Genetic variability and character interrelationships in Strawberry germplasm (Fragaria ananassa. Duch) using Principal Component Analysis and Mahalanobis D² Statistics

Pawandeep Singh
pawandeep-2197010@pau.edu
Punjab Agricultural University
Amardeep Kour
Punjab Agricultural University

Research Article

Keywords: Strawberry, Diversity, Correlation, PCV-GCV, D², PCA

Posted Date: June 3rd, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6674335/v1

Additional Declarations: No competing interests reported.

Abstract

The cultivated strawberry (*Fragaria × ananassa*) is a dicotyledonous, octoploid (2n = 8x = 56) perennial herbaceous plant grown globally. In this study, 20 genotypes were evaluated for key morphological and yield-related traits. Significant phenotypic variation was observed, indicating substantial potential for selection and genetic improvement. For all traits, phenotypic coefficients of variation (PCV) exceeded genotypic coefficients of variation (GCV), suggesting environmental influence on trait expression. The highest PCV (39.13%) and GCV (38.94%) were recorded for fruit yield per plant. Fruit diameter exhibited the highest heritability (H^2 = 99.86%), while fruit yield per plant showed the greatest genetic advance (132.42) and genetic gain (79.85%), indicating additive gene action and effectiveness of selection. Fruit yield was positively and significantly correlated with fruit length, fruit diameter, number of fruits per plant, and berry weight at both phenotypic and genotypic levels. Principal component analysis explained 73.48% of total variability, with PC1 and PC2 contributing 28.40% and 15.90%, respectively. Mahalanobis D^2 analysis grouped genotypes into three distinct clusters, with maximum divergence between Cluster I and III. Crosses between genetically distant clusters are recommended to exploit heterosis and enhance genetic diversity. The observed variability offers valuable potential for breeding and hybridization programs.

Introduction

The strawberry (*Fragaria* x *ananassa* Duch.) a perennial low-growing herb, dicotyledonous, octaploid (2n = 8x = 56), is grown in the majority of agricultural parts of the world. The domesticated strawberry is a french native and one of the 23 species in the genus Fragaria, which is a member of the Rosaceae family (Rousseau-Gueutin et al. 2009). As documented by (Hancock and Luby 1993) and (DiMeglio et al. 2014) strawberries have four naturally occurring ploidy levels, diploid, tetraploid, hexaploid, and octoploid and the basic chromosome number of all these strawberries is 7. Nutritionally, strawberries are notable for their high vitamin C content and also contain a broad spectrum of other essential nutrients, including B-complex vitamins (thiamin, riboflavin, niacin, vitamin B6) and fat-soluble vitamins (A, E, K, and carotenoids) (Alvarez-Suarez et al. 2014; Giampieri et al. 2015). Their characteristic red color is due to the accumulation of anthocyanins, polyphenolic compounds known for their antioxidant and health-promoting properties (PiljacŽegarac et al. 2012; Strathearn et al. 2014). In India, it is grown widely in the geographic regions of Haryana, Punjab, and Uttar Pradesh, J&K, and the low elevations of Himachal Pradesh. Punjab's agricultural climate is favourable for strawberry cultivation and has a lot of potential for commercial strawberry production. However, cultivation of strawberry in Punjab is still at infancy stage with limited varietal options. In different agro-climatic conditions, the performance of different cultivars depends on its interaction with the prevailing environmental conditions and the best cultivars or variety can only be selected by assessing the growths, flowering, yield and other quality contributing characteristics. The cultivation of strawberry gaining impetus in Punjab state and it is a good option towards the diversification initiative of the state. However, to harness maximum benefits from this crop, systematic evaluation of germplasm is the need of hour to identify most suitable and high yielding genotypes under Punjab conditions. Genetic diversity forms the foundation of effective breeding programs, enabling the selection and enhancement of desirable traits. To identify superior high-yielding genotypes, it is essential to evaluate variability parameters such as the genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), and heritability (Dutta et al. 2017). GCV reflects the inherited genetic variation, while PCV accounts for both genetic and environmental influences (Hamidou et al. 2018). Assessing these parameters aids in understanding the extent and nature of variability among genotypes, which is crucial for crop improvement. Heritability estimates, provide insights into the transmissibility of traits from one generation to the next and indicate the stability of phenotypic expression, guiding selection efficiency in breeding programs (Hamidou et al. 2018; Falconer & Mackay 1996). Heritability, when considered alongside genetic advance, helps predict the expected genetic gain and identify traits with high selection potential (Iqbal et al. 2017). Broad-sense heritability represents the ratio of genotypic to phenotypic variance, whereas genetic advance reflects the improvement possible through selection. The combination of both parameters gives a more accurate prediction of genetic improvement than either alone (Devesh et al. 2018).For effective selection of high-yielding genotypes, emphasis should be placed on identifying key traits. Correlation and path coefficient analyses further elucidate the direct and indirect contributions of these traits to yield. A key goal for selection and future development is to produce a superior cultivar with a good yield and high-quality fruit. In order to assess genetic variation and association among morphological and biochemical features, the current study was carried out on 20 strawberry genotypes.

Materials and methods

Experiment details and collection of plant material

The bare-root strawberry plants were obtained from the Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir (SKUAST-K), located in Kashmir, India, during the 2022–2023. A total of 20 strawberry genotypes namely Catskill, Nabila, Ofra, Selva, Red Coat, Oso Grande, Majesty, Camarosa, 13-B/34-B, Jutogh Special, Fairfox, Kaitrain Sweet, Missionary, Dilpasand, Larson, Douglas, Black More, Anthena, Chandler, and Winter Dawn(Table 1). The seedlings were transplanted onto raised beds covered with black polyethylene mulch of 30-micron thickness. Each bed measured 15 meters in length and 1.0 meter in width, with a planting space of 30 cm × 30 cm in three rows per bed. A localized drip irrigation system was employed, providing water at a rate of 1.2 L/h per plant. The experimental layout followed a randomized complete block design (RCBD) with three replications. Observations were recorded on five randomly selected plants per replication.

Table 1
The details of strawberry genotypes and their pedigree.

Sr. No	Name of genotypes	Pedigree
1	Catskill	Unknown
2	Nabila	Ventana and Q6Q8-26
3	Ofra	Parker × 111
4	Selva	Cal 70.3–117 ×Cal 71.98–605
5	Red Coat	Sparkle × Valentine
6	Oso Grande	Parker × Cal 77.3–603
7	Majesty	Unknown
8	Camarosa	Douglas × Cal85.218-605
9	13-B/34-B	Unknown
10	Jutogh Special	Unknown
11	Fairfox	Royal sovereign × Howard 17
12	Kaitrain Sweet	Unknown
13	Missionary	A chance seedling in USA,1900
14	Dilpasand	Unknown
15	Larson	Unknown
16	Douglas	(Tioga sequoia) × Tufts
17	Black More	Missionary× Howard 17
18	Anthena	Unknown
19	Chandler	Douglas × Cal 72-361-105
20	Winter Dawn	FL93-103 and FL 95–316

Morphological characters

Five randomly chosen plants were used to record data on the following characteristics: plant height (cm), crown diameter (mm), number of leaves (cm2), petiole length (cm), flowering duration (days), fruit length (top-bottom) and diameter (maximum diameter) measured with a digital vernier calliper in (cm), days to fruit maturity, number of fruits/plants, fruit weight (g) measured with a digital weighing balance, and fruit yield (g).

Biochemical characters

Biochemical parameters such as total soluble solids (TSS) were determined using hand-held refractometer (ERMA) and vitamin C was analyzed according to (Hussain et al. 2014).Total sugar was determined according to (Dubois et al. 1956). The estimation of antioxidants, including total phenolic content according to (Hatami et al. 2014), total anthocyanin content, (Lee et al. 2005).

Data analysis

R Studio software (Version 4.3.1) was used for Statistical analyses. The experimental data were subjected to analysis of variance (ANOVA) appropriate for a randomized complete block design (RCBD), as outlined by Panse and Sukhatme (1978). Genotypic and phenotypic coefficients of variation (GCV and PCV) were estimated following the methodology described by Burton and De-Vane (1953). Broad-sense heritability (H^2) was computed using the formula proposed by (Allard 1960). The genotypic correlation coefficients were calculated according to (Fisher 1954), while genetic advance was estimated as per the method suggested by (Johnson et al. 1955). Path coefficient analysis was performed following the approach developed by (Dewey and Lu 1959). Principal component analysis (PCA) was conducted and visualized using the "ggplot2", "Factoextra", "gridExtra", "corrplot", and "FactoMineR" packages in R (Wickham 2016; Auguie 2017; Wei &Simko, 2021). Genetic divergence among genotypes was assessed using Mahalanobis' D^2 statistics (Mahalanobis 1936). Based on the magnitude of genetic distances ($D = \sqrt{D^2}$), genotypes were grouped into distinct clusters using Ward’s hierarchical clustering method, employing the squared Euclidean distance as described by (Rao 1952).

Results

Genetic variability for morphological and quality attributes

Among the 20 genotypes, the analysis of variance (ANOVA) for several morphological and qualitative variables in strawberries showed extremely significant differences ($p < 0.01$) for all the traits under study. This indicates that the presence of significant genetic diversity is very important for strawberry improvement programme.The results of the morphological and quality attributes have shown wider variability among the accessions of the strawberry genotypes (Table 3). The average plant spread ranges from 25.17–39.50 (cm), plant height 16.20-27.47 (cm),crown diameter 12.03–18.04 (mm),number of leaves 10.00-20.67,leaf area 62.33–118.00 (cm²),petiole length 6.83–12.27 (cm), days to start flowering 62.51–75.62,duration of flowering 68.47–90.11,fruit length 23.92–39.01 (mm),fruit diameter 16.77–33.90 (mm),days to fruit maturity 34.33–44.67, no of fruits/plants 9.67–16.67, fruit weight 6.46–17.34 (g), fruit yield 60.64-287.12 (g), TSS 6.47–10.27 (brix %),vitamin C 27.39–57.91 (mg/100g),phenolic content 266.43-391.29 (mg/100g), anthocyanin 65.88–89.55 (mg/100g),total sugars 5.23–7.53(mg/100g).

Table 2
ANOVA of different morphological and quality characters in strawberry.

Characters	Mean Sum of Square		
	Replication (df = 2)	Treatment(df = 19)	Error(df = 38)
Plant spread (cm)	56.84	35.04**	5.15
Plant height (cm)	1.3	19.99**	1.87
Crown diameter (mm)	7.17	7.47**	3.81
Number of leaves	14.01	17.99**	5.49
Leaf area (cm²)	1211.52	633.17**	76.55
Petiole length (cm)	2.92	6.40**	0.98
Days to start of flowering	1.44	43.90**	0.03
Duration of flowering	0.21	86.97**	0.41
Fruit length (mm)	0.05	59.33**	0.05
Fruit diameter (mm)	0.009	63.37**	0.03
Days to fruit maturity	40.71	20.75**	4.5
No of fruits/plants	2.2	11.85**	1.21
Fruit weight (g)	1.03	34.61**	0.96
Fruit yield (g)	86.52	12551.49**	38.39
TSS (Brix %)	0.17	3.41**	0.2
Vitamin C (mg/100g)	0.17	229.13**	1.34
Phenolic content (mg/100g)	23.38	4778.83**	14.45
Anthocyanin (mg/100g)	1.85	203.39**	0.49
Total sugars (mg/100g)	0	1.41**	0

Df: Degree of freedom, ** significant at 1.0 percent level, respectively.

Variability, heritability and genetic advance

A significant amount of variability in strawberry germplasm for different traits was observed (Table 3). The PCV and GCV were low (< 10%) for, days to flowering after transplanting, followed by duration of flowering and days to fruit maturity whereas it was moderate (10–30%) for plant spread, plant height, number of leaves/plant, leaf area, petiole length, fruit length, fruit diameter, number of fruits per plant, ascorbic acid, phenolic content, anthocyanin and total sugars and it was high (more than 30%) for both fruit yield per plant and fruit weight. The estimates of heritability (broad sense) ranged from highest (99.86%) for fruit diameter to lowest (39.28%) for crown diameter, respectively. The heritability estimates were moderate (30.1–60.0) for the traits followed by fruit days to maturity and number of leaves per plant. While the higher heritability (> 60%) was observed for the traits like plant spread, plant height, no of leaves, leaf area, petiole length, start of flowering in days, duration of flowering, fruit length, fruit diameter, no of fruits per plant, fruit yield, fruit weight, TSS, ascorbic acid, phenolic content and anthocyanin content. The estimate of genetic advance was highest for yield/plant (132.42) followed by total phenolic content (81.72), whereas low for acidity (0.09) followed by pH (0.42). Additive gene action may be the cause of the higher GA (132.42) and higher heritability for fruit yield/plant. Thus, the selection would have an impact on this character. Genetic advance as percent of mean was lowest values for days to fruit maturity (9.05) and it was highest for fruit yield per plant (79.85) followed by fruit weight (59.16). The high GCV and PCV observed for these characters indicates the presence of adequate genetic variation among the genotypes and suitability of these attributes for further improvement by selection.

Table 3

Analysis of variances for different morphological and bio-chemical traits in strawberry .

Name of traits	GM	Range		PCV (%)	GCV (%)	H^2 (%)	GA	GA(%)
		Min	Max					
Plant spread (cm)	30.51	25.17	39.50	12.75	10.35	65.90	5.28	17.30
Plant height (cm)	18.95	16.20	27.47	14.85	12.97	76.30	4.42	23.34
Crown diameter (mm)	14.75	12.03	18.04	13.87	8.70	39.28	1.66	11.23
No of leaves	15.58	10.00	20.67	19.94	13.10	43.16	2.76	17.73
Leaf area (cm^2)	83.22	62.33	118.00	19.46	16.37	70.79	23.61	28.37
Petiole length (cm)	9.54	6.83	12.27	17.51	14.09	64.74	2.23	23.35
Days to start flowering	69.02	62.51	75.62	5.55	5.54	99.78	7.87	11.40
Duration of flowering	77.44	68.47	90.11	6.99	6.94	98.59	10.99	14.19
Fruit length (mm)	33.01	23.92	39.01	13.48	13.47	99.74	9.15	27.70
Fruit diameter (mm)	23.49	16.77	33.90	19.58	19.56	99.86	9.46	40.27
Days to fruit maturity	39.13	34.33	44.67	8.05	5.95	54.58	3.54	9.05
No of fruit per plant	13.78	9.67	16.67	15.84	13.66	74.43	3.35	24.28
Fruit weight (g)	11.19	6.46	17.34	31.18	29.92	92.11	6.62	59.16
Fruit yield (g)	165.84	60.64	287.12	39.13	38.94	99.07	132.42	79.85
TSS (°Brix)	8.50	6.47	10.27	13.28	12.16	83.86	1.95	22.95
Ascorbic acid mg/100 g	49.11	27.39	57.91	17.90	17.74	98.25	17.79	36.23
Phenols mg/100 g	330.39	266.43	391.29	12.12	12.06	99.10	81.72	24.74
Anthocyanin mg/100 g	79.50	65.88	89.55	10.38	10.35	99.27	16.88	21.23
Total sugars mg/100g	5.96	5.23	7.53	11.61	11.50	98.20	1.40	23.48

Grand mean (GM), range (minimum - maximum), per cent broad sense heritability (H^2) genetic advance (GA), and genetic advance in % (GA %) of parental mean for different morphological and quality traits in Strawberry

Principle component analysis (PCA)

The results of PCA analysis also revealed the presence of variability for different traits. The first five components had extracted Eigen value of > 1 and contributed 84.66% of the total variation (Table 4, Figs. 1 and 2). Principal component 1(PC1) contributed 28.40% of the total variability. The variation on PC1 was positively attributed by important traits like fruit yield (g) (15.50) and fruit weight (g) (12.57). PC2 contributed for 15.90% to the total variability and was depicted mainly by economic traits petiole length, leaf area, and plant height. About 12.20% of the total variability was contributed by PC3 (main attributing traits were days to start flowering and phenolic content) and 9.60% by PC4 (mainly ascribed to TSS and days to fruit maturity). PC5 contributed for 8.20% to the total variability and was depicted by fruit length.

Table 4
Principal components (PCs), corresponding eigenvalues, percentage of explained variance, and cumulative variance contributed by each component

Variables	PC 1	PC 2	PC 3	PC 4	PC 5
Plant spread (cm)	2.383	7.527	8.03	1.788	9.825
Plant height (cm)	0.119	10.649	0.691	16.386	6.011
Crown diameter (mm)	5.06	0.842	7.837	6.955	11.523
Number of leaves	0.172	5.266	11.309	12.06	2.301
Leaf area (cm²)	0.507	15.097	12.558	0.502	0.331
Petiole length (cm)	0.545	21.384	1.066	0.116	2.407
Days to start of flowering	0.267	4.49	18.983	1.641	12.609
Duration of flowering	8.914	2.335	1.817	1.693	2.179
Fruit length (mm)	9.574	0.063	0.322	0.671	18.258
Fruit diameter (mm)	12.293	0.171	0.089	0.156	0.031
Days to fruit maturity	4.546	5.583	0.004	13.13	1.607
No of fruits/plants	12.257	0.038	0.025	0.165	5.258
Fruit weight (g)	12.573	1.744	0.088	1.42	4.857
TSS (Brix %)	3.353	6.912	0.648	13.809	0.122
Vitamin C (mg/100g)	0.216	0.653	12.18	7.297	1.687
Phenolic content (mg/100g)	0.522	4.192	14.626	1.022	1.369
Anthocyanin (mg/100g)	6.354	2.462	0.044	4.383	9.314
Total sugars (mg/100g)	3.665	0.849	7.193	9.87	6.818
Fruit yield (g)	1.172	8.823	2.481	6.648	3.387
Eigenvalue	5.44	3.22	2.36	1.91	1.59
Percentage of variance	28.40	15.90	12.20	9.60	8.20
Cumulative percentage of variance	28.40	44.30	56.50	66.10	74.3

Pearson Correlation Coefficient

To determine the direction and strength of relationships between traits, correlation coefficient analysis is frequently used. Significant correlations were identified among various quantitative morphological and quality traits (Fig. 3.). The yield per plant was significantly and positively correlated with fruit weight (r = 0.88), number of fruits/plant (r = 0.76), fruit diameter (r = 0.66), fruit length (r = 0.65) and duration of flowering (r = 0.62). Non-significant correlation was found between yield and days taken to fruit maturity (r= -0.36).

Path coefficient analysis

The data (Table 5) path coefficient analysis elucidates the influence of distinct independent traits, both in isolation and in conjunction with other traits, on the manifestation of various characteristics related to yield. Path correlation analysis revealed fruit weight (0.517) exhibits the highest positive direct effect on yield per plant, followed by number of fruits per plant (0.342) and fruit diameter (0.144). While negative direct effects observed on yield per plant for and total sugars (-0.113) and leaf area (-0.087). This indicates that there is a real correlation between these traits and yield, and selection for this characteristic directly will increase strawberry yield. Maximum indirect positive effect was found for fruit yield with fruit weight (0.348) via fruit length followed by fruit weight (0.298) via duration of flowering, fruit weight (0.290) via number of fruits. Highest negative indirect effects were recorded for followed by fruit weight (-0.184) via days to start flowering. The reported residual impact was 0.078, reflecting a very low value and just 0.2% contribution from other characters not considered in the study.

Table 5

Path correlation among differeent traits with yield of the plants

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	-0.016	-0.021	-0.001	0.001	0.000	0.010	0.013	0.025	0.003	0.033	0.006	0.023	0.111	0.050	-0.019	-0.0
2	-0.009	-0.037	0.006	0.001	-0.018	0.018	0.011	-0.001	0.000	-0.011	0.001	0.009	0.021	0.012	0.003	-0.0
3	0.000	0.008	-0.026	-0.001	0.015	-0.002	-0.002	-0.026	-0.001	-0.039	-0.018	-0.130	-0.110	-0.019	-0.013	-0.0
4	-0.003	-0.013	0.007	0.003	-0.030	0.012	-0.007	0.003	0.001	0.012	0.009	-0.021	-0.045	0.012	0.010	0.00
5	0.000	-0.008	0.005	0.001	-0.087	0.028	-0.004	-0.031	0.000	0.000	-0.030	-0.024	-0.123	0.026	0.014	0.00
6	-0.003	-0.015	0.001	0.001	-0.054	0.045	0.003	-0.037	0.000	-0.011	-0.023	-0.007	-0.108	0.017	-0.003	-0.0
7	-0.005	-0.010	0.001	0.000	0.007	0.003	0.043	-0.005	-0.005	-0.006	-0.006	0.013	-0.184	-0.002	-0.022	-0.0
8	-0.004	0.000	0.006	0.000	0.024	-0.015	-0.002	0.113	0.007	0.078	0.025	0.128	0.298	-0.004	-0.014	-0.0
9	-0.004	-0.001	0.003	0.000	0.001	0.001	-0.019	0.067	0.012	0.107	0.015	0.127	0.348	0.025	-0.011	-0.0
10	-0.004	0.003	0.007	0.000	0.000	-0.003	-0.002	0.062	0.009	0.144	0.021	0.179	0.271	0.035	-0.014	-0.0
11	0.001	0.001	-0.006	0.000	-0.035	0.014	0.003	-0.038	-0.002	-0.041	-0.073	-0.080	-0.154	0.021	-0.014	-0.0
12	-0.001	-0.001	0.010	0.000	0.006	-0.001	0.002	0.042	0.004	0.075	0.017	0.342	0.290	0.039	-0.003	-0.0
13	-0.003	-0.001	0.006	0.000	0.021	-0.009	-0.015	0.065	0.008	0.075	0.022	0.192	0.517	0.042	-0.002	0.00
14	-0.006	-0.003	0.004	0.000	-0.017	0.006	-0.001	-0.003	0.002	0.037	-0.011	0.100	0.162	0.134	-0.016	-0.0
15	0.004	-0.002	0.004	0.000	-0.016	-0.002	-0.012	-0.020	-0.002	-0.026	0.013	-0.013	-0.013	-0.029	0.076	0.00
16	0.005	0.008	0.002	0.000	-0.012	-0.006	-0.018	-0.016	0.000	-0.025	0.001	-0.059	0.029	-0.010	0.014	0.04
17	-0.003	-0.005	0.011	0.001	-0.007	-0.001	0.005	0.033	0.002	0.086	0.018	0.156	0.103	0.039	0.001	-0.0
18	0.000	0.006	0.009	0.000	-0.012	-0.001	-0.008	0.007	0.001	0.033	0.005	0.122	0.183	0.092	-0.002	0.00

1. Plant spread (cm), 2. Plant height (cm), 3. Crown diameter (mm), 4. No of leaves, 5. Leaf area (cm²), 6. Petiole length (cm), 7. Days to start flowering, 8. Duration of flowering, 9. Fruit length (mm), 10. Fruit diameter (mm),11. Days to fruit maturity, 12. No. of fruit per plant, 13. Fruit weight (g), 14. TSS (°Brix), 15. Ascorbic acid mg/100 g, 16. Phenols mg/100 g, 17. Anthocyanin mg/100 g, 18. Total sugars mg/100g.

Genetic divergence (D^2)

Based on the assessment of performance across multiple traits, (Fig. 4) demonstrates the clustering arrangement of 20 strawberry varieties. These varieties have been categorised into three distinct clusters. Cluster-I encompasses 10 varieties. Cluster-II comprises 3 varieties and Cluster-III comprises 7 varieties. Cluster II is distinguished by their high yields. Whereas cluster III stood out based on their excellent quality attributes like TSS, anthocyanin and total sugars. Cluster I was not remarkable in terms of either yield and their related traits and also quality attributes.

Average D^2 values representing both inter-cluster and intra-cluster divergence are given (Table 6). In this table, the diagonal elements represent the intra-cluster divergence, while the off-diagonal elements signify inter-cluster divergence. Notably, the cluster with the greatest intra-cluster distance is Cluster II, with a value of (113.31), whereas Cluster III exhibits the lowest intra-cluster distance at (59.23). Furthermore, it is noteworthy that the highest inter-cluster distance, quantified at (115.10) occurred between Cluster I and Cluster III, while the lowest inter-cluster distance, (84.08) was observed between Cluster I and Cluster II.

Table 6

Average intra-and inter- cluster divergence (D^2) of 20 genotypes of strawberry

Cluster	I	II	III
I	84.25	84.08	115.10
II		113.31	95.50
III			59.23

Discussions

Germplasm characterization based on the morphological as well as quality traits is of great importance for plant breeders since it is the first step in the initiation of crop breeding programmes and identification of superior genotypes to diversify local production (Pereira-Lorenzo et al. 2012). The significant value of the genotypic mean sum of squares indicating the presence of environmental influence resulted in variation for all the characters among genotypes of strawberry, which can be improved by further breeding techniques. GCV and PCV are the most important parameters to estimate the level of variability present in the population. The present study had (10–30%) for plant height, plant spread, fruit length, fruit diameter, anthocyanin and total sugars, while for fruit yield per plant and fruit weight had higher (> 30%) GCV and PCV, which indicated the presence of wider variability for these traits as previously reported in the strawberry (Mishra et al. 2015; Singh et al. 2018). Further, the values of PCV were higher than their corresponding values of GCV for all the characters under consideration, indicating that the apparent variation is not only due to genotypes but also to the influence of environment in the expression of traits. Heritable variation is important for genetic improvement (Singh 2000). Heritability reflects the percentage of genetic variability that is transferred from parents to offspring. Heritability was classified as suggested by (Robinson et al.1949) into low (0–30%), moderate (30.1–60%) and high (> 60%). The current heritability study reveals that the estimates of heritability (broad sense) range from highest (99.86%) for fruit diameter and lowest (39.28%) for crown diameter respectively. Similarly, findings were also observed by (Mishra et al. 2015; Kumar et al. 2012).Heritability alone is not a good predictor of the exploitable genetic variability, according to (Masood 1986), thus breeders should also take genetic advance into account. The higher heritability coupled with higher GA (132.42) observed for fruit yield/plant might be due to additive gene action so, the selection would be affective for this character. Similar findings were also reported by (Lal et al. 2016; Zareei et al. 2022; Verma et al. 2024).

Principal Component Analysis (PCA) is a multivariate statistical technique employed to elucidate patterns of variation among samples and to identify the variables that predominantly contribute to their similarities and differences. By transforming the original data matrix into a set of orthogonal components, PCA captures the major sources of variance through a complementary representation of score and loading plots. This approach facilitates dimensionality reduction, enables the exploration of inter-object relationships, provides insights into the underlying correlation structure among variables, and determines the optimal number of principal components required to retain the majority of the original information with minimal loss (Rodríguez-Delgado et al. 2002). In this study, PCA was conducted on the complete dataset of mean values, with the first principal component (PC1) accounting for 28.40% and the second principal component (PC2) explaining 15.90% of the total variance (Fig. 2). Collectively, the first two components, both having eigenvalues exceeding 1.0, accounted for 74.30% of the overall variability. Previous studies have successfully applied PCA in the evaluation and classification of strawberry genotypes based on agro-morphological traits and fruit quality parameters (Zareei et al. 2022; Hummer et al. 2022; Šamec et al. 2016; Park et al. 2020; Pinheiro et al. 2021; Chiomento et al. 2021).

From a breeding perspective, correlation studies are crucial because they show the degree to which one or more traits are correlated with one another and also indicate which attribute may be valuable in identifying more crucial ones for a given improvement programme. Path coefficient analysis elucidates the influence of distinct independent traits, both in isolation and in conjunction with other traits, on the manifestation of various characteristics related to yield. The yield per plant was significantly and positively correlated with fruit weight, number of fruits/plant, fruit diameter and fruit length. (Singh et al. 2018; in strawberry, Hummer et al. 2022; Zareei et al. 2022; Verma et al. 2025) in strawberry also employed the path coefficient analysis to investigate the direct and indirect correlations between agro-morphological attributes via the partitioning of correlation coefficients.

In the current study, clusters were constructed using Mahalanobi's D2 statistics (Mahalanobi, 1936), which is an excellent method for assessing genetic divergence. This clustering based on morphological characterization suggests that there is enough genetic variation among them. Numerous clusters were also seen to emerge in *Fragaria ananassa* (Jan et al. 2022; Chhetri et al. 2017), suggesting that the material under study exhibits a great diversity in material. The genotypes from highly distant clusters should be used for the development of high yielding varieties (Patel et al. 2017). Based on these criterion genotypes in clusters I and III would be more fruitful for producing segregates with great heterotic effects as they produced maximum distance between them. Further inter-cluster range observed in the present study was much lower than the results obtained by (Jan et al. 2022) in *Fragaria ananassa.*

The study of different characteristic contribution towards genetic variance is important with regards to selection and choice of parents for hybridization programme and is performed on the basis of D values. The traits like fruit yield per plant and quality parameters can be taken into consideration during selection of parents for hybridization as it showed highest contribution towards genetic divergence.

Conclusion

The present study revealed wider genetic variability among the 20 genotypes of strawberry evaluated for morphological and yield-related traits, demonstrating the substantial potential for genetic enhancement and selection. The very close value of PCV and GCV for all characters indicates least environmental control, with fruit yield showing the highest variability. Because of the high heritability, high genetic gain, and high genetic advance for fruit yield per plant, which indicate that additive genetic influences predominate, making selection more effective. Fruit length, fruit diameter, and quantity of fruits per plant all significantly increase fruit yield, indicating that these characteristics may be important selection criteria. Significant genetic divergence was further validated by the PCA and Mahalanobis D2 analysis, which grouped the genotypes into three different clusters. Inter cluster distance was higher than intra cluster distance which indicates the high genetic diversity in the genotypes. To exploit maximum heterosis and broaden the genetic base the crosses between genetically distant clusters are recommended. So the present study will utilized for direct selection of superior genotypes or in trait specific breeding programs aimed at developing high-yielding, improved *Fragaria ananasa* varieties.

Declarations

Conflict of interest

P. Singh and A. Kour declare that they have no competing interests.

Author Contribution

Singh P: experiment performance, data analysis and writing of the original draft. Kour A: experimental design, supervision, conceptualization, methodology, resources and manuscript drafting.

Acknowledgement

I would like to express my sincere gratitude to Dr. Amit Kumar Khokhar for generously providing the germplasm used in this study. His support and cooperation from SKUAST-K (Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir) were invaluable to the successful completion of this work.

Data availability

The authors declare that the data supporting the findings of this study are available within the article.

References

1. Allard RW (1960) Principles of Plant Breeding. John Wiley and sons. Inc., New York.
2. Alvarez-Suareza JM, Mazzoni L, Forbes-Hernandez TY, Gasparrini M, Sabbadini S, & Giampieri F (2014) The effects of pre-harvest and post-harvest factors on the nutritional quality of strawberry fruits: A review. J. Berry Res. 4, 1–10. https://doi.org/10.3233/JBR-140068.
3. Auguie B, gridExtra: Miscellaneous Functions for & quot; Grid & quot; Graphics (2017) R package version 2.3. Available at: https://CRAN.R-project.org/package=gridExtra&gt.
4. Burton GW, De-Vane EH (1953) Estimating heritability in tall Fescue (Festuca arundinacea) from replicated clonal material. Agronomy. 45:478–481.
5. Chhetri A, Thakur N, Pant SC, Shekhar C, Pandey PK (2017) Assessment of genetic diversity in strawberry (Fragaria× ananassa Duch.) grown under Garhwal Himalayas. Crop Res. 52(1to3), 49-54.
6. Chiomento JLT, Júnior EPL, D'Agostini M, De Nardi FS, dos Santos Trentin T, Dornelles AG, Huzar-Novakowiski J, Calvete EO (2021) Horticultural potential of nine strawberry cultivars by greenhouse production in Brazil: A view through multivariate analysis. Scientia Horticulturae. 279, p.109738.DOI: https://doi.org/10.37256/acbr.2120232332.
7. Devesh P, Moitra PK, Shukla RS, Shukla SS, Pandey S, Arya G (2018) Analysis of variability, heritability and genetic advance of yield, its components and quality traits in wheat. Int. J. Agric. Biotech. 855–859.
8. Dewey DR, Lu KH (1959) A correlation and path coefficient analysis of components of crested wheat grass and seed production. Agron. J. 51 (12), 515–517.
9. DiMeglio L, Staudt G, Yu H, Davis T (2014) A phylogenetic analysis of the genus Fragaria (strawberry) using intron-containing sequence from the ADH-1 gene. Plos One. http://journals.plos.org/plosone/article?id=10.1371/journal. pone.0102237.
10. Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) Colorimetric Method for Determination of Sugars and Related Substances. Anal Chem. 28: 350-56.
11. Dutta S, Munda S, Chikkaputtaiah C, Lal M (2017) Assessment of selection criteria for development of high yielding genoptypes using variability parameters in lemongrass Cymbopogon flexuosus L. J. Essent. Oil-Bear. Plants. 20, 1450–1460. DOI:10.1080/0972060X.2017.1421104
12. Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics, 4th edition. Benjamin Cummings, England.
13. Fisher RA (1954) Statistical methods for research workers. Twelve Edition. Edinburgh: Oliver and Boyd.
14. Giampieri F, Forbes-Hernandez TY, Gasparrini M, Alvarez-Suarez JM, Afrin S, Bompadre S, Battino M (2015) Strawberry as a health promoter: An evidence-based review. Food Funct. 6, 1386–1398. DOI: 10.1039/c5fo00147a.
15. Hamidou M, Souley AKM, Kapran I, Souleymane O, Danquah EY, Ofori K, Ba MN (2018) Genetic variability and its implications on early generation sorghum lines selection for yield, yield contributing traits, and resistance to sorghum midge. Int. J. Agron. 1–10. https://doi.org/10.1155/2018/1864797. DOI:10.1155/2018/1864797.
16. Hancock JF, Luby JJ (1993) Genetic resources at our doorstep: the wild strawberries. BioScience. 43 (3), 141–147. https://doi.org/10.2307/1312017.
17. Hatami T, Emami SA, Miraghaee SS, Mojarrab M (2014) Total Phenolic Contents and Antioxidant Activities of Different Extracts and Fractions from the Aerial Parts of Artemisia biennis Wild. Iran. J. Pharm. Res. 13, 551–559.
18. Hummer KE, Bassil NV, Zurn JD, Amyotte B (2023) Phenotypic characterization of a strawberry (Fragaria ananassa Duchesne ex Rosier) diversity collection. Plants, People, Planet. 5(2) 209–224. https://doi.org/10.1002/ppp3.10316.
19. Hussain M, Ismail M, Ali S (2018) Quantitative Determination of Ascorbic Acid in Commercial Fruit Juices by Redox Titration. Int. J. Pharm. Qual. Assur. 5, 22–25.
20. Iqbal A, Khalil IH, Shah SM, Kakar MS (2017) Estimation of heritability, genetic advance and correlation for morphological traits in spring wheat. Sarhad J. Agric. 33(4) 674–679. http://dx.doi.org/10.17582/journal.sja/2017/33.4.674.679.
21. Jan S, Baba JA, Dar ZA, Ali G, Nazir I (2022) Genetic diversity studies in strawberry (Fragaria × ananassa Duch.) under temperate climatic conditions of North West Himalayas. Bio For. 14:121-25.
22. Johnson HW, Robinson HF, Comstock RE (1955) Estimation of genotype and environmental variability in Soybean. Agron. J. 47, 314–318.
23. Kumar B, Kumar S, Yadav YC (2012) Genetic variability for some metric traits in strawberry (*Fragaria × ananassa* Duch.). Hort Flora Res Spect. **1**:86-88.

24. Lal, Shiv, Singh DB (2016) Genetic divergence among 22 strawberry (*Fragaria´ Ananassa* Duch.) genotypes for bioactive compounds grown in northwest Himalayas. SAARC J Agric.**14**:81. DOI:10.3329/sja.v14i1.29578.
25. Lee J, Durst RW, Wrolstad RE (2005) Determination of total monomeric anthocyanin pigment content of fruit juices, beverages, natural colorants, and wines by the pH differential method: collaborative study. J. AOAC Int. 88, 1269–1278.https://doi.org/10.1093/jaoac/88.5.1269.
26. Mahalanobis PC (1936) On the generalized distance in statistics. Proceedings of the National Institute of Science of India. 2, 49–55.
27. Masood MS, M Y Mujahid, NS Kisana, NI Hashmi (1986) Variability studies in wheat under rainfed conditions. Pak J Agric Res. **7**:244-47.
28. Mishra PK, Ram RB, Kumar N (2015) Genetic variability, heritability, and genetic advance in strawberry (Fragaria× ananassa Duch.). Turk. J. Agric. 39(3). https://doi.org/10.3906/tar-1408-99. DOI:10.3906/tar-1408-99.
29. Panse VG, Sukhatme PV (1978) Statistical methods for research workers. New Delhi: ICAR.
30. Park MJ, Ryu DH, Cho JY, Lee DG, Lee JN, Kang YH (2021) Potential for antioxidant and antihyperglycemic activities of four everbearing strawberry cultivars. Horticulture, Environment, and Biotechnology. 61, 615–623. https://doi.org/10.1007/s13580-020-00240-y.
31. Patel P, Kumar U, Thakur G, Maurya PK (2017) Assessment of genetic diversity through D^2 analysis in tomato (Solanum lycopersicum L.). Bull. Env. Pharmacol. Life Sci. 6:291–224.
32. Pereira-Lorenzo S, dos Santos ARF, Ramos-Cabrer AM, Sau F and Díaz-Hernández MB (2012) Morphological variation in local pears from north-western Spain. Scientia Horticulturae. 138:176–182. https://doi.org/10.1016/j.scienta.2012.02.007
33. Piljac Zˇegarac J, Piljac A, Basha S, Pinter M, Šamec D, Petravic´ Tominac V (2012) Degradation kinetics of total phenols, anthocyanins and antiradical activity of blackcurrant nectars stored at room and refrigerator temperatures. Ital. J. Food. 24, 345–351.
34. Pinheiro DF, Resende JTVD, Constantino LV, Hata FT, Hat NNY, Lustos S BC (2021) Physical, biochemical, and sensory properties of strawberries grown in high-altitude tropical climate. Cienc. E Agrotecnol. 45, e008221. https://doi.org/10.1590/1413-7054202145008221.
35. Rao CR (1952) Advance Statistical Methods in Biometrical Research. J. Willey and Sons, New York p.390.
36. Robinson HF, Comstock RE, Harvey PH (1949) Estimates of heritability and the degree of dominance in corn. Agron J. **41**:353-59.
37. Rodrıguez-Delgado MA, Gonzalez-Hernandez G, Conde-Gonzalez JE, & Perez- Trujillo JP (2002) Principal component analysis of the polyphenol content in young red wines. Food Chem. 78, 523–532.
38. Rousseau-Gueutin M, Gaston A, Arnouche A, Ainouche ML, Olbricht K, Staudt G, Richard L, Denoyes-Rothan B (2009) Tracking the evolutionary history of polyploidy in Fragaria L. (strawberry) New insights from phylogenetic analyses of low-copy nuclear genes. Mol. Phylogenet. Evol. 51: 515-30. https://doi.org/10.1016/j.ympev.2008.12.024.
39. Samec D, Mareti´c M, Lugari´c I, Meˇsi´c A, Salopek-Sondi B, Duralija B (2016) Assessment of the differences in the physical, chemical and phytochemical properties of four strawberry cultivars using principal component analysis. Food Chem. 194, 828–834. https://doi.org/10.1016/j.foodchem.2015.08.095.
40. Singh BD (2000) Plant Breeding: Principles and Methods. New Delhi, India: Kalyani Publishers.
41. Singh G, Kachwaya DS, Kumar R, Vikas G, Singh L (2018) Genetic variability and association analysis in strawberry (Fragaria x ananassa Duch). Electron. J. Plant Breed. 9 (1), 169–182. https://doi.org/10.5958/0975-928X.2018.00021.2.
42. Strathearn KE, Yousef GG, Grace MH, Roy SL, Tambe MA, Ferruzzi MG, Wu QL, Simon JE, Lila MA, Rochet JC (2014) Neuroprotective effects of anthocyanin- and proanthocyanidin-rich extracts in cellular models of Parkinson's disease. **Brain Research**. *1555*, 60–77. https://doi.org/10.1016/j.brainres.2014.01.043
43. Verma P, Singh G, Singh SK, Bakshi M, Mirza AA, Mehandi S, Vijayvargiya V (2025) Correlation, path-coefficient and principal component analysis association among quantitative traits in strawberry to unlock potential of vertical farming system. Kuwait J Sci. ; 52(1). DOI:10.1016/j.kjs.2024.100303
44. Wei T, Simko V (2017) R Package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from: https://github.com/taiyun/corrplot.
45. Wickham H (2016) Ggplot2: Elegant Graphics for Data Analysis. Springer, Houston, TX. Available at: https://ggplot2.tidyverse.org. (Accessed 20 July 2022).
46. Zareei E, Karami F, Aryal R, Saed-Moucheshi A (2023) Genotypic by phenotypic interaction affects the heritability and relationship among quantity and quality traits of strawberry (Fragaria× Ananassa). N. Z. J. Crop Hortic. Sci. 51(4) 594-613. https://doi.org/10.1080/01140671.2022.2039725.

Figures

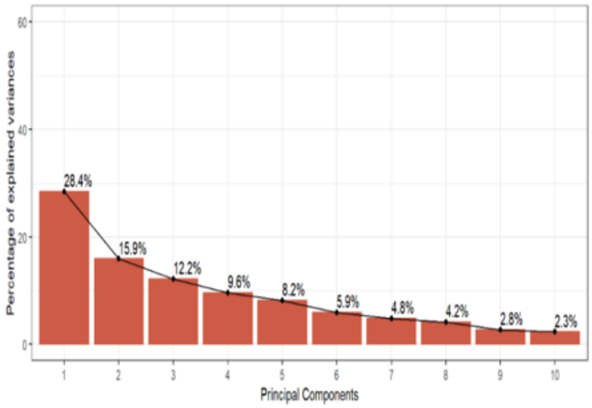

Figure 1

Scree plot

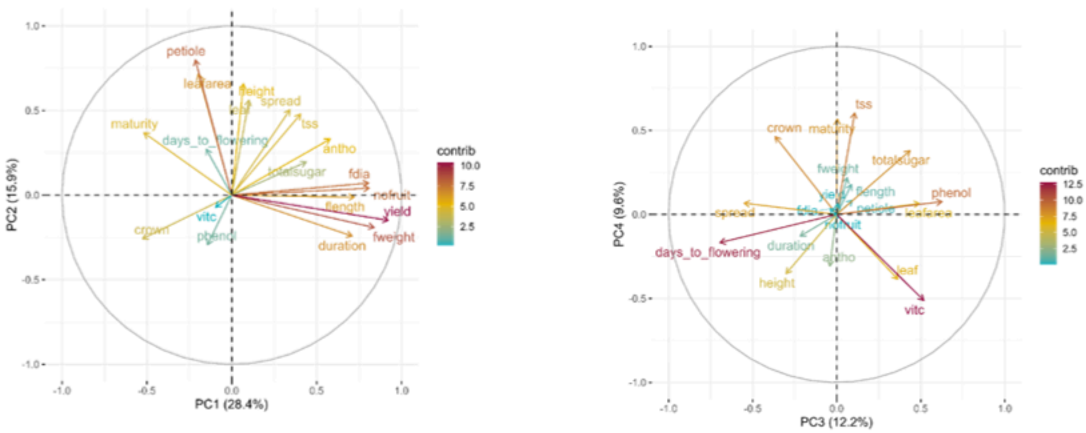

Figure 2

Biplot of principal component analysis (PCA) illustrating the distribution of 20 strawberry genotypes and the contributions of the assessed traits to the total variation.

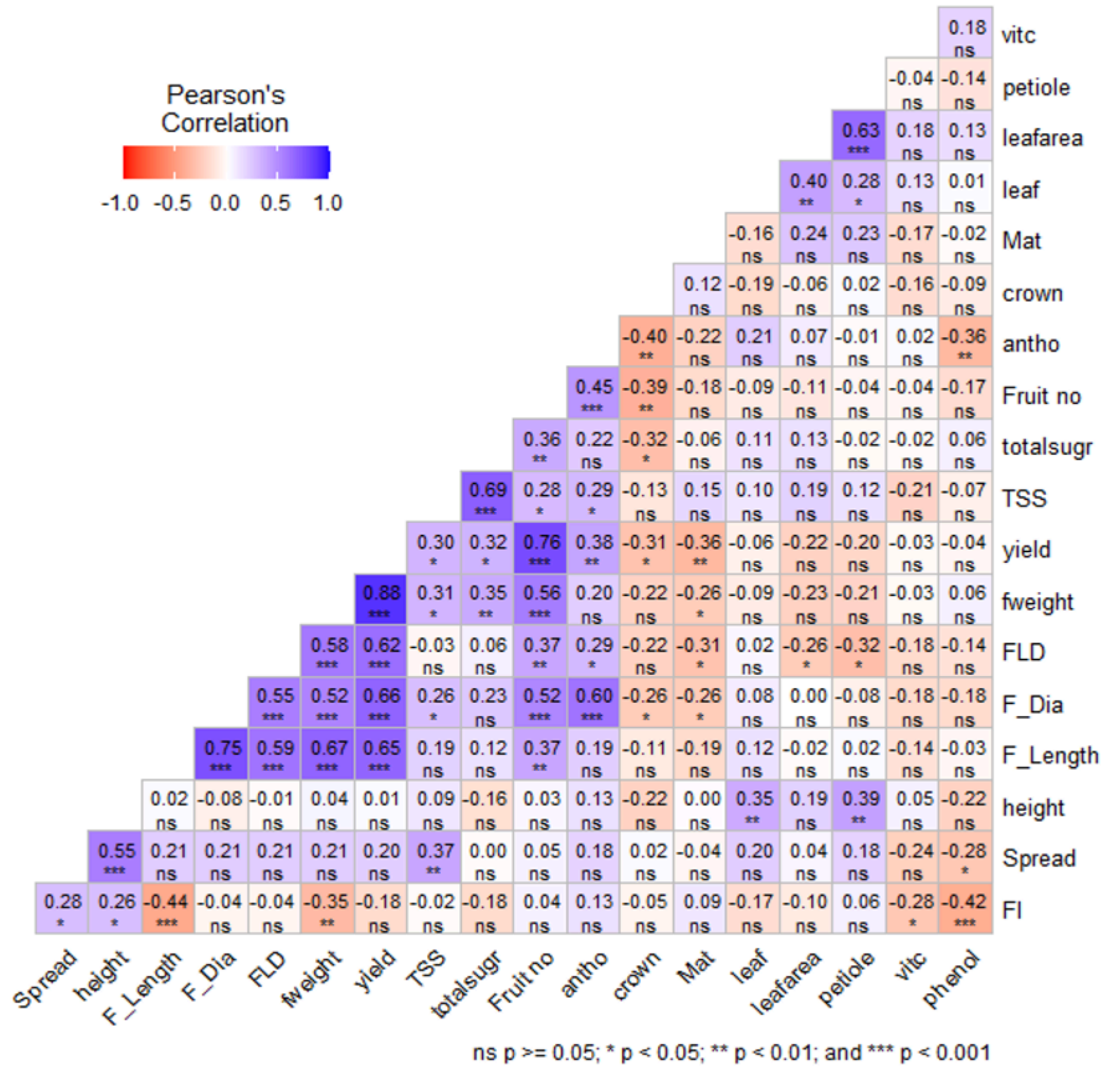

Figure 3

Pearson correlation coefficients among key traits in strawberry

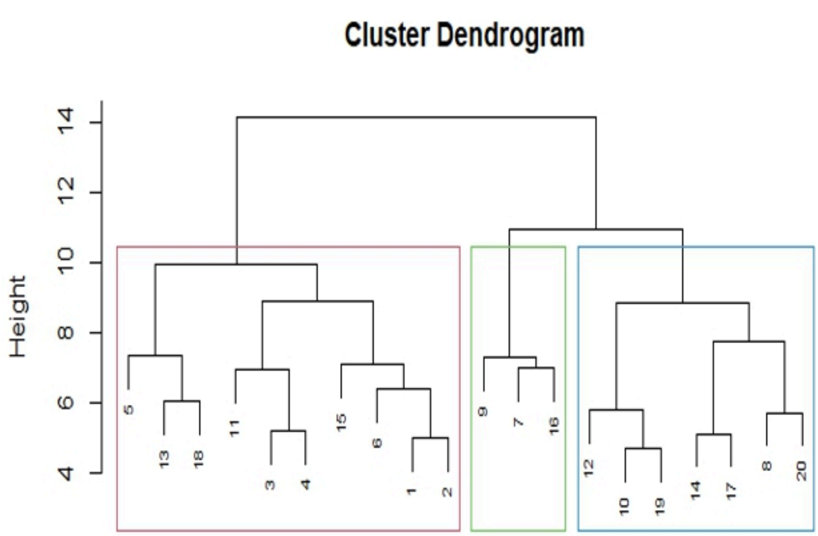

Figure 4

Cluster Dendrogram 1. Catskill, 2. Nabila, 3. Ofra, 4. Selva, 5. Red Coat, 6. Oso Grande, 7. Majesty, 8. Camarosa, 9. 13-B/34-B, 10. Jutogh Special, 11. Fairfox, 12. Kaitrain Sweet, 13. Missionary, 14. Dilpasand, 15. Larson, 16. Douglas, 17. No. of fruit per plant, 18. Black More, 19. Anthena, 20. Winter Dawn.