Received: 03 Apr 2024 | Revised: 21 Jun 2024 | Accepted: 23 Jun 2024

Electronic Journal of Plant Breeding

Research Note

Principal component and correlation analyses study on fruit yield in cucumber (*Cucumis sativus* L.) genotypes

O.A. Umeh[1*], I.S. Umeh[2], J.I. Ulasi[3], E.R. Keyagha[4] and C.O. Cookey[4]

[1]Department of Crop Science and Horticulture, Faculty of Agriculture, Nnamdi Azikiwe University, P.M.B. 5025, Awka, Anambra State.
[2] Department of Measurement and Evaluation, Faculty of Education, Imo State University, P.M.B. 2000, Owerri, Imo State.
[3]Department of Crop Science, Faculty of Agriculture, University of Uyo, P.M.B. 1017, Uyo, Akwa Ibom State.
[4] Department of Crop Science and Technology, Federal University of Technology, P.M.B. 1526, Owerri, Imo State.
***E-Mail:** oa.umeh@unizik.edu.ng

Abstract
The degree of association between yield and its components can be identified using correlation and Principal Component Analyses (PCA). PCA also reveals key characteristics that explain most of the differences between genotypes. A study was formulated to evaluate the relationship between yield and its contributing traits in cucumber. The experiment was conducted with 16 cucumber genotypes in a Randomized Complete Block Design, with three replications. The correlation analysis revealed a strong and statistically significant relationship in number of pistillate flowers (*r = 0.58***), number of branches (*r = 0.43***), vine length (*r = 0.69***), number of leaves (r = *0.73***), leaf area (*r = 0.70***), number of fruits (*r = 0.91***), fruit length (*r = 0.40***), fruit girth (*r = 0.39***), and fruit weight (*r = 0.74***) with fruit yield. PCA revealed that PC1 accounted for 51.53% of the total variation, while PC2 explained 13.91% of the total variability. This study demonstrated that choosing traits such as number of pistillate flowers, number of branches, vine length, number of leaves, leaf area, number of fruits, fruit length, fruit girth, and fruit weight that have a strong positive correlation with fruit yield could be given priority in selection for yield improvement.

Keywords: Correlation, Cucumber, PCA

Cucumber (*Cucumis sativus var. sativus* L.), is a member Cucurbitaceae family, which also encompasses squash (*Cucurbita ssp*.), watermelon (*Citrullus lanatus* (Thunb.) Matsum.), and melon (*Cucumis melo* L.). Cucumber and melon are cultivated more extensively than any other vegetable species, following tomato and watermelon (FAO, 2021). Although cucumber possesses significant diversity for various yield characters, it is currently not fully utilized in terms of its economic potential and has not been extensively explored in breeding efforts (Sharma *et al*., 2018). Hence, it is necessary to evaluate cucumber germplasm in order to identify genotypes with enhanced quality and yield. These genotypes can then be utilized directly as varieties after thorough assessment or as parental lines in hybridization programs (Kumar *et al*., 2011). Estimating the correlation between different yield and yield-related traits will offer valuable information that is essential for cucumber breeders, particularly when selecting based on two or more traits simultaneously. Principal component analysis (PCA) is a statistical technique that reduces the dimensions of multivariate data by eliminating intercorrelations among the traits under study. This allows for the representation of multidimensional relationships on principal axes. Correlation and PCA are used to determine the level of association between yield and its components. It also identifies the main traits that account for most of the variations among the genotypes (Mahalingam *et* *al*., 2020). These analyses could assist breeders in formulating selection strategies to enhance yield. A study was conducted to evaluate association between yield and its contributing characteristics in cucumber which could help in formulating selection indices for efficient crop improvement.

The experimental material consisted of 16 cucumber genotypes obtained from various geographical regions of the Nigeria (**Table 1**). The 16 accessions were sown in a Randomized Complete Block Design, with three replications at the Teaching and Research Farm of Department of Crop Science and Technology, School of Agriculture and Agricultural Technology, Federal University of Technology, Owerri, Nigeria . It is located between 05° 29′ 06s latitude and 07° 02′ 06s longitude at an altitude of 17 m above sea level. The Seeds were planted in rows at a spacing of 0.5cm x 0.5cm, 0.8m between rows, 16 rows per accession per replication in a plot size of 1.5m x 1.0m. All crop management practices were followed for better crop establishment and stand.

The following parameters were recorded: days to emergence, days to 50% emergence, days to male flower initiation, days to first female flower initiation, number of pistillate flowers, days to fruit maturity, number of branches, vine length at 8 weeks after planting, number of leaves at 8 weeks after planting, leaf area, number of fruits , fruit length, fruit girth, and fruit yield per hectare (ton). Correlation and Principal component analysis were carried out using GenStat (2011) software to determine the level of association between yield and its components and to identify the main traits that account for most of the variations among the genotypes.

Correlation studies: An analysis of correlation among the agronomic traits of cucumber showed strong and statistically significant (P<0.01) correlations among number of pistillate flowers (r = 0.58**), number of branches (r = 0.43**), vine length (r = 0.69**), number of leaves (r = 0.73**), leaf area (r = 0.70**), number of fruits (r = 0.91**), fruit length (r = 0.40**), fruit girth (r = 0.39**), and fruit weight (r = 0.74**) with fruit yield. On the other hand, the days to male (r = -0.46**) and female flower initiation (r = -0.49**), as well as the days to maturity (r = -0.52**), showed a negative correlation with fruit yield (**Table 2**). Nevertheless, notable negative correlations were observed between these characteristics (days required for plants to emerge and days required for 50% emergence) and fruit length, fruit girth and fruit weight. Therefore, the number of days for the plants to emerge was found to have a negative correlation with fruit length (r = -0.30*). Similarly, the number of days for 50% of the plants to emerge was negatively correlated with the number of fruits per plant (r = -0.32*), fruit length (r = -0.48**), and fruit girth (r = -0.35*).

Machikowa and Laosuwan (2011) emphasized the significance of correlating specific traits with other traits and with yield when selecting genotypes for yield enhancement. This aligns with the conclusions drawn by Fayeun *et al*. (2012), who noted a notable and favorable correlation between two traits, indicating that these traits can be enhanced concurrently in a selection program. Hence, a symbiotic association between characters and the choice of a particular trait will result in the selection and enhancement of another. The correlation matrix revealed that all traits, with the exception of days to emergence and days to 50% emergence, exhibited a

Table 1. List of cucumber genotypes studied along with source

S. No.	Genotypes	Origin
1.	Songhai local	ADP
2.	Marketer	NIHORT
3.	AOA/Cu	NIHORT
4.	Beit alpha	ADP
5.	Israeli Cu	ADP
6.	Holland POP	NIHORT
7.	Apulia	ADP
8.	Nagano F_1	ADP
9.	Cu 102	THAILAND
10.	Cu 986	THAILAND
11.	Super marketer	NIHORT
12.	OHE/Cu	NIHORT
13.	Pov variety ADP	ADP
14.	Cu 100	THAILAND
15.	Cu 971	THAILAND
16.	Cu 999	THAILAND

NIHORT: National Horticultural Research Institute Kano, Nigeria; ADP: Agricultural Development Programme, Owerri, Nigeria; Thailand Agro Seed Company, Owerri, Niger

https://doi.org/10.37992/2024.1502.065

Table 2. Correlation matrix of the agronomic parameters of the evaluated cucumber genotypes.

Characters	Days to 50% emergence	Days to male flower initiation	Days to female flower initiation	Number of pistillate flowers	Days to maturity	Number of branches at 8WAP	Vine Length at 8WAP	Number of leaves at 8WAP	Leaf Area at 8WAP	Number of fruits	Fruit length	Fruit girth	Fruit weight per plant	Total fruit yield per hectare
Days to emergence	0.443**	0.303*	0.272	-0.017	0.235	0.038	-0.075	-0.067	0.028	-0.031	- 0.300*	-0.119	-0.179	-0.064
Days to 50% emergence		0.440**	0.443**	-0.048	0.374**	-0.21	-0.264	- 0.356*	-0.232	- 0.317*	- 0.476**	- 0.348*	- 0.348*	-0.248
Days to male flower initiation			0.973**	-0.132	0.859**	- 0.380**	- 0.299*	- 0.441**	- 0.305*	- 0.473**	- 0.565**	- 0.475**	- 0.551**	- 0.463**
Days to female flower initiation				-0.139	0.900**	- 0.422**	- 0.340*	- 0.461**	- 0.333*	- 0.482**	- 0.586**	- 0.497**	- 0.567**	- 0.487**
Number of pistillate flowers					-0.191	0.102	0.308*	0.219	0.355*	0.525**	0.005	-0.019	0.26	0.583**
Days to maturity						- 0.491**	- 0.307*	- 0.417**	- 0.378**	- 0.506**	- 0.430**	- 0.406**	- 0.566**	- 0.517**
Number of branches at 8WAP							0.450**	0.581**	0.545**	0.459**	0.566**	0.553**	0.564**	0.427**
Vine Length at 8WAP								0.805**	0.776**	0.647**	0.635**	0.707**	0.807**	0.685**
Number of leaves at 8WAP									0.813**	0.679**	0.688**	0.670**	0.826**	0.728**
Leaf Area at 8WAP										0.660**	0.523**	0.573**	0.775**	0.698**
Number of fruits											0.428**	0.438**	0.720**	0.912**
Fruit length												0.860**	0.683**	0.398**
Fruit girth													0.670**	0.390**
Fruit weight														0.794**

Table 3. Latent vectors for the first four principal components of the genotypes

TRAITS	PC1	PC2	PC3	PC4
Days to emergence	0.08	-0.37	-0.03	-0.64
Days to 50%emergence	0.17	-0.30	0.04	-0.45
Days to female flower initiation	0.27	-0.37	-0.19	0.23
Days to male flower initiation	0.26	-0.39	-0.20	0.20
Days to maturity	0.26	-0.31	-0.27	0.29
Fruit girth(cm)	-0.27	0.04	-0.41	-0.05
Fruit length(cm)	-0.28	0.14	-0.36	0.07
Fruit weight(kg)	-0.33	-0.10	-0.05	0.04
Leaf area at 8weeks after planting(cm)	-0.28	-0.29	-0.08	0.02
Number of fruits	-0.29	-0.20	0.26	0.04
Number of branches at 8 weeks after planting	-0.24	-0.04	-0.18	-0.34
Number of leaves at 8 weeks after planting	-0.31	-0.17	-0.16	0.04
Number of pistillate flowers	-0.12	-0.29	0.53	0.21
Fruit yield per hectare tha^{-1}	-0.29	-0.24	0.30	0.07
Vine Length at 8weeks after planting (cm)	-0.29	-0.25	-0.18	0.16
Total Variation	51.53	13.91	10.08	7.35
Cumulative variation	51.53	65.44	75.52	82.87

strong positive correlation with total fruit yield per hectare. Notable negative correlations were found among days to emergence and days to 50% emergence, with fruit length, number of fruits, and fruit girth. This suggests the presence of direct relationships among these traits. Hence, it is apparent that choosing traits that exhibit a favorable correlation with yield would lead to an increase in the overall fruit yield per hectare. The findings of Praneetha *et al.* (2024), in snap melon *(Cucumis Melo. Var. Momordica*) corroborates with the current research. Golabadi *et al*. (2013), also reported strong positive associations between the quantity of fruits per plant, the length of the stem, and the fruit yield per hectare in cucumber. Ullah *et al*. (2012) reported a positive, and statistically significant correlation between the number of fruits per plant, fruit diameter, flesh thickness, and total fruit yield per hectare in cucumber. The results of the present study revealed a negative correlation between the time taken for male and female flowers initiation and fruit yield per hectare (**Table 2**). This suggested that fruit yield could be affected by reduction in duration of the crop. This outcome corroborated with the findings of Ogbodo *et al*. (2010), in cucumber. The number of fruits per plant and fruit weight showed a significant and positive correlation with fruit yield. This result agreed with the findings of Golabadi *et al*. (2013) in *C. sativus* . Given the significant correlation between fruit yield and many of the traits studied, it was appropriate to utilize PCA, as suggested by Abdi and Williams (2010), to examine the association between characters and measure the genetic diversity. Principal component analysis (PCA) is a statistical technique that reduces the dimensions of multivariate data by eliminating intercorrelations among the traits under study. This allows for the representation of multidimensional relationships on principal axes. The analysis of the genotype's latent vectors for the first four principal components (PCs) axes showed that only PC1 and PC2 had eigenvector values greater than one (**Table 3**). The cumulative variation of the first two PCs was 65.44%, with PC1 accounting for 51.53% of the variation and exhibited a strong positive correlation with the duration between female and male flower initiation as well as the time taken for maturity. PC2 accounted for 13.91% of the total variation and exhibited a positive correlation with fruit length and fruit girth. The average contributions of the reproductive components, including the number of days taken for female flower initiation, male flower initiation, and maturity, as well as the fruit length and fruit girth, were found to be high in PC1 and PC2 respectively. The result suggests that these traits were the major ones that explained most of the variations in the studied cucumber genotypes. The outcome aligns with the conclusions of Geetha and Divya, (2021) and Golabadi *et al*. (2013), who employed PCA in greengram (*Vigna radiata)* and cucumber (*Cucumis sativus*) respectively. Further the results also aligns with the conclusions drawn by Nwangburuka (2010) in Okra. Therefore, in PC1 and PC2, the primary factors contributing to the variability were days to female flower initiation, days to male flower initiation, days to maturity, fruit girth, and fruit length, as indicated by the scree plot (**Fig. 1**)

Based on the study, it is concluded that number of pistillate flowers, number of branches, vine length, number of leaves, leaf area, number of fruits per plant, fruit length, fruit girth, and fruit weight are the major yield

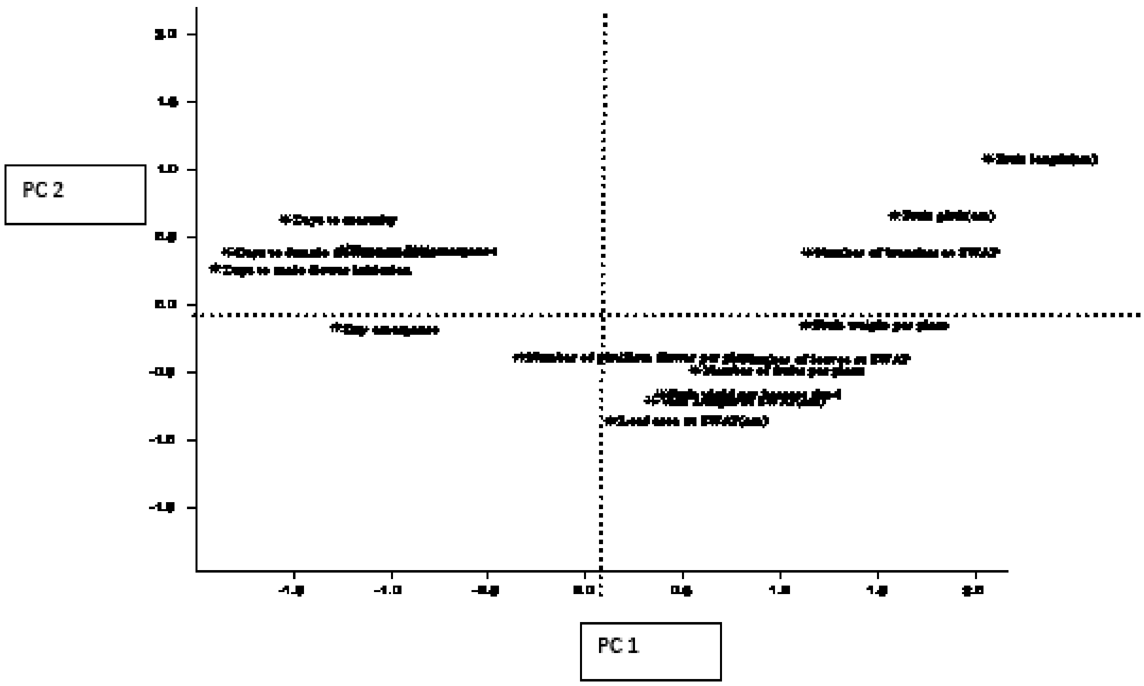

Fig. 1.Scree plot of principal components

PC1, PC2, PC3, and PC4 represent the first, second, third, and fourth principal component axes, respectively.

contributing characters and hence during selection, weightage should be given to these characters for the development of high yielding genotypes of cucumber. The analysis of the latent vectors of the genotypes for the first four principal components showed that only PC1 and PC2 had eigenvector values greater than one. The PCA revealed that the characters with the most significant impact on the variations observed in the studied cucumber genotypes were days to female-male flower initiations, days to maturity in PC1, fruit length and fruit girth in PC2, and hence these characters could be effectively used for selection among the tested entries for yield improvement program in cucumber. Therefore, in order to achieve high yields, it is crucial to prioritize inbreeding and selection based on these characters.

REFERENCES

Abdi, H. and Williams, L.J. 2010. Principle Component Analysis. Wiley Interdisciplinary, Reviews Computational Statistics, **2**: 433-459. [Cross Ref]

FAO, 2021. Production: Crops and livestock products. In: FAO. Rome. Cited March, 2022. https://www.fao.org/faostat/en/#data/QCL

Fayeun, L.S. Odiyi, A.C. Makinde, S.C.O. and Aiyelari, O.P. 2012. Genetic variability and correlation studies in the fluted pumpkin (*Telfairia occidentalis* Hook F.). *Journal of Plant Breeding and Crop Science,* **4**(10):156-160.

Geetha, K. and Divya, S. 2021. Principal component analysis of sorghum landraces for yield contributing traits. *Electronic Journal of Plant Breeding*, **12** (3): 1033 – 1036. [Cross Ref]

GenStat Release 10.3DE. 2011. Discovery Edition 4 VSN International Ltd. Rothamsted Experimental Station, Howel, Hempstead, UK.

Golabadi, M., Eghtedary, A.R. and Golkar, P.P. 2013. Determining relationships between different horticultural traits in (*Cucumis sativus* L.) Genotypes with multivariate analysis. *Sabrao J. Breeding and Genet.*, **45** (3): 447-457.

Kumar, S., Kumar, R., Gupta, R.K. and Sephia, R. 2011. Studies on correlation and path coefficient analysis for yield and its contributing traits in cucumber. *Crop Improvement,* **38**(1):18-23.

Machikowa, T. and Laosuwan, P. 2011. Path coefficient analysis for the yield of early maturing soybean. *Songklanakarin Journal of Science and Technology,* (33) 4:365-368.

Mahalingam, A. Manivanan, N. Bharathi Kumar, K. Ramakrishnan, P. and Vadivel, K. 2020. Character association and principal component analysis for seed yield and its contributing characters in greengram (*Vigna radiata* (L.) Wilczek). *Electronic Journal of Plant Breeding,* **11**(1): 259-262. [Cross Ref]

Nwangburuka, C.C. 2010. Morphological characterization and genetic studies in okra *Abelmoschus esculentus* (L.) Moench. Ph.D. Thesis. The Federal University of Agriculture, Abeokuta, Nigeria.

Ogbodo, E.N. Okorie, P.O. and Utobo, E.B. 2010. Introducing cucumber for cultivation at a new different zone in Ebonyi State, southeastern, Nigeria. *Libyan Agriculture Research Center Journal International,* (16) 1:336-343.

Praneetha, S., Muthuselvi and Kousalya R. 2024. Association studies on yield and yield components in snap melon (*Cucumis Melo. Var. Momordica*). *MadrasAgric.J.*,**111**:1-3. [Cross Ref]

Sharma, S., Kumar, R., Sharma, H.R., Sharma, A. and Gautam, N. 2018. Divergence Studies for Different Horticultural Traits in Cucumber (*Cucumis sativus* L.). *International Journal of Current Microbiology and Applied Sciences*, **7**(2):1733-1741. [Cross Ref]

Ullah, M.Z., Hassan, M.J., Chowdhury, A.Z.M.K.A., Saki, A.I. and Rahman, A.H.M.A. 2012. Genetic variability and correlation in exotic cucumber (*Cucumis sativus* L.) varieties energy PAC Agro Ltd. Manipur, Hotapara, Gazipur. *Bangladesh Journal of Plant Breeding and Genetics,* (25)1:17-23. [Cross Ref]