Plant Genetic Resources: Characterization and Utilization

cambridge.org/pgr

Cladistics groupings of the active breeding cocoa genetic resources of Nigeria for physicochemical and nutraceutical traits

Daniel B. Adewale[1] and Oluwayemisi O. Adeigbe[2]

[1]Department of Crop Science and Horticulture, Federal University Oye-Ekiti, Ikole-Ekiti Campus, Nigeria and [2]Crop Improvement Division, Cocoa Research Institute of Nigeria, Ibadan, Nigeria

Research Article

Cite this article: Adewale DB, Adeigbe OO (2023). Cladistics groupings of the active breeding cocoa genetic resources of Nigeria for physicochemical and nutraceutical traits. *Plant Genetic Resources: Characterization and Utilization* **21**, 454–462. https://doi.org/10.1017/S1479262123000977

Received: 18 August 2022
Revised: 7 November 2023
Accepted: 8 November 2023
First published online: 30 November 2023

Keywords:
clustering; cocoa beans powder; diversity; germplasm; nutraceutical traits

Corresponding author:
Daniel B. Adewale;
Email: d.adewale@gmail.com

Abstract

Preference for functional and nutritious food capable of meeting consumers' demand and health is on the increase. The present preliminary study seeks to assess physico-chemical and nutraceutical diversity in the cocoa bean powder of 77 genotypes present in four Nigerian cocoa field banks. Twenty ripe pods/genotypes in each of the four active breeding field banks at the Cocoa Research Institute of Nigeria (CRIN), Ibadan, Nigeria were utilized. Composite beans from the 20 pods of each genotype were singly fermented, sun-dried and milled. Duplicate samples of the powder of each genotype were analysed for physico-chemical and nutraceutical components. Twenty-one polymorphic variables distinguished the 77 cocoa genotypes. Grouping by dendogram identified four clusters, three differently and uniquely captured 100% of the genotype membership in the local clone, international clone and the regional varieties field bank but 86% of the genotypes in the hybrid trial field bank were grouped in cluster I. Prominent traits with highest values in each clusters were: protein, *p*H, Ca, K and Fe (Cluster I), Zn and Mg (Cluster II), crude fat and P (Cluster III) and crude fibre, ash, theobromine, flavonoids and caffeine (Cluster IV). Exploitable diversity for nutritional quality improvement is present in the active breeding and working collections of Nigerian cocoa field banks.

Introduction

Cocoa (*Theobroma cacao* L.), of the family Malvaceae *sensu lato*, a perennial cash crops (Marita *et al.*, 2001) predominantly grows in the tropics of Central and South America, Asia and Africa. It had an estimated world output of 3.5 million tons in 2006 (de Almeida and Valle, 2007) and 5.2 million metric tons in 2020 (WorldAtlas, 2021). The economic value of cocoa beans exported as whole, broken, raw or roasted was 8.6 billion US Dollars (USD) in 2017 (Eghbal, 2018). It has remained the foremost non-oil source of foreign exchange to the major cocoa-producing countries in West and Central Africa. It is cultivated by more than five million growers in 50 countries and 40–50 million people derive their livelihood from it (CacaoNet, 2012; World Cocoa Foundation, 2012). The chocolate and cocoa powder from cocoa beans are good and nutritional sources of minerals especially potassium, magnesium, copper and iron (Afoakwa *et al.*, 2007; Torres-Moreno, *et al.*, 2015; Adeyeye, 2016).

Aikpokpodion *et al.* (2010) presented the historical chronology of the introduction schemes of cocoa genetic resources into Nigeria as follows: (i). Upper Amazon and 'Trinitario' populations, which were introduced from Trinidad in 1944 (Posnette and Todd, 1951), (ii). 1965–1967 large-scale introduction of Upper Amazon cacao materials from Trinidad and Tabago (Aikpokpodion *et al.*, 2010), (iii). Cacao Introduction Scheme sponsored by the Cocoa Alliance that included quite a large number of intra-Nanay, intra-Parinari, intra-Iquitos and inter-Pound selections (Atanda, 1977; Olatoye and Esan, 1992) and (iv). some cocoa genetic resources from Costa Rica, Indonesia, Fernando Po, Kew Garden, Wageningen and Miami (Jacob *et al.*, 1971). These four introductions have long remained the primary cocoa genetic resources for cultivar development in the breeding programmes and planting materials generation at the Cocoa Research Institute of Nigeria (CRIN), Ibadan, Nigeria. Cocoa genetic resources in the list of the various introductions have greatly dwindled; their erosion has been attributed to: old age of plants and plantations. Furthermore, their neglect has been obvious due to poor record keeping and inconsistent data on extant plants. Urbanization and pressure on land for alternative uses have also contributed to genetic erosion, especially in the out-station breeding field banks of CRIN. For example, the out-station field bank at Ibule-soro, Akure, Ondo state, Nigeria has been lost to urbanization.

CRIN has remained the only agricultural institute in Nigeria with the mandate for cocoa improvement, it therefore hosts cacao germplasm collections of Nigeria. Currently, there are four *ex situ* cocoa breeding field germplasm collections at CRIN headquarters in Ibadan,

Nigeria. The four field banks include: (i) introduced cocoa genetic materials established at the International Clone (IC) field bank, (ii) hybrids generated from crosses among some local clones and established in the hybrid trial (HT) field bank, (iii) hybrids generated for regional (central and west Africa) variety evaluation trial (RVT) field bank and (iv) some old 'C' and 'T' clones that were established at the local clones trial (LC) field bank. The C-clones were mixtures of local Amelonado and red podded local Trinitario (Lockwood and Gyamfi, 1979) and T-clones were the approved Upper Amazon populations whose open pollinated progenies formed the F3-Amazon (Aikpokpodion *et al.*, 2009).

High bean yield and disease (especially black pod) resistance had long been the primary breeding goal of cocoa in Nigeria. A large proportion of the cocoa genetic resources in the Nigerian cocoa field genebanks have remained unutilized for food quality improvement because their potentials for nutritional and functional food values have not been unveiled. The eight cocoa varieties released for Nigerian farmers in 2011 (CRIN, 2011) were promoted for high bean yield and resistance to black pod and pod rot diseases. Record of effort(s) for cocoa bean quality improvement in Nigeria is scarce. The present investigation for physico-chemical and nutraceutical diversity among 77 genotypes could present a platform for the initiation of a bean quality improvement programme for cocoa in Nigeria.

Significant variation exists among West African cocoa genotypes at phenotypic and genomic levels (Opoku *et al.*, 2007; Sonwa *et al.*, 2007; Pokou *et al.*, 2009; Aikpokpodion, 2010; Adewale *et al.*, 2013; Olasupo *et al.*, 2018). The nutraceutical properties of the cocoa germplasm in Trinidad and Tobago and Cameroun have been documented (Bekele and Phillips-Mora, 2019). Quite recently, Adeigbe *et al.* (2021) noted that genotypes within each of the four cocoa field banks in Nigeria showed variation for some physicochemical properties. Utilization of plant genetic resources in breeding programmes is dependent on the understanding of their potentials (CacaoNet, 2012; Bekele and Phillips-Mora, 2019). Cocoa beans have complex raw ingredients. More than 500 flavour compounds have been identified in cocoa products; to the manufacturing industries, understanding the source of each is a science on its own (Reed, 2010). There are documented proximate and mineral composition in cocoa bean in Nigeria (Adeyeye *et al.*, 2010; Ndife *et al.*, 2013), however, the number of genotypes employed in such studies were less than three. The unavailability of nutritional and quality information of the already characterized (phenotypic – Aikpokpodion *et al.*, (2009) and genomic – Olasupo *et al.*, (2018)) Nigerian cocoa germplasm is an evident gap, thus substantiating the essence of the present investigation.

Improved consciousness about health among consumers has greatly increased global promotion of functional foods (Gaikwad *et al.*, 2020). Furthermore, the cocoa-based product manufacturers now seek cocoa with enhanced nutraceutical value (Bekele and Phillips-Mora, 2019). Naturally occurring plant nutraceuticals (which are consumed as food or food parts) deliver benefits to consumers beyond basic nutrition, such as protecting organisms against oxidative stress, and playing other roles, such as: antimicrobial, anti-oxidant, anti-inflammatory and anti-cancer properties (Aguilar *et al.*, 2017; Rosa *et al.*, 2022). Identifying physicochemical and nutraceutical status of the different cocoa genotypes in the active Nigerian cocoa breeding germplasm has become evidently necessary. The present investigation, therefore seeks to identify diversity for some biochemical and bioactive traits among genotypes within the germplasm and superior groups of genotypes for specific nutritional, phytochemical, mineral and physical properties.

Materials and methods

The Active cocoa breeding collections in CRIN

The Breeding and Improvement Division of CRIN, Ibadan, Nigeria, has four active cocoa breeding field banks. They are: LC trial field bank, containing 13 clones; HT field bank containing 23 hybrids; RVT field bank which hosts nine hybrids from Ghana, four from Cote d'Ivoire, three from Cameroon, two common crosses from Nigeria and six local control crosses; and the IC field bank which contain 17 introduced cocoa clones. The genotypes were the treatments in each of the field banks. They were laid out in a randomized complete block design with six replicates. Each field banks contained two rows of five plants. List of the genotypes held in the various active breeding field banks is presented in Table 1.

Materials for the study and analytical protocols

Pods were harvested during a cropping season and bulked from all the trees of a genotype. Twenty uniform fruits were then selected from the bulked pool of fruits for each genotype. Total number of pods/genotypes in the harvest was within 250–450. Pods were broken seeds were extracted and pooled on genotype basis. Pooled seeds of each of the 77 genotypes were separately fermented in trays. Fermented samples were sun-dried and milled into powder using standard procedures (ICCO, 2021). Duplicate samples, each of the 77 genotypes were determined following the recommendations of Fitzpatrick (2013) for powdered food samples. Twenty-one biochemical traits were analysed, including: physical properties (bulk density, dispersibility, re-hydration time, oil and water absorption capacity) and proximate (moisture, crude ash, fat, fibre and protein) contents which were determined following the procedures in Ooi *et al.* (2012). The *p*H was determined following the procedure in Vijayakumar and Adedeji (2017). After high pressure digestion, macro-elements (P, K, Ca and Mg) were analysed by atomic absorption spectrometer and micro-elements (Fe and Zn) were determined by sensitive atomic spectrometric techniques and the results were obtained using the working standards of AOAC (1990); 1000 ppm for each of the genotypes (Poitevin, 2016). Caffeine and theobromine were determined following the procedure in Thomas *et al.* (2004) and flavonoid was analysed following the procedure of Lin and Tang (2007). Variability and character association of 17 of the 21 traits for each of the four breeding field banks have been earlier determined by Adeigbe *et al.* (2021).

Data analysis

Basic descriptive statistics analysis including coefficient of range was carried out on the data and Pearson correlation analysis was conducted on the mean matrix table containing the 77 genotypes and the 21 variables. The replicated data on the 77 treatments and 21 variables was subjected to multivariate analysis of variance using the PROC GLM procedure of SAS version 9.4 (SAS, 2011). A mean data matrix comprising 77 genotypes by 21 variables was submitted to SAS using PROC DISTANCE to generate pairs of Gower genetic distances among the 77 genotypes

https://doi.org/10.1017/S1479262123000977 Published online by Cambridge University Press

Table 1. List of cocoa genotypes in the four active cocoa breeding field banks at the Cocoa Research Institute of Nigeria (CRIN), Ibadan, Nigeria

RVT		HT		IC	LC
RVTG1	(P 7 × PA150) × IMC47	HYB1	T65/7 × T22/28	SPEC 54-1	T12/15
RVTG2	A1/154 × T60/78	HYB2	T12/11 × N38	IFC5	T9/15
RVTG3	C303 × PA120	HYB3	T65/7 × T9/15	APA 4	T53/8
RVTG4	GU144 × EQ × 3338	HYB4	PA150 × T60/887	MAN 15-2	T65/7
RVTG5	GU147A × NA33	HYB5	P7 × T60/887	EET59	T16/17
RVTG6	MAN 15-2 × T85/799	HYB6	P7 × PA150	PA 107	T82/22
RVTG7	PA13 × P19	HYB7	T65/7 × T57/22	PA 150	T58/5
RVTG8	SNK12 × PA150	HYB8	T53/5 × N38	BE10	T57/22
RVTG9	T60/78 × T85/87	HYB9	T65/7 × N38	C77	T12/11
RVTG10	T60/887 × ICS89	HYB10	T53/5 × T12/11	P 7	T101/15
RVTG11	T60/887 × SNK413	HYB11	T65/35 × T30/13	CATIE 1000	T65/35
RVTG12	T63/967 × T17/524	HYB12	T86/2 × T9/15	PA 120	T86/2
RVTG13	T85/799 × PA120	HYB13	T9/15 × T57/22	IMC 47	N38
RVTG14	UPA134 × SNK64	HYB14	T86/2 × T22/28	SCA 6	
RVTG15	PA4 × P7	HYB15	T82/27 × T12/11	Playa Alta	
RVTG16	T65/7 × T101/15	HYB16	T86/2 × T16/17	UF 676	
RVTG17	(SCA6(DRI × DR38)) × PA120	HYB17	T65/7 × T53/8	T85/799	
RVTG18	SNK614 × SCA24	HYB18	T65/7 × T101/15		
RVTG19	C77 × C67	HYB19	T86/2 × T53/8		
RVTG20	T65/7 × T79/501	HYB20	T86/2 × T65/35		
RVTG21	(P 7 × PA150) × Amaz 15-15	HYB21	T101/15 × N38		
RVTG22	A1/154 × T85/185	HYB22	T82/27 × T16/17		
RVTG23	T85/185 × T60/178	HYB23	T86/2 × T57/22		
RVTG24	F3 Amazon				

RVT, regional variety trial; HT, hybrids trial; IC, international clones; LC, local clones.

(Gower, 1971). The paired genetic distances were then submitted for principal component analysis (using PROC PRINCOMP) and clustering analysis (using PROC TREE, WARD minimum variance hierarchical clustering method (Ward, 1963)) in SAS. Mean genetic distances were further generated for each cluster and the mean performances for the 21 traits in each cluster were determined. Significant differences between pairs of clusters were tested by paired *t*-test using SAS. Moreover, inter- and intra-cluster variability was further verified using the cluster means, standard error (SE) and coefficient of variation (CV) for each trait.

Results

Variability and potentials of the different genotypes

In Table 2, the 21 variables significantly ($P \leq 0.05$) differentiated the 77 genotypes. The highest CV (10.8%) was recorded for water absorption capacity, whereas all other traits had low CV within 1.02–9.53% (Table 2). From Table 3, the best genotypes for crude fibre, crude ash, moisture content, theobromine, flavonoid, caffeine and rehydration time were, respectively: T12/11, T65/7, T10/15, T86/2, T12/15, T65/35 and T65/7. The highest proportion of crude protein, calcium, zinc and dispersibility was found in HYB17, HYB19, HYB5 and HYB22, respectively (Table 3). The highest butterfat, crude oil, potassium, iron, phosphorus, bulk density, water absorption capacity and oil absorption capacity were found, respectively, in RVTG6, RVTG5, RVTG19, RVTG10, RVTG20, RVTG10, RVTG18 and RVTG5. RVTG9 had the highest magnesium content (Table 3).

Correlations among the various biochemical traits

In online Supplementary Table S1, 121 pairwise correlations were significant, of which 74 were positive and 47 were negative. Protein content in the cocoa bean had a positive and significant correlation with crude fat, dispersibility, bulk density, calcium, potassium, phosphorus, iron, magnesium, water and oil absorption capacity, but it had a significant negative correlation with crude fibre and theobromine content (online Supplementary Table S1). A significant positive correlation existed between all the mineral elements, except for zinc with phosphorus and potassium, then potassium with magnesium. Furthermore, all the mineral elements had a positive significant correlation with dispersibility (online Supplementary Table S1). The relationships between crude ash, moisture content, theobromine and flavonoid were positive and significant. Moreover, the correlation of bulk

https://doi.org/10.1017/S1479262123000977 Published online by Cambridge University Press

Table 2. Analysis of variance showing the sources of variation, degrees of freedom, coefficient of variation and mean squares of the 21 biochemical traits

		Mean squares						
Sources of variation	DF	*p*H	Prot	Fibre	Fat	Ash	MC	Butter
Genotypes	76	0.083***	2.23***	9.47***	8.86***	0.969***	0.156***	4.38***
Error	154	0.005	0.04	0.009	0.54	0.008	0.037	0.20
CV (%)	–	1.02	1.44	2.22	3.03	2.83	1.88	1.05
		Mean squares						
Sources of variation	DF	Theobr	Flavo	Dispers	WAC	OAC	RehydT	Cal
Genotypes	76	0.139***	0.0004***	84.99***	0.24***	0.058***	29.78***	2.32***
Error	154	0.006	0.00002	0.125	0.009	0.0011	0.58	0.015
CV (%)	–	9.53	3.54	1.11	10.8	5.92	2.48	1.72
		Mean Squares						
Sources of variation	DF	Pota	Phos	Zinc	Iron	Mag	Caffe	BD
Genotypes	76	1.53***	1.96***	0.18***	0.023***	0.06*	0.0025***	0.11***
Error	154	0.009	0.038	0.0002	0.0003	0.04	0.00008	0.0003
CV (%)	–	1.14	2.44	5.82	5.47	6.31	6.46	4.18

Prot, protein; MC, moisture; Theobr, theobromine; Flavo, flavonoids; Caffe, caffeine; Dispers, dispersibility; BD, bulk density; WAC, water absorption capacity; OAC, oil absorption capacity; RehydT, rehydration time; Cal, calcium; Pota, potassium; Phos, phosphorus and Mag, magnesium.
*, ** and *** – significance among the 77 genotypes at $P \leq 0.05$, 0.01 and 0.001 respectively.

density with water and oil absorption capacity was positive and significant (online Supplementary Table S1). Each of the following traits: ash, moisture content, theobromine, flavonoid, caffeine and rehydration time had a negative and significant correlation with crude fat. Moreover, each of the six minerals except potassium equally had a negative and significant correlation with crude fibre (online Supplementary Table S1).

Variance contribution to principal components and genotypes' grouping pattern

In online Supplementary Table S2, the eigenvalues of the first three PC axes were higher than 2.5. The highest variance contribution (37.5%) was in PC1. The first three PC axes accounted for and explained 70% of the total variation among the 77 genotypes (online Supplementary Table S2). The Ward clustering method separated the 77 genotypes into four main clusters at the similarity coefficient above 0.10 (Fig. 1). The grouping pattern was very specific: cluster I contained all genotypes in the hybrid trial field bank and five others: APA 4, Playa Alta, RVTG18, RVTG19 and RVTG20. Clusters II, III and IV, respectively, had 100% members of the genotypes in the international clone field bank, regional variety field bank and local clone field bank (Fig. 1). From the cluster history (data not shown), HYB14 and HYB15 were the first pair of genotypes to unite at 0.000 point of similarity. Cluster II was most compact; the 15 genotypes it captured formed a single cluster at a similarity coefficient of 0.0082, with IMC47 and UF676 as the first pair of genotypes to unite at 0.0003 (Fig. 1). Cluster IV was equally highly compact; the 13 genotypes it contained formed a unit (a cluster) at 0.0232 similarity coefficient, whereas all the 28 genotypes formed Cluster I at 0.0286 similarity coefficient (Fig. 1). The 21 genotypes in Cluster III were loose, their grouping into a cluster was at a lower similarity coefficient (Fig. 1).

Intra and inter-cluster variability

In Table 4, the cluster with the highest (0.933) intra-cluster similarity was Cluster II, the group of genotypes with the least (0.816) intra-cluster similarity was Cluster III (Table 4). Furthermore, similarities between pair of clusters revealed that cluster I and III had the highest inter-cluster similarity (0.63), whereas cluster III and IV had the least similarity (0.261). The six possible paired comparisons of the four clusters by *t*-test statistics revealed significant differences between each pair of the four clusters at $P \leq 0.05$ (Table 4).

The mean performance of the grouped genotypes in each cluster for the 21 biochemical traits was presented in Table 5. Group of genotypes with the highest mean for crude protein, dispersibility, water absorption capacity, rehydration time, calcium, potassium and iron existed in Cluster I (Table 5). Genotypes in Cluster II had the highest mean for zinc and magnesium. Highest values for crude fat, butterfat, bulk density, oil absorption capacity and phosphorus occurred among genotypes in Cluster III, whereas the genotypes in Cluster IV had the highest mean values for crude fibre, ash, moisture content, theobromine, flavonoid and caffeine (Table 5).

Discussion

Existence of significant difference(s) among genotypes for the measured traits reveals the availability of exploitable variation for nutritional improvement of cocoa. Although the present study was based on a seasonal harvest, the result presents inferential platform for subsequent investigation of the same genotypes.

https://doi.org/10.1017/S1479262123000977 Published online by Cambridge University Press

Table 3. Descriptive statistics for the 21 biochemical traits employed for the diversity study of the 77 cocoa genotypes

Variables	Minimum	Maximum	Mean ± SE	CV (%)	CoR (%)
Proximate content					
Crude ash (%)	2.25 (RVTG1)[a]	4.81 (T65/5)	3.21 ± 0.064	2.75	36.17
Crude fat (%)	19.59 (SCA6)	28.66 (RVTG6)	24.34 ± 0.196	2.92	18.80
Crude fibre (%)	2.92 (PA150)	8.73 (T12/11)	4.45 ± 0.202	2.08	49.89
Butterfat (%)	39.77 (T12/15)	46.63 (RVTG6)	42.75 ± 0.146	1.05	7.70
Moisture content (%)	9.65 (RVTG22)	10.70 (T10/15)	10.24 ± 0.026	1.83	5.16
pH	6.22 (N38)	7.03 (RVTG12)	6.64 ± 0.012	1.01	6.16
Crude protein (%)	12.47 (T10/15)	15.15 (HYB17)	13.89 ± 0.098	1.42	9.71
Physical properties					
Bulk density (g cm^3)	0.18 (BE 10)	1.12 (RVTG10)	0.39 ± 0.022	2.87	72.04
Dispersibility (%)	14.99 (T12/15)	35.82 (HYB22)	31.81 ± 0.604	1.08	40.98
Oil absorption capacity (g g^{-1})	0.36 (MAN 15-2)	0.89 (RVTG6)	0.56 ± 0.016	5.86	42.81
Rehydration time (min)	22.00 (RVTG15)	35.33 (T65/5)	30.75 ± 0.358	2.47	23.26
Water absorption capacity (g^{-1})	0.15 (HYB6)	1.19 (RVTG18)	0.88 ± 0.033	10.67	78.14
Phytochemicals content					
Caffeine (mg 100 g^{-1})	0.07 (RVTG3)	0.25 (T65/35)	0.14 ± 0.003	6.36	55.79
Flavonoid (mg 100 g^{-1})	0.12 (RVTG7)	0.18 (T12/15)	0.14 ± 0.001	3.40	18.28
Theobromine (mg 100 g^{-1})	0.13 (RVTG21)	1.21 (T86/2)	0.82 ± 0.024	9.26	81.00
Minerals content					
Phosphorus (mg 100 g^{-1})	6.65 (T86/2)	10.42 (RVTG20)	8.05 ± 0.092	2.08	22.07
Iron (mg 100 g^{-1})	0.13 (RVTG14)	0.52 (RVTG10)	0.31 ± 0.010	5.28	59.79
Magnesium (mg 100 g^{-1})	0.12 (T12/15)	0.41 (RVTG9)	0.26 ± 0.007	6.53	55.97
Calcium (mg 100 g^{-1})	4.86 (T65/35)	8.76 (HYB19)	7.13 ± 0.100	1.22	28.62
Potassium (mg 100 g^{-1})	6.67 (RVTG12)	10.55 (RVTG19)	8.60 ± 0.082	1.13	22.53
Zinc (mg 100 g^{-1})	0.12 (T12/15)	0.41 (HYB5)	0.27 ± 0.009	5.73	54.72

[a]Genotypes are in parenthesis: **RVTG1** – (P7 × PA150) × IMC47, **RVTG3** – C303 × PA120, **RVTG6** – MAN 15-2 × T85/799, **RVTG7** – PA13 × P19, **RVTG10** – T60/887 × ICS89, **RVTG12** – T63/967 × T17/524, **RVTG14** – UPA134 × SNK64, **RVTG15** – PA4 × P7, **RVTG18** – SNK614 × SCA24, **RVTG19** – C77 × C67, **RVTG20** – T65/7 × T79/501, **RVTG21** – (P7 × PA150) × Amaz 15-15, **RVTG22** – A1/154 × T85/185, **HYB5** – P7 × T60/887, **HYB6** – P7 × PA150, **HYB17** – T65/7 × T53/8, **HYB19** – T86/2 × T53/8, **HYB22** – T82/27 × T16/17.
[b]CoR, coefficient of range (X max. – X min.)/(X max. + X min.) × 100.

The 21 studied variables were significantly useful in discriminating the 77 cocoa genotypes in the Nigerian field bank, hence they are polymorphic. The use of polymorphic variables for characterization according to Kaufman and Rousseeuw (1990) do enhances clarity in the grouping pattern of genotypes.

The noted high phenotypic diversity in the studied population for the various biochemical traits makes the cocoa field bank collections of Nigeria a promising valuable resources for nutraceutical breeding and improvement programmes. Aikpokpodion *et al.* (2009) had much earlier reported high morpho-agronomic diversity in the Nigerian cocoa genetic resources. Identification of differences among the various genotypes agrees with the known axiom that different genotypes have unique genetic identities and hosting unique potentials which are primary resources in plant breeding.

Low CV is a good measure of similarity within sample and hence the reliability of generated descriptive statistics from the data, as such values does not imply diversity. In our study, percentage differences (between the minimum and the maximum values) among the 77 cocoa genotypes for the 21 traits ranged from 5.2% in moisture content to 81% in theobromine. Crude fibre, bulk density, water absorption capacity, caffeine, theobromine, iron, magnesium and zinc hosts wider variabilities having coefficient of range values ≥50%. The large variation observed in this study was expected because the different genotypes did not share the same genetic origin as earlier reported for Nigeria cocoa germplasm by Aikpokpodion *et al.* (2009) and Olasupo *et al.* (2018). These arrays of diversity among the extant cocoa genotypes in the field banks of CRIN, Nigeria offer great promises for subsequent cocoa bean nutritional and quality improvements.

Among the 77 genotypes in this study, significant positive correlations existed between some biochemical compounds. This promises to ease independent selection of genotypes for different traits and enhance faster advancement in genetic development of better bean quality in Nigerian cocoa. Summarily from the study, calcium strongly and positively associated with zinc, iron and magnesium, protein improvement in cocoa will positively affect the improvement of calcium, potassium, phosphorus, iron, magnesium and butterfat. In consonance with the report of Jiang *et al.* (2007), phosphorus and potassium were strongly correlated.

https://doi.org/10.1017/S1479262123000977 Published online by Cambridge University Press

Figure 1. The grouping of the seventy seven cocoa genotypes by Ward clustering method.

Among the phytochemicals, selection for caffeine will simultaneously and positively affects flavonoid and theobromine, Meng *et al.* (2009) noted similar relationship.

West African cocoa has long been known to be non-acidic (Wood, 1987), the *p*H range in the fermented beans in our study was 6.21–7.03, and this is in agreement with Jayeola and Oluwadun (2010) who reported near alkaline to alkaline *p*H of 6.4–7.4 for some cocoa genotypes in Nigeria. However, our result differed from that of Ndife *et al.* (2013) who obtained *p*H range of 5.65–6.15, the variation could be due to genotype, processing and component composition. The low acidity to alkalinity status of cocoa bean reported in this study hinted on the inherent health benefit in consuming cocoa products. The nutritional quality of cocoa beans was further revealed in this study, crude protein ranged between: 12.5–15.1%. This is higher than the reported crude protein content of 8.14% in cocoa powder by Ndife *et al.* (2013) for unspecified cocoa genotype(s) from Akure, Ondo state, Nigeria. However, in the review by Rawel *et al.* (2019), a mean crude protein of 11–13% was recorded, however, the work further stated that 11.8–15.7% could be possible depending on the geographical origin of cocoa genotypes. The identified variation among the 77 cocoa genotypes for crude protein in this study present opportunity for genetic improvement of the trait as intercrossing among different genotypes could lead to the evolvement of heterotic progenies.

Diversity is critical to ensuring the future of the world cocoa production and to meeting the challenging and changing climate changes and consumer preferences (CacaoNet, 2012). The classification of the 77 genotypes to the four different groups revealed that there is similarity among genotypes within the same cluster. The hybrids had the best and outstanding performances for 65%

Table 4. Level of similarity within each cluster (in parenthesis), similarities coefficients by Gower genetic distance method between pairs of cluster (**lower diagonal**) and probability showing significance in the paired comparison among the four clusters (**upper diagonal**)

Clusters	Cluster I (0.855)	Cluster II (0.933)	Cluster III (0.816)	Cluster IV (0.855)
Number of genotypes	28	15	21	13
Cluster I (0.855)	–	*	*	***
Cluster II (0.933)	0.629	–	**	**
Cluster III (0.816)	0.631	0.473	–	***
Cluster IV (0.855)	0.342	0.406	0.261	–

*, ** and *** denote significance at 0.05, 0.01 and 0.001 probability level, respectively.

https://doi.org/10.1017/S1479262123000977 Published online by Cambridge University Press

Table 5. Mean and variability of the 21 biochemical traits within each cluster (I to IV)

Biochemical properties	Grand mean ± SE (n = 77; GS = 0.77)	I (n = 28; GS = 0.88)		II (n = 15; GS = 0.93)		III (n = 21; GS = 0.82)		IV (n = 13; GS = 0.86)	
		Mean ± SE	CV (%)	Mean ± SE	CV (%)	Mean ± SE	CV (%)	Mean ± SE	CV (%)
*p*H	6.64 ± 0.012	6.73 ± 0.02	1.34	6.58 ± 0.01	0.66	6.73 ± 0.03	1.90	6.38 ± 0.04	1.99
Crude protein (%)	13.89 ± 0.098	14.49 ± 0.11	4.08	12.95 ± 0.06	1.77	14.17 ± 0.13	4.33	13.22 ± 0.23	6.17
Crude fibre (%)	4.45 ± 0.202	3.50 ± 0.06	9.08	3.21 ± 0.04	5.41	4.31 ± 0.12	13.20	8.15 ± 0.12	5.33
Crude fat (%)	24.34 ± 0.196	24.59 ± 0.17	3.62	23.01 ± 0.26	4.40	26.00 ± 0.37	6.57	22.66 ± 0.21	3.29
Crude ash (%)	3.21 ± 0.064	3.31 ± 0.02	3.38	2.77 ± 0.06	8.88	2.80 ± 0.08	12.92	4.14 ± 0.13	10.92
Moisture content (%)	10.24 ± 0.026	10.25 ± 0.02	1.26	10.31 ± 0.03	0.96	10.06 ± 0.06	2.80	10.39 ± 0.07	2.45
Butterfat (%)	42.71 ± 0.189	42.99 ± 0.14	1.77	41.67 ± 0.12	1.15	44.11 ± 0.17	1.80	41.33 ± 0.24	2.12
Theobromine (mg 100 g^{-1})	0.82 ± 0.024	0.8 ± 0.01	6.32	0.85 ± 0.01	5.30	0.64 ± 0.06	41.83	1.12 ± 0.03	8.16
Flavonoid(mg 100 g^{-1})	0.14 ± 0.001	0.14 ± 0.00	4.13	0.14 ± 0.00	3.76	0.13 ± 0.00	4.01	0.16 ± 0.00	7.43
Caffeine(mg 100 g^{-1})	0.14 ± 0.003	0.14 ± 0.00	10.96	0.13 ± 0.00	2.71	0.12 ± 0.01	23.00	0.17 ± 0.01	23.88
Dispersibility (%)	31.81 ± 0.604	34.96 ± 0.15	2.28	32.98 ± 0.14	1.65	33.63 ± 0.13	1.82	20.70 ± 0.93	16.16
Bulk density(g cm^3)	0.39 ± 0.022	0.39 ± 0.01	8.83	0.26 ± 0.02	36.11	0.52 ± 0.07	58.78	0.37 ± 0.03	29.30
Water absorption capacity(g^{-1})	0.88 ± 0.033	1.0 ± 0.047	21.79	0.42 ± 0.01	9.97	0.93 ± 0.03	16.37	0.94 ± 0.01	5.49
Oil absorption capacity(g g^{-1})	0.56 ± 0.016	0.57 ± 0.02	14.22	0.37 ± 0.00	3.29	0.69 ± 0.03	20.51	0.55 ± 0.01	5.98
Rehydration time(min)	30.75 ± 0.358	32.43 ± 0.14	2.21	32.29 ± 0.06	0.66	26.21 ± 0.52	9.10	32.72 ± 0.38	4.18
Calcium(mg 100 g^{-1})	7.13 ± 0.100	7.86 ± 0.07	4.73	7.51 ± 0.05	2.58	6.79 ± 0.08	5.50	5.65 ± 0.16	10.07
Potassium (mg 100 g^{-1})	8.60 ± 0.082	9.05 ± 0.09	4.99	7.93 ± 0.11	5.41	8.70 ± 0.18	9.48	8.25 ± 0.14	6.08
Phosphorus (mg 100 g^{-1})	8.05 ± 0.092	8.35 ± 0.13	8.41	7.45 ± 0.03	1.55	8.66 ± 0.13	6.80	7.09 ± 0.09	4.63
Zinc(mg 100 g^{-1})	0.27 ± 0.009	0.32 ± 0.01	15.83	0.33 ± 0.01	9.36	0.21 ± 0.01	12.85	0.15 ± 0.01	14.43
Iron(mg 100 g^{-1})	0.31 ± 0.010	0.39 ± 0.01	11.85	0.32 ± 0.01	10.94	0.27 ± 0.02	31.64	0.20 ± 0.01	14.41
Magnesium (mg 100 g^{-1})	0.26 ± 0.007	0.29 ± 0.01	11.92	0.29 ± 0.01	13.38	0.27 ± 0.01	22.95	0.15 ± 0.01	12.43

SE, standard error; CV, coefficient of variation; n, number of genotypes making up a cluster population; GS, genetic similarities.

of the traits. T63/967 × T17/524 was the most alkaline genotype, T65/7 × T53/8 had the highest protein, MAN 15-2 × T85/799 had the highest fat, GU147A × NA33 had the highest oil content, T82/27 × T16/17 structurally was most dispersible, T60/887 × ICS89 had both the highest bulk density and highest iron content, highest calcium, potassium, phosphorus and zinc was obtained in T86/2 × T53/8, C77 × C67, T65/7 × T79/501 and P7 × T60/887 respectively. The present study re-emphasizes the importance of cross breeding, a method which mostly result in heterosis among hybrids. Individual direct selection can be made on these genotypes for advancement or use as parent in di-hybrid cross breeding programme. For each of the clusters, the highest biochemical traits of significance were: high *p*H (i.e. low acidity), protein, Ca, K and Fe (cluster I, containing 82% hybrids, 7% introduced clones and 11% regional varieties), high zinc and magnesium (cluster II, containing introduced clones from international sources), high crude fat and phosphorus (cluster III, containing regional varieties) and high crude ash, fibre, theobromine, flavonoid and caffeine (cluster IV, containing the local clones). This provides a hint that the group of genotypes in each of the various field banks have specific uniqueness to which selection can be directed.

The 13 genotypes in the local clone field bank constitute the earliest introduction of cocoa genetic resources to Nigeria, comprising very few lower Amazon (e.g. N38), mostly upper Amazon and some Trinitario sub-species. The overall superiority in performances of the different genotypes for the various traits was across four field banks. The observed strict grouping of the thirteen cocoa clones in the local clone field bank with the highest content of theobromine, flavonoids and caffeine hints that some percentage of the earliest cocoa germplasm introduced to Nigeria are still intact and further reveal that the Nigerian cocoa germplasm still hold some genetic materials with strong chocolate flavour that is characteristic of the West African Amelonado. The other clusters where the content of theobromine, flavonoids and caffeine were lower must have undergone selections against the quality traits.

Aikpokpodion (2010) informed that Nigerian cocoa farmers hosts significant genetic variability in their fields, genomic divergent within Nigerian cocoa genetic resources is wide (Olasupo *et al.*, 2018), and the present report identified diversity for

https://doi.org/10.1017/S1479262123000977 Published online by Cambridge University Press

physicochemical and nutraceutical in the active breeding field banks of Nigeria cocoa, these presents good insurance against present and future threat to cocoa cultivation, production and quality improvement.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1479262123000977

Acknowledgements. The authors wish to appreciate the staff of the Nursery and Vegetative Propagation Units of the Crop Improvement Division, CRIN, for their support in the course of this work.

Author's contributions. **ABD and AOO** – Conceptualization, investigation and methodology; **AOO** – Data generation; **ABD** – Formal analysis; **ABD** – original draft; **AOO** – review and editing.

Funding statement. None.

Competing interests. None.

Data availability. Data are available on reasonable request from the corresponding author.

Informed consent. None.

References

Adeigbe OO, Adewale BD and Muyiwa AA (2021) Variability and character association of bean biochemical traits of cocoa (*Theobroma cacao*) genotypes in four Nigerian field banks: providing platform for nutrition-based selection. *Journal of Agricultural Food Research* **3**, 1–9 https://doi.org/10.1016/j.jafr.2021.100115

Adewale BD, Adeigbe OO, Adenuga OO, Adepoju AF, Muyiwa AA and Aikpokpodion PO (2013) Descriptive and discriminatory significance of pod phenotypic traits for diversity analysis of cocoa genotypes. *Journal of Plant Breeding and Genetics* **1**, 131–137.

Adeyeye EI (2016) Proximate, mineral and anti nutrient compositions of natural cocoa cake, cocoa liquor and alkalized cocoa powders sourced in Nigeria. *Journal of Advanced Pharmaceutical Science and Technology* **1**, 13–28. https://doi.org//10.14302/issn.2328-0182.japst-15-855

Adeyeye IE, Akinyeye OR, Ogunlade I, Olaofe O and Boluwade OJ (2010) Effect of farm and industrial processing on the amino acid profile of cocoa beans. *Food Chemistry* **118**, 357–363.

Afoakwa EO, Paterson A and Fowler M (2007) Factors influencing rheological and textural qualities in chocolate – a review. *Trends in Food Science and Technology* **18**, 290–298.

Aguilar MSE, Villacorta EMG, Contreras SQ, González JAS and León-Vargas JM (2017) Physicochemical and nutraceutical characterization of sirimbache fruit (*Gaultheria glomerata* (Cav.) Sleumer). *Scientia Agropecuaria* **8**: 411–417. http://dx.doi.org/10.17268/sci.agropecu.2017.04.12

Aikpokpodion PO (2010) Variation in agro-morphological characteristics of cacao, *Theobroma cacao L.*, in farmers' fields in Nigeria. *New Zealand Journal of Crop and Horticultural Science* **38**, 157–170. https://doi.org//10.1080/0028825X.2010.488786

Aikpokpodion PO, Motamayor JC, Adetimirin VO, Adu-Ampomah Y, Ingelbrecht I, Eskes AB, Schnell RJ and Kolesnikova-Allen M (2009) Genetic diversity assessment of sub-samples of cacao, *Theobroma cacao L.* collections in West Africa using simple sequence repeats marker. *Tree Genetics and Genomes* **5**, 699–711. https://doi.org//10.1007/s11295-009-0221-1

Aikpokpodion PO, Kolesnikova-Allen M, Adetimirin VO, Guiltinan MJ, Eskes AB, Motamayor JC and Schnell RJ (2010) Population structure and molecular characterization of Nigerian field gene-bank collections of Cacao, *Theobroma cacao L. Silvae Genetica* **59**, 273–284. https://doi.org//10.1515/sg-2010-0039

AOAC (1990) *Official Methods of Analysis*, 15th Edn. Washington, DC: Association of Official Analytical Chemist.

Atanda OA (1977) The Third Nigerian Cocoa Breeding Programme: A progress report. Paper presented at: Proceedings of the 5th International Cocoa Research Conference (Ibadan, Nigeria).

Bekele F and Phillips-Mora W (2019) Cacao (*Theobroma cacao* L.) breeding. In Al-Khayri JM, Jain S and Johnson D (eds), *Advances in Plant Breeding Strategies: Industrial and Food Crops*. Cham, Switzerland: Springer Nature Switzerland AG, Vol. 1, p. 89. https://doi.org/10.1007/978-3-030-23265-8_12

CacaoNet (2012) A Global Strategy for the Conservation and Use of Cacao Genetic Resources, as the Foundation for a Sustainable Cocoa Economy. Bioversity International, Montpellier, France. p. 186.

CRIN (2011) New Cocoa Varieties for Nigeria: Attributes and Field Management Requirement. Library Information and Documentation Department, Cocoa Research Institute of Nigeria, Ibadan, Nigeria, p. 43.

de Almeida AAF and Valle RR (2007) Ecophysiology of the cacao tree. *Brazilian Journal of Plant Physiology* **19**, 425–448.

Eghbal M (2018) Global Chocolate Industry: From Bean to Bar. Euromonitor International. Available at https://blog.euromonitor.com/global-chocolate-industry/ (accessed 3 September 2023).

Fitzpatrick JJ (2013) Powder properties in food production systems. In Bhandari B, Bansal N, Zhang M and Schuck P (eds), *Handbook of Food Powders*. Sawston, Cambridge: Wood head Publishing Series in Food Science, Technology and Nutrition, pp. 285–308. https://doi.org/10.1533/9780857098672.2.285

Gaikwad KB, Rani S, Kumar M, Gupta V, Babu PH, Bainsla NK and Yadav R (2020) Enhancing the nutritional quality of major food crops through conventional and genomics-assisted breeding. *Frontiers in Nutrition* **7**, 533453. https://doi.org/10.3389/fnut.2020.533453

Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics* **27**, 623–637.

ICCO (2021) Processing Cocoa: summary of the process of transforming cocoa beans into chocolate. Available at https://www.icco.org/processing-cocoa/ (accessed 3 September, 2023).

Jacob VJ, Atanda OA and Opeke LK (1971) Cacao breeding in Nigeria, Progress in Tree Crops Research in Nigeria. CRIN Commemorative Book, Ibadan.

Jayeola CO and Oluwadun AO (2010) Mycoflora and nutritional components of cocoa powder samples in South West Nigeria. *African Journal of Agricultural Research* **5**, 2694–2698.

Jiang C, Gao X, Liao L, Harberd NP and Fu X (2007) Phosphate starvation root architecture and anthocyanin accumulation responses modulated by the gibberellin-DELLA signaling pathway in arabidopsis1 [OA]. *Plant Physiology* **145**, 1460–1470.

Kaufman L and Rousseeuw PJ (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley and Son.

Lin JY and Tang CY (2007) Determination of total phenolic and flavonoid contents in selected fruits and vegetables, as well as their stimulatory effects on mouse splenocyte proliferation. *Food Chemistry* **101**, 140–147.

Lockwood G and Gyamfi MMO (1979) The CRIG cocoa germplasm collection with notes on codes used in the breeding programme at Tafo and elsewhere. Cocoa Research Institute, Ghana Technical Bulletin No. 10.

Marita JM, Nienhuis J, Pires JL and Aitken WM (2001) Analysis of genetic diversity in *Theobroma cacao* with emphasis on witches' broom disease resistance. *Crop Science* **41**, 1305–1316.

Meng CC, Jalil AMM and Ismail A (2009) Phenolic and theobromine contents of commercial dark, milk and white chocolates on the Malaysian market. *Molecules* **14**, 200–209. https://doi.org/10.3390/molecules14010200

Ndife J, Bolaji P, Atoyebi D and Umezuruike C (2013) Production and quality evaluation of cocoa products - plain cocoa powder and chocolate. *American Journal of Food Nutrition* **3**, 31–38.

Olasupo FO, Adewale BD, Aikpokpodion PO, Muyiwa AA, Bhattacharjee R, Gutierrez OA, Motamayor JC, Schnell RJ, Eba S and Zhang D (2018) Genetic identity and diversity of Nigerian cacao genebank collections verified by single nucleotide polymorphisms (SNPs): a guide to field genebank management and utilization. *Tree Genetics and Genomes* **14**, 32–47. https://doi.org/10.1007/s11295-018-1244-2

Olatoye ST and Esan EB (1992) Recent innovation in cacao genetic resources conservation in Nigeria. Paper presented at: International Workshop on Conservation, Characterization and Utilization of Cacao Genetic Resources in the 21st Century, Cocoa Research Unit, Port of Spain, Trinidad, pp. 281–291.

Ooi DJ, Iqbal S and Ismail M (2012) Proximate composition, nutritional attributes and mineral composition of *Peperomia pellucida* L. (Ketumpangan Air) grown in Malaysia. *Molecules* **17**, 11139–11145. https://doi.org/10.3390/molecules170911139

Opoku SY, Bhattacharjee R, Kolesnikova-Allen M, Motamayor JC, Schnell R, Ingelbrecht I, Enu-Kwesi I and Adu-Ampomah Y (2007) Genetic diversity in cocoa (*Theobroma cacao L.*) germplasm collection from Ghana. *Journal of Crop Improvement* **20**, 73–87. https://doi.org/10.1300/J411v20n01_04

Poitevin E (2016) Official methods for the determination of minerals and trace elements in infant formula and milk products: a review. *Journal of AOAC International* **99**, 1–11.

Pokou DN, Ngoran JAK, Lachenaud P, Eskes A, Montamayor JC, Schnell RJ, Kolesnikova-Allen M, Clement D and Sangare A (2009) Recurrent selection of cocoa populations in Côte d'Ivoire: Comparative genetic diversity between the first and second cycles. *Plant Breeding* **128**, 514–520.

Posnette AF and Todd JMA (1951) Virus diseases of cacao in West Africa. VIII. The search for virus resistant cacao. *Annals of Applied Biology* **38**, 785–800.

Rawel HM, Huschek G, Sagu ST and Homann T (2019) Cocoa bean proteins – characterization, changes and modifications due to ripening and post-harvest processing. *Nutrients* **11**, 428. Available at www.mdpi.com/journal/nutrients

Reed S (2010) Sensory Analysis of Chocolate Liquor, Cargill Incorporated. Available at www.wilburchocolate.com

Rosa L, Bonaccorsi di, Patti MC, Vasilyev V and Cutone A (2022) Editorial: nutraceuticals: new perspectives and approaches in human health and disease. *Frontier Pharmacology* **13**, 1003529. https://doi.org/10.3389/fphar.2022.1003529

SAS (2011) *SAS/STAT Software. Version 9.4.* Cary: SAS Institute Inc.

Sonwa DJ, Nkongmeneck BA, Weise SF, Tchatat M, Adesina AA and Janssens MJJ (2007) Diversity of plants in cocoa agroforests in the humid forest zone of Southern Cameroon. *Biodiversity Conservation* **16**, 2385–2400. https://doi.org/10.1007/s10531-007-9187-1

Thomas JB, Yen JH and Schantz MM (2004) Determination of caffeine, theobromine and theophylline in standard reference material 2384, baking chocolate, using reversed-phase liquid chromatography. *Journal of Agricultural Food Chemistry* **52**, 3259–3263.

Torres-Moreno M, Torrescasana E, Salas-Salvadó J and Blanch C (2015) Nutritional composition and fatty acids profile in cocoa beans and chocolates with different geographical origin and processing conditions. *Food Chemistry* **166**, 125–132.

Vijayakumar PP and Adedeji A (2017) Measuring the pH of food products. Cooperative Extension Service, Agriculture and Natural Resources, Family and Consumer Sciences.ID-246. Available at www.ca.uky.edu

Ward JH (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**, 236–244.

Wood GAR (1987) Quality and inspection. In Wood GAR and Lass RA (eds), *Cocoa*. New York: Longman Scientific & Technical, pp. 620–632.

WorldAtlas (2021) The Top Cocoa-Producing Countries. Available at https://www.worldatlas.com/articles/top-10-cocoa-producing-countries.html (accessed 3 September 2023).

World Cocoa Foundation (2012) Cocoa Market Update. Available at https://www.worldcocoafoundation.org/wp-content/uploads/Cocoa-Market-Update-as-of-3.20.2012.pdf (accessed 3 September 2023).

https://doi.org/10.1017/S1479262123000977 Published online by Cambridge University Press