Electronic Journal of Plant Breeding

Research Article

Association analysis and principal component analysis in green gram (*Vigna radiata* (L.) Wilczek) genotypes for yield and its contributing traits

R. Aravinth[1], N.B. John Kingsly[1*], D. Wilson[1], Adlin Pricilla Vasanthi[2] and P. Dinesh Kumar[3]

[1]Division of Genetics and Plant Breeding, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore, India.
[2] Division of Entomology, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore, India.
[3] Division of Agricultural Statistics, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore, India.
*E-Mail: johnkingslyjk@gmail.com

Abstract
A study was carried out to assess character association and genetic diversity among green gram genotypes. The results of correlation coefficient analysis revealed a significantly positive correlation for the traits *viz*., plant height, number of primary branches, number of clusters per plant, pod length, and number of seeds per pod with single plant yield. The path analysis revealed that the number of pods per plant, number of seeds per pod, and hundred seed weight had the highest positive direct contribution to single plant yield. Pod length showed a low indirect contribution to the total variance and all other traits showed a negative contribution to total variance.

Keywords: Green gram, Correlation, Path analysis, Principal component analysis (PCA)

INTRODUCTION

Green gram is a self-pollinated species having 2n=2x=22 chromosomes. Its seeds are rich in protein (25–28%) and contain carbohydrates (60-65%), zinc, phosphorus, vitamins A, magnesium, vitamin C, calcium, iron, potassium and folate (Udayasri *et al.,* 2022). This resilient legume thrives in arid, high-temperature conditions and aids in enhancing soil fertility by providing approximately 30 to 40 kilograms of nitrogen per hectare. As a result, the subsequent crop (often a grain) needs around 25 per cent lesser than the recommended nitrogen. (Mbeyagala *et al*., 2017).

The environment has a significant impact on yield, which is a complicated trait. Green grams' unpredictable growth, asynchronous maturity and losses due to pests and diseases lead to lower yield. Their low yields result from various factors, such as cultivation in less fertile soil, insufficient use of fertilizers, and the absence of improved varieties or hybrids. To meet the increasing demand for green gram, it is crucial to address these yield limitations by developing high-yielding genotypes that are resistant to biotic and abiotic stresses.

India's green gram production spans 5.13 million hectares, yet its productivity remains modest, averaging 601 kg/ha (Bindu *et al*., 2023). This underscores the considerable opportunity to enhance productivity through the development of high-yielding varieties endowed with

resistance to diseases and pests, as well as enriched nutritional profiles. Given that yield is a multifaceted trait shaped by numerous related characteristics like plant height, days to fifty per cent flowering, number of pods per plant and number of clusters per plant, prioritizing the selection of yield component traits exhibiting a robust relationship with overall yield can significantly contribute to its improvement. On the other hand, genetic erosion and susceptibility to pests and diseases mean that these genotypes cannot be utilized for a long time. In order to satisfy the increasing demand for pulses, it is imperative to elevate production and productivity through enhanced management practices, improved variety selection, the utilization of advanced genotypes, and ongoing crop improvement efforts. In biometrical genetics, diversity analysis is a useful technique for evaluating genetic diversity in experimental populations. Estimating the genetic distance within the population for the desired characteristic is crucial for choosing the parents for hybridization.

Establishing a relationship between two variables is called correlation. Phenotypic correlation is the correlation observed between average values of traits across environments. Genotypic correlation refers to the correlation between genetic influences on a trait, specifically the effects they have, without being influenced by the effects of specific experimental plots. Correlation is either, positive or negative, depending on the moment of the variables. When two variables are correlated positively, improving one character simultaneously improves the other. A negative correlation exists when the values of one variable increase while the values of the other variable decreases.

Path analysis is used to estimate the strength and importance of proposed causal relationships between variables, divided into direct and indirect effects. Studies on correlation primarily considers association between two traits that are related but do not alter the impact or cause of the relationship between those two traits. According to Wright (1921), path analysis is a useful statistical tool for quantifying relationships between components and their effects on crop output.

The current study sought to analyse the genetic diversity of green gram genotypes and to choose the best genotypes for the use of parent varieties in future breeding programs.

MATERIALS AND METHODS
Ninety-four green gram genotypes, sourced from diverse locations across India, were employed in this study. The experiment was laid out at Karunya Institute of Technology and Sciences, Coimbatore during October 2022. The site is situated at 11° 56' N latitude and 76° 44' E longitude, with an elevation of 467 meters above Mean Sea Level. The experiment was laid out in Randomized Block Design, replicated three times with a row length of 2.5 meters, with

a spacing of 30 between rows and 10 cm between plants within rows. All recommended agronomic practices were followed for optimal growth of the plants. Observations regarding days to fifty percent flowering, plant height, number of primary branches per plant, number of clusters per plant, number of pods per plant, pod length, number of seeds per pod, hundred seed weight, and single plant yield were recorded from five plants that were randomly chosen from each genotype.

Statistical analysis: Correlations among the quantitative traits were examined following the procedures outlined in Johnson *et al.* (1995), while path analysis was conducted in accordance with Dewey and Lu (1959). These analyses were performed using R software version 4.2.1, with the assistance of the 'Biotools,' 'Agricolae,' and 'Metan' packages, along with the 'STAR (Statistical Tool for Agricultural Research) version 2.0.1,' a toolkit developed by the International Rice Research Institute (IRRI). Principal component analysis was also carried using the STAR software.

RESULTS AND DISCUSSION
Analysis of variance (ANOVA): Significant differences for all the quantitative traits was revealed by the analysis of variance (ANOVA) (**Table 1**). The ANOVA revealed significant differences in all quantitative traits, highlighting the variability within the studied population.

Correlation analysis: The correlation analysis revealed a significant positive association between yield and plant height, number of clusters per plant, number of pods per plant, number of seeds per pod, and hundred seed weight, as presented in **Table 2**. These findings align with prior research conducted by Sandhiya and Saravanan (2018), Majhi *et al.* (2020), and Anand *et al.* (2016), who also reported a positive correlation between the number of pods per plant and single plant yield. Additionally, Hemavathy et al. (2015) and Kandel et al. (2017) have documented similar results regarding the number of clusters per plant, while Mahla et al. (2022) and Jyothsna et al. (2016) have also corroborated these findings. Integrating such traits into selection processes holds the potential to significantly enhance crop improvement.

Furthermore, it was observed that the traits associated with yield, including plant height, number of primary branches, number of clusters per plant, and number of pods per plant, exhibited interrelations among themselves. These findings concur with the research conducted by Singh *et al.* (2018) and Shakeer *et al.* (2022), both of whom identified a positive correlation between the number of primary branches per plant, number of clusters per plant, number of pods per plant, and number of seeds per pod with single plant yield. Moreover, Dawane et al. (2022) noted a positive correlation between the number of seeds per pod and hundred seed weight with single plant yield. Also, the number of pods per plant displayed a significant

Table 1. Analysis of variance (ANOVA)

S. No.	Traits	Mean sum of squares		
		Replication	Genotype	Error
1	Days to fifty per cent flowering	1.9087	36.77**	4.3085
2	Plant height	2.5631	105.15**	6.2427
3	Number of primary branches per plant	0.3872	1.90**	0.4657
4	Number of clusters per plant	0.2581	0.82**	0.1034
5	Number of pods per plant	0.655	136.40**	1.9895
6	Pod length	0.0003	0.47**	0.0223
7	Number of seeds per pod	0.8672	3.36**	1.4297
8	Hundred seed weight	0.0754	1.35**	0.0916
9	Single plant yield	0.4915	42.74**	4.3013

Table 2. Correlation of quantitative traits to single plant yield.

Traits		DFF	PH	NPB	NCP	NPP	PL	NSP	HSW	SYP
DFF	P	0.9245**								
	G	1.0535**								
PH	P	0.0516	0.967**							
	G	0.1905	1.0217**							
NPB	P	0.0455	0.2839**	0.9964**						
	G	0.1273	0.4818**	0.9997**						
NCP	P	-0.0107	0.0831	0.3166**	0.8393**					
	G	0.2259*	0.2363*	0.6171**	1.1181**					
NPP	P	-0.535**	-0.2242**	0.0402	0.315**	0.9573**				
	G	-0.5488**	-0.1802	0.0963	0.5103**	1.0219**				
PL	P	-0.1816**	-0.102	-0.0436	-0.3203**	-0.2329**	0.7336**			
	G	0.044	0.0446	-0.0175	0.0068	-0.0662	1.1534**			
NSP	P	-0.215**	0.0272	-0.062	-0.095	-0.0141	0.0158	0.9607**		
	G	-0.0742	0.1041	-0.0286	0.079	0.0851	0.2693**	1.0638**		
HSW	P	0.3321**	0.2872**	0.1757**	0.1707**	-0.2502**	-0.1201*	-0.0704	0.9908**	
	G	0.5058**	0.3618**	0.2724**	0.2771**	-0.2574*	-0.0569	-0.0899	1.0053**	
SYP	P	-0.0726	0.174**	0.1994**	0.3135**	0.3306**	-0.0817	0.2124**	0.4993**	0.9823**
	G	-0.0295	0.2518*	0.3562**	0.5487**	0.4784**	-0.0002	0.5146**	0.7166**	1.0106**

*, ** Significant at 5% and 1% probability levels respectively.
P: Phenotype; G: Genotype;   DFF: Days to fifty per cent flowering, PH: Plant height (cm), NPB: Number of primary branches per plant, NCP: Number of clusters per plant, NPP: Number of pods per plant, PL: Pod length (cm), NSP: Number of seeds per plant, HSW: Hundred seed weight (g).

negative correlation with the number of days before 50% flowering. Similar results were observed for hundred seed weight, which exhibited an intercorrelation with the number of pods per plant. A deeper understanding of these pathways can lead to targeted breeding strategies to optimize these traits.

Path analysis: Path analysis was utilised in addition to correlation to determine the direction of their association. The path analysis showed that the traits with the greatest positive direct contributions to the single plant

yield were the number of pods per plant, the number of seeds per pod, and the weight of the hundred seeds **(Table 3)**. Direct contribution of the number of pods per plant towards the single plant yield was also reported by Manivelan *et al*., (2019); Ramakrishnan *et al*. (2018); Saini *et al*. (2022); and Gaur *et al*. (2021).

The number of primary branches per plant showed a positive low direct impact on the single plant yield, while the number of clusters per plant indicated a positive moderate direct effect. Similar to this, other traits were

Table 3. Direct and indirect effects of quantitative traits with single plant yield

Traits		DFF	PH	NPB	NCP	NPP	PL	NSP	HSW
DFF	P	**0.27**	0.01	0.00	0.00	-0.40	-0.07	-0.07	0.20
	G	**-0.08**	-0.01	0.02	-0.04	-0.38	0.00	-0.04	0.51
PH	P	0.02	**0.18**	0.00	0.01	-0.17	-0.04	0.01	0.17
	G	-0.01	**-0.07**	0.09	-0.04	-0.13	0.00	0.06	0.36
NPB	P	0.01	0.05	**-0.01**	0.04	0.03	-0.02	-0.02	0.10
	G	-0.01	-0.03	**0.19**	-0.12	0.07	0.00	-0.02	0.27
NCP	P	0.00	0.02	0.00	**0.12**	0.24	-0.12	-0.03	0.10
	G	-0.02	-0.02	0.12	**-0.21**	0.35	0.00	0.04	0.28
NPP	P	-0.16	-0.04	0.00	0.04	**0.72**	-0.09	0.00	-0.15
	G	0.04	0.01	0.02	-0.10	**0.71**	0.00	0.05	-0.26
PL	P	-0.05	-0.02	0.00	-0.05	-0.18	**0.28**	0.01	-0.07
	G	0.00	0.00	0.00	0.00	-0.05	**-0.03**	0.15	-0.06
NSP	P	-0.06	0.01	0.00	-0.01	-0.01	0.01	**0.33**	-0.04
	G	0.01	-0.01	-0.01	-0.01	0.06	-0.01	**0.58**	-0.09
HSW	P	0.10	0.05	0.00	0.02	-0.19	-0.04	-0.02	**0.58**
	G	-0.04	-0.03	0.05	-0.05	-0.18	0.00	-0.05	**1.01**

Residual effect: P: Phenotypic = 0.36   G: Genotypic = 0.2784

also observed to indirectly affect the characteristics that contribute to yield. The results of the path analysis showed that the single plant yield was indirectly influenced by the number of clusters per plant. Pod length showed a moderate indirect contribution to the production of a single plant through the number of seeds per pod, whereas days to 50% flowering demonstrated a large indirect contribution to the yield of a single plant through hundred seed weight. The number of primary branches and the number of clusters per plant both exhibited a moderate indirect contribution towards single plant yield through hundred seed weight, but plant height showed a strong indirect influence towards single plant yield through hundred seed weight. Although these traits show lower direct effects, they contribute indirectly to yield. Hence, these indirect effects can be considered when making selections.

Principal Component Analysis (PCA): For any crop improvement program, genetic variation is important for utilizing any genotype. To effectively utilize the genotype, characterization is important and principal component analysis provides a better screening of different genotypes as suggested by Rabbani *et al.* (1998). Results of the principal component analysis (**Table 4**), revealed that the first four principal components with an eigenvalue of 1.0 accounted for 74 per cent of the total variation. Among the four principal components PC1 with an eigenvalue of 2.49 accounted higher proportion of the total variance (22.73) whereas PC2, PC3, and PC4 with eigenvalues of (1.95), (1.27), (1.02) recorded (21.74), (14.17), (11.34) proportion of the total variance respectively. Raji's (2002) criterion defined the critical limit for proper vector coefficients, with characters having coefficients greater than 0.3 exhibiting significant effects (**Fig. 1**).

Table 4. Eigenvalues and per cent variation and cumulative variation for various principal components among green gram genotypes

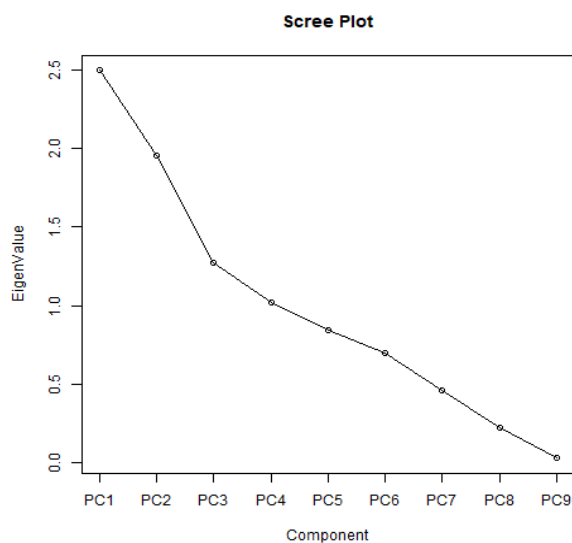| Principal components	Eigenvalues	Percentage of variation	Cumulative Variation
PC1	2.49	22.73	27.73
PC2	1.95	21.74	49.47
PC3	1.27	14.17	63.64
PC4	1.02	11.34	74.98
PC5	0.84	9.39	84.37
PC6	0.69	7.73	92.10
PC7	0.46	5.11	97.21
PC8	0.22	2.45	99.67
PC9	0.03	0.30	100

Fig. 1. Scree plot constructed using the nine principal components

From **Table 5**, PC1 and PC2 traits have a positive and negative impact on the total variance, respectively. From the PC3 the traits like days to fifty per cent flowering, number of primary branches per plant, number of clusters per plant, and number of pods per plant showed positive contribution to the total variance, and all other traits showed negative contribution to the total variance. Plant height, number of primary branches per plant, number of clusters per plant, and pod length contribute to total variance. So, the first four principal components contribute more to the total variance than the other principal components. It's important to select traits such as fifty per cent flowering, plant height, number of primary branches per plant, number of clusters per plant, pod length, and number of seeds per pod for future breeding. Corresponding results for these traits towards the total variance were reported in studies by Mahalingam *et al.* (2020); Mohan *et al.* (2021); Nayak *et al.* (2021) and hence selection of these traits will be beneficial for future breeding programs. This analysis thereby helps us to

identify key traits contributing to total variance, and to select genotypes with favourable combinations of these traits. This can lead to the development of superior crop varieties with improved adaptability and yield potential.

In conclusion, ANOVA, correlation, path analysis, and PCA conducted in this study provided valuable insights into the interrelationships and contributions of various quantitative traits to the overall yield of the crops under investigation. The ANOVA revealed significant differences in all quantitative traits, highlighting the variability within the studied population. Correlation analysis demonstrated positive associations between yield and key traits such as plant height, number of clusters per plant, number of pods per plant, number of seeds per pod, and hundred seed weight. These findings were consistent with previous research, emphasizing the importance of these traits in crop improvement.

Path analysis further elucidated the direct and indirect

Table 5. Contribution of first four principal components to variation

Parameters	PC1	PC2	PC3	PC4
DFF	-0.1612	-0.5509	0.1029	-0.1833
PH	-0.3386	-0.2169	-0.1464	0.4354
NPB	-0.4167	-0.0048	0.1594	0.5533
NCP	-0.4506	0.2097	0.268	0.1081
NPP	-0.1234	0.6461	0.1844	-0.0533
PL	0.0894	-0.0343	-0.5291	0.4272
NSP	-0.0724	0.1946	-0.6912	-0.113
HSW	-0.0097	-0.3133	-0.0863	-0.3843
SYP	-0.5044	0.2255	-0.2676	-0.3365

contributions of different traits to single plant yield. The number of pods per plant, number of seeds per pod, and hundred seed weight emerged as traits with the highest positive direct impacts on yield. Additionally, traits like the number of primary branches and number of clusters per plant exhibited moderate to low direct effects, emphasizing their indirect influence on yield through other pathways. Understanding these relationships can guide targeted breeding strategies to optimize crop traits for enhanced productivity (Geetha *et al.,* 2021).

Principal component analysis identified the key traits contributing to genetic variation, with the first four principal components explaining a substantial proportion of the total variance. Traits such as days to fifty percent flowering, plant height, number of primary branches per plant, number of clusters per plant, pod length, and number of seeds per pod were highlighted as crucial contributors to total variance. Selecting genotypes with favourable combinations of these traits in future breeding programs has the potential to develop superior crop varieties with improved adaptability and yield potential.

REFERENCES

Anand, G., Anandhi, K. and Paulpandi, V. K. 2016. Genetic variability, correlation, and path analysis for yield and yield components in F6 families of green gram (*Vigna radiata* (L.) Wilczek) under rainfed condition. *Electronic Journal of Plant Breeding*, **7**(2): 434-437. [Cross Ref]

Bindu, B. H., Satyanarayana, N. H., Babu, J. D. P. and Ramesh, D. 2023. Genetic divergence study in greengram [*Vigna radiata* (L.) Wilczek]. *Electronic Journal of Plant Breeding*, **14**(1): 336-342.

Dawane, B. K., Patil, D. K., Gite, V. K., Dawane, J. K. and Kale, D. G. 2022. Character association and path analysis studies on seed yield and its yield attributing traits in green gram (*Vigna radiata* (L.) Wilczek).

Dewey, D. R. and Lu, K. 1959. A correlation and path-coefficient analysis of components of crested wheatgrass seed production 1. *Agronomy Journal*, **51**(9): 515-518. [Cross Ref]

Gaur, G. K., Sharma, A.K., Kumar, A. and Meena, P. (2021). Study on Genetic Diversity and correlation in Mungbean [*Vigna radiata* (L.) Wilczek] Under Arid Zone. Biological Forum – An International Journal, 13(3a), 238-242.

Geetha K. and Divya S. 2021. Genetic Studies on Correlation and Path Analysis in Dolichos Bean (*Lablab Purpureus* L.) Genotypes. *Madras Agric. J.*, 108 (7-9): 297-301. [Cross Ref]

Hemavathy, A. T., Shunmugavalli, N. and Anand, G. 2015. Genetic variability, correlation and path co-efficient studies on yield and its components in mungbean [*Vigna radiata* (L.) Wilczek]. *Legume Research-An International Journal*, **38**(4): 442-446. [Cross Ref]

Johnson, H. W., Robinson, H. F. and Comstock, R. E. 1955. Estimates of genetic and environmental variability in soybeans. *Agronomy Journal*, **47**(7): 314-318. [Cross Ref]

Jyothsna, S., Patro, T. S. S. K., Ashok, S., Rani, Y. S. and Neeraja, B. 2016. Character association and path analysis of seed yield and its yield components in green gram (*Vigna radiata*). International Journal of Theoretical and Applied Sciences, **8**(1), 31-36

Kandel, M., Bastola, A., Sapkota, P., Chaudhary, O., Dhakal, P. and Chalise, P. 2017. Association and path coefficient analysis of grain yield and its attributing traits in different genotypes of wheat (*Triticum aestivum* L.). *International Journal of Applied Sciences and Biotechnology*, **5**(4): 449-453. [Cross Ref]

Mahalingam, A., Manivanan, N., Kumar, K. B., 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]

Mahla, Manohar Ram, Dalip, Anju Nehra and Rukoo Chawla 2022. Study of genetic association and path coefficient in greengram [*Vigna radiata* (L.) Wilczek] genotypes. *Biological Forum – An International Journal,* **14**(2): 766-770

Majhi, P. K., Mogali, S. C. and Abhisheka, L. S. 2020. Enhancement of genetic variability for yield and component traits through recombination followed by induced mutagenesis in greengram [*Vigna radiata* (L.) Wilczek]. *Current Journal of applied science and Technology,* **39**(25): 38-48. [Cross Ref]

Manivelan, K., Karthikeyan, M., Blessy, V., Priyanka, A. R., Palaniyappan, S. and Thangavel, P. 2019. Studies on correlation and path analysis for yield and yield-related traits in greengram [*Vigna radiata* (L.) Wilczek]. *The Pharma Innovation Journal*, **8**(9):165-167.

Mbeyagala, E. K., Kwikiriza, N., Amayo, R., Omadi, J. R., Okwang, D. and Obuo, J. E. P. 2017. Participatory selection of mungbean genotypes in Uganda. *African Crop Science Journal*, **25**(2): 253-262. [Cross Ref]

Mohan, S., Sheeba, A. and Kalaimagal, T. 2021. Genetic diversity and association studies in greengram

[*Vigna Radiata* (L.) Wilczek]. *Legume Research-An International Journal*, **44**(7): 779-784.

Nayak, G., Lenka, D., Dash, M. and Tripathy, K. 2021. Character association and principal component analysis for yield and its attributing characters in green gram. *Journal of Crop and Weed*, **17**(3): 169-175. [Cross Ref]

Rabbani, M. A., Iwabuchi, A., Murakami, Y., Suzuki, T. and Takayanagi, K. 1998. Phenotypic variation and the relationships among mustard (*Brassica juncea* L.) germplasm from Pakistan. *Euphytica*, **101**: 357-366. [Cross Ref]

Raji, A. A. 2002. Assessment of genetic diversity and heterotic relationships in African improved and local cassava (*Manihot esculenta* Crantz) germplasm. PhD thesis, University of Ibadan, Ibadan, Nigeria

Ramakrishnan, C. D., Savithramma, D. L. and Vijayabharathi, A. 2018. Studies on genetic variability, correlation, and path analysis for yield and yield related traits in green gram (*Vigna radiata* (L.) Wilczek*). International Journal of Current Microbiology and Applied Sciences*, **7**(3): 2753-2761. [Cross Ref]

Saini, C. K., Patel, N. and Rishabh (2022). Application of Correlation and Path Coefficient to find out Factors which affect total Production and Productivity of Moong. Biological Forum – An International Journal, 14(1), 972-975.

Sandhiya, V. and Shanmugavel, S. 2018. Genetic variability and correlation studies in green gram (*Vigna radiata* (L.) Wilczek). *Electronic Journal of Plant Breeding*, **9**(3): 1094-1099. [Cross Ref]

Shakeer, S., Varma, V. C., Lal, G. M. and Saikumar, S. 2022. Correlation and path coefficient analysis for quantitative traits in greengram [*Vigna radiata* (L.) Wilczek]. *International Journal of Environment and Climate Change*, **12**(11): 3683-3692. [Cross Ref]

Singh, R., Van Heusden, A. W., Kumar, R. and Visser, R. G. 2018. Genetic variation and correlation studies between micronutrient (Fe and Zn), protein content and yield attributing traits in mungbean (*Vigna. radiata* L.). *Legume Research-An International Journal*, **41**(2): 167-174.

Wright, S. 1921. Systems of mating. I. The biometric relations between parent and offspring. *Genetics*, **6**(2): 11. [Cross Ref]