Ghana J. Sci. **63** (1), 2022, 74 - 82 https://dx.doi.org/10.4314/gjs.v63i.5

SSR-BASED GENETIC STRUCTURE STUDY OF SEVENTY-EIGHT COWPEA (*Vigna unguiculata* (L.) Walp) GENOTYPES

P. A. D. Zoryeku, I. K. Asante*, I. K. Blay, E. Ajara Seidu & F. Oppong
(P. A. D. Z., I. K. A., E. A. S. & F. O.: Department of Plant and Environmental Biology, I. K. B.: Department of Crop Science, University of Ghana, Legon).
** Corresponding author's email: asanteisaack57@gmail.com*

ABSTRACT

Seventy-eight cowpea accessions were studied using short sequence repeats (SSR) technique. Genetic structure of these accessions was studied using three SSR polymorphic primers, SSR-6206, SSR-6218 and SSR-6219. A total of eight loci were scored for the three primers with a total of ten alleles. Bayesian clustering method grouped the cowpea genotypes into 4 sub-populations. Ancestral allele frequencies ranged between 0.128 and 0.802, while allele frequencies within sub-populations ranged from 0.001 and 0.997. Allele frequency divergence among sub-populations ranged from 0.145 to 0.406. Expected heterozygosity between individuals in the same sub-population ranged from 0.084 and 0.26, Mean genetic differentiation among sub-populations ranged from 0.374 and 0.687, with a mean geneflow ranging from 0.228 and 0.837. There was relative uniformity within the sub-populations which can be accounted for by independent random genetic drift.

Keywords: SSR, genetic structure, cowpea, heterozygosity, genetic differentiation**.**

Introduction

Natural populations are structured into sub-populations which do not permit total random mating making them genetically distinct from each other. This therefore would result in a set of individuals characterized by some measure of genetic distinction. Such genetic differences may be due to geographic ancestry. Structure in this sense refers to any deviation from random-mating, including inbreeding and assortative mating. It is important to know how much variation exists within each level of structure relative to other levels. Wright (1951) developed the F-statistics, a set of measures for departures from Hardy-Weinberg equilibrium in sub-divided populations. Fixation index (F) a measure of increased homozygosity; Inbreeding coefficient (F_{IS}) a correlation of uniting gametes relative to gametes drawn at random from within a sub-population (Individual within the Sub-population) and F_{ST} a measure of population sub-structure which is beneficial for analysing the overall genetic divergence amongst sub-populations. Elements of population structure assessment include similarity and dissimilarity measure, genetic distances between individuals, clustering and differentiation among sub-populations (Panda *et al.*, 2015).

Genetic diversity plays a key role in the continued existence of species. An insight of the dynamics of genetic diversity within and among populations of species is key to the development of optimum genetic resource management strategies towards conservation, sustainable utilization and genetic

improvement (Panda *et al*., 2015). Genetic diversity assessment is very vital to mitigate threats of environmental fluctuations. For the maximum exploitation of genetic resources in breeding programmes, knowledge about genetic diversity is paramount. Availability of genetic diversity within plant species aids in the selection of superior lines for crop improvement (Ali *et al*., 2007).

Genetic differentiation in populations may be due to a lot of factors, among which are the following suggested by Falconer (1988): random genetic drift occurring independently in different sub-populations, uniformity within sub-populations due to progressive reduction in genetic variation within each sub-population, geographical or ecological causes under natural conditions or from controlled breeding in domesticated or laboratory breeding, genetic drift (Mettler and Gregg, 1969), occurring populations that become small periodically (that is "bottle-neck effect") and populations established by a few emigrants ("founders principle") carrying a small sample of genetic variation from a large population. The study was conducted to study the genetic structure of Ghanaian cowpeas at three SSR gene loci. Such a study will be useful to the cowpea breeders and cowpea germplasm curators.

Experimental

Sample materials

Seventy-eight cowpea accessions were used for the research. These accessions were obtained from the CSIR-Savanna Agricultural Research Institute (SARI) and the Department of Plant and Environmental Biology (DPEB), University of Ghana. The genotypes were: Zaayura, Padi-tuya, Apaagbala, Asontem, IT07K-299-6, SARVX-09-002, IT98K-628, Golinga, SARI-6-2-6, SARVX-09-004, SARVX-07-001, SARI-13-17-2, AS007, Bawuta, Songotra, SARI-5-5-5, IT07K-298-45, IT10K-817-3, IT86D-610, BOTN312, RC-04, BOTN6824, BOTN442, SARI-13-17-2, ASOO5, RC-10, IT08K, SARI-2-50-80, IT08K-150-24, SARI-3-11-88, SARI-1-50-81, IT08K-137-1, Beledi-B, SARI-1-3-90, AGRAC-216, BOTN4223, BOTN003, BOTN007, BOTN002, AS004, BOTN2213, AS008, SARI-1-3-90, Laduni 1B, SARI-3-11-100, IT07K-299-69, BOTN111, AS003, AS009, BOTN6312, EMS30-20, WACCI-TONI, AS006, F2RG009, RC-02, COLMUT10-01, EMS10-01, L69, BOTN212, F2RG011, G41, G35, BOTN422, G25, G84, G18, F2RG011, F2RG004, RC-03, F2RG007, BOTN006, F2RG012, F2RG010, G18, G80, G22, G7, G50 and IT98K-503-1.

SSR analysis

Fifty cowpeas derived SSR primers were screened and three of them, SSR-6206, SSR-6218 and SSR-6219 (Table 1), which produced consistent, reproducible and polymorphic bands were used for DNA amplification.

Collection of leaf samples

The first trifoliate leaves of one plant per accession was harvested two weeks after planting and stored at -80^{o} prior to DNA extraction.

DNA extraction

The Doyle and Doyle (1990) method with modification was used for genomic DNA extraction. Genomic DNA was extracted from one gram of young cowpea leaves harvested on ice. The leaf tissues were ground in liquid nitrogen using a micropestle. The ground leave sample was transferred into Eppendorf tubes and 500 µl of prewarmed (60 ^{o}C) cetyltrimethyl ammonium bromide (CTAB) extraction buffer

was added. The solution was vortexed and incubated at 65 °C for one hour. A solution of phenol/chloroform was added. DNA was precipitated by adding equal volumes of cold isopropanol and NaOAc. DNA pellets were washed using EtOH. The washed DNA pellets were allowed to dry at room temperature. The dried DNA pellets were then dissolved in 1xTE buffer (10 MM Tris (pH 8.0); 0.1 MM EDTA) and treated with RNase (5 µg/µl) at 37 °C for 30 minutes. DNA quality and quantity were checked using Nano-drop spectrophotometer. DNA was stained with a blue dye and run using 1.5% agarose gel electrophoresis.

TABLE 1

List of three cowpea SSR markers

Name	Code	Sequence
SSR-6206 (F)	SSR21	AGGCATGCATTCATCTTTCC
SSR-6206 (R)		GCAGTCATAACCCCAAAACAA
SSR-6218 (F)	SSR34	GTGGAAGGAATGGGTCCAG
SSR-6218 (R)		AGGAAATTTGCATTCCTTTGT
SSR-6219 (F)	SSR35	ACAATGCACAAAATGTGAATCTC
SSR-6219 (R)		GGGAAGCTTAGGAAAAGTTTGA

PCR amplification

The three primers were used to carry out PCR amplification. The cocktail prepared for amplification contained the following: 2 µl template DNA, 6.25 µl master mix, 0.5 µl of primer (forward and backward), and 3.25 µl of double distilled water. The PCR tubes were placed in a thermocycler and programmed for initial denaturation at 95°C for 3 minutes, followed by 30 cycles for 30 seconds at 94°C, 30 seconds at 53°C, 2 minutes at 72°C, and final extension for 15 minutes at 72°C .

Electrophoresis

The PCR product was electrophoresed in 1.5% agarose gel in TAE buffer at a voltage of 100 V for 2 hours 30 minutes. The agarose gel was stained with ethidium bromide and visualized and scored under UV.

Statistical analysis

Band (allele) positions for each cowpea genotype were scored by micro-satellite repeat score approach, where each allele at a given locus was coded by a unique integer and missing allele coded as -999. Marker genotype data was prepared in Microsoft excel and the file saved in Text (tab delimited) type. The software STRUCTURE version 2.2. (Pritchard *et al*., 2000) was used for data analysis. To determine memberships of cowpea genotypes, number and membership of clusters, allele frequency distribution in the clusters, allele frequency divergence among clusters and gene flow.

Results and discussion

Distribution of loci and alleles

Loci and alleles distribution at the 3 SSR markers are presented in Table 2. There was a total of eight loci that gave a total of ten alleles.

TABLE 2

Distribution of loci and alleles at the three SSR markers

Marker	Loci	Number of alleles present
SSR-6206	SSR1	1
	SSR2	2
	SSR3	1
SSR-6218	SSR1	1
	SSR2	1
	SSR3	1
SSR-6219	SSR1	2
	SSR2	1

Proportion of membership

The Bayesian Clustering method grouped the 72 cowpea genotypes into four sub-populations. Proportion of membership of the 72 cowpea genotypes in each of the 4 sup-populations is presented in Tables 3. Sub-population 1 had the lowest proportion of membership of 0.124, while sub-population 3 had the highest proportion of 0.317.

TABLE 3

Overall proportion of membership of the 72 cowpea genotypes in each of the 4 sup-populations

Sub-population	Proportion of membership
1	0.124
2	0.317
3	0.281
4	0.278

Estimated ancestral and sub-populations allele frequencies

Estimated ancestral and sub-populations allele frequencies at the eight SSR gene loci are presented in Table 4. Ancestral allele frequencies ranged between 0.128 (SSR1-2, SSR21-2) and 0.802 (SSR8-1). Within sub-population 1 the allele frequencies ranged from 0.008 (SSR7-2) to 0.993 (SSR8-1), while in sub-population 2 they ranged from 0.001 (SSR2-2, SSR7-2) to 0.998 (SSR8-2). The allele frequencies in sub-population 3 ranged from 0.012 (SSR2-2) to 0.913 (SSR3-1), while in sub-population 4 they ranged between 0.003 (SSR7-2) to 0.997 (SSR5-1).

TABLE 4

Distribution of estimated ancestral and sub-populations allele frequencies at the eight SSR gene loci

Locus	Allele	Ancestral	Sub-population			
			1	2	3	4
SSR1	1	0.295	0.616	0.005	0.130	0.053
*	2	0.705	0.384	0.995	0.870	0.947
SSR2	1	0.452	0.316	0.510	0.311	0.289
	2	0.128	0.656	0.001	0.012	0.001
*	3	0.420	0.028	0.489	0.677	0.710
SSR3	1	0.757	0.240	0.994	0.913	0.992
*	2	0.243	0.760	0.006	0.087	0.008
SSR4	1	0.612	0.978	0.898	0.027	0.995
*	2	0.388	0.022	0.102	0.973	0.005
SSR5	1	0.781	0.984	0.998	0.430	0.997
*	2	0.219	0.016	0.002	0.570	0.003
SSR6	1	0.463	0.952	0.006	0.095	0.993
*	2	0.537	0.048	0.994	0.905	0.007
SSR7	1	0.382	0.972	0.394	0.146	0.024
	2	0.128	0.008	0.001	0.048	0.003
*	3	0.491	0.020	0.605	0.806	0.973
SSR8	1	0.802	0.993	0.998	0.842	0.970
*	2	0.198	0.007	0.002	0.158	0.030

- Frequency of missing allele

Membership probabilities of genotypes

Membership probabilities of each genotype in each of the four sub-populations are shown in Table 5 to Table 8. Membership probabilities ranged from 0.547 (IT0IK-298-45) to 0.985 (F2RG010) in sub-population 1, in sub-population 2 it ranged from 0.662 (F2RG004) to 0.963 (BOTN003, AS006. 0.466 (GENO06, sub-population 3) to 0.985 (GENO69, sub-population 1). In sub-population 3 the range was between 0.466 (Songotra) to 0.980 (2213) and in sub-population 4 the range was between 0.584 (SARI-2-50-80) and 0.969 (F2RG004, 198K-503-1.

TABLE 5

Membership probabilities of Inferred ancestry of individual cowpea genotypes in sub-population 1

Sub-population	Accession name	Genotype	Probability
1	IT01K-298-45	GENO20	0.840
1	6312	GENO52	0.547
1	422	GENO67	0.969
1	F2RG007	GENO68	0.984
1	F2RG010	GENO69	0.985
1	F2RG009	GENO70	0.967

TABLE 6

Membership probabilities of Inferred ancestry of individual cowpea genotypes in sub-population 2

Sub-population	Accession name	Genotype	Probability
2	Padi-tuya	GENO2	0.942
2	Asontem	GENO4	0.952
2	*	GENO7	0.956
2	BELEDI B	GENO25	0.934
2	IT07K-299-69	GENO26	0.913
2	SARI-3-11-180	GENO27	0.917
2	4228	GENO28	0.955
2	IT08K-	GENO36	0.918
2	SARI-6-2-9	GENO37	0.902
2	SARI-3-11-100	GENO39	0.956
2	EMS10-01	GENO40	0.956
2	BOTN003	GENO41	0.963
2	AS006	GENO42	0.963
2	EMS30-02	GENO43	0.901
2	AS010	GENO47	0.915
2	AS009	GENO53	0.954
2	RC-02	GENO54	0.952
2	AS003	GENO55	0.961
2	F2RG004	GENO56	0.662
2	684	GENO57	0.961
2	BON006	GENO58	0.958
2	RC-01	GENO60	0.500
2	G80	GENO61	0.913
2	G5	GENO76	0.961

TABLE 7

Membership probabilities of Inferred ancestry of individual cowpea genotypes in sub-population 3

Sub-population	Accession name	Genotype	Probability
3	Zaayura	GENO1	0.863
3	Apaagbala	GENO3	0.891
3	Bawuta	GENO5	0.975
3	Songotra	GENO6	0.466
3	*	GENO22	0.912
3	*	GENO23	0.959
3	*	GENO24	0.883
3	Agrac-216	GENO33	0.860
3	SARI-1-50-81	GENO34	0.908
3	AS005	GENO35	0.913
3	IT08K	GENO38	0.807
3	WACCI TONI	GENO44	0.913
3	BOTN002	GENO45	0.857
3	2213	GENO46	0.980
3	AS010	GENO48	0.967
3	L69	GENO49	0.867
3	G41	GENO59	0.976
3	G84	GENO62	0.965
3	212	GENO63	0.941
3	442	GENO72	0.943
3	G35	GENO73	0.575
3	G7	GENO74	0.514
3	G22	GENO75	0.847

TABLE 8

Membership probabilities of Inferred ancestry of individual cowpea genotypes in sub-population 4

Sub-population	Accession name	Genotype	Probability
4	*	GENO7	0.941
4	*	GENO8	0.945
4	*	GENO9	0.968
4	*	GENO10	0.820
4	*	GENO11	0.942
4	*	GENO12	0.680
4	SARVX-09-04	GENO13	0.967
4	*	GENO14	0.969
4	*	GENO15	0.946
4	SARI-13-17-2	GENO18	0.970
4	SARI-5-5-5	GENO19	0.944
4	111	GENO29	0.968
4	SARI-2-50-80	GENO30	0.584
4	LUNDI 1B	GENO31	0.969
4	AS004	GENO50	0.968
4	COLMUT10-01	GENO51	0.968
4	Asontem	GENO64	0.820
4	F2RG004	GENO65	0.969
4	RC-03	GENO66	0.588
4	I98K-503-1	GENO71	0.969
4	F2RG002	GENO77	0.846

Nei's nucleotide distance among sub-populations

Results for Nei's nucleotide distance among sub-populations are shown in Table 9. The distances represent allele-frequency divergence among the sub-populations. The distance ranged from 0.145 (sub-populations 2 and 4) to 0.466 (between sub-populations 1 and 3) with an average distance of 0.274.

TABLE 9

Allele-frequency divergence among sub-populations (Net nucleotide distance), computed using point estimates of P

	Sub-population 1	Sub-population 2	Sub-population 3
Sub-population 2	0.314		
Sub-population 3	0.466	174	
Sub-population 4	0.280	0.145	0.265

Average distance (expected heterozygosity) between individual in the same cluster

Results for average distance between individual cowpea genotypes in the same sub-population are presented in Table 10. Average distance ranged from 0.084 (in sub-population 4) to 0.26 (in sub-population 3).

TABLE 10

Average distances (expected heterozygosity) between individual cowpea genotypes in the same cluster

Sub-population	Average distance
1	0.193
2	0.131
3	0.267
4	0.084

Genetic differentiation (F_{st}) and geneflow (N_m)

Table 11 shows results for mean genetic differentiation and gene flow among the four clusters. Mean F_{st} values ranged from 0.374 (sub-population 3) to 0.687 (sub-populations 2 and 4). Mean gene flow ranged from 0.228 (sub-populations 2 and 4) to 0.837 (sub-population 3).

TABLE 11

Estimated mean genetic differentiation (Fst) and gene flow among the four sub-populations

Sub-population	Mean Fst value	Gene Flow
1	0.632	0.291
2	0.687	0.228
3	0.374	0.837
4	0.687	0.228

In a study of genetic diversity and population structure of cowpea from Togo based on Dart marker, Kodjo *et al*. (2021) observed mean Fst value of 0.072. Desalegene *et al*. (2016) also reported mean Fst value of 0.075 among cowpea accessions collected from Ethiopia using SSR markers. In a similar study by Kimaro *et al.* (2020) observed a mean Fst value of 0.032 in 48 pigeon pea germplasm collections using SSR markers. The level of differentiation was higher in the current study as compared to those observed among germplasm lines studied by the previous workers. The difference between the results in the current study and the previous results could be due to different genotypes.

The four sub-populations represent 4 major gene pools of the 78 cowpea genotypes. These sub-populations differed in the distribution of the 8 SSR loci. There was differential allele frequency distribution within the four sub-populations. Allele frequencies for the sub-populations were either lower or higher than their corresponding ancestral allele frequencies. There was therefore an indication of some alleles approaching either extinction or fixation in the respective sub-populations. Allele SSR1-1 is approaching extinction in sub-population 2 with a frequency of 0.005in sub-populations 2, 3 and 4 allele SSR2-2 is approaching extinction with frequencies of 0.001, 0.012 and 0.001 respectively. Similarly, alleles SSR7-2 is approaching extinction in all of the four sub-populations. Allele SSR3-1 is approaching fixation in sub-populations 2 and 4 with frequencies of 0.994 and 0.992, respectively, while allele SSR4-1 is also approaching fixation in sub-populations 1 and 4 with frequencies of 0.978 and 0.995, respectively. Allele SSR5-1 is also approaching fixation in sub-populations 1, 2 and 4 with frequencies of 0.984, 0.998 and 0.997, respectively. In sub-populations 1 and 2, allele SSR6-1is approaching fixation with frequencies of 0.952 and 0.993, respectively. Allele SSR7-2 is also approaching fixation in sub-population 1. Allele SSR8-1 is approaching fixation in sub-populations 1, 2 and 4 with allele frequencies 0.995, 0.998 and 0.970, respectively.

Different levels of genetic differentiation were observed for the 4 gene pools. The levels of genetic differentiation for the 4 sub-populations imply that 31.3% of genetic differentiation resided within genotypes for sub-population 2 and sub-population 4, respectively. In sub-population 1 and sub-population 3 36.8% and 62.6% of genetic differentiation resided within the genotypes. A relatively high level of geneflow of about 85.0% occurred in sub-population 3, while relatively low levels occurred in the remaining three sub-populations. The observations for the current study were at variance with those of Asante *et al*. (2003) who observed 92.5% of genetic differentiation within cowpea genotypes and a geneflow of 100.0% in a study of genetic structure at the isocitrate dehydrogenase and malate dehydrogenase enzyme gene loci of Ghanaian cowpea accessions. These differences could be due to differences in the cowpea genotype and also the genetic markers used. In the current study, there is relative uniformity within the sub-populations which can be accounted for by independent random genetic drift.

Acknowledgements

The authors are grateful to the CSIR-Savanna Agricultural Research Institute for the provision of part of the cowpea genotypes used for the research. They also wish to acknowledge the Biotechnology Centre of the College of Basic and Applied Science, University of Ghana for laboratory space where the SSR analysis was conducted.

References

Asante, I. K., Laing, E., Danquah, E. Y. Danquah & Offei, S. K. (2003) Genetic structure at the isocitrate dehydrogenase and malate dehydrogenase enzyme gene loci in cowpea (*Vigna unguiculata* (L.) Walp) accessions. *Ghana Jnl agric. Sci.* **36**, 53 – 63.

Doyle, D. J. & Doyle, J. L. (1990) A rapid total DNA preparation procedure for fresh plant tissue. *Focus* 12: 13 – 15.

Falconer, D. S. (1988) *Introduction to quantitative genetics*. 3rd edition. English Language Book Society. Longman. England.

Kimaro, D., Melis, R., Sibiya, J., Shimelis, H. & Shayanowako, A. (2020) Analysis of genetic diversity and population structure of pigeonpea [*Cajanus cajan* (L.) Millsp.] accessions using SSR Markers. *Plants* **9**, 1643.

Gbedevi, K. M., Boukar, O., Ishikawa, H., Abe, A., Ongom, P. O., Unachukwu, N., Rabbi, I. & Fatokun, C. (2021) Genetic Diversity and Population Structure of Cowpea [*Vigna unguiculata* (L.) Walp.] Germplasm Collected from Togo Based on DArT Marker. *Genes* **12**, 1451.

Mettler, L. E. & Gregg, T. G. (1969) *Population genetics and evolution.* Prentice Hall Inc. Eaglewood Cliffs. New Jersey.

Panda, S., Naik, D. & Kamble, A. (2015) Population structure and genetic diversity of the perennial medicinal shrub Plumbago AoB PLANT.

Pritchard, J. K., Stephen, M. J. & Donnelly (2000) Inference of population structure using multilocus genotype data. *Genetics **155*** (2), 945 – 59.

Wright, S. (1952) The theoretical variance within and among subdivisions of a population that is in a steady state. *Genetics* 33: 312.

Received 31 May 21; revised 20 Jul 22.