SCIENTIA AGRICOLA

Phenotypic characteristics and genetic diversity of new broccoli (*Brassica oleracea* var. *italica* Plenck) varieties in China

Yiying Zhang[1] , Li Ren[1] , Ruixi Han[2] , Aiai Li[2] , Yu Wang[2] , Ting Lin[3] , Yu Zhang[1] , Hong Zhao[1] , Hairong Chen[1]*

[1]Shanghai Academy of Agricultural Sciences – Institute for Agri-food Standards and Testing Technology, 1000 Jinqi Road – 201403 – Shanghai – China.

[2]Ministry of Agriculture and Rural Affairs – Development Center of Science and Technology, 18 Ronghua South Road – 100176 – Beijing – China.

[3]Jiangxi Agricultural University, 1101 Zhimin Road – 330045 – Nanchang – China.

*Corresponding author: <sh57460009@163.com>

ABSTRACT: Due to its high nutritional value, broccoli (*Brassica oleracea* var. *italica* Plenck) is one of the most popular vegetables worldwide. This study assessed 36 phenotypic characteristics of 111 broccoli varieties to understand the phenotypic diversity of new broccoli varieties and improve their breeding speed with advantages and characteristics in China, including 108 new varieties and three varieties of common knowledge. The genetic diversity, the principal component, and the cluster of phenotypic characteristics of broccoli varieties were further investigated. The results showed that the coefficients of variation of 36 characteristics ranged between 11.18 % and 94.99 %, with their diversity index between 0.26 and 1.82. The 111 broccoli varieties were further classified into eight groups, primarily attributed to the differences in phenotypic characteristics, including curd weight, main stem thickness, plant development degree, plant height, and anthocyanin coloration. The cumulative contribution rate of the first five principal components reached 81.186 %, corresponding to 12 representative phenotypic traits. The analysis indicated that the phenotypic characteristics of broccoli were rich in diversity, especially for several characteristics appreciated by the market, such as weight, curd firmness, and anthocyanin coloration. This study revealed the basic information on the genetic diversity of new broccoli varieties in China from 2017 to 2019 and provided potential breeding strategies for broccoli to meet diverse market demands.

Keywords: phenotypic variability, vegetable quality, plant breeding, yield

Introduction

Broccoli (*Brassica oleracea* var. *italica* Plenck) has become an essential vegetable in China due to its edibility in autumn and winter. Broccoli is rich in vitamins, minerals, and other bioactive compounds, such as glucoraphanin and dietary antioxidants (Eberhardt et al., 2005; Velasco et al., 2007; Soares et al., 2017). Vegetable quality is a comprehensive trait mainly affected by phenotype and multiple internal nutritional elements. As broccoli's curd and tender stems are the main edible part, several phenotypic characteristics are essential factors affecting the broccoli quality, such as the color, weight, shape, curd size, and stem thickness.

A rich genetic diversity of broccoli germplasm resources is the foundation of germplasm innovation and genetic improvement and provides an excellent gene reserve for cultivating new varieties. The analysis of genetic diversity and the evaluation of broccoli germplasm resources are mainly based on morphology, chemical compounds, or molecular markers (Hu and Quiros, 1991; Louarn et al., 2007; Li et al., 2019; Liu et al., 2022b). Conventional phenotypic diversity is the basis and essential content to reveal biodiversity. Phenotypic characteristics can reflect the plant level of genetic variation and diversity in the short term and provide reliable phenotypic data for related studies. Therefore, the description and the evaluation of germplasm resources still rely mainly on phenotypic characteristics. Previous studies found that broccoli cultivars, which have been widespread in China since 1980, had close

relationships and similar genetic diversity (Li et al., 2019). The 165 cauliflower (*Brassica oleracea* var. *botrytis*) inbred lines (primarily from southeast China) presented relatively narrow genetic diversity assessed by the 43 SSR markers (Zhu et al., 2018). However, the genetic diversity level of the new broccoli variety in recent years has yet to be well-documented.

In this study, we collected 108 new varieties of broccoli from 2017 to 2019, representing the breeding advances of broccoli varieties in China. Among them, 63 new broccoli varieties were obtained from the national consortium for seed research, which aims to improve the breeding speed of new varieties with advantages and characteristics. According to the new varieties test guideline, we selected 36 phenotypic characteristics of broccoli and analyzed the phenotypic genetic diversity and the principal components. The genetic characteristics of these new broccoli varieties were obtained from the analysis results, which could provide helpful information for further research to improve new broccoli varieties.

Materials and Methods

Plant materials and field experiments

Seeds of 111 broccoli varieties were collected from 2017 to 2019, composed of 108 new varieties from seven provinces of the Chinese mainland and three varieties of common knowledge from Japan and Chinese Taiwan, including hybrid, the three-sex sterile line, and male-

sterile hybrid. Field experiments were conducted in Shanghai, China (30°98′ N, 121°36′ E, altitude 2.8 m) in 2019. The experiment was conducted in a randomized complete block design replicated three times. Each variety was sown with at least 120 plants in August, where the experiment unit comprised 18 m by 2.31 m with three rows. The row spacing was 55 cm, while the plant spacing was 50 cm. The soils in the experiment site were paddy soils. All trials were managed following standard agricultural practices. The bottom fertilizer was 52500 kg ha^{-1} for organic fertilizer and 375 kg ha^{-1} for compound fertilizer. The treatments were irrigated 2-3 times after field planting. After 15 days of planting, 225 kg ha^{-1} of urea was applied and 450 kg ha^{-1} of compound fertilizer was applied at about 45 days after planting.

Measurement of phenotypic traits and data collection

Thirty-six characteristics were selected from "Guidelines for the conduct of tests for distinctness, uniformity and stability—Broccoli (*Brassica oleracea* var. *italica* Plenck)" (Jian et al., 2015; UPOV, 2016; UPOV, 2018). We assigned criteria for 26 qualitative and pseudo-qualitative characteristics (Table 1). Besides, ten quantitative characteristics, including the degree of plant development, plant height, leaf number, leaf length, leaf width, petiole length, main stem thickness, flower stem length, curd weight, and curd thickness were measured from 20 spikes chosen randomly at each plot. The numerical data for the ten quantitative characteristics were transferred to ordinal scales of 1-9. The specific measurements are as follows: degree of plant development (cm) was the widest plant part; plant height (cm) was measured from ground level to the leaf tip; leaf number was the number of fully-developed leaves in each plant; leaf length (cm) and width (cm) was the longest and widest part of a leaf; petiole length (cm) was measured from one petiole end to the other; main stem thickness (mm) was the longest thickness of main stem by vernier caliper; flower stem length (cm) was measured as the longest branch of curd by vernier caliper; curd weight (g) was measured by electronic balance; curd thickness (cm) was measured as the distance of the widest part of curd.

Statistical analysis

The mean, standard deviation, coefficient of variation, minimum (min), maximum (max), and range for the 36 phenotypic characteristics were calculated. The distribution frequency of 36 characteristics on assignment was analyzed. The numerical values of ten quantitative characteristics were divided into nine intervals, and the distribution frequency of every quantitative characteristic was performed. The Pearson's correlation analysis (Fisher, 1924) was conducted to calculate the relationship among ten quantitative characteristics using SPSS Statistics 23.0 software. The Shannon-Weaver's diversity index ($H′$) was used to evaluate genetic diversity and the calculation formula is $H′ = -\sum P_i \times \ln P_i$ (i = 1, 2, 3, ..., n), where P_i represents the frequency of a specific characteristic in the i-th level (Shannon and Weaver, 1949). The Euclidean distance was used as the distance between germplasms and the UPGMA (Unweighted Pair Group Method with Arithmetic mean) method was used for the cluster analysis (Nei and Li, 1979). The principal component analysis (PCA) (Sneath and Sokal, 1973) was performed on 36 phenotypic characteristics using SPSS Statistics 23.0 software.

Results

Frequency distributions of phenotypic characteristics

The frequency distributions of 36 phenotypic characteristics from 111 broccoli varieties are shown in Table 2. It was found that different phenotypic characteristics had various frequency distributions. Anthocyanin coloration of the hypocotyl (ACH), plant collateral curd (PCC), leaf anthocyanin coloration (LAC), and male sterility (MS) had only two types of expression, including states absent or present. The absent and present state ratios for ACH, PCC, and LAC were approximately 1:1, while MS was about 1:3. Leaf shape was mainly narrow oval and medium oval, accounting for 54.1 % and 45.0 %, respectively. Over half of the broccoli varieties (62.2 %) had no curd anthocyanins. Curd bud uniformity of 74.8 % of broccoli varieties was uniform. Only curd bulb color was light green or medium green, which was approximately 1:1. The maturity was mainly distributed in 115-145 days and 160-190 days, accounting for 53.1 % and 43.2 %, respectively. The degree of plant development, leaf length, leaf width, and curd thickness were distributed within a narrower range than the other quantitively characteristics. The curd thickness of 89.2 % of varieties was distributed in 60.0-79.9 cm and the degree of plant development of 85.6 % was distributed in 12.0-15.9 cm. The leaf length of 82.9 % of varieties was distributed in 30.0-39.9 cm and the leaf width of 90.1 % was distributed in 15.0-19.9 cm.

Diversity of phenotypic characteristics

The mean, standard deviation, coefficient of variation, and diversity index ($H′$) of 36 phenotypic characteristics are shown in Table 3. The coefficient of variation reflects the data discrete and the data variation. A larger coefficient of variation indicates a greater degree of variation. Curd anthocyanin coloration, leaf anthocyanin coloration, and anthocyanin coloration of hypocotyl showed significant variation with coefficients of 94.99 %, 90.50 %, and 89.28 %, respectively. The leaf

Table 1 – Value assignment criteria of 36 phenotypic characteristics assessed in broccoli (*Brassica oleracea* var. *italica*) accessions.

No.	Characteristics	Value assignment criteria								
		1	2	3	4	5	6	7	8	9
1	Anthocyanin coloration of hypocotyl (ACH)	Absent	-	-	-	-	-	-	-	Present
2	Plant collateral curd (PCC)	Absent	-	-	-	-	-	-	-	Present
3	Degree of occurrence of collateral curd (DOCC)	Very weak	Very weak to weak	Weak	Weak to medium	Medium	Medium to strong	Strong	Strong to very strong	Very strong
4	Leaf attitude (LA)	Erect	Erect to semi-erect	Semi-erect	Semi-erect to horizontal	Horizontal	-	-	-	-
5	Leaf foliar wax powder (LFWP)	Very weak	Very weak to weak	Weak	Weak to medium	Medium	Medium to strong	Strong	Strong to very strong	Very strong
6	Leaf color (LC)	Green	Grey green	Blue green	-	-	-	-	-	-
7	Leaf anthocyanin coloration (LAC)	Absent	-	-	-	-	-	-	-	Present
8	Leaf shape (LS)	Narrow oval	Medium oval	Wide oval	-	-	-	-	-	-
9	Undulation of leaf margin (ULM)	Absent or very weak	Weak	Medium	Strong	-	-	-	-	-
10	Leaf edge dentate notch (LEDN)	Absent or very weak	Weak	Strong	-	-	-	-	-	-
11	Leaf apex shape (LAS)	Acute	Obtuse	Rounded	Truncate	Obcordate	-	-	-	-
12	Number of leaf lobes (NLL)	Absent or very few	Very few to few	Few	Few to medium	Medium	Medium to many	Many	Many to very many	Very many
13	Petiole cross-sectional shape (PCS)	Flat	Medium	Circular	-	-	-	-	-	-
14	Curd covered by inner leaves (CCIL)	Absent	Few	Medium	Many	-	-	-	-	-
15	Curd color (CC)	Yellow green	Green	Grey green	Blue green	Purple	-	-	-	-
16	Degree of curd color (DCC)	Very light	Very light to light	Light	Light to medium	Medium	Medium to dark	Dark	Dark to very dark	Very dark
17	Curd anthocyanin coloration (CAC)	Absent or very weak	Very weak to weak	Weak	Weak to medium	Medium	Medium to strong	Strong	Strong to	Very strong
18	Curd convex surface (CCS)	Very weak	Very weak to weak	Weak	Weak to medium	Medium	Medium to strong	Strong	Strong to very strong	Very strong
19	Curd bud size (CBS)	Very small	Very small to small	Small	Small to medium	Medium	Medium to large	Large	Large to very large	Very large
20	Curd bud uniformity (CBU)	Nonuniform	Uniform	-	-	-	-	-	-	-
21	Curd shape in longitudinal section (CSLS)	Circular	Transverse broad elliptic	Transverse medium elliptic	Transverse narrow elliptic	Triangular	-	-	-	-
22	Curd dome shape (CDS)	Truncate	Raised	Obtuse	-	-	-	-	-	-
23	Curd firmness (CF)	Very loose	Very loose to loose	Loose	Loose to medium	Medium	Medium to tight	Tight	Tight to very tight	Very tight
24	Curd bulb color (CBC)	Yellow green	Light green	Medium green	Green with purple strip	-	-	-	-	-
25	Maturity (M)	Very early	Very early to early	Early	Early to medium	Medium	Medium to late	Late	Late to very late	Very late
26	Male sterility (MS)	Absent	-	-	-	-	-	-	-	Present
27	Degree of plant development DPD (cm)	0-29.9	30.0-39.9	40.0-49.9	50.0-59.9	60.0-69.9	70.0-79.9	80.0-89.9	90.0-99.9	100.0-
28	Plant height PH (cm)	0-24.9	25.0-29.9	30.0-34.9	35.0-39.9	40.0-44.9	45.0-49.9	50.0-54.9	55.0-59.9	60.0-
29	Leaf number (LN)	0-11	12	13	14	15	16	17	18	20
30	Leaf length LL(cm)	0-14.9	15.0-19.9	20.0-24.9	25.0-29.9	30.0-34.9	35.0-39.9	40.0-44.9	45.0-49.9	50.0-
31	Leaf width LW (cm)	0-9.9	10.0-12.4	12.5-14.9	15.0-17.4	17.5-19.9	20.0-22.4	22.5-24.9	25.0-27.4	27.5-
32	Petiole length PL (cm)	0-7.9	8.0-9.9	10.0-11.9	12.0-13.9	14.0-15.9	16.0-17.9	18.0-19.9	20.0-21.9	22.0-
33	Main stem thickness MST (mm)	0-29.9	30.0-32.9	33.0-35.9	36.0-38.9	39.0-41.9	42.0-44.9	45.0-47.9	48.0-50.9	51.0-
34	Flower stem length FSL (cm)	0-0.9	1.0-1.9	2.0-2.9	3.0-3.9	4.0-4.9	5.0-5.9	6.0-6.9	7.0-7.9	8.0-
35	Curd weight CW (g)	0-49.9	50.0-99.9	100.0-149.9	150.0-199.9	200.0-249.9	250.0-299.9	300.0-349.9	350.0-399.9	400.0-
36	Curd diameter CD (cm)	0-7.9	8.0-9.9	10.0-11.9	12.0-13.9	14.0-15.9	16.0-17.9	18.0-19.9	20.0-21.9	22.0-

Table 2 – The distribution frequency of the different expression states of the 36 phenotypic characteristics assessed in 111 broccoli varieties.

Characteristics	Frequency of distribution (%)								
	1	2	3	4	5	6	7	8	9
Anthocyanin coloration of hypocotyl (ACH)	56.8	-	-	-	-	-	-	-	43.2
Plant collateral curd (PCD)	45	-	-	-	-	-	-	-	55
Degree of occurrence of collateral curd (DOCC)	3.3	18	34.4	3.3	16.4	6.6	13.1	3.3	1.6
Leaf attitude (LA)	0.9	12.6	33.3	33.3	18	-	1.8		-
Leaf foliar wax powder (LFWP)	-	-	-	1.8	2.7	9	60.3	26.1	-
Leaf color (LC)	3.6	80.2	16.2	-	-	-	-	-	-
Leaf anthocyanin coloration (LAC)	57.7	-	-	-	-	-	-	-	42.3
Leaf shape (LS)	54.1	45	0.9	-	-	-	-	-	-
Undulation of leaf margin (ULM)	-	10.8	75.7	13.5	-	-	-	-	-
Leaf edge dentate notch (LEDN)	84.7	14.4	0.9	-	-	-	-	-	-
Leaf apex shape (LAS)	-	2.7	81.1	15.3	0.9	-	-	-	-
Number of leaf lobes (NLL)	-	1.8	26.1	33.3	32.4	6.3	-	-	-
Petiole cross-sectional shape (PCS)	50.5	48.6	0.9	-	-	-	-	-	-
curd covered by inner leaves (CCIL)	92.8	7.2	-	-	-	-	-	-	-
Curd color (CC)	0.9	30.6	59.5	9	-	-	-	-	-
Degree of curd color (DCC)	-	-	5.4	53.2	35.1	2.7	3.6	-	-
Curd anthocyanin coloration (CAC)	62.2	5.4	4.5	9.9	12.6	3.6	1.8	-	-
Curd convex surface (CCS)	0.9	2.7	40.5	24.3	27.9	2.7	0.9	-	-
Curd bud size (CBS)	1.8	18	17.1	31.5	19.8	5.4	6.3	-	-
Curd bud uniformity (CBU)	25.2	74.8	-	-	-	-	-	-	-
Curd shape in longitudinal section (CSLS)	22.5	53.2	14.4	9.9	-	-	-	-	-
Curd dome shape (CDS)	10.8	86.5	2.7	-	-	-	-	-	-
Curd firmness (CF)	-	1.8	8.1	7.2	32.4	23.4	17.1	8.1	1.8
Curd bulb color (CBC)	-	46.8	53.2	-	-	-	-	-	-
Maturity (M)	-	2.7	20.7	32.4	0.9	20.7	22.5		
Male sterility (MS)	23.4	-	-	-	-	-	-	-	76.6
Degree of plant development DPD (cm)	-	-	0.9	8.1	56.8	28.8	5.4	-	-
Plant height PH (cm)	-	1.8	27.0	37.8	30.6	1.8	0.9	-	-
Leaf number (LN)	1.8	20.7	27.0	26.1	13.5	9.0	0.9	0.9	-
Leaf length LL (cm)	-	-	-	4.5	44.1	38.7	11.7	0.9	-
Leaf width LW (cm)	-	-	3.6	55.0	35.1	6.3	-	-	-
Petiole length PL (cm)	-	0.9	7.2	36.9	32.4	18.0	3.6	0.9	-
Main stem thickness MST (mm)	1.8	0.9	7.2	9.9	37.8	23.4	9.9	5.4	3.6
Flower stem length FSL (cm)	-	0.9	58.6	29.7	5.4	4.5	0.9	-	-
Curd weight CW (g)	-	0.9	1.8	6.3	9.9	35.1	28.8	10.8	6.3
Curd diameter CD (cm)	0.9	-	7.2	51.4	37.8	0.9	0.9	-	0.9

characteristics generally had a small degree of variation, including leaf foliar wax powder, leaf apex shape, and degree of plant development, with coefficients of 11.18 %, 14.08 %, and 13.85 %, respectively. The yield-related characteristics showed a wide range of variability. For instance, curd weight and diameter ranged from 73.72 g to 524.57 g and 7.63 cm to 25.83 cm, respectively. Significant variability was also observed for flower stem length, ranging from 1.69 cm to 6.62 cm.

The Shannon-Weaver diversity index reflects the degree of variety and the non-uniformity of individual distribution among interspecific crossings. A higher diversity index value corresponds to a more uniform distribution of individual characteristics from varieties. The diversity indices of the degree of occurrence of collateral curd, curd bud size, curd firmness, leaf number, main stem thickness, and curd weight were greater than 1.5, which were 1.82, 1.70, 1.75, 1.67, 1.75, and 1.66, respectively. Curd covered by inner leaves and curd dome shape showed characteristics with a low diversity index at 0.26 and 0.46, respectively. The diversity index of 16 characteristics related to curd varied from 0.26 to 1.75, with an average value of 1.11. The diversity index of 13 characteristics related to leaf varied from 0.46 to 1.67, with an average value of 0.93. The genetic diversity of the curd parts was greater than that of the leaf parts.

Correlation of phenotypic characteristics

Correlation coefficients among the 36 phenotypic characteristics were analyzed (data not shown). Five characteristics were significantly correlated with more than 20 characteristics, including curd firmness,

Table 3 – Statistical analysis of the 36 phenotypic characteristics diversity of 111 broccoli varieties.

Characteristics	Mean	Standard deviation	Coefficient of variation	Max	Min	H'
Anthocyanin coloration of hypocotyl (ACH)	4.46	3.98	89.28	9	1	0.68
Plant collateral curd (PCD)	5.40	3.98	73.68	9	1	0.69
Degree of occurrence of collateral curd (DOCC)	4.10	1.99	48.54	9	1	1.82
Leaf attitude (LA)	3.62	1.06	29.33	7	1	1.42
Leaf foliar wax powder (LFWP)	7.06	0.79	11.18	8	4	1.04
Leaf color (LC)	2.13	0.43	20.02	3	1	0.59
Leaf anthocyanin coloration (LAC)	4.39	3.97	90.5	9	1	0.68
Leaf shape (LS)	1.47	0.52	35.35	3	1	0.73
Undulation of leaf margin (ULM)	3.03	0.49	16.34	4	2	0.72
Leaf edge dentate notch (LEDN)	1.16	0.39	33.91	3	1	0.46
Leaf apex shape (LAS)	3.15	0.44	14.08	5	2	0.60
Number of leaf lobes (NLL)	4.15	0.95	22.77	6	2	1.33
Petiole cross-sectional shape (PCS)	1.50	0.52	34.57	3	1	0.74
curd covered by inner leaves (CCIL)	1.07	0.26	24.23	2	1	0.26
Curd color (CC)	2.77	0.62	22.32	4	1	0.93
Degree of curd color (DCC)	4.46	0.80	17.83	7	3	1.08
Curd anthocyanin coloration (CAC)	1.88	1.79	94.99	7	1	1.27
Curd convex surface (CCS)	3.87	1.03	26.54	7	1	1.35
Curd bud size (CBS)	3.91	1.43	36.59	7	1	1.70
Curd bud uniformity (CBU)	1.75	0.44	24.96	2	1	0.56
Curd shape in longitudinal section (CSLS)	2.12	0.87	41.15	4	1	1.18
Curd dome shape (CDS)	1.92	0.36	18.77	3	1	0.46
Curd firmness (CF)	5.60	1.47	26.17	9	2	1.75
Curd bulb color (CBC)	2.53	0.50	19.8	3	2	0.69
Maturity (M)	4.84	1.58	32.69	7	2	1.49
Male sterility (MS)	7.13	3.40	47.76	9	1	0.54
Degree of plant development DPD (cm)	5.30	0.73	13.85	7	3	1.08
Plant height PH (cm)	3.96	1.07	26.93	7	2	1.27
Leaf number (LN)	3.55	1.46	41.07	8	1	1.67
Leaf length LL(cm)	5.52	1.09	19.82	8	4	1.16
Leaf width LW (cm)	4.37	0.89	20.35	6	3	0.99
Petiole length PL (cm)	4.65	1.24	26.54	8	2	1.44
Main stem thickness MST (mm)	5.35	1.72	32.11	9	1	1.75
Flower stem length FSL (cm)	3.57	1.03	28.95	7	2	1.06
Curd weight CW (g)	6.26	1.61	25.77	9	2	1.66
Curd diameter CD (cm)	4.31	1.16	26.84	9	1	1.07

petiole length, main stem thickness, flower stem length, and curd weight. However, the undulation of leaf margin, leaf apex shape, number of leaf lobes, and curd covered by inner leaves only had significant correlations with less than five characteristics. Furthermore, the most significant positive correlations were found for anthocyanin coloration of hypocotyl, leaf, and curd. Flower stem length was significantly negatively correlated with 17 characteristics in this study. As the flower stem length increases, the shape of the curd tends to become a transverse narrow elliptic, significantly affecting yield. Maturity and male sterility were all significantly positively correlated with the degree of plant development, main stem thickness, and curd weight. Curd weight had significant correlation coefficients with most quantitative characteristics except leaf number. The results suggested that plant growth affects yield-related characteristics.

Cluster analysis

Based on the data of 36 phenotypic characteristics, 111 varieties of broccoli were classified into eight groups (Figure 1). Pictures of typical broccoli varieties in eight groups are also shown in Figure 2. Group I included three varieties (BO24, BO79, and BO83), whose degree of plant development, leaf length, main stem thickness, flower stem length, and the weight and diameter of curd were significantly smaller than that in the other varieties. For example, curd weight of the three varieties was 200-250 g. Group II contained 19 varieties with no collateral curd, anthocyanin coloration of hypocotyl, leaf, and curd, and male sterility. Group III contained seven varieties, such as BO49, BO104, BO23, and BO54, etc., whose quantitative characteristics were generally less than average. Group IV, composed of 16 varieties (BO12, BO68, BO90, BO110, etc.), had collateral curd and anthocyanin coloration

of hypocotyl, leaf, and curd. Group V contained four varieties (BO2, BO4, BO40, BO84) with collateral curd and anthocyanin coloration of the leaf. Group VI only
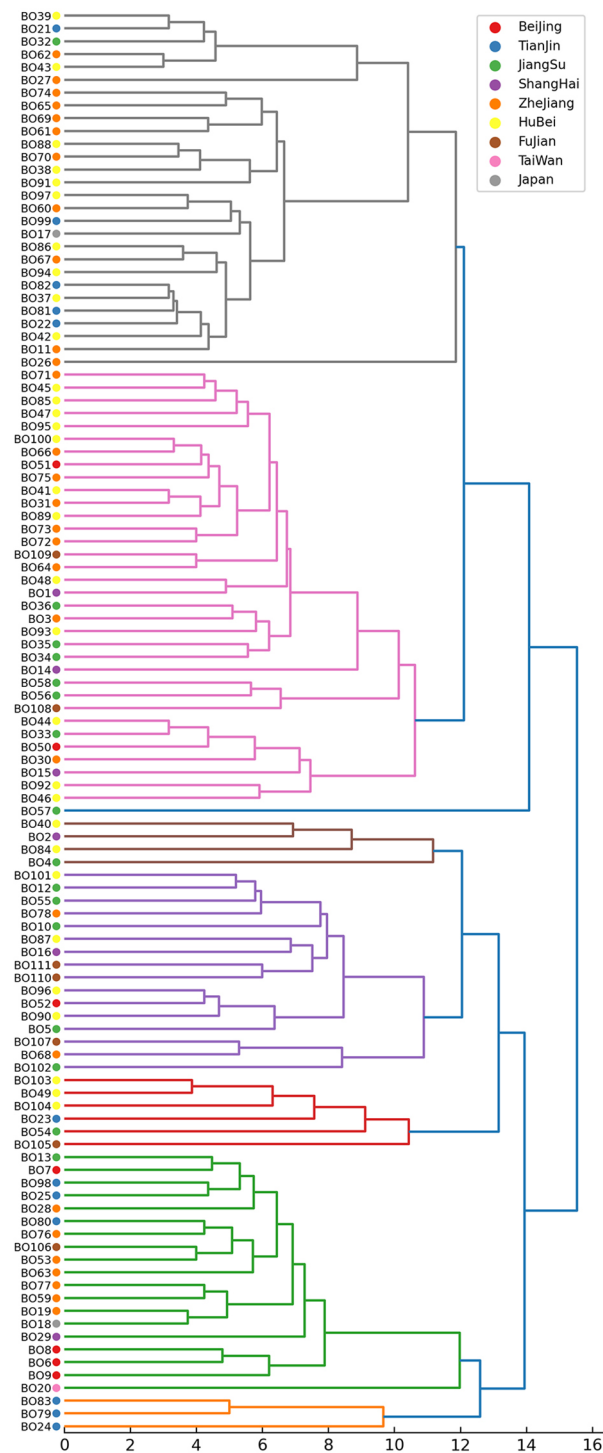
had one variety (BO57), which was more significant than the overall average on quantitative characteristics. Group VII included 34 varieties with an intermediate level of quantitative characteristics (BO1, BO34, BO51, BO89, etc.). Group VIII had no anthocyanin coloration of hypocotyl, leaf, curd, and no collateral curd covered by inner leaves.
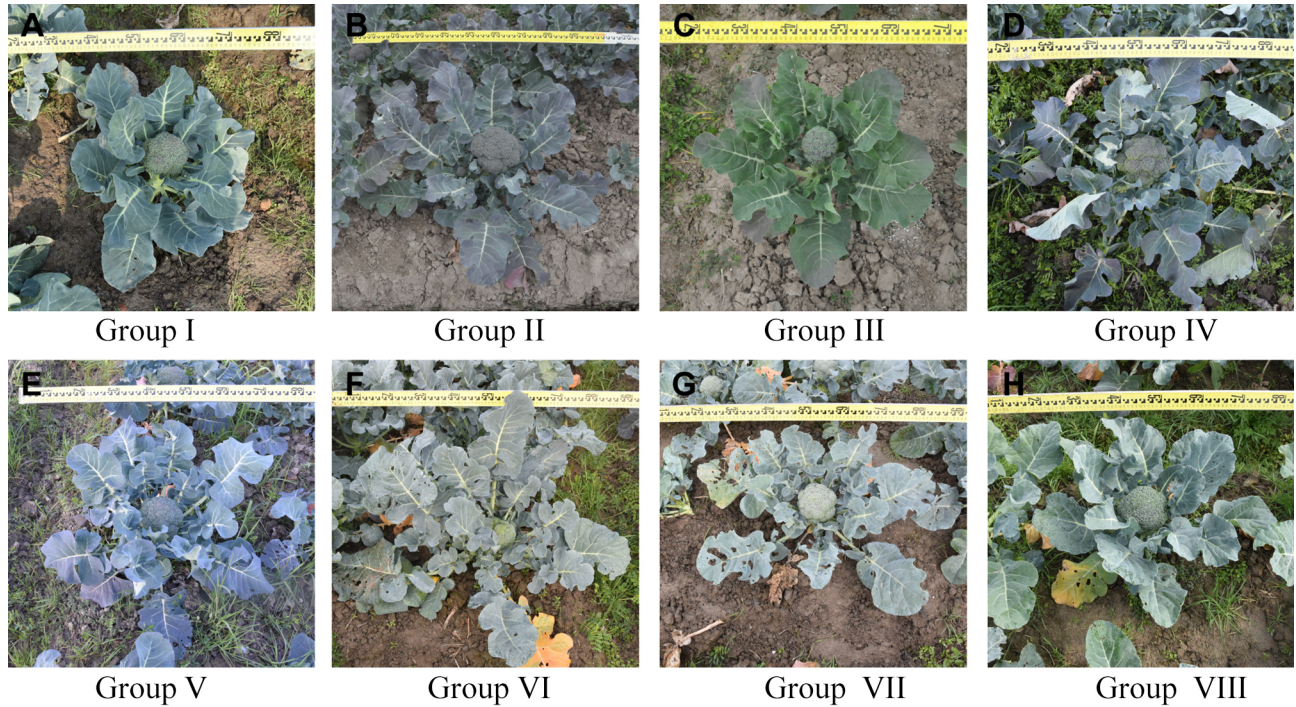
Principal component analysis

The principal component analysis (PCA) can obtain the most critical information because it is one of the methods to analyze multi-factor complex problems using the dimension reduction idea effectively. Each group selected the two most significant values of the absolute value among the eigenvector to reflect the contribution rate of its corresponding numerical trait. The results showed that the characteristic values of the five principal components were higher than 1. At the same time, the cumulative contribution rate reached 81.186 %, which could reflect the basic information of 36 phenotypic traits of broccoli (Table 4). The first principal component (PC1) accounted for 34.362 % of the total variance, primarily attributed to the anthocyanin coloration of hypocotyl, leaf, and curd. The second principal component (PC2) accounted for 22.873 % and was attributed primarily to collateral curd and its degree of occurrence. The third principal component (PC3) was represented by male sterility, responsible for an additional 13.672 % of the total variance. The fourth principal component (PC4) accounted for 5.875 % of the total variance and was attributed primarily to maturity, main stem thickness, curd firmness, and curd weight. The fifth principal component (PC5) mainly reflected curd bud size, leaf number, and anthocyanin coloration of hypocotyl and leaf, accounting for 4.404 %.

Projecting the accessions onto the first two PCs showed apparent separations corresponding well to the eight clustered groups (Figure 3). The varieties of Group VII were mainly distributed on the upper left of the origin, corresponding to hypocotyl, leaf, and curd have no anthocyanin coloration. Group V and Group VI were distributed above the origin, while the main stem thickness of their corresponding varieties was higher than the other five groups. Group III and Group IV were mainly distributed on the upper right of the origin, reflecting the characteristics of the lightest curd weight. Group VIII was mainly distributed on the lower left of the origin, reflecting heavy curd weight characteristics. Group I was mainly distributed below the origin with lower plant height. The germplasm resources of Group II were mainly distributed on the lower right of the origin, which had anthocyanin coloration on hypocotyl, leaf, and curd.

Figure 1 – Cluster dendrogram based on 36 phenotypic characteristics from 111 broccoli varieties. The different colors of the cluster dendrogram represent different groups, and the color dots on the left indicate the source of the corresponding resource.

Discussion

Information on many new varieties can reflect the situation of variety breeding in a region. Therefore, it

Figure 2 – Typical broccoli varieties of eight cluster groups.
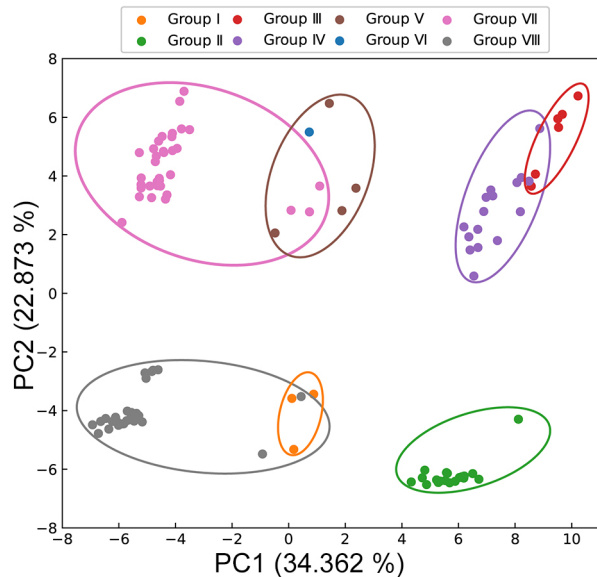
Figure 3 – Principal coordinate two-dimensional plot of 111 broccoli varieties. Different colors correspond to the different groups in Figure 1. PC1 = The first principal component; PC2 = The second principal component.

is necessary to assess the new indigenous varieties. However, relevant reports have yet to be on the phenotypic diversity analysis of new broccoli varieties based on large samples in China. The 111 experimental materials used in this study contained much information regarding the germplasm resources. Abundant

germplasm resources can broaden the choice of the public and enrich the market. The commercial value of the broccoli depends mainly on the quality of its curd, including the curd weight (CW), curd firmness (CF), curd covered by inner leaves (CCIL), curd anthocyanin coloration (CAC), and maturity (M). In this study, the diversity index of CCIL only was 0.26. Besides, only 7.2 % of broccoli varieties had curd covered by inner leaves, indicating that most breeders used mature breeding technology to avoid breeding varieties with CCIL, as this characteristic could significantly affect the quality of the curd. In contrast, the Shannon-Weaver diversity indices of CW, CF, CAC, and M were 1.66, 1.75, 1.27, and 1.49, respectively. The results indicated that these quality-related characteristics had rich variations. There is still room for optimization of these four characteristics in the subsequent breeding.

The CW is directly related to the broccoli yield. In this study, the CW positively correlated with the degree of plant development, plant height, leaf length, and leaf width. Thus, these characteristics, indicative of high-quality varieties, could be the critical target for future breeding.

The purple color of broccoli is due to the accumulation of anthocyanins (Liu et al., 2020). Anthocyanin pigments are synthesized in different parts of broccoli plants, such as hypocotyl, leaf, and curd. It is a valuable trait to characterize broccoli genetic resource, since it positively correlates with tolerance to abiotic stresses (Chalker-Scott, 1999). Some broccoli varieties with green curd produce anthocyanins in flower buds

Table 4 – Principal component analysis of phenotypic characteristics of 111 broccoli varieties.

Properties	PC1	PC2	PC3	PC4	PC5
Anthocyanin coloration of hypocotyl (ACH)	3.651	−0.607	0.131	0.068	1.418
Plant collateral curd (PCD)	0.507	3.832	0.811	0.115	0.073
Degree of occurrence of collateral curd (DOCC)	0.760	2.158	0.170	-0.448	−0.091
Leaf attitude (LA)	−0.222	0.136	0.107	0.185	−0.054
Leaf foliar wax powder (LFWP)	−0.142	−0.054	0.082	0.064	0.107
Leaf color (LC)	−0.126	−0.037	0.059	0.155	−0.010
Leaf anthocyanin coloration (LAC)	3.663	−0.552	0.435	0.469	−1.201
Leaf shape (LS)	−0.085	−0.043	0.120	0.046	−0.012
Undulation of leaf margin (ULM)	−0.015	0.044	0.007	0.065	−0.056
Leaf edge dentate notch (LEDN)	0.005	−0.007	0.053	0.001	−0.083
Leaf apex shape (LAS)	−0.026	−0.088	0.006	0.046	−0.012
Number of leaf lobes (NLL)	−0.118	−0.012	0.173	0.036	−0.205
Petiole cross-sectional shape (PCS)	0.044	0.059	0.027	−0.077	−0.014
curd covered by inner leaves (CCIL)	0.027	0.020	0.021	0.036	0.069
Curd color (CC)	0.212	0.045	0.092	0.018	0.082
Degree of curd color (DCC)	0.315	0.010	0.151	0.026	0.046
Curd anthocyanin coloration (CAC)	1.468	−0.124	0.215	0.135	−0.221
Curd convex surface (CCS)	−0.094	0.094	0.445	0.525	−0.002
Curd bud size (CBS)	0.567	−0.085	−0.226	−0.225	−0.349
Curd bud uniformity (CBU)	−0.014	−0.040	−0.014	−0.032	0.102
Curd shape in longitudinal section (CSLS)	0.399	0.255	−0.073	−0.165	0.090
Curd dome shape (CDS)	−0.117	−0.065	0.068	0.091	−0.022
Curd firmness (CF)	−0.630	−0.177	0.588	0.768	0.152
Curd bulb color (CBC)	0.125	0.111	−0.056	−0.075	0.037
Maturity (M)	−0.085	0.265	0.728	0.913	−0.047
Male sterility (MS)	−0.368	−0.978	3.042	−1.109	−0.005
Degree of plant development DPD (cm)	−0.138	0.005	0.333	0.372	0.164
Plant height PH (cm)	0.040	0.077	0.323	0.360	0.067
Leaf number (LN)	−0.094	−0.244	0.259	0.216	−0.461
Leaf length LL (cm)	−0.167	−0.011	0.280	0.303	0.126
Leaf width LW (cm)	−0.212	−0.157	0.060	0.060	0.101
Petiole length PL (cm)	−0.388	0.179	0.457	0.404	0.217
Main stem thickness MST (mm)	−0.579	−0.301	0.616	0.818	0.182
Flower stem length FSL (cm)	0.405	0.080	−0.279	−0.368	−0.001
Curd weight CW (g)	−0.673	−0.169	0.547	0.695	0.020
Curd diameter CD (cm)	−0.086	−0.048	0.273	0.311	0.140
Eigenvalue	32.259	21.472	12.835	5.515	4.134
Variability (%)	34.362	22.873	13.672	5.875	4.404
Cumulative of variability (%)	34.362	57.235	70.908	76.782	81.186

under low winter temperatures (Rahim et al., 2019). In our study, among 36 characteristics of 111 broccoli varieties, the coefficient of variation of anthocyanin coloration was the highest, indicating a large significant variation with anthocyanin coloration in the materials tested. Richness in anthocyanins has attracted more and more attention as an essential trait in purple broccoli breeding (Liu et al., 2022a). Anthocyanin-rich varieties of broccoli could be used as the primary material for anthocyanin research in the future.

Moreover, the quality and nutrition of broccoli had a close genetic relationship with maturity and the same cultivar maturity was usually different in various regions (Coupe et al., 2003; Dallaire et al., 2006; Tian et al., 2018; Lu et al., 2009). In this study,

the diversity at maturity was significant among the broccoli varieties bred in China in recent years. In previous studies, maturity was generally considered the most reliable indicator for categorizing broccoli varieties (Quiros et al., 2011). Our study showed that maturity was significantly correlated with ten characteristics of leaf and curd, meaning that the growth of a large volume of vegetative organs took time to accumulate, resulting in later maturity. In addition, by comparing the information on each variety maturity and geographical area, we found that most varieties bred in North China (e.g., Beijing and Tianjin) have an early maturing period. In contrast, the varieties bred in South China (e.g., Fujian) were mainly late-maturing. The varieties bred in East and Central China (e.g., Shanghai and Hubei) were

either early-maturing or late-maturing. However, Zhu et al. (2018) found that most broccoli inbred lines were not tightly clustered according to maturity. The results above suggested that a certain introgression occurs in the gene pool of different mature broccoli varieties.

Li et al. (2019) pointed out that 95 broccoli cultivars have been separated into five primary subgroups since 1980 in China, including primary cultivars and innovative resources. Zhu et al. (2018) found that the genetic diversity of the inbred lines was relatively low, which derived from the primary areas for cauliflower breeding in China (Fujian, Zhejiang, and Taiwan). In the current study, 108 new broccoli varieties were mainly bred from the Northeast, East, and Southeast China. The cluster analysis showed that the varieties tested were not clustered strictly according to the region but were dispersed in each group. It was also found that the two broccoli resources from Japan, BO17 and BO18, were clustered into Group VIII and Group II, respectively. This result may be related to the cross-utilization of germplasm resources and the germplasm derived from the same regions with different genetic backgrounds. An unobvious correlation between materials and regions indicated that the joint regulation of the internal genetic and external environment regulated the plant phenotypic traits.

Several studies used cluster and principal component analysis methods to comprehend broccoli genetic diversity and relationships. Desirable hybridization combinations of plant materials were obtained through the analysis method above (Sun et al., 2001; Adeyemo et al., 2011; Zhao et al., 2014). In our study, some phenotypic characteristics affected the taxonomy of broccoli germplasm from the PCA, mainly including collateral curd, male sterility, maturity, curd weight, curd firmness, anthocyanin coloration of hypocotyl, leaf, and curd. Curd firmness is an essential characteristic that breeders usually consider in broccoli breeding. We also found that most broccoli varieties were not tightly clustered according to the curd firmness or curd weight. A similar result was also found in cauliflower germplasms (Zhu et al., 2018).

Phenotypic studies are an essential foundation of biological research. Without the support of phenotypic data, it is impossible to rely solely on bioinformatics data, such as those from the genome and transcriptome, to understand the breeding characteristics and mechanisms (Zhang et al., 2022). Therefore, our study provided phenotypic information on new broccoli varieties, which could help screen genotypes in broccoli breeding and advance breeding development.

Conclusion

This study analyzed the genetic diversity in 108 new and three varieties of common knowledge of broccoli on 36 phenotypic characteristics. Results show that the variation coefficient of the 36 characteristics ranged from 11.18 % to 94.99 %, indicating the apparent differences among phenotypic characteristics of broccoli varieties. A higher diversity index of broccoli characteristics was present mainly in the curd characteristics, suggesting that the curd-related characteristics were the main factors in the phenotypic differences of broccoli. Moreover, the broccoli varieties were further clustered into eight groups, primarily attributed to the phenotypic difference in anthocyanin coloration, degree of plant development, plant collateral curd, main stem thickness, flower stem length, and curd weight of broccoli. Considering the commercial value of the broccoli, several curd-related traits which primarily affected the quality of broccoli, such as the curd weight, firmness, anthocyanin coloration, and maturity, were rich in diversity. Curd weight was significantly positively correlated with main stem thickness, plant degree of development, and leaf length. Thus, these characteristics could guide target traits for breeding high-yield broccoli varieties in the future.

Acknowledgments

Authors' Contributions

Conceptualization: Chen HR, Han RX. **Data curation:** Zhang YY, Lin T, Zhang Y, Zhao H. **Funding acquisition:** Chen HR. **Investigation:** Ren L. **Methodology:** Ren L, Lin T. **Resources:** Li AA, Wang Y. **Software:** Zhang YY. **Supervision:** Chen HR, Han RX. **Visualization:** Zhang YY. **Writing-original draft:** Zhang YY, Ren L. **Writing-review & editing:** Han RX. Yiying Zhang and Li Ren contributed equally to this work.

References

Adeyemo OA, Menkir A, Melaku G, Omidiji O. 2011. Genetic diversity assessment and relationship among tropical yellow endosperm maize inbred lines using SSR markers. Maydica 56: 43-50.

Chalker-Scott L. 1999. Environmental significance of anthocyanins in plant stress responses. Photochemistry and Photobiology 70:1-9. https://doi.org/10.1111/j.1751-1097.1999.tb01944.x

Coupe SA, Sinclair, BK, Watson LM, Heyes JA, Eason JR. 2003. Identification of dehydration-responsive cysteine proteases during post-harvest senescence of broccoli florets. Journal of Experimental Botany 54: 1045-1056. https://doi.org/10.1093/jxb/erg105

Dallaire R, LeBlanc DI, Tranchant CC, Vasseur L, Delaquis P, Beaulie C. 2006. Monitoring the microbial populations and temperatures of fresh broccoli from harvest to retail display. Journal of Food Protection 69: 1118-1125. https://doi.org/10.4315/0362-028x-69.5.1118

Eberhardt MV, Kobira K, Keck AS, Juvik JA, Jefery EH. 2005. Correlation analyses of phytochemical composition, chemical, and cellular measures of antioxidant activity of broccoli (*Brassica oleracea* L. var. italica). Journal of Agricultural and Food Chemistry 53: 7421-7431. https://doi.org/10.1021/jf051495k

Fisher RA. 1924. The distribution of the partial correlation coefficient. Metron 3: 329-332.

Hu JG, Quiros CF. 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. Plant Cell Reports 10: 505-511. https://doi.org/10.1007/BF00234583

International Union for the Protection of New Varieties of Plants [UPOV]. 2016. Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability: TG/45/7. UPOV, Geneva, Switzerland.

International Union for the Protection of New Varieties of Plants [UPOV]. 2018. Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability: TG/151/5. UPOV, Geneva, Switzerland.

Jian YC, Kang JG, Ding YH, Huang ZC, Chen HR. 2015. Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability: Broccoli (*Brassica oleracea* var. *italica* Plenck): NY/T 2757-2015. Ministry of Agriculture, Beijing, China. (in Chinese)

Li Z, Mei Y, Liu Y, Fang Z, Yang LM, Zhuang M, et al. 2019. The evolution of genetic diversity of broccoli cultivars in China since 1980. Scientia Horticulturae 250: 69-80. https://doi.org/10.1016/j.scienta.2019.02.034

Liu C, Yao X, Li G, Huang L, Xie Z. 2020. Transcriptomic profiling of purple broccoli reveals light-induced anthocyanin biosynthetic signaling and structural genes. PeerJ 8: e8870. https://doi.org/10.7717/peerj.8870

Liu C, Yao X, Li G, Huang, L, Liu C, Xie Z. 2022a. Development of novel markers and creation of non-anthocyanin and anthocyanin-rich broccoli (*Brassica oleracea* var. *italica*) cultivars. Applied Sciences 12: 6267.

Liu SH, Wu QY, Xu T, Wang HC, Wang JW, Zhang ZL, et al. 2022b. Profiling of sulforaphane content in forty-seven Chinese broccoli germplasms. Agronomy Journal 114: 1599-1614. https://doi.org/10.1002/agj2.21065

Louarn S, Torp AM, Holme IB, Andersen SB, Jensen BD. 2007. Database derived microsatellite markers (SSRs) for cultivar differentiation in *Brassica oleracea*. Genetic Resources and Crop Evolution 54: 1717-1725. https://doi.org/10.1007/s10722-006-9181-6

Lu X, Liu L, Gong Y, Zhao L, Song X, Zhu X. 2009. Cultivar identification and genetic diversity analysis of broccoli and its related species with RAPD and ISSR markers. Scientia Horticulturae 122: 645-648. https://doi.org/10.1016/j.scienta.2009.06.017

Nei M, Li W-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences 76: 5269-5273. https://doi.org/10.1073/pnas.76.10.5269

Quiros CF, Farnham MW. 2011. The Genetics of *Brassica oleracea*. p. 261-289. In: Schmidt, R., Bancroft, I., eds. Genetics and genomics of the brassicaceae: plant genetics and genomics; crops and models. Springer, New York, NY, USA. https://doi.org/10.1007/978-1-4419-7118-0_9

Rahim MA, Afrin KS, Jung HJ, Kim HT, Park JI, Hur Y, et al. 2019. Molecular analysis of anthocyanin biosynthesis-related genes reveal BoTT8 associated with purple hypocotyl of broccoli (*Brassica oleracea* var. *italica* L.). Genome 62:253-266. https://doi.org/10.1139/gen-2018-0173

Shannon CE, Weaver W. 1949. The Mathematical Theory of Communication. University of Illinois, Urbana, IL, USA (1949: 3224-3226).

Sneath PHA, Sokal RR. 1973. Numerical Taxonomy: The Principles and Practices of Numerical Classification. W.F. Freeman, San Francisco, CA, USA.

Soares A, Carrascosa C, Raposo A. 2017. Influence of different cooking methods on the concentration of glucosinolates and vitamin C in broccoli. Food and Bioprocess Technology 10: 1387-1411. https://doi.org/10.1007/s11947-017-1930-3

Sun GL, William M, Liu J, Kasha KJ, Pauls KP. 2001. Microsatellite and RAPD polymorphisms in Ontario corn hybrids are related to the commercial sources and maturity ratings. Molecular Breeding 7: 13-24. https://doi.org/10.1023/A:1009680506508

Tian M, Xu X, Liu F, Fan X, Pan, S. 2018. Untargeted metabolomics reveals predominant alterations in primary metabolites of broccoli sprouts in response to preharvest selenium treatment. Food Research International 111: 205-211. https://doi.org/10.1016/j.foodres.2018.04.020

Velasco P, Cartea ME, Gonzalez C, Vilar M, Ordas A. 2007. Factors affecting the glucosinolate content of kale (*Brassica oleracea acephala* group). Journal of Agricultural and Food Chemistry 55: 955-962. https://doi.org/10.1021/jf0624897

Zhang Y, Cao YF, Huo HL, Xu JY, Tian LM, Dong XG, et al. 2022. An assessment of the genetic diversity of pear (*Pyrus* L.) germplasm resources based on the fruit phenotypic traits. Journal of Integrative Agriculture 21: 2275-2290. https://doi.org/10.1016/S2095-3119(21)63885-6

Zhao Z, Gu H, Sheng X, Yu H, Wang J, Zhao J, et al. 2014. Genetic diversity and relationships among loosecurd cauliflower and related varieties as revealed by microsatellite markers. Scientia Horticulturae 166: 105-110. https://doi.org/10.1016/j.scienta.2013.12.024

Zhu S, Zhang X, Liu Q, Luo T, Tang Z, Zhou Y. 2018. The genetic diversity and relationships of cauliflower (*Brassica oleracea* var. *botrytis*) inbred lines assessed by using SSR markers. PLoS One 13: e0208551. https://doi.org/10.1371/journal.pone.0208551