Guijarro-Real *et al.* 2020. Euphytica 216:53
https://doi.org/10.1007/s10681-020-02586-x

Large scale phenotyping and molecular analysis in a germplasm collection of rocket salad (*Eruca vesicaria*) reveal a differentiation of the gene pool by geographical origin

Carla Guijarro-Real[1], Alejandra Navarro[2], Salvatore Esposito[2], Giovanna Festa[2], Rosaria Macellaro[2], Carlo Di Cesare[2], Ana Fita[1], Adrián Rodríguez-Burruezo[1], Teodoro Cardi[2], Jaime Prohens[1], Pasquale Tripodi[2]*

[1]Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, 46022 Valencia, Spain

[2]CREA Research Centre for Vegetable and Ornamental Crops, Pontecagnano (SA), Italy

*Corresponding author:

Pasquale Tripodi

CREA Research Centre for Vegetable and Ornamental Crops

Via dei Cavalleggeri 25, 84098 Pontecagnano-Faiano (SA), Italy

Phone: +39-089-386217

Fax: +39-089-384170

e-mail:pasquale.tripodi@crea.gov.it

Abstract

Cultivated rocket (*Eruca vesicaria*) is a leafy vegetable highly appreciated for its health-promoting virtues and consumed both raw and cooked as ready-to-use vegetable. Despite *Eruca* being cultivated worldwide, only a few cultivars are available and limited breeding activities have been carried out so far. Therefore, the genetic resources available represent an unexploited potential source of variation for breeding. In the present study, 155 *E. vesicaria* accessions from 30 countries across Europe, Asia, Africa, and America have been characterized for 54 qualitative and quantitative morphological and quality traits. Conventional descriptors and automated tools for the determination of the quality, morphology, and colour of leaves have been used. Genetic diversity was assessed using 15 inter simple sequence repeat (ISSR) and simple sequence repeat (SSR) markers. A high level of diversity was evidenced in the collection. Significant differences were found in most of the traits with the exception of five pseudo-qualitative descriptors. The first and second dimensions of a principal component analysis (PCA) with phenotypic traits accounted for 25.69 % of total variation showing a stratification of the genotypes according to the European and Asian origins. In total, 75% of the variation was contained in the first 15 components having eigenvalues higher than 1.0. Also, the population structure divided the collection into two main clusters separating European genotypes from the rest. Furthermore, hierarchical cluster analysis confirmed a geographical separation, grouping the accessions into three major clusters, which were differentiated by plant architecture, leaf and flower colour, leaf water status, leaf blade shape and hairiness of the leaves and stem. Our approach has broadened the knowledge of the diversity within the *Eruca* gene pool, thus contributing to identify sources of variation and to select the best candidates for cultivated rocket breeding programs, as well as to determine the genetic basis of plant and leaf traits in future genome-wide association studies.

Keywords: Cultivated rocket, Eruca, Genetic resources, Morphological Descriptors, Phenotyping, Leaf imaging, Genetic diversity

Introduction

Eruca vesicaria (L.) [*Eruca sativa* (Mill.)] is a leafy vegetable crop, part of the Brassicaceae family, which encloses species of interest for the content of bioactive compounds and fibers. The crop Known as 'arugula' or 'cultivated rocket' is a diploid (2n = 22), annual plant grown in spring-summer cycle and highly adapted to very marginal conditions which make it tolerant to several biotic and abiotic stresses (Padulosi and Pignone, 1996; Tripodi et al. 2017). The taxonomy of *Eruca* has been controversial till the early 2000s and there are still some unresolved issues regarding the number of species recognized (The Plant List, 2019). According to Gómez-Campo (2003) the scientific name for cultivated rocket is *E. vesicaria* and four subspecies are recognized: *vesicaria*, *sativa* Mill., *longirostris* Uechtr., Maire and *pinnatifida* Desf. Among them, *E. vesicaria* subsp. *sativa* Mill. commonly known as *E. sativa* is the most consumed species and it has spread to different parts of the world. The other subspecies are circumscribed to the Southwest Mediterranean basin including Spain, Italy. Algeria and Morocco (Gómez-Campo, 2003). According to Vavilov (1926), the genus *Eruca* originated in a large area including the Northwest region of India and Southwest Asia which are considered the primary centres of origin, while an important secondary centre of diversification is located in the Mediterranean temperate regions.

The use of cultivated rocket dates back to ancient times and is well documented in old Greek and Roman literature as a food condiment, oil seeds crop and medicinal plant (Hall et al. 2012; Padulosi and Pignone, 1996). The interest in the cultivation has increased since mid-1990'due to novel trends such the development of high added-value products addressed to the processed market, the greater demand for ready-to-use food products by the customers, and greater attention to a healthy balanced diet. In fact, rocket salad is highly recommended due to its very low-calorie content (25 cal 100 g^{-1} of fresh leaves) and high levels of health-promoting compounds with important nutraceutical and anticancer properties (Bell and Wagstaff, 2014; Higdon et al. 2007). Moreover, the peculiar pungent taste and strong flavour, due to the presence of glucosinolates (Pasini et al. 2012; Taranto et al. 2016), make this crop highly preferred in ready-to-use mixed salad packages or as topping of many dishes. Secondary uses are also recognized in cosmetics and medicine (Padulosi and Pignone, 1996). Although cultivated rocket can be considered as an underutilized or neglected crop, there is a growing interest in fresh herbs and more natural seasoning in the larger Central and Northern European markets (e.g., UK, Germany) (CBI report, 2019). In this respect, Italy and Spain are the main producers and exporting countries thanks to their favourable geographical position and agroclimatic conditions for the cultivation of rocket salad. These two countries are also main growers of 'wild rocket' (*Diplotaxis* spp.), a crop cultivated in both winter and spring season and characterized by lower plants and smaller leaves. Despite there is a large market for the latter species, the cultivated form is reported to hold a lower nitrate content in the leaves and a greater amount of vitamin C, phenolic

compounds and glucosinolates (Lenzi et al. 2000; Pasini et al. 2012; Taranto et al. 2016), and consequently with a considerable growth potential. Although the main Brassica species have been extensively studied at the genomic and phenomic levels, most of the investigations in cultivated rocket have been performed at the biochemical level (e.g., Bennett et al. 2002; Bell et al. 2015; D'Antuono et al. 2008; Pasini et al. 2012; Taranto et al. 2016; Warwick et al 2007). By contrast, studies aimed at evaluating the morphological, agronomic and genetic characterizations are scarce and mostly limited to small collections (Bozokalfa et al. 2010, 2011; Egea-Gilabert et al. 2009; Taranto et al. 2016; Warwick et al. 2007). Therefore, the *E. sativa* gene pool represents an unexploited potential source of genetic variability to be used for genetic improvement of agronomic and quality-related traits, as well as for disease resistance (Gilardi et al. 2007; Pane et al. 2017). However, unlike other crops, few efforts have been conducted for the development of novel varieties. Major constraints such as the biological and reproductive barriers occurring within this species, as well as the absence of seed certification system such as a national register of varieties in some countries (e.g., Italy) have limited breeding activities in both public and private sectors (Tripodi et al. 2017).

Assessment of morpho-agronomic features of crops is a first step for managing, preserving and promoting the use of germplasm resources. In *Eruca*, a comprehensive list of descriptors has been established for the characterization of the vegetative and reproductive parts of the plant (IPGRI, 1999). Although easy to assess, IPGRI descriptor they do not cover traits related to the quality and physiological response of plants, and moreover, leaf features (size, shape, colour) are not precisely assessed resulting in suboptimal characterization information. For covering this gap, automated devices and freeware for deep phenotyping can be applied (Awada et al. 2018). Understanding genetic diversity is a key step for selection and breeding of germplasm resources. In cultivated rocket, genomic resources for in-depth diversity investigation has not yet been established; however, different types of molecular markers have been already validated and can be applied within this species (Egea-Gilabert et al. 2009; Taranto et al. 2016; Thakur et al. 2018).

The purpose of this study was to evaluate a collection of 155 *Eruca sativa* accessions retrieved from main gene-banks using a comprehensive phenotyping strategy enclosing common descriptors for the estimation of agro-morphological and phenological traits, and automated devices for the assessment of morphological and qualitative features of leaves. In order to complement the morphological, agronomic and quality traits analysis, molecular markers were surveyed for population structure and phylogenetic relationships investigations. The multidisciplinary approach and the integration of the various sources of data can be considered the main attempt in cultivated rocket and a first step toward the enhancement of *E. sativa* germplasm, involving a large collection and several traits, some of which have never been assessed in this crop. The gained information will contribute to increasing the knowledge of the phenotypic and genetic diversity of the cultivated rocket salad gene-pool. Data could be used to select the best parent lines for breeding programs and/or for direct consumption, as example selecting lines with a leaf morphology similar to the wild

rocket, which is highly appreciated particularly in Southern European markets, or identifying new types to use in mixed salad bags. Moreover, better integration with genomics will allow the exploiting of phenotypic data in genome-wide association studies.

Materials and Methods

Plant material

A set of 155 accessions of *E. vesicaria* (151 subsp. *sativa* and 4 subsp. *pinnatifida*) sampled from 30 countries from around the world were included in this study (**Figure 1, Supplementary Table 1**). Genotypes were retrieved from the CGN (Wageningen, The Netherlands), IPK (Gatersleben, Germany) and USDA-GRIN (USA) seed banks. The study was performed at the experimental station of the Research Centre for Vegetable and Ornamental Crops located at Pontecagnano (SA, Italy). Seeds were sown on the third week of March 2018 in pots (diameter 23 cm) using a mixture of peat and perlite (5:1) as soil medium, and plants were grown until late July. Plants were grown under a glasshouse with controlled environmental conditions in a completely randomized design using three plants for each accession. The cultivation method followed the standard agronomic practices for the crop.

Morphological descriptors

Plants were phenotyped for 33 qualitative/pseudo-qualitative descriptors selected from the standardized descriptors for *Eruca* spp. (IPGRI, 1999). Three categories of descriptors were analysed: i) traits related to seedling and leaf characteristics (15); ii) vegetative descriptors related to the characteristics of the plant and the stem (11); and, iii) flowering traits (7). Full details of the traits scored and method of assessment can be found in the *Eruca* descriptors list (IPGRI, 1999) and in Table 1. A new descriptor ('Lateral lobe shape') was included, ranging from 0 (crenate) to 3 (entire). Details are reported in Supplementary Figure 1.

Leaf qualitative traits

Leaf fresh weight was obtained on a sample of 10 leaves for each accession; dry weight was obtained at a temperature of 70 °C until constant weight. The plant water status of rocket accessions was determined by the leaf relative water content (RWC_l; %) and the osmotic potential at full turgor (Ψ_{100s}; MPa). Ψ_{100s} and RWC_l were performed in excised leaves harvested at midday (10:30 - 12:30 h solar time) using three developed leaves per plant rocket accession. RWC_l, was calculated according to the formula (Barrs and Weatherley, 1962) as follows:

$$RWC_l = \frac{FW-DW}{TW-DW} \times 100 \quad (1)$$

which consider for the whole leaf the fresh (FW), dry (DW) and turgid (TW) weights (g). After harvesting and weighting (FW), for each leaf, the cut end was placed in distilled water and kept in dim light at 4 °C for 24-48 h till the reaching of TW. DW was assessed after drying for 48 h at 70 °C. The leaves used for determining Ψ_{100s} were placed by their petiole into flasks of distilled water and kept overnight in dim light at 4 °C to reach full saturation. Leaves were then dried by filter paper, wrapped in aluminum foil and immediately frozen in liquid nitrogen before storing at -30 °C. Before the measurements, samples were thawed and leaf sap was extracted for immediate determination of osmolality (mOsmol kg^{-1}) using a freezing point osmometer (Osmomat 3000, Gonotec GmbH, Berlin, DE), and Ψ_{100s} (MPa) was obtained by multiplying the osmolality with –2.479 (conversion factor at 25 °C; Taiz and Zeiguer, 2006 and Pariyar et al. 2013). CIELAB colour coordinates (L*a*b*) were measured using a CR-210 Chroma Meter (Minolta Corp., Osaka, Japan) on a sample of three leaves per accession. Measurements were done in duplicate on the two opposite lobes excluding the central rib and expressed as L*, a*, b* values, which indicate darkness/lightness (close to 0/close to 100), red/green (positive values/negative values) and yellow/blue (positive values/negative values) intensity, respectively. Chroma (C) and hue angle (h) were estimated by the a* and b* values using the following equations:

$$C = [(a^*)^2 + (b^*)]^{1/2} \quad (2)$$

$$h = \tan^{-1} \frac{b^*}{a^*} \quad (3)$$

Chroma indicates colour saturation, while hue is a measure of the angle in the CIELAB colour chart (0° or 360° indicates red hue, while angles below 90°, 180°, and 270° indicate yellow, green and blue hue, respectively). Relative chlorophyll content in leaves at vegetative and flowering stage was determined using a SPAD-502 meter (Konica-Minolta, Japan).

Total phenolics content was determined in three independent replicates per accession according to the Folin–Ciocalteu procedure (Singleton and Rossi, 1965) as indicated in Guijarro-Real et al. (2019). For each replicate, 125 mg of lyophilised powder were extracted with 5 mL of 70% (v/v) acetone containing 0.5% (v/v) glacial acetic acid, for 24 h under continuous stirring. An aliquot of 65 µL reacted with 0.5 mL of diluted (10%, v/v) Folin–Ciocalteu reagent (Sigma-Aldrich, St. Louis, MO, USA) and left at room temperature for 5 min. The solution was then incubated with 0.5 mL of sodium carbonate (60 g L^{-1}, Sigma-Aldrich) for 90 min. Absorbance was measured at 750 nm with a Jasco spectrophotometer (Bio-Rad, Herts, UK). Gallic acid was used for standard calibration, and results were expressed as grams of gallic acid equivalents (GAE) per 100 g of dry matter (DM).

Leaf shape traits

Three leaves per accession at the maturity stage were harvested and subjected to automated phenotyping using a CanoScan LiDE 210 photo scanner (Canon, Tokyo, Japan) at a resolution of 300 dpi. Scans were conducted in a dark

room using a white panel as a contrast. Morphometric analysis was performed using the MorphoLeaf v.1.14 plug-in for Free-D v.1.14 software (Andrey and Maurin, 2005). Nine quantitative descriptors were recorded as described in Table 2. These descriptors included: i) traits related to the leaf size directly obtained from the software used, and ii) traits indirectly determined from the measurements. Thus, total leaf length and width (including petiole), blade length, perimeter and area (without petiole), and petiole length and width were directly determined. Traits indirectly measured included the leaf shape index (>1 indicates narrow shape, <1 indicates broad shape) and petiole shape index.

Genetic analysis and population structure

DNA was isolated from young leaves of each genotype using NucleoSpin® Plant II Midi kit (MACHEREY-NAGEL GmbH & Co. KG., Düren, Germany). Concentration and quality parameters (adsorbance at 260/280 and 260/230) were measured before the molecular analyses using a UV–Vis spectrophotometer (ND-1000; NanoDrop, Thermo Scientific, Wilmington, DE, USA). Genetic diversity was assessed using twelve Simple Sequence Repeats (SSR) and three Inter Simple Sequence Repeats (ISSR) (Table 3) previously reported to be polymorphic in *E. sativa* (Taranto et al. 2016, Thakur et al. 2018). PCR amplifications were performed following the conditions described in Taranto et al. (2016) and Thakur et al. (2018) for ISSR and SSR, respectively. The reactions were amplified using a C-1000 Touch™ Thermal Cycler (Bio-Rad). Electrophoretic separation was performed in 2.5% methaphor-agarose gels (Lonza, USA) using gene-ruler DNA ladder (Life Technologies™) for fragment sizes estimation. The amplicons were visualized using SYBR® safe (Life Technologies™) staining and the fluorescence was viewed using Gel Doc™ XR (Biorad). Polymorphic patterns were then scored for their absence/presence and combined into a separate rectangular binary matrix. Only robust bands across all samples were considered. Polymorphic Information Content (PIC), Effective Multiplex Ratio (EMR), Marker Index (MI) and Resolving power (RP) were calculated. PIC value was calculated over the bands for each primer according to the following equation:

$$PIC_x = 2 \times F_x \times (1 - F_x) \quad (4)$$

where x is the marker, Fx and (1-Fx) are the frequencies of the present and absent marker bands, respectively. EMR (Powell et al. 1996) is defined as:

$$EMR = n_p \times \left(\frac{n_p}{n}\right) \quad (5)$$

where n_p is the fraction of polymorphic loci and n_p/n is the number of polymorphic loci for an individual assay. MI is a measure of the marker efficiency defined as the product of the average PIC for the polymorphic bands in any assay and EMR for that assay (Powell et al. 1996).

$$MI = PIC \times EMR \quad (6)$$

RP for each primer is defined as follows:

$$RP = \sum Ib \quad (7)$$

in which "Ib" represents the informativeness of a band taking the values of 1-(2x [0.5-p]), being "p" the proportion of each genotype containing the band (Prevost et al. 1999).

The population structure of the rocket salad collection was estimated using the Bayesian model implemented in STRUCTURE v 2.3.4 (Pritchard et al. 2008). This method infers the allele frequencies for each genotype estimating the best subgroups (K) of a population on the basis of the MCMC (Markov Chain Monte Carlo) algorithm. Correlation among alleles frequencies was performed using the admixture model analysis. For each value of K (ranging from 2 to 10), five runs have been carried out using burn-in period of 50,000 steps and 50,000 MCMC replicates as parameters. For each K, a membership coefficient (qi) was calculated. The assignment of a genotype to a certain group was considered if its membership coefficient (qi) was ≥ 0.50. For the definition of the best number of subgroups, we used the Evanno's method implemented in Structure Harvester software (Earl and vonHoldt, 2012). The algorithm calculates the maximum value of Δ associated with each subgroup, estimating the best K value for the population.

Data analyses

Analysis of variance (ANOVA) was performed using JMP v7.0 (SAS Institute, Cary, NC). Means, standard errors, range values and coefficient of variation were used for descriptive analysis of traits.

Coefficient of variation was calculated as follows:

$$CV = \frac{Std}{M}\ x\ 100 \quad (8)$$

In which "Std" is the standard deviation and "M" is the mean for a trait.

Significant differences among species means were detected using Tukey HSD (honest significant difference) test ($P<0.05$).

Broad-sense heritability (H^2) was estimated as as follows:

$$H^2 = VG/VP \quad (9)$$

where VG is genotypic component of variance and VP is the total phenotypic variance.

Hierarchical Cluster Analysis (HCA) using the Ward's coefficient and Principal Component Analysis (PCA) were performed with the computer package XLSTAT 2012.1. Correlations across the genotypes for phenotypic traits were calculated using the Pearson's test at $P<0.01$ after Bonferroni's correction for multiple comparisons (Curtin and Schultz, 1998). The correlogram was constructed and visualized using the Corrplot package implemented in R. The relationship between phenotypic and molecular data matrices was computed by the Mantel test (Mantel, 1967) using the Pearson's r-value.

Results

Phenotypic variation

The characterization with morphological descriptors revealed a high degree of diversity within the *Eruca* collection (Table 1, Figure 2). Out of the 33 conventional descriptors used, highly significant differences ($P < 0.001$) were found for 26 traits including all seedling and leaf characteristics. As exceptions, the width of the plant and stem and the related shape indexes did not show significant differences among individuals of the collection, and the flower sepal was caducous in all genotypes. For five traits, the range of variation did not cover the entire descriptors scale, in fact, the collection did not include individuals with white colour of hypocotyls, leaves and flower raceme, very thick stem, and vary hairy leaf mid-vein.

A coefficient of variation of ~ 90% and above 155% was observed for growth habit and the hairiness of leaves, respectively. The remaining traits exhibited a CV lower than 55%. In general, high heritability values were observed ranging from 0.91 to 1.00 (Table 1).

As occurred for the plant descriptors, a wide diversity was found for qualitative and leaf shape traits (Table 2). In all traits, the analysis of variance revealed highly significant differences among accessions ($P<0.001$). A wide range of variation was observed for several quantitative traits, with the coefficient of variation being over the 90% for stem shape index (4.71-150), leaf area (19.98 - 196.81) and leaf perimeter (36.78 - 221.81). Except for leaf fresh and dry weight and petiole length and shape, the remaining traits showed a CV lower than 35%. The broad-sense heritability ranged from 0.55 (Leaf blade length with petiole) to 0.90 (Ψ_{100s}). Qualitative traits exhibited on average a greater heritability (Table 2). CIELAB coordinates showed for all accessions analysed, an average L* value of 41.18, a negative a* value of −17.34, and a positive b* value of 22.58. Average values for chroma and hue angle were 28.53 and 128.03, respectively. Overall, the collection showed vivid colours with a predominant amount of the greenness and yellowness components, while hue angle fell in the green, green-blue range. The value of chlorophyll as SPAD unit was on average higher at the flowering stage, with 10% of increment with respect to the vegetative stage (Table 2). Total polyphenols ranged from 3.79 to 8.72 g GAE 100 g^{-1} with a mean of 5.94 g GAE 100 g^{-1}. Leaves length ranged from a minimum of 14.71 cm to a maximum of 25.9 cm. The petiole represented from 1/100 to 1/3 of the total leaf length. Leaf width was on average 9.53 cm, reaching 13.41 cm as the highest value (Table 2). The average leaf shape index indicated leaves more obovate or rounded than lanceolate. The average leaf area was 84.17 cm^2, reaching the leaves of accession PI 650819 the greatest area with a value of 4.6-fold times higher than the lowest one (PI 426198, 31.72 cm^2). Finally, the blade perimeter without petiole had an average value of 91.57 cm, with a 5-fold difference between the lowest and the highest values.

Genetic diversity

The fifteen markers tested on *E. sativa* accessions allowed the detection of 57 polymorphic loci (Table 3), for a total of approximately 8,800 data points. On average, 3.8 polymorphic alleles were produced, ranging from 1 (BrgMS4533) to 10 (BrgMS490). Among the markers used, the mean PIC was 0.20 with values ranging from 0.01 for BrgMS421 to 0.50 for UBC-811. EMR values ranged from 1.00 (BrgMS4533) to 10.00 (BrgMS490) with an average of 3.80. RP had an average value of 0.27 ranging from 0.01 for BrgMS421 to 0.96 for UBC-811. MI ranged from 0.03 for BrgMS421 to 2.00 for UBC-811 with an average value of 0.78. The best markers for the discrimination of individuals within the population were the ISSRs UBC-811, UBC-815 and UBC-876 which showed the best PIC and RP values (Table 3). Germplasm groups were defined by STRUCTURE with the number of ancestral populations (K) ranging from 2 to 10. Mean LnP(K) and ΔK values were retrieved from the STRUCTURE HARVESTER freeware (Figure 3). At K = 2, a cluster of accessions with European origins was observed. K = 3 highlighted the clustering of accessions mostly according to the centre of origin (Supplementary Table 2).

Multivariate analysis and correlation between traits

The principal component analysis (PCA) in the first two dimensions explained 25.69% of the total variance (Figure 4A). The first component, accounting for 16.43% of the total variance, was positively correlated to 17 plant, leaf, flower and colour traits, and negatively correlated with remaining traits, including RWC, antioxidants and chlorophyll content (Figure 4B). The second component which explained 9.25% of the total variance was positively correlated to several leaf shape traits, RWC and antioxidants and negatively correlated to chlorophyll content. IPGRI (1999) traits and leaf morphology were distributed in both negative and positive parts of the biplot. The eigenvalues and the variable contribution for the first two components are reported in Supplementary Table 3. The first fifteen components showed eigenvalues higher than 1.0 and explained up to 75% of the variation. The 100% of variation was explained by the first 35 components (Supplementary Figure 2). The projection of the accessions on the two-dimensional PCA graph evidenced a discrete intraspecific variability within the *E. sativa* collection under study revealing a geographical differentiation between genotypes. Indeed, the European accessions which clustered with American ones were quite separated from the Asian accessions. On the contrary, the African accessions were interspersed between the European and Asian groups. The Pearson rank correlation coefficients after Bonferroni correction calculated for qualitative and quantitative traits revealed a variable correlation stringency between pairs of traits (Figure 5). Strong positive significant correlations were observed between traits related to plant architecture and leaf morphology, and between flowering time and leaf weight. Negative significant correlations were found among CIELAB coordinates and between leaf shape and width. Total polyphenols were negatively correlated to petal colour and leaf perimeter and positively

correlated to chlorophyll content and RWC, leaf weight and sepal hairiness although these correlations were not tight. For the identification of the most important traits which determine the variability of the collection, we further selected four morphological descriptors, four qualitative traits and three leaf shape traits having a high correlation to the first two principal components and mainly contributing to the variance explained (Supplementary Table 3). A new PCA was inferred using these traits, in this case explaining the 59.98% of total variation with the first two components (Supplementary Figure 3). The derived two-dimensional PCA graph allowed better discrimination of the accessions by geographical origin. The degree of relationship between elements of phenotypic and molecular matrices showed a significant correlation between plant descriptors and leaf quantitative traits ($r = -0.035$, $P < 0.001$) while no significant correlations were found between plant descriptors and genetic data ($r = -0.017$, $P = 0.073$) and quantitative traits and genetic data ($r = -0.0002$, $P = 0.982$).

Cluster Analysis

Hierarchical clustering based on plant descriptors, agronomic, qualitative and leaf shape traits separated the collection into three main clusters (Figure 6). The first cluster (C1) included 48 genotypes, most of which (42) were retrieved from Middle Eastern countries and Pakistan, while the remaining accessions corresponded to the United States (2), Egypt (1) and Germany (1). The second cluster (C2) included 81 genotypes, most of them represented by European accessions (58), followed by Asian (13), African (8), and American (2) accessions. The third cluster (C3) included a miscellany of 26 accessions, 16 retrieved from Europe and the rest from other world regions. The analysis of variance among clusters revealed highly significant differences ($P < 0.001$) for 38 traits, and significance at $P < 0.01$ and $P < 0.05$ for the other ten traits. Accessions from C1 evidenced a predominant stem axis elongation, thicker leaves, more intense flower petal colour, and leaf pigments, as well as a moderate increase in average values for L, a* and chroma, and the lowest values of RWC and chlorophyll content (Table 4). Accessions from C2 had the highest average values for many traits, in particular for plant height and width, leaf shape traits and total phenolics content. Moreover, hypocotyl colour, leaf and sepal hairiness, and lateral lobes shape were more intense in C2 respect the other two clusters. Accessions included in C3 were characterized by late flowering and a more marked pubescence in leaves and stems respect the other two clusters. Furthermore, accessions belonging to this cluster had higher values of RWC, Ψ_{100s} and chlorophyll content (SPAD).

Discussion

Assessing crop variability is of paramount importance for the management and breeding of genetic resources. In the present study, a large collection of *E. vesicaria* has been investigated for its molecular diversity and phenotyped for 54

plant and leaf traits by means of common descriptors and semi-automatic high-throughput techniques. So far, no large-scale genotyping and phenotyping studies have been carried out in cultivated rocket and this research was aimed at covering this gap, representing an attempt to deeply evaluate a broad collection of this crop in terms of number of accessions.

In agreement with previous findings (Egea-Gilabert et al. 2009; Taranto et al. 2016), the analysis with molecular markers in this work confirmed that ISSRs are more informative than SSRs to discriminate cultivated rocket genotypes. The set of microsatellites analysed in the present study was previously used in a *Brassica* species genetic diversity study, which included few *E. sativa* genotypes for phylogenesis purposes (Thakur et al. 2018). Although the markers selected were those that amplified two or more alleles in *Eruca* (Thakur et al. 2018), the information was related to only 4 accessions and not representative for a larger collection. In our work we confirmed that microsatellites developed in *Brassica* species can be used to investigate the genetic diversity of cultivated rocket, although more specific markers or genomic resources could give better insights. Population structure revealed a stratification of the accessions according to geographical origin. In agreement with our findings, Warwick et al. (2007), testing 46 accessions with a combination of four AFLP primer pairs, reported a distinction of *Eruca* genotypes according to the Mediterranean and Asian origins. The collection analysed by these authors was partially common to our core set, although the separation of the *E. vesicaria* subsp. *sativa* and *E. vesicaria* subsp. *pinnatifida* subspecies, reporting 8 out of 234 AFLP polymorphic bands unique to the *E. vesicaria* subsp. *pinnatifida* cluster, was not found in our study. Although different types of *E. vesicaria* subsp. *pinnatifida* accessions were used in the two studies, the contrasting results could be due both to the different type of markers and number of polymorphic sites developed, as well as the probable mis-classification occurring within gene banks. Therefore, a deeper investigation is required to clarify this aspect.

Most morpho-agronomic traits exhibited significant differences among genotypes suggesting the abundant diversity within the *Eruca* core-set analysed. On the basis of principal components and cluster analysis results, the accessions were separated according to their geographical origin, highlighting differences between the Asian and European genotypes. Asian accessions displayed a more intense yellow colour of flowers and less intense leaves colour, while European ones stood out mainly in relation to the earliness in flowering, greater plant height and higher content of chlorophyll and polyphenols. Regarding leaf morphological parameters, European accessions had longer leaves and a greater surface area, while Asian ones had a greater perimeter, due to a more marked lobation. The observed variation in leaf morphology is probably related to selection factors leading to a reduced evapotranspiration surface in accessions originating from warmer countries. The rate of variation observed in the present study agrees with previous findings (Bozokalfa et al. 2011; Egea Gilarbert et al. 2009; Warwick et al. 2007). Bozokalfa et al. (2011) characterized 35 *Eruca* and *Diplotaxis* accessions for 51 agro-morphological traits. Although the accessions studied belonged to different

species, the authors highlighted how the Turkish cultivated rocket plant gene-pool was diverse from the rest. In our study, the whole phenotypic dataset showed about 26% of the variation in the first two principal components while the identification of the minimal set of 11 qualitative and quantitative traits maximized the variability among accessions, highlighting the separation of the European and the Asian gene pools. The whole variation observed in the collection studied is similar to previous essays aimed at the characterization of leaf morphological traits in Spanish local cultivated rocket materials (Egea Gilarbert et al. 2009) and the characterization of agronomic and seed quality traits in 159 *E. sativa* accessions from diverse genebanks (Warwick et al. 2007). Despite these latter investigations assessing diverse *Eruca* germplasm resources in terms of site sampling, an unclear separation following the geographical origin was observed. This was probably due to the lower number of traits recorded and countries represented in comparison to our study. As we observed, the differentiation of the gene pools was due to the several plant and leaf traits whose differentiation may be the sum of ecogeographical and anthropic factors, as well as the natural selection occurring in the corresponding areas of origin and centres of domestication (Meyer and Purugganan, 2013; Fine, 2015).

In cultivated rocket, leaf morphology and colour are attributes of particular importance to be considered for market types definition and consumers acceptance (Egea-Gilabert et al. 2009; Bell et al. 2017), thus being major objectives to pursue for varietal selection. The imaging analysis used in this work allowed to precisely assess the size and shape of leaves, reinforcing conventional descriptors. In fact, several correlations were found between the two types of measures. The collection included very different types, from small to large leaves with different thickness, shape apex and lobation. The perimeter obtained through leaf scans has been revealed a good predictor for the morphology of leaves since it considered the degree of lobation and the shape of lobes. Perimeter was indeed more correlated to leaf lobation than blade area. We also introduced a new trait, 'lateral lobe shape', which according to our results can be suggested as an effective descriptor for leaf characterization. The scale proposed for this new descriptor allowed a better discrimination of the accessions and ranked among the first 15 traits contributing to the variation in the second component.

The collection was also analyzed for traits related to quality and resistances to biotic and abiotic stress. Antioxidant capacity is an important trait indicating the polyphenols content and other reducing compounds able to reduce deleterious oxidative stress and prevent chronic diseases (Ninfali et al. 2005). Furthermore, polyphenols can be involved in defence mechanisms against various pathogens and confer tolerance to different environmental stresses (Cheruiyot et al. 2007; Daayf et al. 2012). Relative water content and osmotic potential of leaves are parameters related to the plant water status and can be used for screening the tolerance to abiotic stresses (drought and salinity) of different genotypes (Tanentzap et al. 2015). Leaf colour coordinates have been reported to be related with the total glucosinolates content in leaves (Taranto et al. 2016). The presence of glandular trichomes, which contribute to the hairiness of stem

and leaves, play a key role in defence against herbivores and insect, and represent an anatomical modification to increase light reflectance, and thus reducing the water losses (Wagner, 1991; Hauser, 2014). Various evidences are reported in Brassicaceae (Dalin et al. 2008; Handley et al. 2005).

Different genotypes displayed desirable scores for the above-mentioned traits, and several correlations were found with the agro-morphological characteristics. The knowledge retrieved from the germplasm studied will facilitate the selection of novel lines for cultivation, allowing to identify potential candidates for breeding programmes. Efforts in *Eruca* genomics could enhance the broad phenotypic information by combining it with genome-wide association, giving in this way insight into the understanding of the genetic basis of plant and leaf traits in cultivated rocket.

Conclusions

In the present study, a large collection of cultivated rocket has been comprehensively investigated for a broad range of traits and for its genetic diversity. Both phenomic and molecular data demonstrated how *E. vesicaria* subsp. *sativa* has been diversified according to selection factors probably occurred in the main respective centres of origin. For genetic improvement, the multidisciplinary strategy used represent an essential step toward a better use of the potentiality stored in cultivated rocket. Indeed, the wide variability observed can be considered as a reservoir of traits to be exploited for future breeding programmes. Furthermore, from the gained phenotypic information, it will be possible to select potential lines for both cultivation and different end-uses. Future direction toward the dissection of the genetic basis underlying plant and leaf traits will be possible by means of integration of phenomic and genomic data.

Acknowledgements

The work was supported by 'RGV-FAO' project funded by the Italian Ministry of Agriculture, Food and Forestry. C. Guijarro-Real thanks the Ministerio de Educación, Cultura y Deporte of Spain (MECD) for its financial support by means of a predoctoral FPU grant (FPU14/06798), and for the specific grant for mobility (EST17/00354) from the same Organization.

Confict of interest

The authors declare that they have no confict of interest.

References

Andrey P, Maurin Y (2005) *Free-D*: an integrated environment for three-dimensional reconstruction from serial sections. J Neurosci Methods 145:233-244

Awada L, Phillips PWB, Smyth SJ (2008) The adoption of automated phenotyping by plant breeders. Euphytica 214: 148

Barrs HD, Weatherley PE (1962). A re-examination of the relative turgidity technique for estimating water deficits in leaves. Aust J Biol Sci 15(3):413–428

Bell L., Methven L., Wagstaff C. 2017. The influence of phytochemical composition and resulting sensory attributes on preference for salad rocket (*Eruca sativa*) accessions by consumers of varying TAS2R38 diplotype. Food Chemistry 222:6-17

Bell L, Oruna Concha MJ, Wagstaff C (2015) Identification and quantification of glucosinolate and flavonol compounds in rocket salad (*Eruca sativa*, *Eruca vesicaria* and *Diplotaxis tenuifolia*) by LC-MS: Highlighting the potential for improving nutritional value of rocket crops. Food Chem 172:852-861

Bell L, Wagstaff C (2014) Glucosinolates, myrosinase hydrolysis products, and flavonols found in rocket (*Eruca sativa* and *Diplotaxis tenuifolia*). J Agric Food Chem 62:4481-4492

Bennett RN, Mellon FA, Botting NP, Eagles J, Rosa EAS, Williamson G (2002) Identification of the major glucosinolate (4-mercaptobutyl glucosinolate) in leaves of *Eruca sativa* L. (salad rocket). Phytochemistry 61:25-30

Bozokalfa MK, Eşiyok D, Ilbi H, Kaygisiz AT (2010): Estimates of genetic variability and association studies in quantitative plant traits of *Eruca* spp. landraces. Genetika 42(3):501-512.

Bozokalfa MK, Eşiyok D, İlbi H, Kavak S, Kaygisiz AT (2011) Evaluation of phenotypic diversity and geographical variation of cultivated (*Eruca sativa* L.) and wild (*Diplotaxis tenuifolia* L.) rocket plant. Plant Genet Resour C 9:454−563

CBI Centre for the Promotion of Imports from developing countries (2019) https://www.cbi.eu/market-information/fresh-fruit-vegetables/fresh-herbs/europe/. Accessed 21 September 2019

Cheruiyot EK, Mumera LM, Ngetich WK, Hassanali A, Wachira F (2007) Polyhenols as potential indicators for osmotic tolerance in tea (*Camellia sinensis* L.). Biosci Biotechnol Biochem 71:2190–2197. doi: 10.1271/bbb.70156

Curtin F, Schulz P (1998) Multiple correlations and Bonferroni's correction. Biol Psychiatry 44:775–777

Daayf F, El Hadrami A, El-Bebany AE, Henriquez MA, Yao Z, Derksen H, El-Hadrami I, Adam LR (2012) Phenolic compounds in plant defense and pathogen counter-defense mechanisms. Rec Adv Polyphen Res 3:191–208

Dalin P, Agren J, Björkman C, Huttumen P, Kärkkäinen K (2008) Leaf trichome formation and plant resistance to herbivory. In: Schaller A, ed. *Induced plant resistance to herbivory*. Springer Science+Business Media. pp 89-105

D'Antuono LF, Elementi S, Neri R (2008) Glucosinolates in *Diplotaxis* and *Eruca* leaves: Diversity, taxonomic relations and applied aspects. Phytochemistry 69:187−199

Earl DA and vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genet Resour 4:359–361

Egea-Gilabert C, Fernandez JA, Migliaro D, Martinez Sanchez JJ, Vicente MJ (2009) Genetic variability in wild vs. cultivated *Eruca vesicaria* populations as assessed by morphological, agronomical and molecular analyses. Sci Hortic 121:260−266

Fine PVA (2015) Ecological and evolutionary drivers of geographic variation in species diversity. Annu Rev Ecol Evol Syst 46:369–392. doi: 10.1146/annurev-ecolsys-112414-054102

Gilardi G, Chen G, Garibaldi A, Zhiping C, Gullino ML (2007) Resistance of different rocket cultivars to wilt caused by strains of *Fusarium oxysporum* under artificial inoculation conditions. J Plant Pathol 89:113–117

Gómez-Campo C (2003) Morphological characterization of *Eruca vesicaria* (Cruciferae) germplasm. Bocconea 16:615–624

Guijarro-Real, C., Prohens J, Rodriguez-Burruezo A, Adalid-Martínez AM, López-Gresa MP, Fita A (2019) Wild edible fool's watercress, a potential crop with high nutraceutical properties. PeerJ, 7:e6296

Hall MKD, Jobling JJ, Rogers GS (2012) Some perspectives on rocket as a vegetable crop: A review. Veg Crop Res Bull 76:21–41

Hanldey R, Ekbom B, Ågren J (2005) Variation in trichome density and resistance against a specialist insect herbivore in natural populations of *Arabidopsis thaliana.* Ecol Entomol 30:284–292

Hauser MT (2014) Molecular basis of natural variation and environmental control of trichome patterning. Front Plant Sci 5:320

Higdon JV, Delage B, Williams DE, Dashwood RH (2007) Cruciferous vegetables and human cancer risk: Epidemiologic evidence and mechanistic basis. Pharmacol Res 55:224−236

IPGRI (1999) Descriptors for rocket (*Eruca* spp.). International Plant Genetic Resources Institute, Rome, Italy.

Lenzi A, Tesi R, 2000. Effect of some cultural factors on nitrate accumulation in rocket [*Diplotaxis tenuifolia* (L.) D.C. – *Eruca sativa* Mill.]. Riv. Agronomia 34(4):419–424

Mantel NA (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. Nat Rev Genet 14:840–852

Ninfali P, Mea G, Giorgini S, Rocchi M and Bacchiocca M (2005) Antioxidant capacity of vegetables spices and dressings relevant to nutrition. Br J Nutr 93:257–266

Padulosi S, Pignone D (1996) Rocket: A Mediterranean crop for the world. Report of a workshop. – IPGRI International Plant Genetic Resources Institute, Bioversity International Rome, Italy, pp 101

Pane C, Sigillo L, Caputo M, Serratore G, Zaccardelli M, Tripodi P (2017) Response of rocket salad germplasm (*Eruca* and *Diplotaxis* spp.) to major pathogens causing damping-off, wilting and leaf spot diseases. Arch Phytopathology Plant Protect 50:167-177

Pariyar S, Eichert T, Goldbach HE, Hunsche M, Burkhardt J (2013) The exclusion of ambient aerosols changes the water relations of sunflower (*Helianthus annuus*) and bean (*Vicia vaba*) plants. Environ Exp Bot 88:43–52

Pasini F, Verardo V, Caboni MF, D'Antuono LF (2012) Determination of glucosinolates and phenolic compounds in rocket salad by HPLC-DAD−MS: Evaluation of *Eruca sativa* Mill. and *Diplotaxis tenuifolia* L. genetic resources. Food Chem 133:1025-1033

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225−238

Prevost A, Wilkinson MJ (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theor Appl Genet 98:107-112

Pritchard JK, Stephens M and Donnelly P (2005) Inference of population structure using multilocus genotype data. Genetics 155:945-959

Singleton VL, Rossi JA Jr (1965) Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. Am J Enol Viticult 16:144-158

Tanentzap FM, Stempel A, Ryser P (2015) Reliability of leaf relative water content (RWC) measurements after storage: Consequences for in situ measurements. Botany 93:535–541

Taiz L, Zeiger E (2006) Plant Physiology. 4th Edition, *Sinauer* Associates Inc., *Sunderland,* Massachusetts, USA.

Taranto F, Francese G, Di Dato F, D'Alessandro A, Greco B, Onofaro Sanajà V, Pentangelo A, Mennella G, Tripodi P (2016) Leaf metabolic, genetic, and morphophysiological profiles of cultivated and wild rocket salad (*Eruca* and *Diplotaxis* spp.) J Agric Food Chem 64:5824-5836

The Plant List (2019). http://www.theplantlist.org/tpl1.1/search?q=Eruca. Accessed 21 September 2019

Thakur AK, Singh KH, Singh L, Nanjundan J, Khan YJ, Singh D (2018) SSR marker variations in Brassica species provide insight into the origin and evolution of *Brassica amphidiploids*. Hereditas 155:6. doi: 10.1186/s41065-017-0041-5

Tripodi P, Francese G, Mennella G (2017) Rocket salad: Crop description, bioactive compounds and breeding perspectives. Adv Hortic Sci 31:107-113

Vavilov NI (1926) Centers of origin of cultivated plants. Bull Appl Bot Genet Plant Breed 16:1-248

Wagner GJ (1991) Secreting glandular trichomes: more than just hairs. Plant Physiol 96:675-679

Warwick SI, Gukel RK, Gomez-Campo C and James T (2007) Genetic variation in *Eruca vesicaria* (L.) Cav. Plant Genet Resour C 5:142–153

Table 1. List of IPGRI plant qualitative pseudo-qualitative descriptors for *Eruca* and their variation parameters.

Descriptor name	Descriptor number[a]	Acronym	Scale	Mean Square	Mean	Range[b]	CV%	H^2
Seedling and leaf characteristics								
dd for germination	4.6.1	ddg	n° of days	2.29	7.92	(15.00 - 5.00)**	19.84	0.91
Hypocotyl colour	6.1.1	Hc	6 = purple; 1 = white	2.03	4.37	(6.00 - 2.00)**	32.43	1.00
Leaf colour	6.1.2 & 6.1.18	Lc	7 = purple; 1 = white green	0.39	3.75	(5.00 - 1.00)**	16.60	0.99
Leaf margin	6.1.3	Lm	6 = undulate; 1 = entire	0.26	1.68	(4.00 - 1.00)**	30.37	0.99
Leaf growth attitude (<30 gg)	6.1.11	Lg-	3 = erect; 1 = prostrate	0.04	2.93	(3.00 - 2.00)**	6.86	0.96
Leaf growth attitude (> 30 gg)	6.1.11	Lg+	3 = erect; 1 = prostrate	0.22	2.21	(3.00 - 1.00)**	21.03	1.00
Leaf lobation	6.1.13	Ll	2 = markedly present; 0 = absent	0.29	1.79	(2.00 - 0)**	29.82	1.00
Leaf apex shape	6.1.14	La	4 = bradly rounded; 1 = narrowly acute	0.36	3.17	(4.00 - 1.30)**	19.14	0.96
Leaf blade thickness	6.1.15	Lt	7 = thick; 3 = thin	2.70	4.87	(9.00 - 2.00)**	33.64	0.99
Leaf lamina attitude	6.1.17	Lla	3 = convex; 1 = concave	0.31	2.25	(3.00 - 1.00)**	24.84	0.99
Leaf hairiness	6.1.19	Lh	7 = dense; 0 = absent	4.85	1.30	(7.00 - 0)**	167.89	1.00
Leaf Vein hairiness	6.1.19	Lv	7 = dense; 0 = absent	2.54	1.01	(5.00 - 0)**	157.40	1.00
Lateral lobe shape	-	Lls	3 = entire with big lobes; 0 = crenate with small lobes	0.76	2.30	(4.00 - 0)**	37.71	1.00
Petiol and or miedwein enlargment	6.1.20	Pe	3 = enlarged; 1 = narrow	0.28	1.54	(3.00 - 1.00)**	34.45	0.99
Petiole and or midwein colour	6.1.26	Pc	6 = red; 1 = white	2.14	2.65	(6.00 - 2.00)**	54.96	1.00
Vegetative characteristics								
Plant growth habit	6.2.2	Pg	9 = other*; 3 = branching terminating in pre-floral; 1 = branching supporting leaves	2.34	1.70	(9.00 - 1.00)**	89.94	1.00
Stem axis elongation or enlargment	6.2.12	Sa	2 = elongated; 1 = enlarged	0.24	1.59	(2.00 - 1.00)**	31.17	0.99
Plant height PH	6.2.3	Ph	cm	357.45	46.92	(101.00 - 9.00)*	40.04	1.00
Plant width PW	6.2.4	Pw	cm	170.69	58.01	(96.00 - 24.00)^{NS}	22.39	1.00
Plant shape index (PH/PW)	6.2.5	Px	index	0.11	0.83	(2.25 - 0.15)^{NS}	39.12	1.00
Stem thickening	6.2.13	St	7 = thick; 3 = thin	0.45	1.62	(3.00 - 1.00)**	41.38	0.99
Stem lenght SL	6.2.14	Sl	cm	342.00	43.14	(98.00 - 4.00)*	42.73	1.00
Stem width SW	6.2.15	Sw	cm	0.10	1.04	(2.33 - 0.50)^{NS}	30.04	0.99
Stem shape index (SL/SW)	6.2.16	Sx	index	384.56	43.32	(150 - 4.71)^{NS}	45.33	0.99
Stem colour	6.2.17	Sc	5 =red or purple; 1 = light green	0.69	1.62	(5.00 - 1.00)**	50.93	1.00
Stem pubescens	6.2.19	Sp	7 = dense; 0 = absent	4.20	4.37	(8.00 - 0)**	46.98	0.98
Flowering characteristics								
Flowering time (days after sowing)	6.3.1	Ft	n° of days	137.14	53.42	(91.00 - 34.00)**	21.85	1.00
Raceme colour	6.3.3	Frc	6 = purple; 1 = white	1.86	4.53	(6.00 - 2.00)**	30.10	0.99
Flower sepal persistency	6.3.12	Fsp	1 = persistent; 0 = caduceus	-	0.00	(0.00 - 0.00)^{NS}	-	-
Flower sepal swelling	6.3.13	Fss	1 = swollen; 0 = not swollen	0.01	0.99	(1.00 - 0)**	11.70	1.00
Flower sepal hairiness	6.3.14	Fsh	7 = dense; 0 = absent	3.17	4.80	(7.00 - 0)**	37.13	0.99
Flower Petal colour	6.3.15	Fpc	4 = yellow; 1 = white	1.22	2.49	(4.00 - 1.00)**	44.20	1.00
Flower petal veins	6.3.16	Fpv	1 = present; 0 = absent	0.07	0.90	(1.00 - 0)**	29.60	1.00

[a]Number according to the IPGRI descriptors for *Eruca* [b]SE standard error [c]Significance of differences between accessions means: ** $P < 0.001$; * $P < 0.05$; *NS* not significant

Table 2. Analysis of variance, mean, range, coefficient of variation and heritability for quantitative leaf traits in 155 cultivated rocket accessions.

Descriptor	Acronym	Scale	MS[a]	Fratio[b]	Mean	Range	CV%	H^2
Leaf Quality Traits								
Leaf fresh weight	LF	g	3.45	6.90**	2.26	5.23 - 0.56	53.83	0.77
Leaf dry weight	LD	g	0.02	8.74**	0.20	0.5 - 0.05	51.04	0.81
Relative water content	RWC	%	59.43	9.56**	85.96	95.37 - 73.03	5.67	0.82
Osmolarity	Yst	Mpa	0.05	19.52**	-0.71	0.00 - (-1.30)	18.88	0.90
SPAD measure at vegetative stage	SPADv	nmol chl/cm^2	60.53	10.33**	45.93	55.75 - 32.35	10.62	0.83
SPAD measure at flowering stage	SPADf	nmol chl/cm^2	106.96	13.01**	50.34	63.65 - 34.5	12.66	0.86
L*	L*	0-100	15.61	6.03**	41.18	48.37 - 35.54	6.36	0.75
a*	a*	+100;-100	8.88	4.65**	-17.34	-12.93 - (-21.92)	13.52	0.69
b*	b*	+100;-100	28.66	6.52**	22.58	30.4 - 15.12	15.56	0.76
Chroma	Chroma		35.84	5.97**	28.53	37.47 - 20.05	13.92	0.74
Hue	Hue	°degree	6.76	5.67**	128.03	131.11 - 123.73	2.99	0.73
gGAE/100g DW	GAE	g/100g	1.69	8.23**	5.94	8.72 - 3.79	13.99	0.80
Leaf shape traits								
Leaf blade lenght with petiole (in cm)	LBLp	cm	11.80	2.55**	19.94	22.27 - 10.77	13.25	0.55
Leaf blade lenght without petiole (in cm)	LBL	cm	9.76	3.02**	17.44	25.9 - 14.71	13.28	0.60
Leaf blade width (cm)	LBW	cm	9.49	4.33**	9.53	13.41 - 4.75	22.47	0.68
Leaf shape index (Ll-/Lbw)	LS	index	0.61	4.99**	1.92	4.17 - 1.31	27.68	0.71
Petiol lenght	PL	cm	5.01	3.77**	2.49	6.96 - 0.23	63.86	0.65
Petiol width	PEW	cm	0.03	6.32**	0.45	0.76 - 0.27	24.16	0.75
Petiol shape index PL/PW	PS	index	43.50	4.04**	6.17	22.76 - 0.37	75.18	0.66
Leaf blade area without petiole	LBA	cm^2	1613.06	4.25**	84.17	147.51 - 31.72	33.26	0.67
Leaf perimeter	LBP	cm	2498.37	9.93**	91.57	195.15 - 39.78	34.32	0.83

[a]*MS* mean squares, *SE* standard error [b]*Indicate significance of differences between accessions means at $P < 0.001$

Table 3. List of ISSR primers used, nucleotide sequence, alleles per locus, polymorphic index content value (PIC), effective multiple ratio (EMR), resolving power (RP), and marker index (MI) in *Eruca* genotypes.

ISSR Name	Motif	Total loci	Polymorphic loci	PIC	EMR	Rp	MI
UBC-811	$(GA)_8C$	4	4	0.50	4.00	0.96	2.00
UBC-815	$(CT)_8G$	4	4	0.44	4.00	0.66	1.77
UBC-876	$(GATA)_2(GACA)_2$	2	2	0.40	2.00	0.56	0.81
BrgMS75	(GAA)11	2	2	0.03	2.00	0.03	0.05
BrgMS316	(CT)14	5	5	0.37	5.00	0.49	1.85
BrgMS334	(TC)10	7	7	0.14	7.00	0.15	0.98
BrgMS372	(TC)11	5	5	0.12	5.00	0.13	0.60
BrgMS418	(GT)18	3	3	0.02	3.00	0.02	0.06
BrgMS421	(ATG)8	2	2	0.01	2.00	0.01	0.03
BrgMS457	(AT)19	3	3	0.06	3.00	0.06	0.18
BrgMS490	(TA)11	10	10	0.22	10.00	0.26	2.24
BrgMS777	(AT)12	4	4	0.13	4.00	0.15	0.54
BrgMS4533	(AG)9	1	1	0.33	1.00	0.42	0.33
BRMS016	(TC)20	3	3	0.02	3.00	0.02	0.05
sORA26	$(GA)_5$	2	2	0.13	2.00	0.14	0.26

Table 4. Variance analysis and Tukey's for the three identified clusters according to hierarchical cluster analysis.

Y	Cluster 1	Cluster 2	Cluster 3	F Ratio	Prob > F
Seedling and leaf characteristics					
ddg	8.3 ± (1.68)a	7.56 ± (0.71)b	8.35 ± (2.71)ab	4.6105	<0.05
Hc	3.86 ± (1.47)b	4.81 ± (1.31)a	3.92 ± (1.21)ab	9.0976	<0.001
Lc	3.56 ± (0.68)b	3.8 ± (0.6)ab	3.98 ± (0.48)a	4.2501	<0.05
Lm	1.68 ± (0.53)a	1.66 ± (0.45)a	1.73 ± (0.64)a	0.1868	NS
Lg-	2.88 ± (0.24)b	2.99 ± (0.06)a	2.83 ± (0.31)b	10.4421	<0.001
Lg+	1.86 ± (0.37)b	2.31 ± (0.37)a	2.52 ± (0.51)a	28.3263	<0.001
Ll	1.94 ± (0.19)a	1.93 ± (0.27)a	1.07 ± (0.9)b	45.5010	<0.001
La	2.75 ± (0.5)b	3.39 ± (0.56)a	3.28 ± (0.53)a	21.7189	<0.001
Lt	5.38 ± (1.21)a	4.69 ± (1.66)a	4.5 ± (2.06)a	3.5371	<0.05
Lla	2.06 ± (0.38)b	2.37 ± (0.61)a	2.23 ± (0.59)ab	4.8138	<0.001
Lh	1.16 ± (2.06)b	0.87 ± (1.75)b	2.92 ± (2.9)a	9.8299	<0.001
Lv	0.31 ± (0.87)b	1.56 ± (1.82)a	0.58 ± (1.14)b	11.9029	<0.001
Lls	2.1 ± (0.54)b	2.72 ± (0.6)a	1.33 ± (1.13)c	42.1446	<0.001
Pe	1.58 ± (0.5)a	1.48 ± (0.54)a	1.65 ± (0.56)a	1.3694	NS
Pc	2.44 ± (1.24)a	2.88 ± (1.64)a	2.31 ± (1.09)a	2.2521	NS
Vegetative characteristics					
Pg	2.15 ± (2.63)a	1.5 ± (0.29)a	1.43 ± (0.27)a	3.1636	<0.05
Sa	1.73 ± (0.45)a	1.54 ± (0.5)ab	1.43 ± (0.51)b	3.5062	<0.05
St	1.32 ± (0.51)b	1.8 ± (0.67)a	1.59 ± (0.77)ab	8.5168	<0.001
Sc	1.59 ± (0.54)a	1.68 ± (0.95)a	1.5 ± (0.86)a	0.5008	NS
Sp	3.13 ± (1.81)b	4.75 ± (1.94)a	5.5 ± (1.75)a	17.1075	<0.001
Flowering characteristics					
Ft	42.81 ± (5.28)b	57.86 ± (9.4)a	60.48 ± (13.44)a	48.8320	<0.001
Frc	3.4 ± (0.79)c	5.31 ± (1.06)a	4.26 ± (1.48)b	49.0653	<0.001
Fsp	1 ± (0)	1 ± (0)	1 ± (0)	-	-
Fss	1 ± (0)a	1 ± (0)a	0.91 ± (0.29)b	5.8792	
Fsh	3.27 ± (1.03)b	5.68 ± (1.52)a	5 ± (1.76)a	42.7199	<0.001
Fpc	3.6 ± (0.57)a	1.88 ± (0.82)b	2.22 ± (1)b	73.4309	<0.001
Fpv	0.9 ± (0.23)ab	0.94 ± (0.23)a	0.76 ± (0.4)b	4.3162	<0.05
Plant Traits					
PH	33.35 ± (11.5)b	58.87 ± (15.41)a	35.69 ± (14.86)b	57.9645	<0.001
PW	49.19 ± (11.69)b	63.16 ± (11.3)a	58.64 ± (11.63)a	22.1929	<0.001
PX	0.72 ± (0.29)b	0.96 ± (0.3)a	0.63 ± (0.26)b	16.6480	<0.001
SL	30.35 ± (10.3)b	54.52 ± (15.87)a	31.22 ± (14.6)b	53.2824	<0.001
SW	0.9 ± (0.21)b	1.13 ± (0.28)a	1.02 ± (0.46)ab	8.9279	<0.001
SX	34.87 ± (13.36)b	50.48 ± (19.38)a	36.42 ± (22.82)b	12.7353	<0.001
Leaf Quality Traits					
LF	1.43 ± (0.77)b	2.56 ± (1.17)a	2.89 ± (1.22)a	64.67	<0.001
LD	0.12 ± (0.06)b	0.23 ± (0.1)a	0.24 ± (0.1)a	84.45	<0.001

RWC	83.33 ± (4.95)b	86.99 ± (4.52)a	87.74 ± (3.67)a	36.01	<0.001
Yst	-0.71 ± (0.07)ab	-0.72 ± (0.08)b	-0.67 ± (0.27)a	4.69	<0.001
SPADv	44.37 ± (5.16)b	46.39 ± (4.79)a	47.41 ± (3.81)a	12.60	<0.001
SPADf	46.02 ± (5.21)b	52.25 ± (5.22)a	52.86 ± (7.55)a	61.32	<0.001
L*	42.24 ± (2.33)a	40.34 ± (2.26)b	41.85 ± (3.23)a	30.14	<0.001
a*	-17.72 ± (2.05)b	-17.41 ± (1.94)b	-16.42 ± (3.5)a	8.34	<0.001
b*	23.28 ± (3.73)a	22.32 ± (3.22)b	22.09 ± (3.82)b	4.36	<0.05
Chroma	29.27 ± (4.16)a	28.32 ± (3.67)ab	27.81 ± (4.36)b	4.21	<0.05
Hue	127.47 ± (1.8)b	128.09 ± (1.65)ab	128.9 ± (8.52)a	3.63	<0.05
GAE	5.58 ± (0.88)b	6.11 ± (0.74)a	6.08 ± (0.8)a	21.18	<0.001
Leaf shape traits					
LBLp	19.13 ± (2.88)b	20.4 ± (2.41)a	20 ± (2.55)a	10.83	<0.001
LBL	16.92 ± (2.74)b	17.87 ± (2.01)a	17.08 ± (2.12)b	9.16	<0.001
LBW	9.62 ± (2.38)a	10.03 ± (1.82)a	7.79 ± (1.71)b	37.58	<0.001
LS	1.87 ± (0.64)b	1.83 ± (0.33)b	2.31 ± (0.64)a	27.88	<0.001
PL	2.21 ± (1.51)b	2.52 ± (1.6)ab	2.92 ± (1.64)a	5.20	<0.001
PW	0.39 ± (0.09)b	0.48 ± (0.1)a	0.46 ± (0.13)a	30.99	<0.001
PS	6.15 ± (4.62)a	5.85 ± (4.53)a	7.16 ± (4.9)a	2.37	NS
LBA	72.11 ± (25.15)b	94.83 ± (26.8)a	73.24 ± (23.22)b	43.75	<0.001
LBP	102.44 ± (36.8)a	94.09 ± (24.78)b	63.67 ± (21.85)c	48.32	<0.001

* *ns* not significative

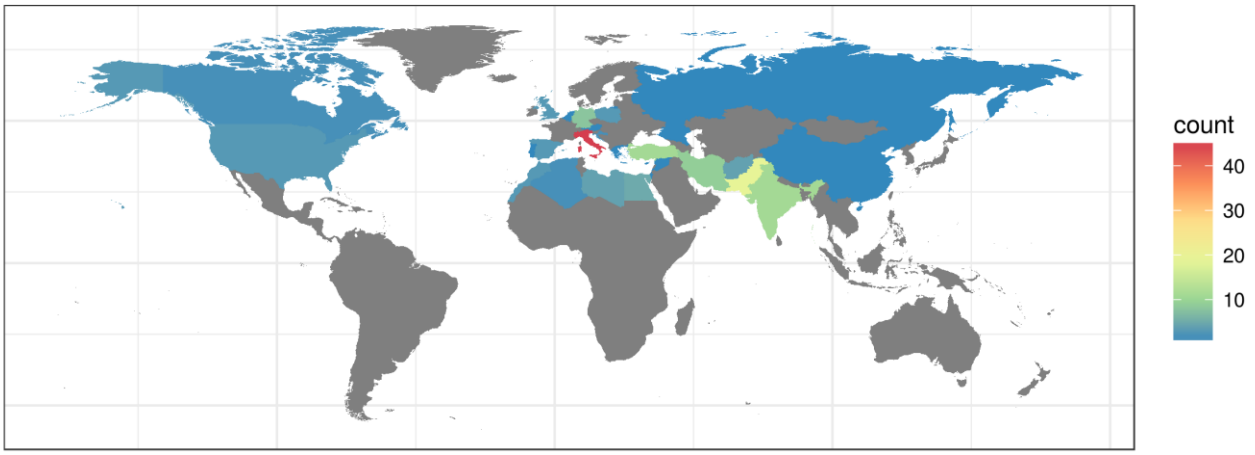

Figure 1: Geographical provenance of the *Eruca* accessions studied. Europe: Austria (1), Belgium (1), Croatia (2), Cyprus (1), Czechoslovakia (3), England (4), Germany (8), Greece (1), Italy (45), Netherlands (1), Poland (3), Portugal (1), Spain (3), Yugoslavia (1). Asia: Afghanistan (4), China (1), India (11), Iran (9), Israel (1), Pakistan (19), Russia (1), Syria (1), Turkey (11). America: Canada (2), United States (3). Africa: Algeria (2), Egypt (5), Libia (4), Morocco (3).

Figure 2: Examples of differences observed in the collection for: a) the morphology of the leaves, b) the morphology of the flower (left) and colour of the petals (right), c) the colour of the hypocotyl, and d) the flowering stem and pubescence of the sepals.

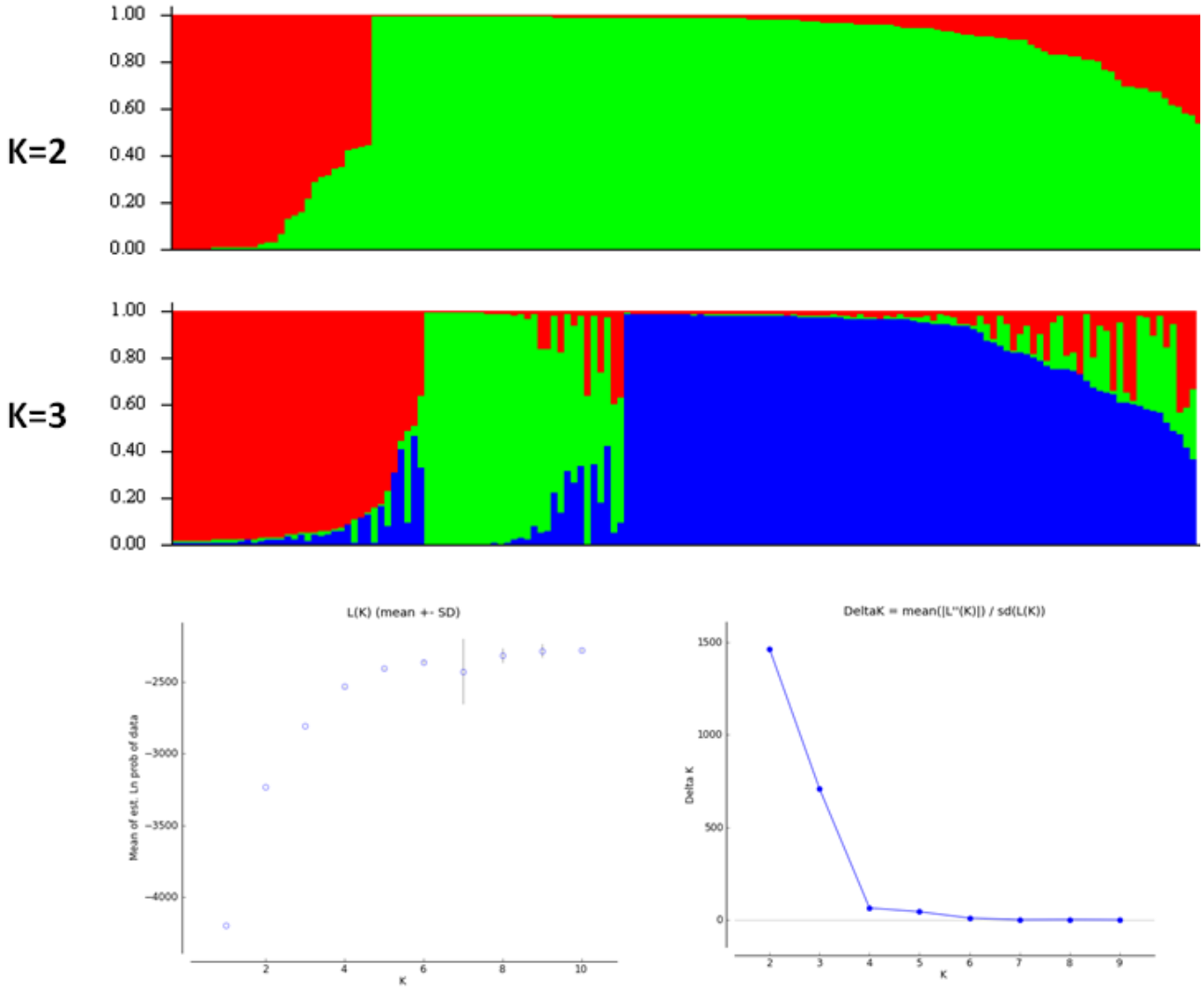

Figure 3: Bar-plot describing the population structure estimated by the Bayesian clustering. Each individual is represented by a thin vertical line, which is partitioned into K coloured segments whose length is proportional to the estimated membership coefficient (q). The population divided into two (K=2) and three (K=3) groups according to the most informative K value using Evanno's method. At K2, the red block corresponds to European accessions with high proportion in Italian genotypes. At K3, European (red and green blocks) and Asian (blue block) accessions are indicated.

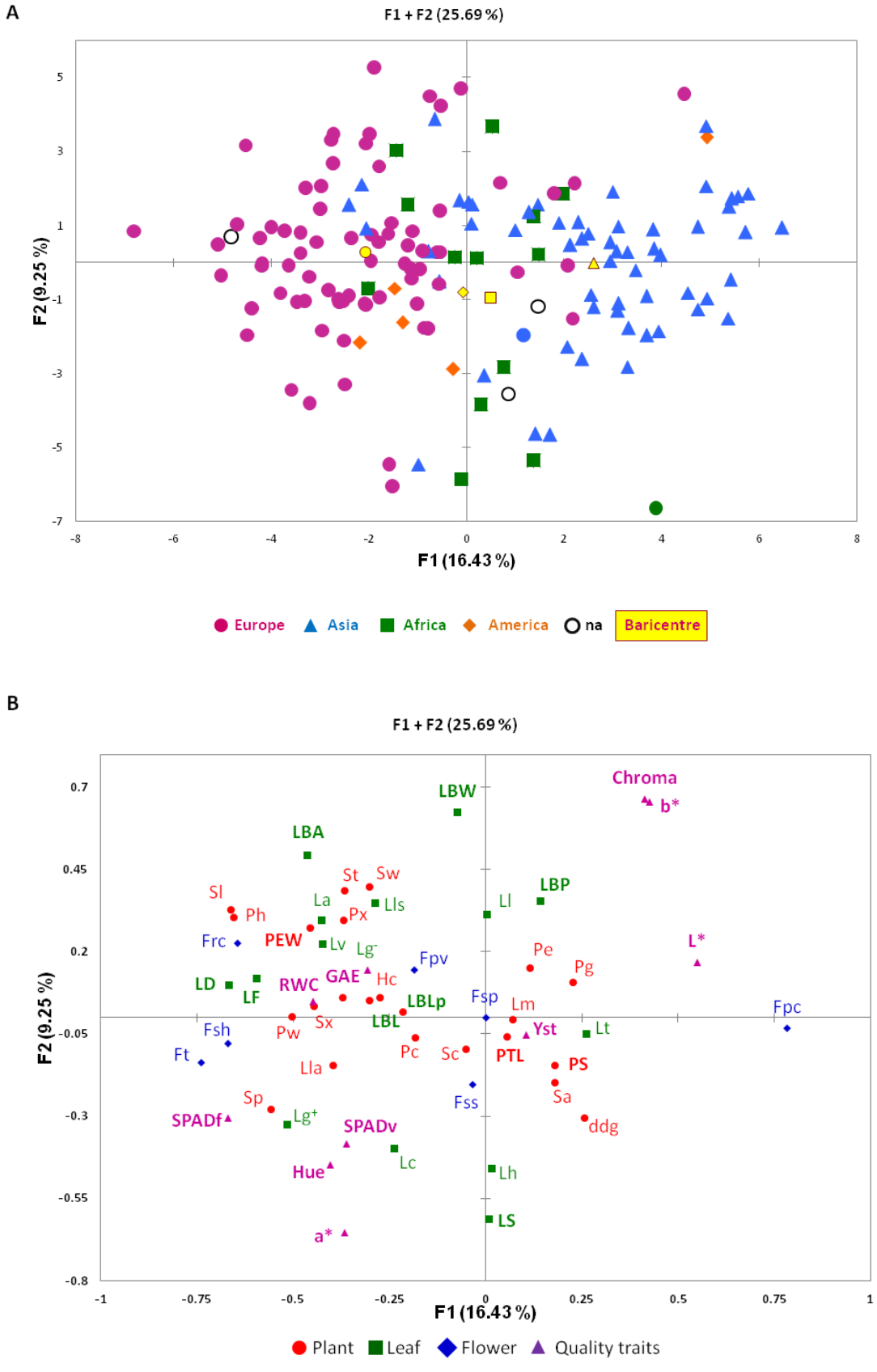

Figure 4: a) Loading plot of the first (X) and second (Y) principal components of 155 cultivated rocket genotypes based on 54 phenotypic traits. Based on provenance, accessions are represented by purple dots (European), by blue triangles (Asian), by green squares (African), and by orange rhombus (American). First and second component centroids for each

continent are indicated in yellow. b) Distribution of the traits scored. The direction and distance from the centre of the biplot indicate how each OTU contributes to the first two components. The different category of traits is indicated using different colour codes as following: a) plant traits with red dots; b) leaf traits with green squares; c) flower traits by blue rhombus; d) qualitative traits by purple triangles. Quantitative traits are in bold, while the remaining are plant descriptors.

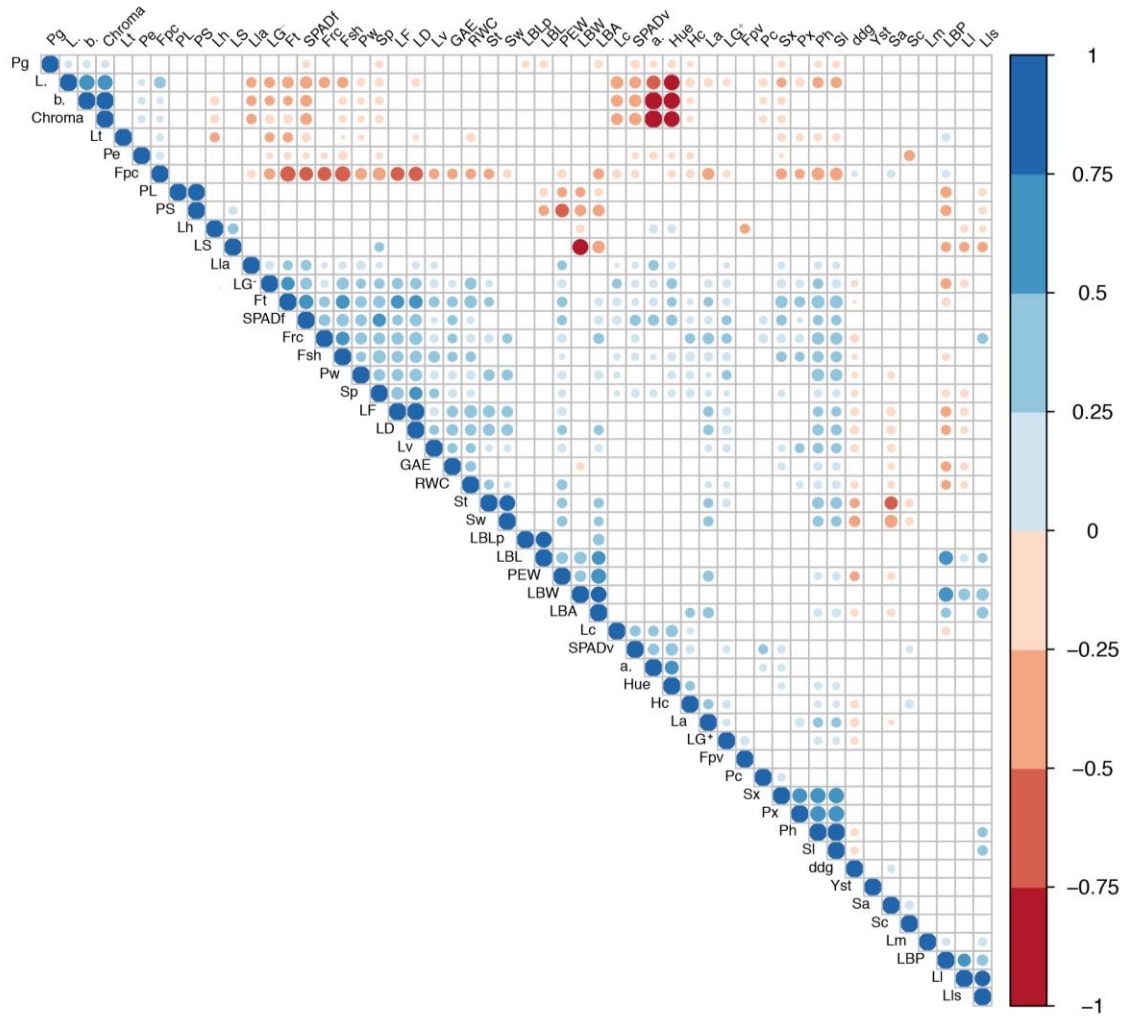

Figure 5: Pearson's rank correlation coefficients between pairs of phenotypes. Only correlation coefficients with *P* value < 0.01 after Bonferroni correction are shown. Colour intensity is proportional to the correlation coefficients. On the right side of the correlogram, the legend colour shows the correlation coefficients and the corresponding colours.

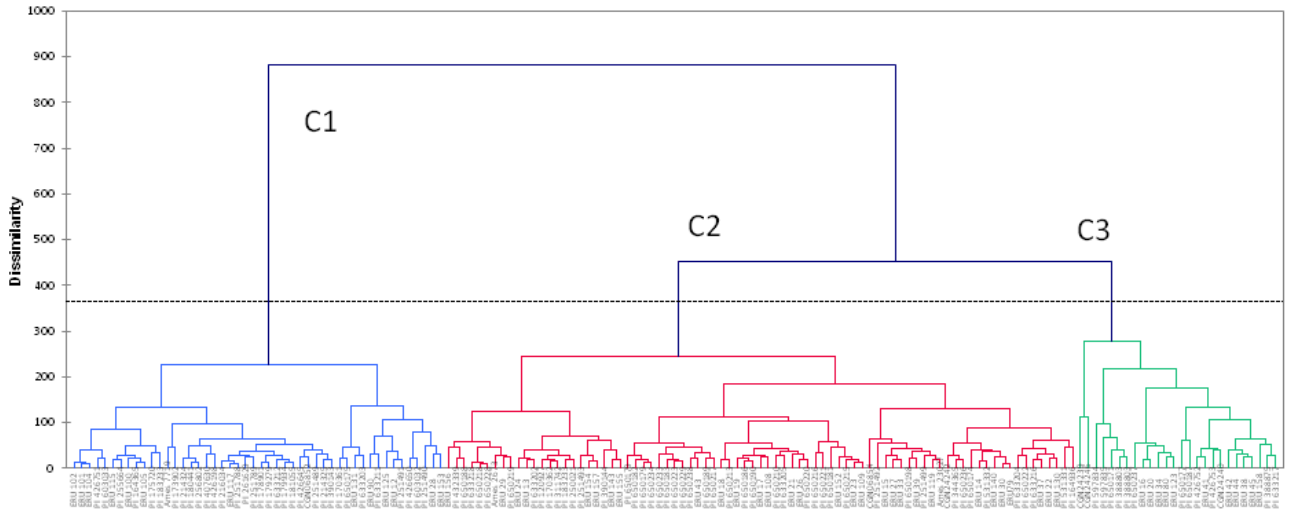

Figure 6: Cluster analysis (Ward coefficient) based on plant descriptors and quantitative traits for the 155 *Eruca sativa* genotypes evaluated in the present study.