The Population Structure of a Globe Artichoke Worldwide Collection, as Revealed by Molecular and Phenotypic Analyzes

Domenico Rau1*, Giovanna Attene1, Monica Rodriguez1, Limbo Baghino2, Anna Barbara Pisanu2, Davide Sanna2, Alberto Acquadro3, Ezio Portis3* and Cinzia Comino3

1 Dipartimento di Agraria, Sezione di Agronomia, Coltivazioni Erbacee e Genetica (SACEG), Università degli Studi di Sassari, Sassari, Italy, 2 Agenzia AGRIS Sardegna (Servizio Ricerca sui Sistemi Colturali Erbacei, Settore Innovazione dei Modelli Gestionali e Studio Della Biodiversità Nelle Colture Intensive), Oristano, Italy, 3 Dipartimento di Scienze Agrarie, Forestali ed Alimentari (DISAFA), Genetica Vegetale (Plant Genetics), Università degli Studi di Torino, Turin, Italy

The knowledge of the organization of the domesticated gene pool of crop species is an essential requirement to understand crop evolution, to rationalize conservation programs, and to support practical decisions in plant breeding. Here, we integrate simple sequence repeat (SSR) analysis and phenotypic characterization to investigate a globe artichoke collection that comprises most of the varieties cultivated worldwide. We show that the cultivated gene pool of globe artichoke includes five distinct genetic groups associated with the major phenotypic typologies: Catanesi (which based on our analysis corresponds to Violetti di Provenza), Spinosi, Violetti di Toscana, Romaneschi, and Macau. We observed that 17 and 11% of the molecular and phenotypic variance, respectively, is between these groups, while within groups, strong linkage disequilibrium and heterozygote excess are evident. The divergence between groups for quantitative traits correlates with the average broad-sense heritability within the groups. The phenotypic divergence between groups for both qualitative and quantitative traits is strongly and positively correlated with SSR divergence ($F_{ST}$) between groups. All this implies a low population size and strong bottleneck effects, and indicates a long history of clonal propagation and selection during the evolution of the domesticated gene pool of globe artichoke. Moreover, the comparison between molecular and phenotypic population structures suggests that harvest time, plant architecture (i.e., plant height, stem length), leaf spininess, head morphology (i.e., head shape, bract shape, spininess) together with the number of heads per plant were the main targets of selection during the evolution of the cultivated germplasm. We emphasize our findings in light of the potential exploitation of this collection for association mapping studies.

Keywords: Cynara cardunculus var. scolymus L., divergent selection, germplasm collection, microsatellite, $Q_{ST}$, quantitative traits, simple sequence repeats, UPOV traits
INTRODUCTION

The study of the organization of the gene pool of crops is an essential prerequisite for both theoretical and practical reasons. This provides insight into the evolutionary history of a crop, strengthens our knowledge for the implementation of correct conservation strategies, and rationalizes operational choice in plant breeding.

For major crops, many studies on population structure are available, such as *Triticum aestivum* (Würschum et al., 2013), *Oryza sativa* (Jin et al., 2010), *Zea mays* (Vigouroux et al., 2008) and *Hordeum vulgare* (Malysheva-Otto et al., 2006; Tanto Hadado et al., 2010; Rodrigues et al., 2012), *Lycopersicon esculentum* (Allitos et al., 2014; Rodrigues et al., 2020), *Solanum melongena* (Cericola et al., 2013, 2014; Portis et al., 2015), *Phaseolus spp.* (Angioi et al., 2010; Bitocchi et al., 2013; Rodrigues et al., 2013), and *Glycine max* (Bandillo et al., 2015). However, for other minor crops, such as *Cynara cardunculus* var. *scolymus* L., the cultivated globe artichoke, information remains limited.

Globe artichoke (2n = 2x = 34) is a perennial open-pollinated crop (De Vos, 1992; Pecaut, 1993) that is probably native to the Mediterranean basin (Sonnante et al., 2007; Gatto et al., 2013) where the cultivation of this crop is also particularly relevant. Based on data from the Food and Agriculture Organization Corporate Statistical Database1 (FAOSTAT, 2020), the total world artichoke production in 2020 was 1,516,955 tons. In total, three circum-Mediterranean countries, namely Italy (367,080 tons), followed by Egypt (308,844 tons), and Spain (196,970 tons), were the most important world producers, accounting for 57.5% of the global production.

The edible portion of the artichoke plant is the head, or capitulum, which is formed as a composite inflorescence that is harvested before the flowers bloom. Furthermore, globe artichoke is considered a functional food and a source of nutraceutical ingredients, such as bioactive phenolic compounds, inulin, fiber, and minerals; in addition, artichoke leaf extracts have long been used in folk medicine, particularly for liver complaints (Comino et al., 2007, 2009; Lattanzio et al., 2009; Moglia et al., 2009, 2016; Menin et al., 2010, 2012; Eljounaidi et al., 2014, 2015).

Many different clonal varietal groups best adapted to different local environments have been described in this species. Cultivated germplasm has traditionally been classified into four main groups based on the capitulum traits (Dellacecca et al., 1976; Porceddu et al., 1976; Vanella et al., 1981): (1) The "Spinosi" group has long spines on the bracts and leaves; (2) "Violetti" has medium-sized, violet-colored, and less spiny heads; (3) "Romaneschi" has spherical or sub-spherical non-spiny heads; and (4) "Catanesi" has relatively small, elongated, and non-spiny heads. Elia and Miccolis (1996) conducted a phenotypic characterization in a common garden experiment. They conducted a multivariate analysis considering eight quantitative traits across 104 accessions, and they distinguished five main morphogenetical clusters. These were early; medium-early; late with small heads; late-violet; and late with large heads. Furthermore, Lanteri et al. (2004) studied a wide collection of cultivated globe artichoke using amplified fragment length polymorphism (AFLP) markers and found concordances between molecular and the proposed phenotypic classifications. Indeed, the performed multivariate analysis showed four main clusters that corresponded to the phenotypical typologies of cultivated globe artichoke: “Catanesi” (together with “Violet de Provence”), “Romaneschi,” “Violetti di Toscana,” and “Spinosi.”

Later, by analyzing leafy cardoon and globe artichoke accessions using SSR and inter-SSR markers, Pagnotta et al. (2017) observed that “Catanesi” resulted well separated from the other accessions, and the “Spinosi” types formed a well-defined group together with some “Romaneschi” and “Violetti” types. The “Romaneschi” types showed low uniformity, as they were distributed into several genetic subgroups. Thus, except for the “Catanesi” group, the association between molecular groups and phenotypic types was not clear. Moreover, as noted by Pagnotta et al. (2017), the accessions clustered not only based on their typology but also based on the geographical origin. They argued that the environmental conditions of the different field gene banks or the methods used to conserve plant materials might have affected the germplasm characteristics.

More recently, Pavan et al. (2018) used genotyping by sequencing to analyze samples of globe artichoke, cultivated cardoon (*C. cardunculus* var. *altilis*), and wild cardoon (*C. cardunculus* var. *sylvestris*), and showed a clear-cut genetic separation between globe artichoke and the cultivated and wild cardoon. When the analysis was repeated assuming three genetic groups, the authors observed the separation of “Catanesi” from all of the other accessions and that the group of green-headed landraces typical of the Apulia region, Italy (Green Apulian), also formed a well-defined group. Furthermore, similarly to Pagnotta et al. (2017) and Pavan et al. (2018) observed that “Romaneschi” were attributed to a variable group that consisted of admixed individuals, which suggested that they should not be considered a genetically uniform varietal typology.

To the best of our knowledge, no studies are available that compare the molecular and phenotypic population structures for the same collection of globe artichokes. Thus, in this study, we characterized a germplasm collection comprising 110 of the most important globe artichoke varieties cultivated worldwide by conducting a deep phenotypic characterization in a common garden and under field conditions. A total of thirty-five quantitative phenological, morphological, and yield traits and 23 qualitative descriptors from the International Union for the Protection of New Varieties of Plants (UPOV) were considered to describe the plants in the stages of the first, secondary, and tertiary heads. Thus, the phenotypic data were compared with the results of SSR analysis, to investigate the relationships between molecular and phenotypic variability.

This study advances our understanding of the evolution of globe artichoke, strengthens the knowledge of the organization of its gene pool, and can help in rationalizing the operational choices when applying association mapping and defining breeding strategies in this species.
MATERIALS AND METHODS

Plant Material
A total of 110 globe artichoke accessions were analyzed using both SSR and quantitative traits, as listed in Supplementary Table 1. All of these came from the living artichoke collection maintained by AGRIS at the experimental farm located at Oristano (Sardinia, Italy). The collection comprises most of the globe artichoke varieties cultivated worldwide, with the accessions subdivided into seven main typologies based on their different morphogenotypic traits (primarily related to head characteristics). These typologies are named as “Catanesi” (CAT), “Violetti di Provenza” (VPR), “Violetti di Toscana” (VTO), “Romaneschi” (ROM), “Macau” (MAC), “Spinosi” (SPI), and “Green et al.” (GEA), as reported by Comino et al. (2016).

Nine clones of “Spinoso sardo” selected by AGRIS based on morphological and production traits were also analyzed (Supplementary Table 1).

Simple Sequence Repeat Assay
Total genomic DNA was extracted from young leaves (three plants per accession) and characterized using 26 genomic SSR (gSSR) markers targeting loci distributed through all of the 17 chromosomes of globe artichoke (Supplementary Table 2A). The markers were selected according to their good resolution power and reliability (Scaglione et al., 2009, 2016; Portis et al., 2012, 2016).

Chloroplast-SSR (SR) variation was also investigated using 10 primer pairs designed by Weising and Gardner (1999) and 23 designed by Chung and Staub (2003) (Supplementary Table 2B); the markers have been already used to study the population structure of wild cardoon, the progenitor of globe artichoke (Rau et al., 2016), and other plants (see, e.g., Angioi et al., 2009, 2010; Desiderio et al., 2013).

Individual-Based Clustering
To infer the population structure of the globe artichoke collection using multilocus genetic profiles, we first used the model-based (Bayesian) clustering algorithm implemented in the Structure software (Pritchard et al., 2000; Falush et al., 2003). The data were analyzed by setting the admixture model and the correlated allele frequencies options. All runs were carried out with 50,000 burn-ins and 100,000 iterations (Falush et al., 2003). Phenotypic typology was not used as prior information to assist clustering. A total of twenty independent runs were carried out for each number of assumed ancestral populations (K). A range of K from 1 to 10 was explored (Evanno et al., 2005). The K value that best explained the data was determined using the ad hoc statistics (ΔK; Evanno et al., 2005) with the CLUMPAK software (Kopelman et al., 2015).

To complement the population structure analysis, discriminant analysis of principal components (DAPC) was used as a multivariate unsupervised method implemented in the R package adegenet (Jombart, 2008; Jombart et al., 2010). DAPC divides genetic variation into between- and within-group components, to maximize the former, and builds linear combinations of alleles (linear discriminants) that best separate the groups. The find.clusters function was used to detect the number of clusters in the collection adopting the K-means clustering method and determining the most likely number of subpopulations based on the Bayesian Information Criterion (BIC). The number of principal components to be retained was determined using the cross-validation function Xval.dapc, based on the root mean square error.

Variation Among Genetic Groups
The total SSR genetic variance (σ²T) was partitioned into two components: among-population (σ²S) and within-population (σ²b). The divergence between populations was quantified as FST = σ²a/σ²T, where σ²T = σ²a + σ²b. The partition of molecular variance was accomplished by the AMOVA implemented in Arlequin v. 3.1.5.2 (Excoffier and Lischer, 2010). The level of group divergence was also quantified by calculating the RST. While FST is based on the infinite allele model (IAM), RST is based on the stepwise mutation model (SMM), i.e., it also considers the variance of SSR allele size (Gaggiotti et al., 1999). The FST and RST statistics have the same expectations when group divergence is caused only by drift, whereas RST is larger than FST if stepwise-like mutations are also contributing to divergence; thus, comparing the FST and RST values can provide insights into the relative contribution of drift and mutation to group divergence (Hardy et al., 2003). This comparison can be done by testing whether the observed RST is significantly larger than its value after randomizing allele sizes (pRST; Hardy et al., 2003). If the RST is not significantly different from pRST, it is possible to conclude that RST = FST (Hardy et al., 2003). The p-values were obtained after 999 random permutations using the SPAGeDI software (Hardy and Vekemans, 2002).

Diversity and Linkage Disequilibrium Within Genetic Groups
The number of alleles per locus (na), the observed heterozygosity (H0), and the gene diversity (He, Nei, 1978) were estimated for each genetic group identified by the Structure and DAPC analyzes. When referring to the qualitative traits variation, to avoid confusion with SSR, we indicate the number of classes per trait as n and the gene diversity of Nei (1978) as I (instead of He). To determine the coefficient of inbreeding (FIS), total variance (σ²T) was partitioned into three components: among populations (σ²S), among individuals within-populations (σ²b), and within individuals (σ²a). The total FIS was calculated as σ²a/(σ²S + σ²a). Group-specific FIS was also calculated. This coefficient compares the expected and observed heterozygosities and the expected value under the Hardy-Weinberg equilibrium is 0. Thus, FIS > 0 indicates heterozygote deficit compared to the Hardy-Weinberg equilibrium (as a consequence of inbreeding), while FIS < 0 indicates heterozygote excess (as in clonally propagated crops or when selection in favor of heterozygotes is applied).

The BOTTLENECK ver. 1.2 software (Piry et al., 1999) was used to determine whether the globe artichoke groups experienced recent founder events or population bottlenecks. In this case, both a reduction of the number of alleles (na) and of expected heterozygosity (He) is expected, with the former that
reduces faster than the latter (Nei, 1975). Given \(n_a\) and the sample size, the program BOTTLENECK computes, for each sample and for each locus, the distribution of the heterozygosity expected under the assumption of mutation-drift equilibrium (\(H_{EQ}\)). As \(n_a\) reduces faster than \(H_E\), it is expected that, on average, after a bottleneck \(H_E > H_{EQ}\) (Cornuet and Luikart, 1996). The number and percentage of pairs of SSR loci in linkage disequilibrium (LD; \(p < 0.05\)) were calculated using the Arlequin software (Excoffier and Lischer, 2010).

**Phenotypic Characterization and Statistical Analysis**

The collection was characterized in the AGRIS experimental field (Azienda Palloni, Oristano, Sardinia, Italy). Individual plants were obtained by vegetative propagation using “ovoli” in July. The plants were managed following standard agronomic practices for the crop, both for mineral nutrition and water supply.

Phenotypic data were recorded under field conditions, measuring three plants for each variety (single plant = replicate) in a completely randomized design. A total of twenty-one traits were recorded at the beginning of the first head harvest (Supplementary Table 3A). Among these, 11 traits were also recorded for the secondary and tertiary heads (Supplementary Table 3B). The traits “Total number of heads” and “Total weight of heads” were measured for secondary and tertiary heads. The collection was also characterized by recording 23 qualitative phenotypic descriptors defined by UPOV (Supplementary Table 3C).

The total phenotypic variance (\(\sigma^2_T\)) was decomposed into two components, within SSR groups (\(\sigma^2_W\)) and between SSR groups (\(\sigma^2_B\)), applying Nested Analysis of Variance (NANOVA). For each trait, the variance associated with each component was estimated considering a model with groups, varieties within groups, and replicates as random factors and applying the restricted maximum likelihood (REML) method implemented in the JMP software (ver. 10.0.0, SAS, Institute, NC, United States).

Variance components were first used to calculate the heritability of traits. In globe artichoke, cross-fertilization is promoted by protandry (Mauromicale and Lerna, 2000), and when the progenies are tested, they show wide segregation of morphological and production traits (Basnitzki and Zohary, 1987). Thus, high heterozygosity was expected a priori. As the globe artichoke varieties considered in this study were obtained by vegetative propagation (clonally), it is assumed that the three replicates of each variety have the same genotype. This means that genetic differences among varieties, both within groups (\(\sigma^2_W\)) and between groups (\(\sigma^2_B\)), can be partially due to maternal and non-additive effects (dominance and epistasis), rather than to additive effects. Thus, the sum \(\sigma^2_W + \sigma^2_B\) is equal to the total genetic variance (\(\sigma^2_G\)) and the ratio between genetic variance (\(\sigma^2_G\)) and total variance (\(\sigma^2_T\)) represents the broad-sense heritability (\(h^2_B\)); i.e., \(h^2_B = \sigma^2_G/\sigma^2_T\), where \(\sigma^2_T = \sigma^2_W + \sigma^2_B + \sigma^2_E\), with \(\sigma^2_E\) representing the environmental variance.

To explore the pattern of phenotypic divergence between genetic groups of globe artichoke varieties, variance components were used to calculate the degree of differentiation for quantitative traits between groups by applying the formula

\[
Q_{ST} = \frac{\sigma^2_B}{\sigma^2_G + 2\sigma^2_W}
\]

(Spitze, 1993; Whitlock, 1999; Leinonen et al., 2008). The \(Q_{ST}\) statistic was introduced by Spitze (1993) as an analog to \(F_{ST}\) for molecular markers. This equation assumes that \(F_{IS} = 0\) and that \(\sigma^2_B\) is the additive genetic variance due to differences between populations, and \(\sigma^2_W\) is the mean additive genetic variance within populations. Successively, Bonnin et al. (1996) derived a generalized relation for any possible value of \(F_{IS}\), for which

\[
Q_{ST(i)} = \frac{(1 + F_{IS})\sigma^2_B/[(1 + F_{IS})\sigma^2_B + 2\sigma^2_W]}{\sigma^2_G + 2\sigma^2_W}
\]

To calculate the \(Q_{ST}\), we also applied this formula by entering the estimated total \(F_{IS}\) (see previous paragraph) and assuming that the \(F_{IS}\) of a quantitative trait locus is the same as the \(F_{IS}\) of the neutral markers. We presented two \(Q_{ST}\) values calculated as in Spitze (1993) and Bonnin et al. (1996). In our case, \(\sigma^2_B\) and \(\sigma^2_W\) are components of the total genetic variance, as they can include both non-additive and maternal effects. To mark this difference compared to Spitze’s \(Q_{ST}\) and Bonnin’s \(Q_{ST(i)}\), we used the notation \(Q_{ST_B}\) and \(Q_{ST(i)_B}\), where the subscript “B” means “Broad.”

The qualitative descriptors of the UPOV were statistically treated as molecular markers, considering each trait as a single SSR marker and the states of the trait as different alleles. Thus, for each group, statistics of genetic diversity (\(n_a, H_E\)) and divergence (\(F_{ST-qtl}\)) were calculated as for the SSR markers (Tanto Hadado et al., 2009). The correlations between the three divergence measures (i.e., \(F_{ST}, Q_{ST_B}, F_{ST-qtl}\)) were determined. The values of \(F_{ST-qtl}\) were compared with the \(F_{ST}\) values obtained for SSRs to flag qualitative traits putatively under selection, using the Arlequin software v. 3.5.1.2 (see below).

The levels of molecular and quantitative variations were compared following the rationale of the \(F_{ST} - Q_{ST}\) approach (Spitze, 1993; Whitlock, 1999; Leinonen et al., 2008). This is based on the relative magnitude of the \(F_{ST}\) and \(Q_{ST}\) and allows inferences to be made about the relevance of selection in the evolution of a specific quantitative trait (see “Discussion” section). To exclude possible selection effects on \(F_{ST}\) values calculated for the SSR loci, the (putatively) neutral “benchmark” \(F_{ST}\) was calculated by conducting the outlier tests as implemented in Arlequin v. 3.5.1.2 (100,000 reps). This test is similar to that implemented in FDIST computer program (Beaumont and Nichols, 1996). We ran simulations reiteratively, eliminating after each simulation the markers that showed the signature of divergent selection (\(p < 0.05\)). After obtaining a dataset without markers with signatures of divergent selection, the average \(F_{ST}\) across loci was assumed as the putatively neutral benchmark. The confidence intervals (C.I.) of the putatively neutral \(F_{ST}\) were determined by bootstrapping over loci (100,000 replicates).

**RESULTS**

**Population Structure Based on Genomic SSRs**

All SSR markers tested were polymorphic (Minimum Allele Frequency, MAF, > 1%). A total of 152 alleles were observed: two to 13 alleles per locus (\(n_a\)) were detected, giving a mean of 5.85. Gene diversity (\(H_E\)) per locus varied from 0.231 to 0.822
When the 26 SSRs were combined, each accession showed a unique multilocus profile.

Based on Evanno’s AK statistics, the highest hierarchical level of the population structure was K = 2 (Supplementary Figure 1A). However, a secondary peak at K = 5 was also evident, which indicated substructures (Supplementary Figure 1). The genetic subdivisions resembled phenotypic typologies quite well (Figure 1A). Indeed, at K = 2, most of the Catanesi and Violetti di Provenza types clustered together and were clearly separated from the remaining accessions; most of the individuals (> 95%) were assigned to their group with qi > 0.80 (Figure 1B). The two additional clustering steps first separated most Spinosi (K = 3) and then a group that included all the Violetti di Toscana and most green-headed types (K = 4; Figure 1A). Finally, at K = 5, most of the Romanschi were split from Macau. At K = 5, Catanesi and Violetti di Provenza still clustered together, while Violetti di Toscana grouped with many Green-headed accessions (Figure 1A). This remained true for K up to 10 (not shown). The phenotypic typology “Green et al.” was not genetically well-defined. Indeed, it comprised varieties assigned to different genetic groups that were typical of Violetti di Toscana, Macau, and Catanesi-Violetti di Provenza (Figure 1A).

Based on these data, from here on, we adopted the label CAT-VPR to designate the nuclear SSR genetic group related to the Catanesi-Violetti di Provenza phenotypic typology; similarly, the labels SPI, VTO-GEA, ROM, and MAC indicate the genetic groups associated with the phenotypic typologies Spinosi, Violetti di Toscana-Green et al., Romanschi, and Macau, respectively. Overall, at K = 5, most of the accessions were assigned to their group with high probability (qi > 0.8); Figure 1B). Overall, there was a strong association between phenotypic typologies and genetic classifications (χ² test: p < 10⁻⁴). The highest frequency (~30%) of admixed individuals (qi < 0.80) was observed within the VTO-GEA group.

CAT-VPR and SPI showed the highest net-nucleotide distance (0.171), while the lowest distances were observed between MAC and ROM (0.084), VTO-GEA and MAC (0.073), and VTO-GEA and ROM (0.071) (Figure 1C). According to AMOVA, the average genetic divergence (FST) between these five groups was moderate, being 0.170 (Table 1) while putatively neutral FST was slightly lower (0.134).

According to RST, the average divergence among the genetic groups was 0.148 (Table 1). Furthermore, the RST and FST values were not statistically different (Table 1) because the observed RST (0.148) and the RST calculated by randomizing allele size were not different (pRST = 0.179; p = 0.364; Table 1). Therefore, in this study, stepwise-like mutations are not contributing to group divergence. The neighbor-joining tree emphasizes the proximity between ROM, MAC, and VTO-GEA, and their distance from CAT-VPR and SPI (Figure 1D).

When applying the DAPC analysis and the function find.cluster with the K-mean clustering method, the BIC values plot showed a knee when the number of groups was five, although the minimum BIC value was observed for K = 9 (Figure 2A). Thus, based on the BIC, five was a parsimonious number of clusters for the modeling of our data. DAPC was carried out considering the five groups detected by the K-means clustering method. Four linear discriminants were maintained for the analysis (Figure 2A). The first two linear discriminants were sufficient to identify five groups that essentially corresponded to those previously identified by Structure (Figure 2A). At opposite extremes of the first linear discriminant, there were...
CAT-VPR and SPI; the second linear discriminant mainly separated VTO-GEA from both ROM and MAC. The third and the fourth linear discriminants refined the relationships among the groups (Supplementary Figure 2). For further details on the classification of the individual accessions provided by DAPC see Supplementary Figure 3.

The correspondence between the Structure and DAPC classifications is reported in Figure 2B. Only four varieties were classified into different groups by the two methods (Figure 2B). These were Testa di Ferro, Gross Camus, Spinoso di Gonnos, and Spinoso di Gela that, based on Structure, were all admixed, with the coefficient of membership (q) to the prevalent genetic groups between 0.489 and 0.644. In two of the four cases, the “mismatches” between the two clustering methods involved varieties that were moved between closely related groups. Indeed, Spinoso di Gonnos was assigned to ROM by Structure, while it was assigned to MAC by DAPC; on the contrary, Testa di Ferro was assigned to MAC by Structure, while it was assigned to ROM by DAPC.

Therefore, both the model-based (Structure) and the unsupervised (DAPC) methods agreed when identifying five groups.

Genomic SSR Polymorphism Within Groups

There was a significant difference between the groups for the number of alleles per locus (nA) (Wilcoxon’s test: χ² = 25.381; d.f. = 4; p = 0.00004; Table 2) with VTO-GEA that showed the highest value. The groups also showed significantly different levels of gene diversity (Hg) (Wilcoxon’s non-parametric test: χ² = 18.69, d.f. = 4; p = 0.0009). The VTO-GEA group showed the highest Hg and was significantly more diverse than SPI and CAT-VPR, which showed the lowest Hg values; ROM and MAC were in intermediate positions (Table 2).

AMOVA analysis performed considering the hierarchical level “within-individual” showed a global negative and significant coefficient of inbreeding, Female = −0.317, p < 10⁻⁵; Table 2; C.I.99% from −0.408 to −0.153) that indicated an excess of heterozygotes within the five genetic groups. In all cases except MAC, the group-specific FIS were negative and statistically significant (p < 10⁻⁵). The FIS value varied significantly between groups (Wilcoxon’s non-parametric test: χ² = 34.61; d.f. = 4; p = 5.58 × 10⁻⁷). CAT-VPR, SPI, and ROM showed stronger excess of heterozygotes than MAC and VTO-GEA (Table 2).

The correspondence between the Structure and DAPC classifications is reported in Figure 2B. Only four varieties were classified into different groups by the two methods (Figure 2B). These were Testa di Ferro, Gross Camus, Spinoso di Gonnos, and Spinoso di Gela that, based on Structure, were all admixed, with the coefficient of membership (q) to the prevalent genetic groups between 0.489 and 0.644. In two of the four cases, the “mismatches” between the two clustering methods involved varieties that were moved between closely related groups. Indeed, Spinoso di Gonnos was assigned to ROM by Structure, while it was assigned to MAC by DAPC; on the contrary, Testa di Ferro was assigned to MAC by Structure, while it was assigned to ROM by DAPC.

Therefore, both the model-based (Structure) and the unsupervised (DAPC) methods agreed when identifying five groups.

Genomic SSR Polymorphism Within Groups

There was a significant difference between the groups for the number of alleles per locus (nA) (Wilcoxon’s test: χ² = 25.381; d.f. = 4; p = 0.00004; Table 2) with VTO-GEA that showed the highest value. The groups also showed significantly different levels of gene diversity (Hg) (Wilcoxon’s non-parametric test: χ² = 18.69, d.f. = 4; p = 0.0009). The VTO-GEA group showed the highest Hg and was significantly more diverse than SPI and CAT-VPR, which showed the lowest Hg values; ROM and MAC were in intermediate positions (Table 2).

AMOVA analysis performed considering the hierarchical level “within-individual” showed a global negative and significant coefficient of inbreeding, Female = −0.317, p < 10⁻⁵; Table 2; C.I.99% from −0.408 to −0.153) that indicated an excess of heterozygotes within the five genetic groups. In all cases except MAC, the group-specific FIS were negative and statistically significant (p < 10⁻⁵). The FIS value varied significantly between groups (Wilcoxon’s non-parametric test: χ² = 34.61; d.f. = 4; p = 5.58 × 10⁻⁷). CAT-VPR, SPI, and ROM showed stronger excess of heterozygotes than MAC and VTO-GEA (Table 2).

Chloroplast SSR Analysis

Universal cpSSR primers revealed low levels of cytoplasmic variability. Indeed, of the 27 assayed loci, only one (cpSSR-7) was polymorphic, with two alleles hereafter named “I” and “II.” The “I” allele was the most frequent (0.67) (Supplementary Figure 4). However, the chloroplast and nuclear patterns of variation were correlated (Supplementary Figure 4; χ² test: χ² < 0.01). The two most divergent nuclear SSR groups (CAT-VPR, SPI) also show alternative cpSSR alleles: CAT-VPR was monomorphic for allele “I,” while SPI was monomorphic for allele “II” (Figure 2).

Patterns of Phenotypic Variation Within and Between the SSR-Derived Groups

Table 3 reports the results of the NANOV A for the 21 quantitative traits measured at the harvest of the first head. Within the SSR groups, the varieties showed different means for all of the quantitative traits (p < 10⁻⁵). The broad-sense heritability (h²B) within groups varied widely depending on the traits considered, from 0.48 (stem diameter) to 0.98 (harvest time) (Table 3), with an average of 0.65. The five
FIGURE 2 | Discriminant analysis of the principal components (DAPC). (A) Five groups of cultivated globe artichoke varieties as depicted by the first two discriminant functions. Insets, left: eigenvalues for the first four discriminant functions. The small square embedded in the lower left part of the figure represents the BIC values as a function of the numbers of assumed genetic groups. Roman numbers (I and II) indicate the chloroplast alleles observed within each nuclear SSR group. (B) Correspondence between the Structure and DAPC classifications.
SSR-derived groups showed different mean $h^2_B$ across traits. Indeed, these ranged from 0.56 (CAT-VRP) to 0.77 (VTO-GEA), with ROM (0.60), MAC (0.63), and SPI (0.70) in intermediate positions. These differences between groups were statistically significant based on ANOVA ($p < 0.01$), while the Tukey-Kramer HD test separated CAT-VRP from the other four groups ($p < 0.05$).

Based on ANOVA and Tukey-Kramer HD multiple comparison test, 10 out of 21 quantitative traits differed between the five genetic groups (with $p$-values from $< 0.01$ to $< 10^{-5}$; Table 3). The strongest group divergence was seen for plant height ($Q_{ST,B} = 0.279$), polar/equatorial head-length ratio ($Q_{ST,B} = 0.223$), harvest time ($Q_{ST,B} = 0.200$), and length of stem ($Q_{ST,B} = 0.183$) (Table 3). Among the traits that
defined the size and shape of the bracts and the receptacle, the width of the bract and bract height/length ratio was the most diverse among the groups (Q\textsubscript{ST_B} = 0.163, 0.151, respectively; p < 10\textsuperscript{-4} in both cases; Table 3). The genetic groups showed low divergence for weight of first head (Q\textsubscript{ST_B} = 0.00), diameter of stem (Q\textsubscript{ST_B} = 0.011), number of leaves on stem (Q\textsubscript{ST_B} = 0.010), and receptacle height (Q\textsubscript{ST_B} = 0.014).

When considering the stages of the harvest of the secondary and tertiary heads, these globe artichoke varieties showed statistically different means (p < 10\textsuperscript{-5}) for all the traits examined (Table 4). Again, the SSR-derived groups were strongly divergent for head shape (polar/equatorial head-length ratio, Q\textsubscript{ST_B} = 0.26 and 0.28, at the two considered stages, respectively) and length of the stem (Q\textsubscript{ST_B} = 0.28, 0.22, respectively), and to a lesser extent, for harvest time (Q\textsubscript{ST_B} = 0.17, 0.16, respectively). Furthermore, this analysis showed differences in yield performances between ST\textsubscript{B} for harvest time (Q\textsubscript{ST_B} = 0.17, 0.16, 0.12, respectively) and length of the stem (Q\textsubscript{ST_B} = 0.15, 0.11, respectively). The genetic groups showed low divergence for weight of first head (Q\textsubscript{ST_B} = 0.010), number of leaves on stem (Q\textsubscript{ST_B} = 0.011), and receptacle height (Q\textsubscript{ST_B} = 0.014).

The correlation between the degree of quantitative-genetic subdivision among populations (Q\textsubscript{ST}) and the average heritability within populations (h\textsuperscript{2}_B) did not reach statistical significance (p = 0.118, 0.573, respectively). Based in Q\textsubscript{ST(0)_B}, the traits “plant height” and “head shape” remain statistically significant (Figure 3B) exceeding the C.I.95% for the neutral F\textsubscript{ST}.

The highest level of population structure that separated CAT\_VPR from the other groups (K = 2; Figure 1) was associated with an abrupt variation in the harvest time (Figure 4A). Indeed, the CAT\_VPR group was 35–42 days earlier, on average, than the other groups. The SPI group, which split at K = 3 (Figure 1), tended to be the earliest among the late groups, for both the secondary and tertiary heads (Figure 4A). Furthermore, the differences between the SSR-derived groups for yield performance paralleled those for harvest time (Figure 4B).

### Relationship Between Molecular and Quantitative Trait Variation

At the within-group level, the correlation between H\textsubscript{E} and the average h\textsuperscript{2}_B was positive but not significant (Pearson r = 0.72; n = 5; p > 0.05).

There was a strong and positive correlation between Q\textsubscript{ST_B} and the average within-groups h\textsuperscript{2}_B (REML method: r = 0.736, n = 21, p = 0.0001; 95% C.I.: 0.446–0.887; Figure 3A). Furthermore, of the 21 quantitative traits measured at the harvest of the first head, four had Q\textsubscript{ST_B} values that exceeded the C.I. for the molecular, putatively neutral F\textsubscript{ST} (Figure 3A and Table 3). These were plant height, head shape (as polar/equatorial head-length ratio), harvest time, and stem length, which exceeded the upper C.I.99.9%. Furthermore, the bract width was also significant as it was outside the C.I.95%. Among these traits, head shape, stem length and, to a lesser extent, harvest time, confirmed their outlying behaviors also at the stages of the secondary and tertiary heads (Table 4).

For the “total weight of heads” Q\textsubscript{ST} > F\textsubscript{ST} for the secondary heads, but not for the tertiary heads. The correlation between CAT\_VPR and the other groups (K = 2; Figure 1) was associated with an abrupt variation in the harvest time (Figure 4A). Indeed, the CAT\_VPR group was 35–42 days earlier, on average, than the other groups. The SPI group, which split at K = 3 (Figure 1), tended to be the earliest among the late groups, for both the secondary and tertiary heads (Figure 4A). Furthermore, the differences between the SSR-derived groups for yield performance paralleled those for harvest time (Figure 4B). Indeed, the correlation between these two variables was positive and significant (REML: r = 0.940; C.I.95% = 0.340–0.997; p = 0.0174), i.e., the yield increased with the lateness. Thus, the CAT\_VPR yield was 41.7% lower than that of VTO-GEA, n = 21, p = 0.0001; 95% C.I.: 0.446–0.887; Figure 3A). Furthermore, of the 21 quantitative traits measured at the harvest of the first head, four had Q\textsubscript{ST_B} values that exceeded the C.I. for the molecular, putatively neutral F\textsubscript{ST} (Figure 3A and Table 3). These were plant height, head shape (as polar/equatorial head-length ratio), harvest time, and stem length, which exceeded the upper C.I.99.9%. Furthermore, the bract width was also significant as it was outside the C.I.95%. Among these traits, head shape, stem length and, to a lesser extent, harvest time, confirmed their outlying behaviors also at the stages of the secondary and tertiary heads (Table 4).
FIGURE 3 | Relationships between the level of quantitative-genetic subdivision among the SSR groups ($Q_{ST,B}$) and the average heritability within the SSR groups ($h^2_B$) for the 21 phenotypic traits recorded at the harvest of the first head, as listed in Table 3. $Q_{ST,B}$ was calculated assuming $F_{IS} = 0$ (A) or for the neutral $F_{IS} = –0.315$ estimated from the SSR data (B). Orange continuous line, regression line; green continuous line: average level of genetic subdivision among groups ($F_{ST}$) based on putatively neutral SSRs; green, yellow, and red dashed lines: lower and upper limits (95%, 99%, 99.9%, respectively) of the neutral $F_{ST}$ based on bootstrap.

TABLE 5 | Average numbers and weights of the secondary and tertiary heads for the five SSR-derived groups of cultivated globe artichokes.

| Group      | Number (n) | Weight (g) |
|------------|------------|------------|
|            | Secondary  | Tertiary   | Secondary  | Tertiary   |
| CAT-VPR    | 2.9 a      | 2.6 a      | 157 a      | 113 a      |
| ROM        | 3.7 ab     | 3.7 b      | 181 ab     | 125 b      |
| MAC        | 3.4 ab     | 3.6 b      | 209 b      | 123 b      |
| SPI        | 3.2 a      | 3.5 b      | 196 b      | 123 b      |
| VTO-GEA    | 4.1 b      | 4.8 c      | 203 b      | 129 b      |
| Range (%)  | 41.3       | 84.6       | 29.3       | 14.1       |

Data are means, where those with different letters are significantly different ($p < 0.05$; Tukey’s honestly significant difference multiple comparisons test).

TABLE 6 | Differences between SSR-derived groups for plant height, stem length, and head and bract shape.

| SSR-derived group | Plant height (cm) | Stem length (cm) | Head shape (p/E) | Bract shape Width (mm) | Height/Width |
|-------------------|-------------------|------------------|------------------|------------------------|-------------|
| CAT-VPR           | 67.5 b            | 28.6 ab          | 1.17 ab          | 42.7 a                 | 0.43 b      |
| SPI               | 82.9 ab           | 29.8 ab          | 1.2 ab           | 41.6 ab                | 0.46 ab     |
| VTO-GEA           | 100.5 a           | 29.7 ab          | 1.19 ab          | 38.3 b                 | 0.48 a      |
| MAC               | 103.5 a           | 36.1 a           | 1.06 bc          | 38.0 ab                | 0.49 a      |
| ROM               | 79.1 b            | 22.6 b           | 1.01 c           | 35.3 b                 | 0.49 a      |

p/E, polar/equatorial length ratio. Data are means, where those with different letters are significantly different ($p < 0.05$; Tukey’s honestly significant difference multiple comparisons test).

The most productive group (Figure 4B). The ROM, MAC, and SPI groups tended to have intermediate positions. In all of the groups, the secondary heads accounted for ~50% of the total yield (from 46% to 52%). The variation between groups for the yield components was also notable (Table 5). This was particularly evident for the number of secondary and tertiary heads (Table 5), where the top-ranking group was VTO-GEA (4.1 and 4.8, respectively), while the bottom-ranking group was CAT-VPR (2.9 and 2.6, respectively). The CAT-VPR group also occupied the last position for the weights of the heads (157 g and 113 g, respectively), while the MAC group tended to have larger heads, particularly for the primary and secondary heads (253 and 209 g; Figure 4B and Table 5, respectively).

Bract shape (as width and height/width ratio) varied according to clustering history (Table 6). Indeed, CAT-VPR showed the most oblong bracts, while VTO-GEA, MAC, and ROM showed the narrowest and most rounded bracts. The SPI group occupied a transitional position. For head shape (as polar/equatorial length ratio), the SPI showed the most oblong head (Table 6). Two close groups based on SSRs, MAC, and ROM, both showed the most spherical heads while VTO and CAT showed intermediate characteristics. Plant height and stem length did not follow the history of clustering based on Structure (Table 6). The average plant height was positively correlated with the average inbreeding coefficient, $F_{IS}$ (Figure 5). As the group-specific $F_{IS}$s were all negative, this meant that the lower the heterozygote excess, the taller the plants. Furthermore, when the plant stature increased, $H_E$ and $h^2_B$ also increased, while LD tended to decrease (Figure 5).

Comparison Between Molecular and Qualitative Trait Variations

Among the 23 UPOV descriptors applied, 20 were polymorphic throughout the entire collection (Table 7). Within each group,
**FIGURE 4** Differences between the SSR groups for the average harvest time (A) and yields of primary, secondary, and tertiary heads (B). Group means are compared referring to the same head type (i.e., primary, secondary, and tertiary) and means with different letters are significantly different ($p < 0.05$; Tukey-Kramer HD test). In panel (A) green boxes = interquartile range; red bar = mean; white bar = median.
TABLE 7 | Diversity statistics (Nei's diversity index, \(I_{Nei}\), and number of phenotypic classes, \(n_c\)) for the 23 UPOV descriptors applied within and between the five SSR groups of the cultivated globe artichoke.

| Trait                      | Particular                                      | CAT-VPR | ROM | MAC | SPI | VTO-GEA | \(n_c\) | \(I_{Nei}\) | Average | s.d. | Total | Average | s.d. | Total |
|----------------------------|-------------------------------------------------|---------|-----|-----|-----|---------|---------|-----------|---------|-----|-------|---------|-----|-------|
| Lobe                       | Shape of tip (excluding terminal lobe)          | 1       | 0.00| 1   | 0.00| 1       | 0.00    | 1         | 0.00    | 1   | 0.00  | 1       | 0.00| 1     |
|                           | Shape of tip of secondary lobes                 | 1       | 0.00| 1   | 0.00| 1       | 0.00    | 1         | 0.00    | 1   | 0.00  | 1       | 0.00| 1     |
| Leaf                       | Attitude                                        | 2       | 0.11| 2   | 0.33| 2       | 0.39    | 1         | 0.00    | 1   | 0.00  | 2       | 0.57| 3     |
|                           | Spininess                                       | 1       | 0.00| 1   | 0.00| 1       | 0.00    | 2         | 0.06    | 1   | 0.00  | 1.8     | 0.84| 3     |
|                           | Incisions                                       | 3       | 0.53| 1   | 0.00| 1       | 0.00    | 1         | 0.00    | 2   | 0.06  | 1.6     | 0.89| 3     |
|                           | Hairiness on upper side                         | 1       | 0.00| 1   | 0.00| 1       | 0.00    | 1         | 0.00    | 1   | 0.00  | 1.0     | 0.00| 1     |
| Leaf blade                 | Intensity of green color (upper side)           | 2       | 0.51| 3   | 0.69| 2       | 0.39    | 2         | 0.33    | 3   | 0.68  | 2.4     | 0.55| 3     |
|                           | Hue of green color                              | 3       | 0.41| 1   | 0.00| 2       | 0.39    | 2         | 0.33    | 2   | 0.50  | 2.0     | 0.71| 3     |
|                           | Intensity of gray hue                           | 2       | 0.51| 2   | 0.51| 3       | 0.67    | 2         | 0.60    | 3   | 0.57  | 2.4     | 0.55| 3     |
|                           | Blistering                                      | 3       | 0.38| 3   | 0.65| 3       | 0.60    | 3         | 0.73    | 4   | 0.51  | 3.2     | 0.45| 4     |
| Petiole                    | Anthocyanin coloration at base                  | 4       | 0.69| 4   | 0.75| 2       | 0.56    | 2         | 0.60    | 4   | 0.75  | 3.2     | 1.10| 4     |
| Central flower head        | Shape in longitudinal section                   | 3       | 0.54| 4   | 0.75| 3       | 0.73    | 3         | 0.73    | 5   | 0.74  | 3.8     | 0.84| 5     |
|                           | Shape of tip                                    | 2       | 0.36| 3   | 0.64| 3       | 0.73    | 3         | 0.73    | 3   | 0.39  | 2.8     | 0.45| 4     |
|                           | Time of appearance                              | 5       | 0.72| 3   | 0.64| 4       | 0.73    | 3         | 0.73    | 4   | 0.60  | 3.8     | 0.84| 5     |
|                           | Anthocyanin coloration of inner bracts          | 3       | 0.42| 2   | 0.51| 2       | 0.36    | 2         | 0.60    | 3   | 0.55  | 2.4     | 0.55| 3     |
|                           | Density of inner bracts                         | 3       | 0.46| 2   | 0.51| 3       | 0.69    | 3         | 0.80    | 3   | 0.64  | 2.8     | 0.45| 3     |
| Outer bract                | Main shape                                      | 1       | 0.00| 2   | 0.44| 2       | 0.20    | 1         | 0.00    | 3   | 0.17  | 1.8     | 0.84| 3     |
|                           | Apex shape                                      | 2       | 0.16| 1   | 0.00| 3       | 0.38    | 2         | 0.33    | 3   | 0.59  | 2.2     | 0.84| 3     |
|                           | Color (external side)                           | 3       | 0.26| 3   | 0.56| 3       | 0.64    | 2         | 0.33    | 4   | 0.72  | 3.0     | 0.71| 4     |
|                           | Reflexing (of tip)                              | 3       | 0.11| 2   | 0.44| 2       | 0.20    | 1         | 0.00    | 3   | 0.49  | 2.2     | 0.84| 3     |
|                           | Spine length                                    | 2       | 0.11| 1   | 0.00| 2       | 0.20    | 3         | 0.60    | 3   | 0.54  | 2.2     | 0.84| 3     |
|                           | Mucron                                          | 1       | 0.00| 2   | 0.51| 2       | 0.36    | 1         | 0.00    | 2   | 0.45  | 1.6     | 0.55| 2     |
| Receptacle                 | Shape in longitudinal section                   | 2       | 0.21| 3   | 0.56| 2       | 0.56    | 3         | 0.33    | 2   | 0.30  | 2.2     | 0.45| 3     |
| Mean                       |                                                | 2.30    | 0.28| 2.09| 0.37| 2.26    | 0.39    | 1.96     | 0.37    | 2.78 | 0.43  | 2.28    | 0.31| 4     |
| s.d.                       |                                                | 1.06    | 0.24| 1.00| 0.29| 0.86    | 0.26    | 0.83     | 0.31    | 1.04 | 0.26  | 0.96    | 0.10| 8     |

The number within each bubble indicates the average \(h^2_B\) across the quantitative traits.

![Figure 5](image-url) Relationships between coefficient of inbreeding (\(F_{IS}\)), plant height (cm), gene diversity (\(H_e\)), linkage disequilibrium (LD), and average heritability (\(h^2_B\)) for the five SSR groups of globe artichokes. Bubble size is proportional to the level of LD within the groups (as% of pair of loci with \(p < 0.05\); Table 2). Bubble shading (darkness) is proportional to the level of \(H_e\) (Table 2). The number within each bubble indicates the average \(h^2_B\) across the quantitative traits.

The average divergence for qualitative traits (\(F_{ST_{-qlt}}\)) among the five SSR-derived groups was 0.20. This is close to the estimated value for the SSR markers (\(F_{ST} = 0.17\)) and quantitative traits (\(Q_{ST_B} = 0.18\)). However, \(F_{ST_{-qlt}}\) varied widely among the traits (Figure 6). In five cases, \(F_{ST_{-qlt}}\) was large enough to fit outside the “neutral envelope” obtained by simulation using SSRs.

Interestingly, there was strong correlation between the qualitative traits and SSR distances between the pairs of groups (REML: \(r = 0.918; n = 10\); \(L_{95\%} = 0.685; U_{95\%} = 0.981; p = 1.76 \times 10^{-4}\)) (Figure 7A). This was stronger than the positive and marginally significant correlation observed between quantitative-genetic divergence (\(Q_{ST_B}\)) and SSR genetic divergence (\(F_{ST}\)) between the pairs of SSR-derived groups (REML: \(r = 0.670; n = 10\); \(L_{95\%} = 0.070; U_{95\%} = 0.914; p = 0.034\); Spearman-rank \(\rho = 0.636; p = 0.048\)) (Figure 7).
and Macau than to Spinosi and Violette di Toscana. Differently, in our study, Romaneschi and Macau are closer to Violette di Toscana and to several varieties classified as “Green et al.”. As the collections analysed in Lanteri et al. (2004) and in our study are similar, the observed differences could be inherent to the different markers used (AFLP vs. SSR). Indeed, AFLPs are biallelic and dominant markers, while SSRs are multiallelic and codominant markers. Moreover, AFLPs and SSR markers have different mutational models and rates. The mutation rates of SSRs are between $10^{-3}$ and $10^{-4}$, as inferred from genetic data (Estoup and Angers, 1998; Mariette et al., 2001; Garcia et al., 2007) and estimated from observed mutations (Vigouroux et al., 2002; Thuillet et al., 2005). The mutation rate for AFLPs is lower being between $10^{-6}$ and $10^{-5}$ (Mariette et al., 2001; Gaudeul et al., 2004; Kropf et al., 2009). On this basis, different marker types can lead to different precision when assessing relationships between single individuals, as well as to different estimations of within- and between-population components of genetic diversity (see, e.g., Gaudeul et al., 2004). Moreover, high evolutionary rates necessarily result in low phylogenetic signals, and conversely, low evolutionary rates result in high phylogenetic signals (Revell et al., 2008); thus, with their slower mutation rate, AFLP markers can highlight more ancient population splits (Flanders et al., 2009). Therefore, the population genetic structure described by Lanteri et al. (2004) might represent a more ancient one than that depicted in this study.

The Signature of Long-Term Clonal Propagation

The chloroplast and nuclear genetic polymorphisms as well as the molecular and phenotypic population structures, were associated. This is expected if there are multiple germplasm sources contemporarily differentiated for nuclear, organellar loci, and phenotypic traits. Furthermore, LD was detected within each of the five groups, and the diversity statistics based on SSRs and quantitative and qualitative traits are correlated. These results indicate that the five groups of globe artichoke varieties were genetically relatively isolated, and that the recombination accumulated through crop evolution was overall not sufficiently high (or the time elapsed was too low) to break up the cytonuclear and marker-trait LD. The presence of LD in the whole dataset is also indicated by the correlation between the divergence of the group for SSR markers ($F_{ST}$) and the divergence for the quantitative traits ($Q_{ST,B}$) and the qualitative traits ($F_{ST,qt}$).

Within the groups, in addition to LD, we observed negative $F_{IS}$, which indicates an excess of heterozygotes compared to the Hardy-Weinberg expectation. Detecting LD is not sufficient to call for clonal reproduction (de Meûs and Balloux, 2004; Halkett et al., 2005). However, negative $F_{IS}$ is the ultimate signature of clonal diploid populations (Halkett et al., 2005). These results are consistent with the expected structure of genetic diversity of a collection of artichoke cultivars where clonal propagation is a common practice; in this species, propagation is carried out with “carducci” (basal shoots) or “ovoli” (semi-dormant shoots with a limited root system). Furthermore, our study also shows that some groups have a higher excess of heterozygotes than others.

**DISCUSSION**

The Globe Artichoke Cultivated Gene Pool Clusters Into Five Subgroups

This study shows that the analyzed globe artichoke accessions clustered into two main genetic groups. The first comprises the Catanesi and Violette di Provenza types and is quite uniform. The second is more complex and includes four additional subgroups related to Spinosi, Green et al., Violette di Toscana, Romaneschi, and Macau.

Based on the phenotypic traits and multivariate analysis, Elia and Miccolis (1996) distinguished five main morphophenological clusters: “early,” “medium-early,” “late with small head,” “late-violet,” and “late with large head.” This genetic classification based on SSRs shows good agreement with that of Elia and Miccolis (1996) based on the phenotypic data. Indeed, their “early” and “medium-early” clusters correspond to our CAT-VPR group, which is composed of Catanesi and Violette di Provenza, and is characterized by long-shaped and small main and secondary heads, low-medium yield, early harvest, and plant with low vigour (short, with low numbers of heads). Their “late with small head” cluster corresponds to our SPI group, which includes Spinosi types characterized by long-shaped and medium-small heads and late harvest. The “late-violet” cluster corresponds to VTO-GEA that includes “Violetti di Toscana” and several other varieties assigned to the category “Green et al.”. On average, these plants were vigorous (tall and high-yielding) and showed less elongated heads than the previous groups. Finally, the “late with large head” cluster correlates with our ROM and MAC groups, which are characterized by lateness and large sizes of the principal and secondary heads, with a shape that tends to be spherical.

Based on AFLP markers, Lanteri et al. (2004) reported that Catanesi and Violette di Provenza were closer to Romaneschi
FIGURE 7 | Relationships between the degree of phenotypic divergence for qualitative (A) and quantitative (B) traits for the phenotypic traits listed in Supplementary Table 3 and the pairwise genetic distances between groups for the SSR markers.

and this suggests that different breeding histories and various degrees of historical “clonality” characterize the evolution of the five identified genetic groups.

Albeit selfing is not precluded, cross-fertilization in globe artichoke is promoted by protandry (Mauromicale and Ierna, 2000). In a recent re-sequencing study Acquadro et al. (2017) documented the significant heterozigosity of some cultivated clonal varieties. Consistent with this finding, Basnitzki and Zohary (1987) showed that the (self) seed progenies of clonal varieties showed wide segregation of morphological and production traits. Thus, it can be hypothesized that the excess of heterozygotes at SSR loci might reflect the selection of plants with superior hybrid vigor because of the heterozygote advantage (overdominance). However, genomic surveys have suggested that loci with heterozygote advantage must be considered only a small minority of all loci in a species (Hedrick, 2012). Furthermore, complete clonality is comparable to a long-term population bottleneck with a population size of one, which implies that heterozygote excess is also expected if the selection is not acting (Balloux et al., 2003). We also observed that the groups with a strong excess of heterozygotes (and strong LD, low $H_E$, low $h^2_B$) were shorter and produced fewer and smaller heads (i.e., were less vigorous) than groups with low heterozygotes (and lower LD, higher $H_E$, higher $h^2_B$). This suggests that selection against deleterious mutations might be less effective in groups with high clonality (and high heterozygote excess) than in those with low clonality (with low or null heterozygote excess). Indeed, when there is no recombination, selection against deleterious mutations is less effective, and deleterious mutations have the potential to accumulate (i.e., the mutation load) (McKey et al., 2010; Gaut et al., 2015). The accumulation of mutations in a purely clonal line ultimately leads to lower fitness, which is associated with lower agronomic performance (McKey et al., 2010). However, in globe artichoke, this remains to be demonstrated with a dedicated study.

**Comparison Between Molecular and Phenotypic Population Structures Indicates a Role for Divergent Selection**

Under neutral evolution (i.e., absence of selection), $Q_{ST} = F_{ST}$ is expected, such that the proportion of total variance allocated among populations (or genetic groups, in this case) should be equal or similar for both quantitative traits and molecular markers (Leinonen et al., 2013). Therefore, assuming that $F_{ST}$ has been estimated considering putatively “neutral” SSR loci, the inequality $Q_{ST} > F_{ST}$ indicates divergent selection among the populations. Indeed, this suggests that for a given quantitative trait, populations (or genetic groups) show different mean values not only because of random demographic processes (e.g., genetic drift, migration, mutation), but also due to the selection that exerts differential pressure among the populations. On the contrary, the inequality $Q_{ST} < F_{ST}$ suggests uniform selection across populations: populations tend to have the same trait means because the selection has the same direction in all of the populations.

Following this rationale, we compared the $F_{ST}$ and $Q_{ST,B}$ obtained among the genetic groups of the globe artichokes identified by the analysis of the population structure. This comparison suggests that some quantitative traits have an “excess” of divergence between groups ($Q_{ST,B}$) compared to the putatively neutral markers ($F_{ST}$) ($Q_{ST,B} > F_{ST}$). For these traits, divergent selection, rather than demographic factors, might be responsible for the phenotypic differences observed between the genetic groups. These traits are harvest time, plant height, head shape, and stem length. Among the qualitative traits, divergent selection can be suggested for leaf spininess,
spine length, apex shape, and reflexing of the tip of the outer bracts as they showed an $F_{ST}$-qtl that exceeded the “neutral” $F_{ST}$. In general, it is possible that these traits were under the farmer’s attention when selecting plants for clonal propagation, as they strongly condition both agronomic practice (e.g., harvest time) and consumer preference (e.g., stem length, head shape, spininess). However, caution should be exercised when interpreting these data. Indeed, deviations from the Hardy-Weinberg equilibrium might affect the $Q_{ST}$-$F_{ST}$ comparison. Inbreeding ($F_{IS} > 0$) can lead to $Q_{ST} < F_{ST}$ under neutrality (instead of $Q_{ST} = F_{ST}$), which results in low power when detecting divergent selections. On the contrary, when, as in the present case, there is an excess of heterozygotes ($F_{IS} < 0$), it is possible that under neutrality $Q_{ST} > F_{ST}$ (instead of $Q_{ST} = F_{ST}$), which leads to a higher risk of false positives (Cubry et al., 2017). As expected, when the $Q_{ST}$ is calculated considering our negative neutral $F_{IS}$ estimate, the $Q_{ST}$ values are reduced. However, the traits “plant height” and “head shape” remain statistically significant.

For the trait “plant height,” $Q_{ST} > F_{ST}$ is often observed in both trees (Yang et al., 1996) and herbaceous plants (Waldmann and Andersson, 1998; Kawakami et al., 2011; Eriksen et al., 2012; Michalski et al., 2017). Interestingly, very high $Q_{ST}$s were observed for plant height, flowering date, and capitulum size (diameter) in the sunflower species Helianthus maximiliani, which like C. cardunculus var. scolymus belongs to the Asteraceae family, is perennial and is diploid (Kawakami et al., 2011). Variations in these traits are relevant to local adaptation to spatially heterogeneous environments. Furthermore, as might also be suggested for our study, Kawakami et al. (2011) noted that plant height and flowering time might have both been found under divergent selection because of the typical life-history trade-off between growth and reproductive timing (Obeso, 2002). The existence of this trade-off in cultivated artichoke was suggested by Crinò et al. (2008). Alternatively, plant height and flowering time might at least partially share the same genetic control, as has been suggested in other plant species (Tondelli et al., 2014; Bustos-Korts et al., 2019).

The $Q_{ST}$-$F_{ST}$ comparison assumes that the genetic control of the quantitative traits is additive (Spitze, 1993). Non-additive (dominance or epistasis) and maternal effects might have different weights for different traits and might confound any inference (Leinonen et al., 2013). It is possible that traits with simple genetic control, a strong dominance variance component, and under certain demographic models show $Q_{ST}$ inflated over the neutral expectation, mimicking divergent selection (López-Fanjul et al., 2003, 2007). However, such inflation is unlikely for traits with polygenic control (Goudet and Martin, 2007), and, in general, with non-additive inheritance, the $Q_{ST}$-$F_{ST}$ approach to look for diversifying selection would be conservative (Whitlock, 1999, 2008; Cubry et al., 2017). Furthermore, the $Q_{ST}$-$F_{ST}$ approach can be suggested as an exploratory tool (Leinonen et al., 2008; Whitlock, 2008), particularly when detailed information on phenotypic traits is not available. In this case, as Whitlock (2008) noted, “the difficult statistical properties of $Q_{ST}$ become less important, and if the $Q_{ST}$ of several traits is measured, then the traits with the highest $Q_{ST}$ values may be good candidates for further study of selection.” In other words, non-additive and maternal effects probably did not significantly alter the conclusions of this study, and plant height and head shape followed by harvest time, stem length, and the number of heads are good candidates for having been subjected to divergent selection during globe artichoke evolution.

Another observation indicates a role for divergent selection among the genetic groups. Indeed, we found that although there is a substantial amount of variation among the traits for $Q_{STB}$, much of this variation appears to be associated with variation in $h^2_B$. Under neutrality, $Q_{ST}$ does not correlate with $h^2_B$ (Lande, 1992). In contrast, the response to directional selection on a quantitative trait (and hence the $Q_{STB}$ between populations), ceteris paribus, is proportional to $h^2_B$ (Falconer and Mackay, 1996). Therefore, the positive correlation between $Q_{ST}$ and $h^2_B$ observed in this study is qualitatively consistent with the hypothesis that selection contributed to the phenotypic differences observed between the groups. In previous studies on Pinus contorta (Yang et al., 1996), Medicago truncatula (Bonnin et al., 1996), and Clarkia dudleyana (Podolsky and Holtsford, 1995), the correlation between $Q_{ST}$ and $h^2_B$ was also positive; a similar situation has also been observed in animals, such as in the microcrustacean Daphnia pulex (Lynch et al., 1999).

**CONCLUSION**

This study shows that a globe artichoke collection that comprises most of the varieties cultivated worldwide has a clonal population structure and is moderately structured. The data consistently show that five genetic groups with different degrees of “historical” clonality characterize the domesticated gene pool of the globe artichoke. Clonal propagation was probably accompanied by clonal selection for a set of different phenotypic characteristics. The traits that affect phenology (harvest time), plant architecture (plant height, stem length), head shape, and spininess (of both leaves and bracts) are suggested as the main targets of selection. The overall covariance between the molecular and phenotypic population structures indicates that genome-wide association mapping strategies must deal with the identified population structure to minimize the risk of false positives.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in this study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

**AUTHOR CONTRIBUTIONS**

DR, GA, MR, and EP conceived and designed the experiments. LB, AP, and DS performed the field trial. LB, MR, and GA carried out field data curation. MR carried out DNA extraction. AA and CC performed SSR analysis. DR performed statistical analyzes.
Structure analysis. Evanno’s Supplementary Figure 1 | Supplementary Figure 1 | Structure analysis. Evanno’s ΔK statistic as a function of the number “true” genetic populations assumed.

REFERENCES

Acquadro, A., Barchi, L., Portis, E., Mangino, G., Valentino, D., Mauromicale, G., et al. (2017). Genome reconstruction in *Cynara cardunculus* taxon gains access to chromosome-scale DNA variation. *Sci. Rep.* 7:5617. doi: 10.1038/s41598-017-05065-5

Affitos, S., Schijlen, E., de Jong, H., de Ridder, D., Smit, S., Finkers, R., et al. (2014). Exploring genetic variation in the tomato (*Solanum section Lycopersicon*) clade by whole-genome sequencing. *Plant J.* 80, 136–148. doi: 10.1111/tjp.12161

Angioi, S. A., Rau, D., Attene, G., Nanni, L., Bellucci, E., Loizzo, G., et al. (2010). Beans in Europe: origin and structure of the European landraces of *Phaseolus vulgaris* L. *Theor. Appl. Genet.* 121, 829–843. doi: 10.1007/s00122-010-1353-2

Angioi, S. A., Rau, D., Rodriguez, M., Loizzo, G., Desiderio, F., Papa, R., et al. (2009). Nuclear and chloroplast microsatellite diversity in *Phaseolus vulgaris* L. from Sardinia (Italy). *Mol. Breed.* 23, 413–429. doi: 10.1007/s11032-008-9245-8

Balloux, F., Lehmann, L., and de Meeds, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics* 164, 1635–1644. doi: 10.1093/genetics/164.4.1635

Bandillo, N., Jarquin, D., Song, Q., Nelson, R., Cregan, P., Specht, J., et al. (2010). A Population structure and genome-wide association analysis on the USDA soybean germplasm collection. *Plant Genome* 3:110. doi: 10.3835/plantgenomev2015040024.00024.018

Basnitzki, Y., and Zohary, D. (1987). A seed-planted cultivar of globe artichoke. *HortScience* 22, 678–679.

Beaumont, M. A., and Nichols, R. A. (1996). Evaluating loci for use in the genetic differentiation of *Phaseolus* clonal and partially clonal diploids. *Proc. R. Soc. Lond. B Biol. Sci.* 263, 1619–1626. doi: 10.1098/rspb.1996.0237

Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., Biagetti, E., et al. (2013). Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. *New Phytol.* 197, 300–313. doi: 10.1111/nph.12477

Bonnin, I., Huguet, T., Gherardi, M., Prosperi, J.-M., and Olivieri, I. (1996). High level of polymorphism and spatial structure in a sowing plant species, *Medicago truncatula* (Leguminosae), shown using RAPD markers. *Am. J. Bot.* 83, 843–855. doi: 10.1002/(ISSN)1537-2197.1996.tb12776.x

Bustos-Korts, D., Dawson, I. K., Russell, J., Tondelli, A., Guerra, D., Ferrandi, C., et al. (2019). Exome sequences and multi-environment field trials elucidate the genetic basis of adaptation in barley. *Plant J.* 99, 1172–1191. doi: 10.1111/tjp.14414

Cericola, F., Portis, E., Lanteri, S., Toppino, L., Barchi, L., Acciarri, N., et al. (2014). Linkage disequilibrium and genome-wide association analysis for anthocyanin pigmentation and fruit color in eggplant. *BMC Genomics* 15:896. doi: 10.1186/1471-2164-15-896

Cericola, F., Portis, E., Toppino, L., Barchi, L., Acciarri, N., Ciarcia, T., et al. (2013). The population structure and diversity of eggplant from Asia and the Mediterranean basin. *PloS One* 8:e73702. doi: 10.1371/journal.pone.0073702

Chung, S.-M., and Staub, J. E. (2003). The development and evaluation of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa. *Theor. Appl. Genet.* 107, 757–767. doi: 10.1007/s00122-003-1131-3

Comino, C., Hahn, A., Moglia, A., Menin, B., Bourgaud, F., Lanteri, S., et al. (2009). The isolation and mapping of a novel hydroxycinnamoyltransferase in the globe artichoke chlorogenic acid pathway. *BMC Plant Biol.* 9:30. doi: 10.1186/1471-2229-9-30

Comino, C., Lanteri, S., Portis, E., Acquadro, A., Romani, A., Hahn, A., et al. (2007). Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus*. *BMC Plant Biol.* 7:14. doi: 10.1186/1471-2229-7-14

Comino, C., Portis, E., Giongloso, A., Acquadro, A., Pisanu, A. B., Baghino, L., et al. (2016). Toward a genome-wide association (GWA) mapping approach in globe artichoke. *Acta Hortic.* 1147, 51–56. doi: 10.17660/ActaHortic.2016.11477

Cornuet, J. M., and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001–2014. doi: 10.1093/genetics/144.4.2001

Cubry, P., Scotti, I., Oddou-Muratorio, S., and Lefèvre, F. (2017). Generalization of the QST framework in hierarchically structured populations: impacts of inbreeding and dominance. *Mol. Ecol. Res.* 17, e67–e83. doi: 10.1111/1755-0998.12693

De Meeds, T., and Balloux, F. (2004). Clonal reproduction and linkage disequilibrium in diploids: a simulation study. *Infect. Genet. Evol.* 4, 345–351. doi: 10.1016/j.meegid.2004.05.002

De Vos, N. E. (1992). Artichoke production in California. *HortTechnology* 2, 438–444.

Dellaccea, V., Magnifico, V., Marzi, V., Porceddu, E., and Scarascia Mugnozza, G. T. (1976). “Contributo alla conoscenza delle varietà di carciofo coltivate nel mondo,” in *Proc. II Int. Congr. Artichoke*, (Turin: Minerva Medica) 2, 199–316.

Desiderio, F., Bitocchi, E., Bellucci, E., Rau, D., Rodriguez, M., Attene, G., et al. (2013). Chloroplast microsatellite diversity in *Phaseolus vulgaris*. *Front. Plant Sci.* 3:312. doi: 10.3389/plants.2012.00312

Elia, A., and Miccolis, V. (1996). Relationships among 104 artichoke (*Cynara scolymus*) L accessions using cluster analysis. *Adv. Hortic. Sci.* 10, 158–162.

Eljounaidi, K., Cantar, K., Comino, C., Moglia, A., Hahn, A., Bourgaud, F., et al. (2014). Cytochrome P450s from *Cynara cardunculus* L. CYP71AV9 and CYP71B15, catalyze distinct hydroxylation in the sesquiterpene lactone biosynthetic pathway. *Plant Sci.* 223, 59–68. doi: 10.1016/j.plantsci.2014.03.007

Eljounaidi, K., Comino, C., Moglia, A., Cantar, K., Genere, A., Hahn, A., et al. (2015). Accumulation of cyanaropicrin in globe artichoke and localization of enzymes involved in its biosynthesis. *Plant Sci.* 239, 128–136. doi: 10.1016/j.plantsci.2015.07.020

Eriksen, R. L., Desrondvil, T., Hierro, J. L., and Kesseli, R. (2012). Morphological differentiation in a common garden experiment among native and non-native specimens of the invasive weed yellow starthistle (*Centaurea solitissilis*). *Biol. Invasions* 14, 1459–1467. doi: 10.1007/s10530-012-0172-6

Estoup, A., and Angers, B. (1998). “Microsatellites and minimisatellites for molecular ecology? theoretical and empirical consideration,” in *Advances in molecular ecology*, ed. G. Carvalho (Amsterdam, Netherlands: IOS Press), 55–86.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
Pavan, S., Curci, P. L., Zuluaga, D. L., Blanco, E., and Sonnante, G. (2018). Genotyping-by-sequencing highlights patterns of genetic structure and domestication in artichoke and cardoon. *PLoS One* 13:e0205988. doi: 10.1371/journal.pone.0205988

Pecaut, P. (1993). “Globe artichoke: Cynara scolymus L.” in *Genetic Improvement of Vegetable Crops*, eds G. Kallou and B. O. Bergh (Amsterdam: Pergamon), 737–746. doi: 10.1016/B978-0-08-040826-2.50054-0

Piry, S., Luikart, G., and Cornuet, J.-M. (1999). Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *J. Hered.* 90, 502–503. doi: 10.1093/jhered/90.4.502

Podolsky, R. H., and Holltford, T. P. (1995). Population structure of morphological traits in *Clarksia duldyana* I. Comparison of F2 by among allozymes and morphological traits. *Genetics* 140, 733–744. doi: 10.1093/genetics/140.2.733

Porceddu, E., Dallacecca, V., and Bianco, V. V. (1976). “Classificazione numerica di cultivar di carciofo,” in *Proceedings II International Congress of Artichoke*, (Turin: Minerva Medica), 1105–1119.

Portis, E., Cericola, F., Barchi, L., Toppino, L., Acciarri, N., Pulcini, L., et al. (2015). Association mapping for fruit, plant and leaf morphology traits in eggplant. *PLoS One* 10:e0135200. doi: 10.1371/journal.pone.0135200

Portis, E., Portis, F., Valente, M., Moglia, A., Barchi, L., Lanteri, S., et al. (2016). A genome-wide survey of the microsatellite content of the globe artichoke genome and the development of a web-based database. *PLoS One* 11:e0162841. doi: 10.1371/journal.pone.0162841

Portis, E., Scaglione, D., Acquadro, A., Mauromicale, G., Mauro, R., Knapp, S. J., et al. (2012). Genetic mapping and identification of QTIL for earliness in the globe artichoke/cultivated cardoon complex. *BM. Res. Notes* 5:252. doi: 10.1186/1756-0500-5-252

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093 genetics/155.2.945

Rau, D., Rodriguez, M., Rapposelli, E., Murgia, M. L., Papa, R., Brown, A. H. D., et al. (2016). Spatial genetic structure in wild cardoon, the ancestor of cultivated globe artichoke: limited gene flow, fragmentation and population history. *Plant Sci.* 253, 194–205. doi: 10.1016/j.plantsci.2016.09.013

Revell, L. J., Harmon, L. J., and Collar, D. C. (2008). Phylogenetic Signal, evolutionary process, and rate. *Syst. Biol.* 57, 591–601. doi: 10.1080/106351508023427

Rodriguez, M., Rau, D., Angioi, S. A., Bellucci, E., Bitocchi, E., Nanni, L., et al. (2013). *European Phaseolus coccineus* L. landraces: population structure and adaptation, as revealed by cpSSRs and phenotypic analyses. *PLoS One* e57337. doi: 10.1371/journal.pone.0057337

Rodriguez, M., Rau, D., O’Sullivan, D., Brown, A. H. D., Papa, R., and Attene, G. (2012). Genetic structure and linkage disequilibrium in landrace populations of barley in Sardinia. *Theor. Appl. Genet.* 125, 171–184. doi: 10.1007/s00122-012-2124-8

Rodriguez, M., Scintu, A., Posadinu, C. M., Xu, Y., Nguyen, C. V., Sun, H., et al. (2020). GWAS based on RNA-Seq SNPs and high-throughput phenotyping combined with climatic data highlights the reservoir of valuable genetic diversity in regional tomato landraces. *Genes* 11:1387. doi: 10.3390/genes11131387

Scaglione, D., Acquadro, A., Portis, E., Taylor, C. A., Lanteri, S., and Knapp, S. J. (2009). Ontology and diversity of transcript-associated microsatellites mined from a globe artichoke EST database. *BM. Genomics* 10:454. doi: 10.1186/1471-2164-10-454

Scaglione, D., Reyes-Chin-Wo, S., Acquadro, A., Froenicke, L., Portis, E., Beitel, C., et al. (2016). The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-able low-pass sequencing strategy of F1 progeny. *Sci. Rep.* 6:19427. doi: 10.1038/srep19427

Sonnante, G., Pignone, D., and Hammer, K. (2007). The domestication of artichoke and cardoon: from roman times to the genomic age. *Ann. Bot.* 100, 1095–1100. doi: 10.1093/aob/mcm127

Spitz, K. (1993). Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* 135, 367–374. doi: 10.1093/genetics/135.2.367

Tanto Hadado, T., Rau, D., Bitocchi, E., and Papa, R. (2009). Genetic diversity of barley (*Hordeum vulgare* L) landraces from the central highlands of Ethiopia: comparison between the Belg and Meher growing seasons using morphological traits. *Genet. Resour. Crop Evol.* 56, 1131–1148. doi: 10.1007/s10722-009-9437-2

Vigouroux, Y., Glubiz, J. C., Matsuoka, Y., Goodman, M. M., Sanchez, G. J., and Doebely, J. (2008). Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *Am. J. Bot.* 95, 1240–1253. doi: 10.3732/ajb.0800997

Vigouroux, Y., McMullen, M., Hittinger, C. T., Houchnis, K., Schulz, L., Kresovich, S., et al. (2002). Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *Proc. Natl. Acad. Sci. U. S. A.* 99, 9650–9655. doi: 10.1073/pnas.1122429

Waldmann, P., and Andersson, S. (1998). Comparison of quantitative genetic variation and allozyme diversity within and between populations of *Scabiosa canescens* and *S. columbaria*. *Hereditas* 81, 79–86. doi: 10.1046/j.1365-2540.1998.00379.x

Weising, K., and Gardner, R. C. (1999). A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* 42, 9–19.

Whitlock, M. C. (2008). Evolutionary inference from QTL: evolutionary inference from QTL. *Mol. Ecol.* 17, 1885–1896. doi: 10.1111/j.1365-294X.2008.03712.x

Würschum, T., Langer, S. M., Longin, C. F. H., Korzun, V., Akhunov, E., Ebmeyer, E., et al. (2013). Population structure, genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. *Theor. Appl. Genet.* 126, 1477–1486. doi: 10.1007/s00122-013-2065-1

Yang, R.-C., Yeh, F. C., and Yanchuk, A. D. (1996). A comparison of isozyme and quantitative genetic variation in *Paninis contorta* ssp. *latifolia* by *FAST* Genetics/155.2.945

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Rau, Attene, Rodriguez, Baghino, Pisano, Sanna, Acquadro, Portis and Comino. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.