



Yuzuncu Yil University  
Journal of Agricultural Sciences  
(Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi)

<https://dergipark.org.tr/en/pub/yyutbd>



ISSN: 1308-7576

e-ISSN: 1308-7584

Research Article

**Genetic Variation, Heritability, Principal Component Analysis, Correlation and Path Coefficient Analysis in the Fruit Samples of *Sechium edule* (Jacq.) Sw. Genotypes**

Sanjay KUMAR<sup>\*1</sup>, Asikho KISO<sup>2</sup>, Lima ASENLA<sup>3</sup>

<sup>1</sup>Banaras Hindu University, Department of Botany, Varanasi 221005, India

<sup>2,3</sup>Nagaland University, Department of Botany, Lumami 798627, India

<sup>1</sup><https://orcid.org/0000-0003-4463-9159>, <sup>2</sup><https://orcid.org/0000-0003-1976-5751>, <sup>3</sup><https://orcid.org/0000-0002-8610-258X>

\*Corresponding author e-mail: [sanjaykumarnew.2021@gmail.com](mailto:sanjaykumarnew.2021@gmail.com)

**Article Info**

Received: 18.11.2021

Accepted: 02.03.2022

Online published: 15.03.2022

DOI:10.29133/yyutbd.1025466

**Keywords**

Correlation,  
Heritability,  
Path coefficient,  
Principal coefficient analysis

**Abstract:** Genetic diversity, heritability, the genetic advance of yield, and associated traits are some important criteria to generate some basic information related to the genetic improvement of crops. Some characters of *Sechium edule* (Jacq.) Sw. genotype fruits have been evaluated for their improvement purpose. Genotypes and fruit samples of *Sechium* were randomly collected for fruit traits such as length (FL), width (FW), circumference/girth (FC), and the number of ridges (FR) from the various parts of village Kigwema of Kohima district, Nagaland at a mean value of latitude (25.60690 N), longitude (94.34250 E) and altitude (1538 masl) for the purpose. Genotypes and fruit samples collection for trait study normally distributed in histogram plot and normality test. Analysis of variance (ANOVA) estimated significant differences in fruit sample traits. The phenotypic coefficient of variation (PCV) was greater than the genotypic coefficient of variation (GCV) for all the traits. The phenotypic and genotypic coefficient of variation was recorded maximum for trait fruit length, while maximum heritability was recorded for trait fruit circumference. High heritability and high genetic advance estimates for fruit circumference suggest that it could be considered for further improvement through various breeding programs. Principle component analysis (PCA) showed that fruit length and fruit ridges numbers are responsible for most of the variations observed in the fruit morphology and could be considered for its improvement. Fruit width recorded maximum for correlation coefficient direct value indicated towards effect on the fruit circumference and direct selection of the trait for its improvement.

**To Cite:** Kumar S, Kiso A, Asenla, L, 2022. Genetic Variation, Heritability, Principal Component Analysis, Correlation and Path Coefficient Analysis in the Fruit Samples of *Sechium edule* (Jacq.) Sw. Genotypes. *Yuzuncu Yil University Journal of Agricultural Sciences*, 32(1): 164-174.

DOI: <https://doi.org/10.29133/yyutbd.1025466>

***Sechium edule* (Jacq.) Sw Genotiplerinin Meyve Örneklerinde Genetik Varyasyon, Kalıtım, Temel Bileşen Analizi, Korelasyon ve Yol Katsayısı Analizi**

**Makale Bilgileri**

Geliş: 18.11.2021

Kabul: 02.03.2022

Online yayınlanma: 15.03.2022

DOI: 10.29133/yyutbd.1025466

**Özet:** Genetik çeşitlilik, kalıtsallık, verim ve verime bağlı özelliklerde genetik ilerleme, ürünlerin genetik iyileşmesi ile ilgili bazı önemli bilgileri üretmek için önemli kriterlerdir. Bu çalışma, *Sechium edule* (Jacq.) Sw. genotiplerine ait meyvelerin bazı özelliklerini iyileştirme amacıyla yapılmıştır. *Sechium edule* (Jacq.) Sw. genotiplerine ait meyveler, Nagaland Kohima iline bağlı Kigwema

**Anahtar Kelimeler**

Korelasyon,  
Kalıtım,  
Yol katsayısı,  
Temel katsayı analizi

köyünün (25.60690 N, 94.34250 E) çeşitli yerlerinden rastgele toplanmıştır. Genotiplerde; uzunluk (FL), genişlik (FW), çevre/çevre (FC) ve sırt sayısı (FR) özellikleri değerlendirilmiştir. Genotipler ve meyve örnekleri histogram grafiğinde ve normallik testinde normal olarak dağıtılmıştır. Varyans analizi (ANOVA) sonucunda, meyve özelliklerinde önemli farklılıklar tespit edilmiştir. Tüm özellikler için, fenotipik varyasyon katsayısı (PCV), genotipik varyasyon katsayısından (GCV) daha yüksek bulunmuştur. Fenotipik ve genotipik varyasyon katsayısı, meyve uzunluğunda, kalıtım derecesi ise meyve çevresinde maksimum değere ulaşmıştır. Meyve çevresi için yüksek kalıtsallık ve yüksek genetik ilerleme tahminleri, bunun çeşitli ıslah programları yoluyla daha fazla iyileştirme için kullanılabileceğini düşündürmektedir. Temel bileşen analizi (PCA), meyve morfolojisinde gözlemlenen varyasyonların çoğunda meyve uzunluğu ve meyve sırt sayılarından sorumlu olduğunu ve bunların gelişimi için değerlendirilebileceğini göstermiştir. Meyve eni, meyve çevresi ve gelişimi için bu özelliğin seçimine yönelik olarak belirtilen korelasyon katsayısı doğrudan değer için maksimum kaydedilmiştir.

**1. Introduction**

Nature fulfills every need of human beings and grows various plants with economic importance. Most of the plants in nature are unexplored and gathering information on nature and its variation is an important criterion to select a genotype with desirable traits for a successful plant breeding program (Dyulgerova and Valcheva, 2014; Ipek and Balta, 2020). *Sechium edule* (Jacq.) Sw., Cucurbitaceae, is a nutritious crop, mostly uses matured fruit as vegetables and boiled salads. The demands of various cucurbitaceous vegetables by the people in the society caused to increase the production of such important crops. Therefore to achieve the target, information on genetic variability, heritability, the genetic advance of yield, and associated traits is essentially important.

Regression is an estimation of the linear relationship between a dependent variable and one or more independent variables (Vrbik, 2018). Also, it is a test method to observe data is approximately normally distributed (Dimitrova et al., 2020). Plotting a histogram of the variables of interest indicated the shape of normally distributed data, and test of normality is performed and confirmed through Q-Q plots and Kolmogorove-Smirnov and Shapiro-Wilk's W test in SPSS ver. 16 (Kolmogorove, 1933; Smirnov, 1948; Shapiro and Wilk, 1965).

Analysis of variance (anova) is generally used to ensure whether the means of two or more groups of samples are significantly differ or not from each other (Turkheimer, 2015). Heritability and agronomic characteristics are important traits that may be utilized to estimate and improve the yield of crops successfully. Falconer and Mackey (1996) defined heritability as the measure of correspondence between breeding values and phenotypic values. Heritability plays a predictive role in breeding and expresss its dependability on phenotype, which guides its breeding value (Wray and Visscher, 2008). Genetic advance is a direct relationship between heritability and response to selection. High genetic advance with high heritability estimates the most effective condition for selection (Moore and Shenk, 2017). Heritability is, therefore important, and its application increases when a genetic advance is calculated as it indicates the gain in character obtained under particular selection pressure. Therefore, genetic advance is an important selection parameter that helps in the selection breeding program.

The correlation procedure calculates the correlation between variables and measures the strength of the linear relationship between two variables. The knowledge of the relationship between genetic and phenotypic data is important and valuable when the traits are considered for selection.

Path coefficient analysis allows an effective means of partitioning correlation coefficients into the unidirectional pathway and alternative pathways. This analysis permits a critical examination of specific factors that produce a given correlation and can be successfully employed in formulating an effective strategy (Okuyama et al., 2004).

Principal component analysis (PCA) is a mathematical procedure to transform, possibly a large number of correlated variables into a smaller number of uncorrelated variables, usually known as principal components (Chatfield and Collis, 1980). The first principal component (PC1) accounts for as much of the variability in the data as possible, and each succeeding component (PC2, PC3, PC4 etc.) accounts for as much of the remaining variability as possible. The objective of the PCA is to discover

or reduce the dimensionality of the data set and to identify new meaningful causal variables (Jolliffe, 2002).

The purpose of cluster analysis is to discover a system of organizing observations where members of the group share specific properties in common. Cluster analysis is a class of techniques that classifies cases into groups that are relatively homogeneous within themselves and relatively heterogeneous between each other (Yim and Ramdeen, 2015).

So, the purpose of the present study is to generate some basic information, using standard statistical methods such as analysis of variation, correlation, regression, path correlation coefficient analysis, principal component analysis, cluster, and factor analysis to differentiate and assess the variations, heritability (broad sense), genetic advance, yield and contribution through direct and indirect effects of some important traits on the total variation of genotypes for the further genetic improvement of the crop.

## 2. Material and Methods

### 2.1 Plant materials

The twelve genotypes of *Sechium* were randomly collected from randomized complete block design at Kigwema village of Kohima district, Nagaland, with a mean value of latitude (25.60690 N), longitude (94.34250 E), and altitude (1538 masl). The mean values of 5 fruits from each genotype of 3 randomized complete block designs were used for quantitative traits such as length, width, circumference/girth, and the number of ridges for further analysis.

### 2.2 Methods

#### 2.2.1 Regression and data normality test

Regression analysis is a statistical technique applied to draw a relationship among interrelated variables and analyzed using SPSS ver. 16 as suggested by Landau and Everitt (2004). The genotypes fruit samples quantitative traits were regressed against the genotypes, and the histogram plot suggested the normal distribution and shape of data. Test of normality was performed in SPSS ver. 16 and suggested the Shapiro-Wilk test, which is sensitive and more appropriate for smaller samples than the Kolmogorove-Smirnov test.

#### 2.2.2 Analysis of variance

Analysis of variance (ANOVA) was carried out using procedures explained for SPSS ver. 16 by Landau and Everitt (2004). Genotypic and phenotypic coefficients of variability were estimated using the formula given by Burton and Devane (1953).

$$PCV = \frac{\sqrt{\sigma^2 p}}{\bar{x}} \times 100 \quad (1)$$

$$GCV = \frac{\sqrt{\sigma^2 g}}{\bar{x}} \times 100 \quad (2)$$

Where,

PCV=phenotypic coefficient of variation

GCV=genotypic coefficient of variation

$\sigma^2 p$ = phenotypic variance=  $\sigma^2 g + \sigma^2 e$

$\sigma^2 g$ = genotypic variance= (mean sum of squares due to genotypes – error mean sum of squares) ÷ replications

$\sigma^2 e$ =environmental variance= (error mean sum of squares)÷replications

$\bar{x}$ =general mean

PCV and GCV values were classified low (<10%), moderate (10-20%), and high (>20%) as suggested by Sivasubramanian and Menon (1973).

### 2.2.3 Heritability (broad sense), genetic advance, and genetic advance mean percent

Heritability in the broad sense ( $H^2$ ) was estimated as per the formula suggested by Allard (1960).

$$H^2 = \frac{\sigma^2_g}{\sigma^2_p} \times 100 \quad (3)$$

The heritability was categorized as low (0-30%), medium (31-60%), and high (61% and above) according to Johnson et al., 1955.

Genetic advance estimated according to Allard (1960).

$$GA = K \times \sigma_p \times H^2 \quad (4)$$

Where,

$K$ =selection differential at 5% selection intensity, which accounts for a constant value 2.06

$\sigma_p$ =phenotypic standard deviation

Genetic advance over mean (GAM) is expressed in percentage and calculated using the following formula.

$$GAM = \frac{GA}{\bar{x}} \times 100 \quad (5)$$

Genetic advance as percent over mean is categorized as low (<10%), moderate (10-20%) and high (>20%) according to Johnson et al., 1955.

### 2.2.4 Correlation and Path coefficient analysis

The Pearson correlation was performed for the quantitative traits using SPSS ver. 16 and followed Okuyama et al., 2004. Path coefficient analysis for the quantitative traits was performed in MS-Excel spreadsheet using an inbuilt statistical software package and followed the procedure of Akintunde (2012). Correlation coefficients were calculated for their direct and indirect component parts and represented in a graphical form.

### 2.2.5 Principal component analysis or Factor analysis

Principal component analysis (PCA) or factor analysis was performed using SPSS ver. 16, and followed the procedure suggested by Chatfield and Collis (1980). PCA is used to calculate the initial and extracted communalities suggested the total variation by 2 major components in the traits.

### 2.2.6 Cluster analysis

Hierarchical cluster analysis is performed for grouping the fruit sample traits and followed the procedure of Everitt et al., 2011. Cluster analysis using ward method dendrogram suggested the 2 groups of quantitative traits of fruit samples

## 3. Results

The quantitative traits data of *Sechium* fruit samples are normally distributed as suggested by the histogram plot with normal bell-shape and test of normality where Shapiro-Wilk statistics are more towards normal distribution than Kolmogorov-Smirnov statistics at  $p \leq 0.05$  (Karney, 2016). Shapiro-Wilk test is more sensitive towards outliers and smaller data samples (Shore, 2012 and 2011). Therefore, normality could be assumed for the data; as a result, any other test assumptions may be satisfied, and finally, an appropriate parametric test can be used (Figure 1 and Table 1).

Analysis of variance (ANOVA) was performed for the quantitative traits of *Sechium* genotypes fruit traits, and mean value was recorded high for FC ( $21.7 \pm 0.32$ ) followed by FL ( $11.5 \pm 0.31$ ), Fwd ( $8.51 \pm 0.15$ ), and FR ( $4.73 \pm 0.11$ ) respectively are significantly differs from each other at the probability level of  $p \leq 0.05$  (Table 2).

Table 1. Test of normality for quantitative data set of fruit samples

	FL	Fwd	FR	FC
<b>Mean</b>	11.5	8.51	4.73	21.7
<b>Standard Error</b>	2.43	1.21	0.88	2.53
<b>Kolmogorove-Smirnov</b>	0.473	0.260	0.212	0.378
<b>p-value<sub>K-S</sub></b>	0.001	0.007	0.009	0.019
<b>Shapiro-Wilk</b>	0.552	0.788	0.867	0.644
<b>p-values<sub>W</sub></b>	0.000	0.003	0.006	0.002

FL=fruit length, Fwd=fruit width, FR=fruit ridges number, FC=fruit circumference/girth.

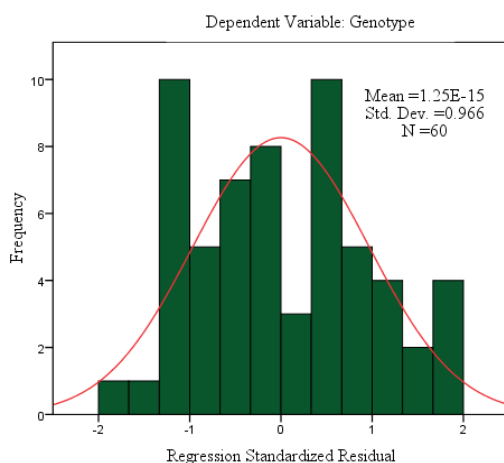


Figure1. Genotype and fruit characters histogram plot of approximately normally distributed data.

Table 2. Analysis of variance of *Sechium* fruit samples traits

Quantitative traits of fruit samples	FC	FL	Fwd	FR
Mean square within groups (error)	0.71	0.88	0.28	0.22
Mean square between groups (genotypes)	31.29	27.87	6.67	3.17
Replicates	5	5	5	5
Standard deviation	0.32	0.31	0.15	0.11
Minimum	18	6	6	3
Maximum	26	17	11	7
Range	8	11	5	4
Variance	6.417	5.915	1.474	0.775

FC=fruit circumference/girth, FL=fruit length, Fwd=fruit width, FR=fruit ridges number.

The estimates of the phenotypic and genotypic variances showed similar trend as of the mean value from ANOVA (Table 3).

Table 3. Estimates of variance, heritability and related components

Fruit traits	Vp	Vg	Ve	Vg×e	PCV%	GCV%	ECV%	H <sup>2</sup>	GA	GAM%
FC	6.25	6.11	0.14	0.85	11.52	11.39	1.72	0.97	5.05	23.27
FL	5.56	5.39	0.17	0.91	20.50	20.18	3.58	0.96	4.80	41.73
Fwd	1.32	1.27	0.05	0.06	13.50	13.24	2.62	0.96	2.39	28.08
FR	0.63	0.59	0.04	0.02	16.78	16.23	4.22	0.93	1.68	35.51

Vp=phenotypic variance, Vg=genotypic variance, Ve=environmental variance, Vg×e=interaction of genotype and environment variance, PCV%=phenotypic coefficient of variance percent, GCV%= genotypic coefficient of variance percent, ECV%= environment coefficient of variance percent, H<sup>2</sup>= heritability (broad sense), GA=genetic advance, GAM%=genetic advance mean percent, FC=fruit circumference/girth, FL=fruit length, Fwd=fruit width, FR=fruit ridges number.

Pearson's correlation of quantitative traits of fruit samples is significant at the level of probability of significance  $p \leq 0.01$  (2-tailed), and all traits are positively associated with each other (Table 4).

Table 4. Pearson's correlation for the association of quantitative fruit sample traits

Fruit traits	FC	FL	Fwd	FR
FC	1			
FL	0.600**	1		
Fwd	0.949**	0.646**	1	
FR	0.609**	0.554**	0.591**	1

\*\*Correlation is significant at the 0.01 level (2-tailed).

FC=fruit circumference/girth, FL=fruit length, Fwd=fruit width, FR=fruit ridges number.

For the estimation of the partial regression coefficients or direct path coefficients, fruit trait, FC was considered as a dependent (resultant) and FL, Fwd, and FR independent (causal) variables (Table 5).

Table 5. Partial regression coefficients or direct path coefficients of causal factors

	Coefficients	S.E.	t-Stat	P-value	Lower 95%	Upper 95%
FC(Intercept)	-0.00544	0.040600752	0.133987677	0.893892696	-0.08677308	0.075893079
FL(X1)	-0.048746146	0.055804222	0.873520747	0.386109837	-0.160535435	0.063043143
Fwd(X2)	0.927212549	0.057432332	16.14443509	9.99748E-23	0.812161764	1.042263333
FR(X3)	0.087487328	0.052839165	1.655728817	0.103370677	-0.018362238	0.193336894

Correlation coefficients into its direct and indirect component parts and path analysis and contributions of related components towards fruit sample trait, FC were estimated. Pearson correlation values are the combination of direct and indirect components of the variable relationship. The direct or indirect relationship of variables is separated out and used to draw a direction of the path and their contributions to the overall growth of the crop (Table 6 and Figure 2).

Table 6. Correlation coefficients into its direct and indirect component parts

<b>FC vs. FL</b>	
Direct effect of FL on FC	-0.048
Indirect effect of FL via FL on FC	-0.048 [1.000×(-0.048)]
Indirect effect of FL via Fwd on FC	0.598 (0.646×0.927)
Indirect effect of FL via FR on FC	0.048 (0.554×0.087)
Total indirect effects	0.598
Total correlation (direct+indirect effects)	0.550
<b>FC vs. Fwd</b>	
Direct effect of Fwd on FC	0.927
Indirect effect of Fwd via FL on FC	-0.031 [0.646×(-0.048)]
Indirect effect of Fwd via Fwd on FC	0.927 (1.000×0.927)
Indirect effect of Fwd via FR on FC	0.051 (0.591×0.087)
Total indirect effects	0.02
Total correlation (direct+indirect effects)	0.947
<b>FC vs. FR</b>	
Direct effect of FR on FC	0.087
Indirect effect of FR via FL on FC	-0.026 [0.554×(-0.048)]
Indirect effect of FR via Fwd on FC	0.547 (0.591×0.927)
Indirect effect of FR via FR on FC	0.087 (1.000×0.087)
Total indirect effects	0.521
Total correlation (direct+indirect effects)	0.608

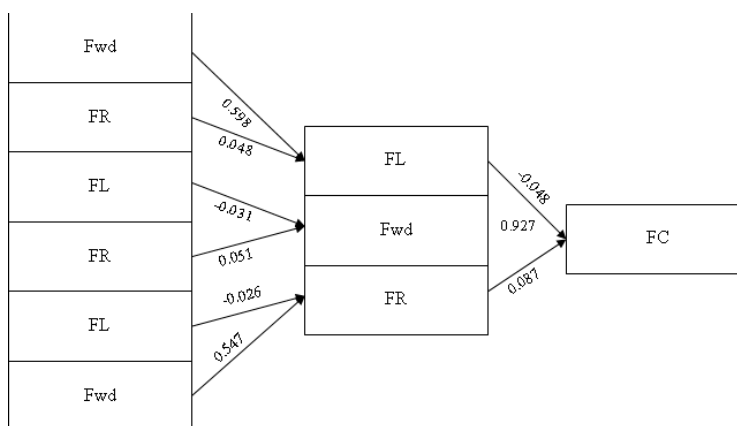


Figure 2. Path analysis and contributions of related components towards fruit sample trait, FC (MS publisher of MS word ver. 2010 was used to draw the figure).

Table 7. Anti image correlation and KMO and Bartlett’s test for the measure of sampling adequacy

Anti-image correlation				KMO and Bartlett’s test		
	FC	FL	Fwd	FR	Kaiser-Meyer-Olkin Measure of Sampling Adequacy	
					0.718	
<b>FC</b>	<b>0.649<sup>a</sup></b>	0.116	-0.907	-0.216	Bartlett's Test of Sphericity	Approx. Chi-Square 194.307
<b>FL</b>	0.116	<b>0.849<sup>a</sup></b>	-0.303	-0.297		df 6
<b>Fwd</b>	-0.907	-0.303	<b>0.645<sup>a</sup></b>	0.045		Sig. 0.000
<b>FR</b>	-0.216	-0.297	0.045	<b>0.883<sup>a</sup></b>		

Test of normality suggested that data are normally distributed and fit for test assumptions and parametric tests (Table 1). Also, in order to whether the present data are suitable for principle component analysis/factor analysis or not, anti image correlation, as well as Kaiser-Meyer-Olkin (KMO) and Bartlett’s test, was performed. anti-image correlation suggested the value more than 0.600<sup>a</sup> (diagonal value) and suitable for the measure of sample adequacy. Similarly, results are obtained in KMO and Bartlett’s test (0.718 and *p*-value=0.000) for the measure of sample adequacy (Table 7).

The principal component is an important multivariate analysis technique used to examine the association between characters and measures genetic diversity (Table 8-9 and Figure 3). Hierarchical cluster dendrogram using ward method for fruit characters were estimated (Figure 4).

Table 8. Principal component analysis and initial and extracted communalities

	Communalities		Component matrix <sup>a</sup>		Sum of squared loadings		Communalities Extracted (sum)
	Initial	Extraction	1	2	1	2	
<b>FC</b>	1	0.966	0.927	-0.326	0.859=(0.927) <sup>2</sup>	0.106=(-0.326) <sup>2</sup>	0.965
<b>FL</b>	1	0.719	0.801	0.278	0.641=(0.801) <sup>2</sup>	0.077=(0.278) <sup>2</sup>	0.718
<b>Fwd</b>	1	0.975	0.935	-0.318	0.874=(0.935) <sup>2</sup>	0.101=(-0.318) <sup>2</sup>	0.975
<b>FR</b>	1	0.847	0.784	0.481	0.614=(0.784) <sup>2</sup>	0.231=(0.481) <sup>2</sup>	0.845
		<b>3.507</b>			<b>2.988</b>	<b>0.515</b>	<b>3.503</b>

Extraction method: a principal component analysis.  
<sup>a</sup>2 components extracted.

Table 9. Factor analysis (2 components) and total variation explained

Component	Initial Eigenvalues			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
<b>1</b>	2.991	74.767	74.767	<b>2.991</b>	74.767	74.767
<b>2</b>	0.516	12.902	87.669	<b>0.516</b>	12.902	87.669
<b>3</b>	0.446	11.141	98.810			
<b>4</b>	0.048	1.190	100.00			

Extraction method: Principal component analysis.

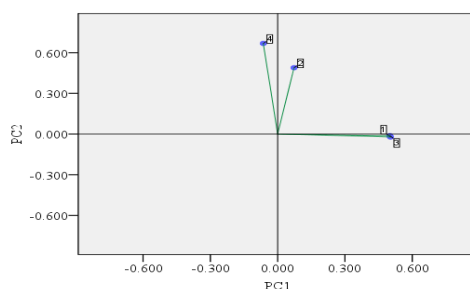


Figure 3. Principal component axes for fruit characters Rescaled Distance Cluster Combine.

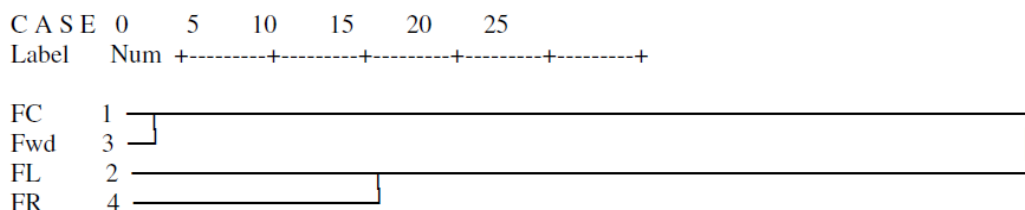


Figure 4. Hierarchical cluster dendrogram using ward method for fruit characters.

#### 4. Discussion

Data are normally distributed and could be assumed for an appropriate parametric test as revealed from the normality test (Table 1 and Figure 1). The significant differences among and between the traits from ANOVA suggest the presence of genetic variation and provide an opportunity for the consideration of fruit samples in plant breeding improvement programs (Gelman, 2005; Dag et al., 2018). The maximum phenotypic ( $V_p$ ) and genotypic ( $V_g$ ) variances was estimated for FC ( $V_p=6.25$ ;  $V_g=6.11$ ) followed by FL ( $V_p=5.56$ ;  $V_g=5.39$ ), Fwd ( $V_p=1.32$ ;  $V_g=1.27$ ) and FR ( $V_p=0.63$ ;  $V_g=0.59$ ) respectively. The minimum  $V_p$  was recorded for the trait FR. As expected, the phenotypic coefficient of variation (PCV) was greater than the genotypic coefficient variation (GCV) for all the traits. A larger difference between PCV and GCV values suggests the greater environmental effects on the traits. The PCV and GCV estimate the nature and magnitude of variation and explain the variation either due to genetic or environmental causes. It was estimated maximum for FL (pcv%=20.50; gcv%=20.18) followed by FR (pcv%=16.78; gcv%=16.23), Fwd (pcv%=13.50; gcv%=13.24) and FC (pcv%=11.52; gcv%=11.39) respectively. All the traits showed smaller differences, and minimum PCV and GCV value was estimated for FC. Heritability is the proportion of genetic variance in phenotypic variance that gives information about the inheritance of traits. Traits with high heritability are easy to improve through selection, and it is recorded high for all traits in the study. The highest heritability was estimated for FC (97%), followed by FL (96%), Fwd (96%), and FR (93%), respectively. Similar results have been reported in previous studies (Turkheimer, 2011; Tester and Langridge, 2010). It is also considered that high heritability does not indicate high genetic gain always. Therefore, both heritability and genetic advance are considered together for prediction and their final effects on traits for selection (Luby et al., 2015; Johnson et al., 2011). Fruit circumference, FC showed high heritability coupled with high genetic advance. High heritability coupled with high genetic advance suggests the additive effects on the control of the particular trait (Hartung and Schiemann, 2014; Lipi et al., 2020). On the other hand, high heritability with low genetic advance indicates the non-additive effects on the control of a particular trait (Heckerman et al., 2016; Ning et al., 2020). High heritability and high genetic advance for FC suggest a good trait to consider for selection and improvement in plant breeding programs (Table 3).

Quantitative traits of fruit samples are significant at the level of probability of significance  $p \leq 0.01$  (2-tailed), and all traits are positively associated with each other and indicate towards possible contribution in trait improvement. The dependent variable and independent variables in the correlation matrix for fruit sample traits are represented as Y and X1, X2, and X3, respectively (Mahdavi, 2013; Székely et al., 2007) (Table 4).



For the estimation of the partial regression coefficients, fruit sample trait, FC was considered as a dependent (resultant) and FL, Fwd, and FR independent (causal) variables. FL showed a small but negative (-0.048) and negligible direct effect on the FC. The indirect effect on FC through Fwd and FR is positive, where the Fwd effect is in considerable magnitude than FR. The total sum value of the correlation (0.550) is positive and significant at  $p \leq 0.01$ , which suggests that indirect selection of the trait through Fwd could be useful. Fwd showed the highest and positive (0.927) direct effect on FC. The indirect effects are negative (-0.031) through FL and positive (0.051) for FR. Both the indirect effects are small and negligible in magnitude. The total sum value of correlation (0.947) is positive and highly significant at  $p \leq 0.01$ , which suggests that direct selection of the trait could be useful. FR showed a positive (0.087) but small and negligible direct effect on the FC. The indirect effect is positive (0.547) and negative (-0.026) through Fwd and FL, respectively. The indirect effect through Fwd is considerable magnitude than FL. The total sum value of correlation (0.521) is positive and significant at  $p \leq 0.01$ , which suggests that indirect selection of the trait could be useful (Tarka, 2017; Bentler and Chih-Ping, 2016). The partial regression coefficients or direct path coefficients, direct and indirect component parts are presented in Table 5-6 and Figure 2.

The principal component is an important multivariate analysis technique used to examine the association between characters and measures the genotype/varieties genetic diversity (Esposito et al., 2007; Muradoglu et al., 2021). The result of PCA for the fruit traits of *Sechium* genotypes showed that the first two components explained 74.76% and 87.66% of the cumulative variation for fruit populations, respectively. The first component (PC1) accounts for 74.76% of total variation and is positively and highly associated with all the traits in the study called the fruit component. The second component (PC2) explained 12.90% of total variation and was positively associated with FL and FR while negatively associated with FC and Fwd (Table 7-8). The genetic variation of quantitative traits based on multivariate analysis using PCA suggests that FL and FR as the most important trait for explaining variability in the fruits of *Sechium* genotypes followed by FC and Fwd. The contribution of FL and FR was observed high in the principal component axes (Figure 3). The observation suggests that FL and FR are the major traits that explain the total variation in the fruit morphology of the *Sechium* genotypes and may be considered for further improvement. Similar results were reported in other crops by various authors (Rosso and Pagano, 2005; Chandran and Padya, 2000).

Hierarchical Cluster or dendrogram analysis of the fruit traits using ward method measured the interval with squared euclidean distance showed traits classified into 2 groups suggest the contribution towards other groups (Figure 4).

## Conclusion

Fruit circumference (FC) was recorded with maximum heritability (broad sense), and genetic advance could be recommended for further improvement of the trait. Fruit length (FL) and fruit ridges (FR) numbers are responsible for fruit morphological trait variations as indicated by principal component analysis.

## Acknowledgements

Both the authors are thankful to the people of Kigwema village, Kohima, for fruit sample collection and its study. Author SK conceived, designed, analyzed, performed, and wrote the paper. Author AK collected and contributed to data collection.

## References

- Akintunde, A. N. (2012). Path analysis step by step using Excel. *Journal of Technical Science and Technologies*, 1(1), 9-15.
- Allard, R. W. (1960). *Principles of plant breeding*. New York: John Willey and Sons Inc.
- Bentler, P. M., & Chih-Ping, C. (2016). Practical Issues in Structural Modeling. *Sociological Methods and Research*, 16(1), 78–117.
- Burton, C. W., & Devane, E. H. (1953). Estimating heritability in tall Festuca (*Restuca arundinaceae*) from donar material. *Agron J.*, 45, 1476-1481.

- Chandran, K., & Padya, S. M. (2000). Morphological characterization of *Arachis* species of section *Arachis*. *Plant Genetic Resources Newsletter*, 121, 38-41.
- Chatfield, C., & Collis. (1980). *Introduction to multivariate analysis*. Boca Raton: CRC Press.
- Dag, O., Dolgun, A., & Konar, N. M. (2018). one way tests: An R Package for One-Way Tests in Independent Groups Designs. *The R Journal*, 10 (1): 175–199.
- Dimitrova, D. S., Kaishev, V. K., & Tan, S. (2020). Computing the Kolmogorov–Smirnov distribution when the underlying cdf is purely discrete, mixed or continuous. *Journal of Statistical Software*, 95 (10), 1–42.
- Dyulgerova, B., & Valcheva, D. (2014). Heritability, variance components and genetic advance of yield and some yield related traits in Barley doubled haploid lines. *Turkish Journal Agricultural Natural Science*, 1 (special issue), 614-617.
- Esposito, M. A., Martin, E. A., Cravero, V. P., & Cointy, E. (2007). Characterization of pea accessions by SRAPs markers. *Scientia Horticulturae*, 113, 329-335.
- Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis: Wiley series in probability and statistics*. John wiley and sons. p.352.
- Falconer, D. S., & Mackey, T. F. C. (1996). *Introduction to quantitative genetics*. New York: Longman.
- Gelman, A. (2005). Analysis of variance? Why it is more important than ever. *The Annals of Statistics*, 33, 1–53.
- Hartung, F., & Schiemann, J. (2014). Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU. *The Plant Journal*, 78 (5), 742–752.
- Heckerman, D., Gurdasani, D., Kadie, C., Pomilla, C., Carstensen, T., Martin, H., Ekoru, K., Nsubuga R. N., Ssenyomo, G., Kamali, A., Kaleebu, P., Widmer, C., & Sandhu, M. S. (2016). Linear mixed model for heritability estimation that explicitly addresses environmental variation. *Proceedings of the National Academy of Sciences of the United States of America*, 113(27), 7377–7382.
- Ipek, P., & Balta, F. (2020). Fruit properties of rose hip (*Rosa* spp.) genotypes selected from Akkus, Ordu Province. *Yuzuncu Yil University Journal of Agricultural Sciences*, 30 (2), 338-344. DOI:10.29133/yyutbd.680453
- Johnson, H. W., Robinson, H. F., & Comstock, R. E. (1955). Estimation of genetic and environmental variability in soybean. *Journal of Agronomy*, 47, 314-318.
- Johnson, W., Penke, L., & Spinath, F. M. (2011). Understanding Heritability: What it is and What it is Not. *European Journal of Personality*, 25 (4), 287–294.
- Jolliffe, I. (2002). *Principal component analysis*. Springer series in statistics (2<sup>nd</sup> edition). New York, USA: Springer.
- Karney, C. F. F. (2016). Sampling exactly from the normal distribution. *ACM Transactions on Mathematical Software*, 42, 1–14.
- Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari.*, 4, 83–91.
- Landau S, & Everitt B S. (2004). *A handbook of statistical analysis using SPSS*. Boca Raton: Chapman and Hall/CRC- A CRC press company, London, New York, p.300.
- Lipi, L. F., Hasan M. J., Akter A., Quddus M. R., Biswas P. L., Ansari A., & Akter S. (2020). Genetic variation, heritability and genetic advance in some promising rice hybrids. *SAARC Journal Agriculture*, 18(2), 39-49.
- Luby, C. H., Kloppenburg, J., Michaels, T. E., & Goldman, I. L. (2015). Enhancing Freedom to Operate for Plant Breeders and Farmers through Open Source Plant Breeding. *Crop Science*, 55 (6), 2481–2488.
- Mahdavi, D. B. (2013). The Non-Misleading Value of Inferred Correlation: An Introduction to the Cointelation Model. *Wilmott Magazine*, 67, 50–61.
- Moore, D. S., & Shenk, D. (2017). The heritability fallacy. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1–2), e1400.
- Muradoglu, F., Gursoy, S., & Guler, E. (2021). Multivariate analysis revealed the morphological variability among *Crataegus* species. *Yuzuncu Yil University Journal of Agricultural Sciences*, 31 (4), 961-972. DOI:10.29133/yyutbd.9745388
- Ning, Z., Pawitan, Y., & Shen, X. (2020). High-definition likelihood inference of genetic correlations across human complex traits. *Nature Genetics*, 52 (8), 859–864.

- Okuyama, L. A., Federizzi, L. C., & Neto, J. F. B. (2004). Correlation and path analysis of yield and its components and plant traits in wheat. *Ciencia Rural* (Santa Maria), 34(6), 1701-1708.
- Rosso, B., & Pagano, E. (2005). Evaluation of introduced and naturalized populations of red clover (*Trifolium pretense* L.) at Pergamino EEA-INTA, Argentina. *Genetic Resources and Crop Evolution*, 52, 507-511.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika*, 52(3-4), 591-611.
- Shore, H. (2011). Response Modeling Methodology. *WIREs Comput Stat.*, 3 (4), 357–372.
- Shore, H. (2012). Estimating Response Modeling Methodology Models. *WIREs Comput Stat.*, 4 (3), 323–333.
- Sivasubramanian, S., & Menon, M. (1973). Heterosis and inbreeding depression in rice. *Madras Agric. J.*, 60, 1139.
- Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19 (2), 279–281.
- Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing independence by correlation of distances. *Annals of Statistics*, 35 (6), 2769–2794.
- Tarka, P. (2017). An overview of structural equation modeling: Its beginnings, historical development, usefulness and controversies in the social sciences. *Quality & Quantity*, 52 (1), 313–354.
- Tester, M., & Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science*, 327 (5967), 818–822.
- Turkheimer, E. (2011). Still missing. *Research in Human Development*, 8 (3–4), 227–241.
- Turkheimer, E. (2015). Genetic Prediction. *The Hastings Center Report*, 45 (5 Suppl.), 32–38.
- Vrbik, J. (2018). Small-Sample Corrections to Kolmogorov–Smirnov Test Statistic. *Pioneer Journal of Theoretical and Applied Statistics*, 15 (1–2), 15–23.
- Wray, N., & Visscher, P. (2008). Estimating Trait Heritability. *Nature Education*, 1 (1), 29.
- Yim, O., & Ramdeen, K. T. (2015). Hierarchical cluster analysis: comparison of three linkages measures and application to psychological data. *The quantitative methods for Psychology*, 11(1), 8-21.