Research Article

Journal of Plantation Crops, 2013, 41(3): 307-313

Principal component analysis and association studies for yield component traits in plus trees of *Theobroma cacao*

R.B. Karthik Kumar*, P. Jansirani, K. Iyanar, P. Paramaguru and N. Kumar

Department of Spices and Plantation Crops, Horticultural College and Research Institute
Tamil Nadu Agricultural University, Coimbatore - 3

(Manuscript Received: 12-07-13, Revised: 26-08-13, Accepted: 02-09-13)

Abstract

A survey conducted in the major cocoa growing tracts of Tamil Nadu could identify 27 plus trees. These trees were observed for a period of two years to study the variability in yield and quality traits and to generate knowledge on the correlation between the ten different traits. Principal component analysis was used to identify the distinguishing traits and grouping them in clusters. The selected plus trees showed large variability in yield and biochemical traits. The number of pods per tree per year ranged from 57.6 to 93.9 with an average pod yield of 73.6. The average dry bean weight was 1.06 g which is a desirable trait in selection of the individuals having high processing efficiency in factory. The dry bean yield per tree per year was 2590 to 3665 g and the average yield was 3131 g. GRest analysis scored the plus trees based on the weightage of the desirable traits of individual trees. The first four PC axes explained 81.3 per cent of the total variation among the 27 plus trees of cocoa and all the traits studied were grouped in five basic clusters. The number of pods per tree (NP) was positively correlated with the pod value (PV) and dry bean weight per tree (DBYT).

Keywords: Association, cocoa, GRest analysis, PCA

Introduction

Cocoa (*Theobroma cacao* L.), is being cultivated in more than 58 countries in an area of about 8.88 million ha mostly grown in the tropical regions from 10 to 20 degrees latitude north or south of the equator (Motamayor *et al*., 2002). The annual production of cocoa beans is 4.22 million tonnes and the global productivity of cocoa is around 475 kg of dry beans. Around 69 per cent of world's production is from African countries (Ivory Coast, Cameroon, Ghana and Nigeria), 13 per cent from Central and South America (Brazil, Colombia, Ecuador, Venezuela and Mexico) and 18 per cent from Asian countries (Indonesia, India, Malaysia and Philippines). The contribution from India is 0.34 per cent of world production (http://dccd.gov.in).

Characterization of the trees and eventual selection (*i.e*., the mass selection of ortets) based on individual phenotypic values is a classic first step in cocoa breeding programme (Paulin, 1995). The variability of 96 wild cocoa trees showed a highly significant tree effect enabling ortets selection and hence these ortets could be used as clones, or as parents in crosses with other genetic groups. The average bean weight can be used as an important selection index for selecting the ortets. The variability found for average fresh bean weight among the 96 wild trees was substantial, with a coefficient of variation of 10.5 per cent and extremes of 2.2 and 3.8 g (Lachenaud and Oliver, 2005).

Morpho-agronomic characteristics of the pods, seeds and flowers have also been used to evaluate cocoa genotypes (Engels, 1986; Bekele and Bekele, 1996; Lachenaud *et al*., 1999; Lachenaud and Oliver, 2005). Earlier diversity studies have often concentrated on field gene bank collections (Bekele *et al*., 2006; Engels, 1986; Lachenaud and Oliver, 2005), however, the study on cocoa

*Corresponding Author: arthikhorts@gmail.com

accessions and diversity maintained on-farm in field plantations by farmers is also carried out by Aikpokpodion (2010).

Cocoa is an introduced crop in India and is predominantly grown in mixed stands and backyards of the humid tropics of Kerala, whereas in Tamil Nadu, it is grown as an intercrop in the coconut and arecanut plantations of tropical region. Large population of cocoa trees of Forastero type exists in the regions comprising, Pollachi, Sethumadai, Vettaikaranpudur, Kulathupudur and Sathyamangalam of Tamil Nadu. The parents of these seed materials are not known and the trees have adapted and acclimatized to the local climatic conditions of the region for many years. High degree of genetic variability has been observed in these plantations (Thondaiman, 2011). The previous studies on the survey, identification and molecular characterization of plus trees of cocoa in these areas resulted in selection of 27 plus trees with good genetic potential.

The objective of the study was set to determine the amount of variability in quantitative and bio-chemical traits among the plus trees of cocoa selected in plantations, to select the plus trees based on GRest analysis for further crop improvement programme and to generate knowledge on the correlation between the traits studied.

Materials and methods

The experimental material consists of 27 plus trees of cocoa identified in the plantations of Coimbatore district of Tamil Nadu through the previous work on 'Survey and identification of plus trees of cocoa' (Thondaiman, 2011). The plus trees were more than 12 years old and located in adjacent fields with similar agronomic management.

The data was collected for two years (July 2010 to June 2012). The number of pods harvested per tree (NP) was counted at an interval of 15 days throughout the study period. Only healthy, mature pods were selected for characterization. Fruits were considered mature when their colour had turned (or were about to turn) yellow. For observing the bean characters, five pods were selected from each tree and the wet bean mass separated from the central placenta was weighed and the average fresh bean weight per pod (FBWP) was calculated. The average number of beans per pod (NBP) was counted. The wet bean mass separated from five pods were fermented and dried to moisture content of 6-8 per cent. The dried beans were weighed and the mean weight of dry bean per pod (DBWP) was calculated.

Pod value (PV) is the number of pods required to make one kilogram of dried cocoa beans. As the number of pods available at single harvest is not sufficient for calculating the pod value, a conversion ratio was used. A general theoretical estimate of conversion ratio of 40 as derived from preliminary observations (Gregory, 1983) was used to compare pod values for all the trees in the present study. The dry bean yield per tree (DBYT) was calculated by multiplying the number of pods harvested and dry bean weight per pod.

Fat was estimated by petroleum ether extraction method using Soxhlet apparatus and expressed in percentage (Elain Apshara *et al*., 2009). Folin-Ciocalteau reagent method was used for estimating the total phenol (Bray and Thorpe, 1954).

Statistical analysis

The four major traits that have to be considered in selection of trees are dry bean yield per tree per year, pod value, shell per cent and fat content (Chesney, 2007). These characters were statistically scored by a software *viz.*, GRest (Genetic Resource Selection and Evaluation Tool) designed by Cocoa Research Association (CRA), United Kingdom. The GRest is a flexible and dynamic web-based management tool that allows accessions to be impartially and unambiguously ranked, based on information on the traits the individuals possess and the genetic diversity it represent (Turnbull and Hadley, 2012).

All the morphological quantitative traits were subjected to Principal Component Analysis (PCA) using the correlation matrix to define the pattern of variation. PCA axes with eigen values $\geq$0.8 were selected to define the variation among accessions for agronomic and morphological traits. The statistical package *viz.*, SPSS 16 was used for analysis.

Results and discussion

The selected plus trees showed large variability in all the yield and biochemical traits (Table 1). The number of pods per tree per year was in the range from 57.6 to 93.9 with average pod yield and CV of 73.6 and 11.7 per cent respectively. The average number of beans observed in the trees was 41.0 with a maximum of 48.4. Assemat *et al*. (2006) reported that the number of beans per pod was variable and ranged from 88-110 in Guyana population while Efombagn *et al.* (2009) reported a mean seed number of 40.5 per pod in the farm accessions of Cameroon.

Single dry bean weight was found to be in the range of 0.97-1.16 g. Efombagn *et al.* (2009) reported that the mean individual dry seed weight of cocoa beans as 0.92 g in farm accessions.

(2009) reported that the cocoa accessions from farmers field had mean pod value of 26.3.

The average content of fat (%), total phenol (mg equivalent for pyrocatechol per g) and total carbohydrates (%) was 76.44, 46.85 and 22.03 respectively. Pires *et al.* (1998) obtained an average fat content in the dry unfermented beans of 490 accessions of the Centro de Pesquisa do Cacau germplasm collection in Ilhéus, Bahia, Brazil as 53.2 per cent, ranging from 45.4 per cent in CC 57 to 60.3 per cent in NA 312. The traits like number of pods per tree per year, fresh bean weight per pod, pod value and total phenol are having coefficient of variation above 10 per cent which is medium and hence these trees can be further utilized as a clonal material or parents in crop improvement programme after assessing their compatibility reactions.

Table 1. Variability in plus trees of cocoa

Traits	Mean	Maximum	Minimum	Difference	CV(%)
Number of pods tree^{-1} year^{-1} (NP)	73.59	93.90	57.50	36.4	11.7
Number of beans pod^{-1} (NBP)	40.99	48.40	35.00	13.4	8.7
Single dry bean weight (SDBW) (g)	1.06	1.16	0.97	0.2	5.7
Fresh bean weight per pod (FBWP) (g)	117.30	145.74	93.39	52.3	12.7
Dry bean weight per pod (DBWP) (g)	44.92	52.13	38.67	13.5	7.8
Dry bean yield tree^{-1} year^{-1} (DBYT) (g)	3131	3665	2590	1075	8.7
Pod value (PV)	26.78	35.96	22.26	13.7	11.8
Total phenol (TP) (mg equivalent for pyrocatechol per g)	76.44	94.63	57.13	37.5	12.1
Fat content (F) (%)	46.85	53.93	37.75	16.2	7.6
Total carbohydrate (CHO) (%)	22.03	25.60	18.10	7.5	9.1

The average dry bean weight was found to be 1.06 g which is a desirable trait in selection of the individuals having high processing efficiency in factory. Bekele (1993) found the average dry bean weight for 53 clones of varied origins (Trinitarios and Forasteros, Trinidad) as 1.1 g with a CV of 17 per cent.

The average fresh bean yield per pod was 117.3 g with a maximum of 145.7 g and CV of 12.7 per cent. The dry bean yield per tree per year was in the range of 2590-3665 g and average yield was 3131.4 g. The trees exhibited a pod value with 22.26 as minimum and 35.96 as maximum with CV of 11.8 per cent. Chesney (2007) also reported a mean pod value of 61 trees of Guyana as 36.3 with ten accessions having pod value ≥21 while Efombagn *et al.*

GRest analysis

The GRest analysis scored the accessions based on the information of the traits possessed by the individual trees on a 0-9 scale for all the traits. The combination of the weighted scores is presented in the Table 2. The maximum points of 165.2 was scored by VPS 15 and it was followed by 144.9 (SME 26). The least score of 39.0 was scored by the tree, SME 2.

The maximum score for individual trait of dry bean yield per tree per year was scored by VPS 15 (9.0) while SMJ 15 scored the maximum value for pod value (8.9). KUL 2 scored the maximum values (8.9) for fat content and shell per cent. The cocoa trees ranked by the tool in the top ten places can be used for further crop improvement programmes like

Table 2. Scores by accessions in GRest analysis (Scale: max-9 and min-0)

Rank	Score	Plus tree	Individual trait score			
			Average dry bean yield tree^{-1} year^{-1} (kg)	Pod value	Fat (%)	Shell (%)
1	165.19	VPS15	9.00	4.96	7.98	7.41
2	144.90	SME 26	6.78	4.65	7.98	7.41
3	142.71	VPS13	7.53	5.76	7.04	5.82
4	130.68	SME 9	6.91	4.96	6.50	5.03
5	130.58	SME 24	6.40	7.48	6.69	5.29
6	127.53	VPS 12	4.86	8.60	7.59	6.88
7	121.70	SMJ 3	5.96	5.16	6.46	4.76
8	118.76	SMJ 4	7.37	4.96	5.01	2.38
9	116.83	SMJ 15	5.44	8.99	6.07	4.24
10	113.83	KUL 2	2.30	3.13	8.90	8.90
11	108.73	SME 21	4.14	3.54	6.93	5.56
12	107.54	SMJ 25	4.52	1.72	6.65	5.29
13	107.42	SME 28	4.81	5.26	6.07	4.24
14	104.90	SME 29	6.61	0.00	4.81	2.12
15	104.58	KUL 18	2.34	4.65	7.94	7.41
16	100.68	SME 6	3.31	7.99	6.46	4.76
17	94.59	SMJ 21	6.29	0.20	4.11	0.79
18	93.49	KUL 25	2.80	6.47	6.34	4.76
19	92.18	SMJ 37	4.40	3.13	5.20	2.65
20	85.16	SME 16	1.59	5.97	6.65	5.03
21	80.67	SMJ 33	2.40	0.20	6.07	4.24
22	78.00	SMJ 10	5.07	8.09	2.58	1.06
23	72.55	SME 5	2.43	2.22	5.09	2.65
24	71.99	SMJ 50	0.88	0.81	6.50	4.76
25	66.30	SMJ 18	3.60	2.22	3.52	0.00
26	53.01	SMJ 34	0.00	0.51	5.48	3.18
27	39.04	SME 2	3.77	4.85	7.00	0.26

clonal multiplication and evaluation or hybridization.

Principal component analysis

Eigen values represent the amount of variation accounted in each PC and the eigen vectors for each trait under the study for the four principle components are tabulated below.

Table 3. Principal component showing the eigen values, proportion of variation and total variation across axis

Principal component	Eigen values	Variation (%)	Total variation explained across axis
1	3.493	34.93	34.92
2	2.337	23.37	58.30
3	1.341	13.41	71.71
4	0.961	9.61	81.32
5	0.726	7.26	88.59
6	0.488	4.88	93.47
7	0.369	3.69	97.16
8	0.177	1.77	98.93
9	0.081	0.81	99.75
10	0.025	0.24	100.00

The first four PC axes explained 81.3 per cent of the total variation among the 27 plus trees of cocoa. The first PC contributed 34.9 per cent to the total variation while the second and third PC contributed 23.3 and 13.4 per cent respectively. The contribution by the fourth PC to the total variation was 9.6 per cent.

The 27 plus trees were significantly distinguished by the eigen vector (≥0.8) in the different principal components (non-rotated values).

The number of pods per tree and number of beans per pod was distinguished by the eigen vector in the principal component 1. The second principal component had single bean dry weight and dry bean yield per tree which discriminated the 27 plus trees. Total phenol and fat per content significantly distinguished the plus trees in third and fourth principal components respectively.

The tri-dimensional spatial figuration shows five basic clusters of the traits studied (Fig. 1). The

Table 4. Principal component analysis of traits of cocoa plus trees - non-rotated loadings

Trait	Principal components (non-rotated values) (eigen vector ≥0.8) 1	2	3	4
NP	0.891	-0.055	0.378	0.085
NBP	0.862	0.049	0.217	0.351
SDBW	0.062	0.909	-0.025	-0.229
FBWP	0.665	0.368	-0.400	0.216
DBWP	-0.647	0.611	0.081	0.201
PV	0.678	0.370	0.246	0.094
TP	0.261	-0.644	0.899	0.598
F	0.158	0.461	-0.576	0.810
CHO	-0.613	0.230	0.335	0.211
DBYT	0.400	0.826	0.686	0.142

NP - Number of pods tree^{-1} year^{-1}; NBP - Number of beans pod^{-1}; SDBW - Single dry bean weight(g); FBWP Fresh bean weight per pod(g); DBWP - Dry bean weight per pod (g); DBYT - Dry bean yield tree^{-1} year^{-1} (g); PV- Pod value; TP- Total phenol (mg equivalent for pyrocatechol per g); F- Fat content (%); CHO - Total carbohydrate (%)

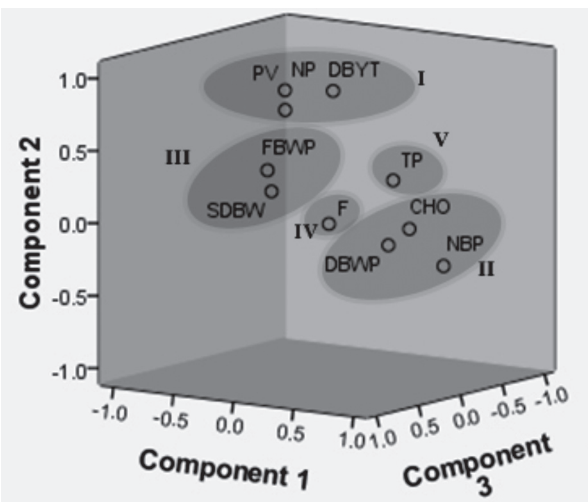

Fig. 1. Tri-dimensional graph showing the grouping of the ten traits studied

cluster I had three traits (NP, PV, NBYT) and cluster II had three traits (DBWP, CHO, NBP) grouped together. Two traits (FBWP and SDBW) were found grouped in cluster III. The cluster IV and V has Fat (F) and total phenol (TP) respectively. The dry bean yield per tree is determined by the number of pods (NP) as well as the pod value (PV). The traits PV and NP together contribute to dry bean yield and are clustered together.

Correlation

The number of beans per pod, dry bean weight per pod and carbohydrate content are highly correlated and hence grouped in one cluster. The fresh bean weight per pod and single dry bean weight are correlated as observed in the correlation table and hence they occupied together in cluster III. The fat content and total phenol are biochemical components and their association with other traits is less established and hence occupies separately in different clusters.

The number of pods per tree (NP) was positively and significantly ($P \leq 0.01$) correlated with the pod value (PV) and dry bean weight per tree (DBYT) with *r* values of 0.636 and 0.623 respectively (Table 5). The number of beans per pod (NBP) had positive and significant correlation with dry bean weight per pod (DBWP) and total carbohydrate content (CHO).

The correlation between single dry bean weight (SDBW) and fresh bean weight per pod (FBWP), dry bean weight per pod (DBWP), pod value (PV), dry bean yield per tree was positive and

Table 5. Correlation co efficient of different traits of cocoa

	NP	NBP	SDBW	FBWP	DBWP	PV	TP	F	CHO	DBYT
NP	-	-0.663**	-0.075 ^{ns}	0.438 *	-0.578 **	0.636 **	0.290 ^{ns}	-0.015 ^{ns}	-0.383 *	0.623**
NBP			-0.108 ^{ns}	-0.558 **	0.692**	-0.510 **	-0.047 ^{ns}	-0.062 ^{ns}	0.551**	-0.069 ^{ns}
SDBW				0.337 *	0.479**	0.363 *	-0.588 **	0.265 ^{ns}	0.090 ^{ns}	0.337*
FBWP					-0.124 ^{ns}	0.381 *	0.059 ^{ns}	0.467**	-0.414 *	0.212 ^{ns}
DBWP						-0.199 ^{ns}	-0.364 *	0.124 ^{ns}	0.449**	0.094 ^{ns}
PV							0.047 ^{ns}	0.198 ^{ns}	-0.119 ^{ns}	0.408*
TP								-0.118 ^{ns}	-0.199 ^{ns}	-0.059 ^{ns}
F									-0.027 ^{ns}	-0.032 ^{ns}
CHO										-0.039 ^{ns}
DBYT										-

NP - Number of pods tree^{-1} year^{-1}; NBP - Number of beans pod^{-1}; SDBW- Single dry bean weight (g); FBWP Fresh bean weight per pod (g); DBWP - Dry bean weight per pod (g); DBYT- Dry bean yield tree^{-1} year^{-1} (g); PV- Pod value; TP- Total phenol (mg equivalent for pyrocatechol per g); F- Fat content (%); CHO - Total carbohydrate (%)

significant. The fresh bean weight per pod (FBWP) had significant and positive correlation (r = 0.381) with pod value. Oyedokun *et al.* (2011) reported that the weight of the individual bean is positively and significantly correlated with the length, width, bean weight and the ratio of the bean length to its width. There was also a negative and significant correlation between the number of pods per tree (NP) and number of beans per pod (NBP) and dry bean weight per pod.

The correlation between number of pods (NP) and number of beans per pod (NBP), dry bean weight per pod (DBWP) was negative and significant. Similarly the number of beans per pod had negative and significant correlation with fresh bean weight per pod (FBWP) and pod value (PV). The fresh bean weight per pod had negative and significant correlation with total carbohydrates while dry bean weight per pod had negative and significant correlation with total phenol.

High level of variability is documented in the plus trees of cocoa. The first ten ranking plus trees obtained from the scores of GRest analysis may be used in the crop improvement programmes like clonal multiplication and evaluation in various environments or hybridization. The principal component analysis distinguished the traits in different clusters based on the correlation possessed by the individual traits with one another. The correlation of the different yield and biochemical traits in cocoa is documented. Positive and significant correlation was noticed between number of pods per tree, single dry bean weight, pod value and dry bean yield per tree per year.

References

Aikpokpodion, P.O. 2010. Variation in agro-morphological characteristics of cacao, *Theobroma cacao* L., in farmers' fields in Nigeria, *New Zealand Journal of Crop and Horticulture* **38**(2): 157-170.

Assemat, S., Lachenaud, P., Ribeyre, F., Davrieux, F., Pradon, J.L and Cros, E. 2006. Bean quality traits and sensory evaluation of wild Guianan cocoa populations (*Theobroma cacao* L.). *Genetic Resources and Crop Evolution* **53**(6): 1303.

Bekele, F.L. and Bekele, I. 1996. A sampling of the phonetic diversity in the International Cocoa Genebank of Trinidad. *Crop Science* **36**(1): 57-64.

Bekele, F.L., Bekele, I., Butler, D.R. and Bidaisee, G.G. 2006. Patterns of morphological variation in a sample of cacao (*Theobroma cacao* L.) germplasm from the International Cocoa Genebank, Trinidad. *Genetic Resources and Crop Evolution* **53**: 933-948.

Bekele, F.L. 1993. Use of botanical descriptors for cocoa characterization: CRU experiences. In: *International Workshop Conservation, characterisation and utilisation of cocoa genetic resources in the 21st century*, CRU, Trinidad, September 13th - 17th, 1992. pp. 77-91.

Bray, H.G., and Thorpe, M.U. 1954. Analysis of phenolic compounds of interest in metabolism. *Methods of Biochemical Analysis* **9**: 27-52.

Chesney, P.E.K. 2007. Preliminary characterisation and evaluation of cocoa (*Theobroma cacao* L.) on-farm genetic diversity in the northwest region of Guyana. In: *Proceedings of International Workshop on Cocoa Breeding for Farmers' Needs*. San José, Costa Rica: 15th - 17th October 2006. pp. 85.

Efombagn, M.I.B., Sounigo, O., Nyassé, S., Manzanares-Dauleux, M. and Eskes, A.B. 2009. Phenotypic variation of cacao (*Theobroma cacao* L.) on farms and in the gene bank in Cameroon. *Journal of Plant Breeding and Crop Science* **1**(6): 258-264.

Elain Apshara, S., Bhat, V.R., Ananda, K.S., Nair, R.V. and Suma, D. 2009. Evaluation and identification of high yielding trees in Nigerian cocoa germplasm. *Journal of Plantation Crops* **37**(2): 111-116.

Engles, J.M.M. 1986. The identification of cacao cultivars. *Acta Horticulturae* **182**: 195-202.

Gregory, Z. 1983. Genetic variability and Correlation studies in cocoa (*Theobroma cacao*. L). MSc. thesis. Kerala Agricultural University, Kerala, India.

Lachenaud, P. and Oliver, G. 2005. Variability and selection for morphological bean traits in wild cocoa trees (*Theobroma cacao* L.) from French Guiana. *Genetic Resources and Crop Evolution* **52**: 225-231.

Lachenaud, P., Bonnot, F. and Oliver, G. 1999. Use of floral descriptors to study variability in wild cocoa trees (*Theobroma cacao* L.) in French Guiana. *Genetic Resources and Crop Evolution* **46**: 491-500.

Motamayor, J.C., Risterucci, A.M., Lopez, P.A., Ortiz, C.F., Moreno, A. and Lanaud, C. 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* **89**: 380-386.

Oyedokun, A.V., Omoloye, A.A., Adewale, B.D., Adeigbe, O.O., Adenuga, O.O. and Aikpokpodion, P.O. 2011. Phenotypic variability and diversity analysis of bean traits of some cocoa hybrids in Nigeria. *Asian Journal of Agricultural Sciences* **3**(2): 127-131.

Paulin, D. 1995. Les me´thodes de se´lection du cacaoyer. In: *Traitements statistiques des essais de se´lection. Actes* *du se´minaire de ge´ne´tique quantitative*, Montpellier, France 12th - 14th September 1994. pp. 243-257.

Pires, J.L., Cascardo, J.C.M., Lambert, S.V. and Figueira, A. 1998. Increasing cocoa butter yield through genetic improvement of *Theobroma cacao* L.: Seed fat content variability, inheritance, and association with seed yield. *Euphytica* **103**: 115-121.

Thondaiman, V. 2011. Studies on genetic diversity and molecular characterization of cocoa (*Theobroma cacao* L.). Ph.D thesis, Tamil Nadu Agricultural University, Coimbatore-3.

Turnbull, C.J. and Hadley, P. 2012. *International Cocoa Germplasm Database CD-ROM*. CRA Ltd./NYSE Liffe/University of Reading, UK.