (Ed.), *Atti 2° Congresso Internazionale di Studi sul Carciofo* (pp.199-315). Turin, Italy.
- Diwan, N., McIntosh, M. S. and Bauchan, G. R. (1995). Methods of developing a core collection of annual Medicago species. *Theoretical and Applied Genetics*, 90 (6), 755-761. doi: 10.1007/BF00222008.
- Erksine, W. and Muehlbauer, F. J. (1991). Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. *Theoretical and Applied Genetics*, 83 (1), 119-125. doi: 10.1007/BF00229234.
- Fernández, J., Curt, M. D. and Aguado, P. L. (2006). Industrial applications of *Cynara cardunculus* L. for energy and other uses. *Industrial Crops and Products*, 24 (3), 222–229. doi: 10.1016/j.indcrop.2006.06.010.
- Frankel, O. H. (1984). Genetic perspectives of germplasm conservation. In: W. K. Arber, K. Llimensee, W. J. Peacock and P. Stralinger (Eds.), *Genetic Manipulation: Impact on Man and Society* (161-170). Cambridge, U. K. Cambridge University Press.
- Gatto, A., De Paola, D., Bagnoli, F., Vendramin, G. G. and Sonnante, G. (2013). Population structure of *Cynara cardunculus* complex and the origin of the conspecific crops artichoke and cardoon. *Annals of Botany*, 112 (5), 855–865. doi: 10.1093/aob/mct150.
- Gouesnard, B., Bataillon, T. M., Decoux, G., Rozale, C., Schoen, D. J. and David, J. L. (2001). MSTRAT: An Algorithm for Building Germ Plasm Core Collections by Maximizing Allelic or Phenotypic Richness. *Journal of Heredity*, 92 (1), 93–94. doi: 10.1093/jhered/92.1.93.
- Gower, J. C. (1975). Generalised procrustes analysis. *Psychometrika*, 40 (1), 33-51. doi: 10.1007/BF02291478.
- Hatz, B. G., Jahoor, A. and Fischbeck, G. (1996). RFLP-polymorphism among European accessions of the barley core collection. In: G. Scoles and B. Rosnagel (Eds.), *Proceeding of the VII International Barley Genetics Symposium* (176-178). Saskatoon, Canada.
- Hu, J., Zhu, J., Xu, X. M. (2000). Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. *Theoretical and Applied Genetics*, 101 (1-2), 264-268. doi: 10.1007/s001220051478.
- Igartua, E., Gracia, M. P., Lasa, J. N., Medina, B., Molina-Cano, J. L., Montoya, J. L. and Romagosa, I. (1998). The Spanish barley core collection. *Genetic Resources of Crop Evolution*, 45 (5), 475–481. doi: 10.1023/A:1008662515059.
- Kim, T. H., Jung, S. H. and Cho, K. H. (2007). Interlinked mutual inhibitory positive feedbacks induce robust cellular memory effects. *FEBS Letters*, 581(25), 4899-904. Doi: 10.1016/j.febslet.2007.09.020.
- Knüpffer, H. and van Hintum, T. J. L. (1995). The Barley Core Collection: an international effort. In: T. Hodgkin, A. H. D. Brown, T. J. L. van Hintum and E. A. V. Morales (Eds.), *Core Collections of Plant Genetic Resources* (171-178). New York, USA. John Wiley and Sons.
- Lanteri, S., Di Leo, I., Ledda, L., Mammeli, M. G. and Portis, E. (2001). RAPD variation within and among populations of globe artichoke (*Cynara scolymus* L.) cv "Spinoso sardo". *Plant Breeding*, 120 (3), 243-246. doi: 10.1046/j.1439-0523.2001.00605.x.
- Li, G. and Quiros, C. F. (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. *Theoretical and Applied Genetics*, 103 (1-2), 455-461. doi: 10.1007/s001220100570.
- Li, Z., Zhang, H. L., Cao, Y. S., Qiu, Z. E., Wei, X. H., Tang, S. X., Yu, P. and Wang, X. K. (2003). Studies on the sampling strategy for primary core collection of Chinese ingenious rice. *Acta Agronomica Sinica*, 29, 20–24.
- Li, Y., Shi, Y., Cao, Y. and Wang, T. (2005). Establishment of a core collection for maize germplasm preserved in Chinese National Genebank using geographic distribution and characterization data. *Genetic Resources and Crop Evolution*, 51 (8), 845-852. doi: 10.1007/s10722-005-8313-8.
- Li, X., Yan, W., Agrama, H., Hu, B., Jia, L. and Jia, M. (2010). Genotypic and phenotypic characterization of genetic differentiation and diversity in the USDA rice mini-core collection. *Genetica*, 138 (11-12), 1221-1230. doi: 10.1007/s10709-010-9521-5.
- Mauro, R., Portis, E., Acquadro, A., Lombardo, S., Mauromicale, G. and Lanteri, S. (2009). Genetic diversity of globe artichoke landraces from Sicilian small-holdings: implications for evolution and domestication of the species. *Conservation Genetics*, 10 (2), 431-440. doi: 10.1007/s10592-008-9621-2.
- Nagy, S., Poczai, P., Cernák, I., Gorji, A.M., Hegedus, G. and Taller, J. (2012). PIC calc: an online program

- to calculate polymorphic information content for molecular genetic studies. *Biochemical Genetics*, 50 (9-10), 670-672. doi: 10.1007/s10528-012-9509-1.
- Noirot, M., Hamon, S. and Anthony, F. (1996). The principal component scoring: a new method of constituting a core collection using quantitative data. *Genetic Resources and Crop Evolution*, 43 (1), 1-6. doi: 10.1007/BF00126934.
- Ortiz, R., Ruiz-Tapia, E. N. and Mujica-Sanchez, A. (1998). Sampling strategy for a core collection of Peruvian quinoa germplasm. *Theoretical and Applied Genetics*, 96 (3-4), 475-483. doi: 10.1007/s001220050764.
- Pagnotta, M. A., Fernández, J. A., Sonnante, G. and Egea-Gilabert, C. (2017). Genetic diversity and accession structure in European *Cynara cardunculus* collections. *PLoS ONE* 12 (6):e0178770. doi: 10.1371/journal.pone.0178770.
- Portis, E., Barchi, L., Acquadro, A., Macua, J. I. and Lanteri, S. (2005). Genetic diversity assessment in cultivated cardoon by AFLP (Amplified Fragment Length Polymorphism) and microsatellite markers. *Plant Breeding*, 124 (3), 299-304. doi: 10.1111/j.1439-0523.2005.01098.x.
- Raccuia, S. A., Mainolfi, A., Mandolino, G. and Melilli, M. G. (2004). Genetic diversity in *Cynara cardunculus* revealed by AFLP markers: comparison between cultivars and wild types from Sicily. *Plant Breeding*, 123 (3), 280-284. doi: 10.1111/j.1439-0523.2004.00983.x.
- Ruiz Valcárcel, M., Giraldo Carbajo, P., Royo, C. and Carrillo Becerril, J. M. (2013). Creation and validation of the Spanish durum wheat Core Collection. *Crop Science*, 53 (6), 2530-2537. doi: 10.2135/cropsci2013.04.0238.
- Schoen, D. J. and Brown, A. H. D. (1993). Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. *Proceedings of the National Academy of Sciences USA*, 90 (22), 10623-10627. doi: 10.1073/pnas.90.22.10623.
- Sokal, R. R. and Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *Taxon* 11 (2), 33-40. doi: 10.2307/1217208.
- Sonnante, G., Carluccio, A., De Paolis, A. and Pignone, D. (2008). Identification of artichoke SSR markers: molecular variation and patterns of diversity in genetically cohesive taxa and wild allies. *Genetic Resources and Crop Evolution*, 55 (7), 1029-1046. doi: 10.1007/s10722-008-9310-5.
- Tivang, J., Skroch, P. W., Nienhuis, J. and De Vos, N. (1996). Randomly amplified polymorphic DNA (RAPD) variation among and within artichoke (*Cynara scolymus* L.) cultivars and breeding populations. *Journal of the American Society for Horticultural Science*, 121 (5), 783-788. doi: 10.21273/JASHS.121.5.783.
- Tohme, J., Jones, P., Beebe, S. and Iwanaga, M. (1995). The combined use of agroecological and characterization data to establish the CIAT *Phaseolus vulgaris* core collection. In: T. Hodgkin, A. H. D. Brown, T. J. L. van Hintum and E. A. V. Morales (Eds.), *Core Collections of Plant Genetic Resources* (95-107). New York, USA. John Wiley and Sons.
- van Hintum, T. J. L. (1995). Hierarchical approaches to the analysis of genetic diversity in crop plants. In: T. Hodgkin, A. H. D. Brown, T. J. L. van Hintum and E. A. V. Morales (Eds.), *Core Collections of Plant Genetic Resources* (23-34). New York, USA. John Wiley and Sons.
- van Hintum, T. J. L., Brown, A. H. D., Spillane, C. and Hodgkin, T. (2003). Colecciones núcleo de recursos fitogenéticos. *Boletín Técnico No. 3 del IPGRI*, 42 pp. Instituto Internacional de Recursos Fitogenéticos, Rome, Italy.
- Wang, J. C., Hu, J., Zhang, C. F. and Zhang, S. (2007). Assessment on evaluating parameters of rice core collections constructed by genotypic values and molecular marker information. *Rice Science*, 14 (2), 101-110. doi: 10.1016/S1672-6308(07)60015-8.
- Zewdie, Y., Tong, N. K. and Bosland, P. (2004). Establishing a core collection of capsicum using a cluster analysis with enlightened selection of accessions. *Genetic Resources and Crop Evolution*, 51 (2), 147-151. doi: 10.1023/B:GRES.0000020858.96226.38.
- Zhang, X., Zhao, Y., Cheng, Y., Feng, X., Guo, Q., Zhou, M. and Hodgkin, T. (2000). Establishment of sesame germplasm core collection in China. *Genetic Resources and Crop Evolution*, 47 (3), 273-279. doi: 10.1023/A:1008767307675
- Zhang, X., Wang, J. Y., Zhang, X. Z., Li, T. H., Wang, K., Xu, X. F. and Han, Z. H. (2010). Sampling strategy to develop a primary core collection of apple cultivars based on fruit traits. *African Journal of Biotechnology*, 9 (2), 123-127. doi: 10.5897/AJB09.300.