

## Article

# Analysis of genetic diversity and population structure of cowpea [*Vigna unguiculata* (L.) Walp] genotypes using single nucleotide polymorphism markers

Mbali Thembi Gumede <sup>1,2\*</sup>, Abe Shegro Gerrano <sup>2,3</sup>, Assefa Beyene Amelework <sup>2</sup> and Albert Thembinkosi Modi <sup>1</sup>

<sup>1</sup>School of Agricultural, Earth and Environmental Sciences, College of Agriculture, Engineering and Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville, 3209, Pietermaritzburg, Republic of South Africa.

<sup>2</sup>Agricultural Research Council- Vegetables, Industrial and Medicinal Plant Institute, Private Bag X293, Pretoria 0001, Republic of South Africa.

<sup>3</sup>Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717-3150, United States of America.

\* Correspondence: mbaliphakathwayo@gmail.com; ORCID ID: <https://orcid.org/0000-0002-1414-4510>; ORCID ID: <http://orcid.org/0000-0001-7472-8246>

**Abstract:** Cowpea [*Vigna unguiculata* (L.) Walp] is an important legume crop with immense potential for nutritional and food security, income generation and livestock feed in Sub-Saharan Africa. The crop is highly tolerant to heat and drought stresses that makes it an extremely important crop for improving resilience in crop production in the face of climate change. This study was carried out to assess the genetic diversity and population structure of 90 cowpea accessions using Single Nucleotide Polymorphism (SNP) markers. Out of 11 940 SNPs used, 5864 SNPs were highly polymorphic and maintained for genome diversity analysis. Polymorphic information content (PIC) values ranged from 0.22 to 0.32 with a mean value of 0.27. The model based population structure analysis classified the 90 cowpea accessions into four sub populations at K=4, while the distance-based cluster analysis grouped the accessions into three distinct clusters. The analysis of molecular variance (AMOVA) revealed that 59% and 69% of the total molecular variation was attributed to among individual variation for model-based and distance-based population, respectively and 18% was attributed to within individual variations. Furthermore, the low heterozygosity among cowpea accessions and the high inbreeding coefficient observed in this study suggests that the accessions reached acceptable level of homozygosity. This study would serve as a reference for future selection and breeding programs of cowpea with desirable traits and for systematic conservation of these plant genetic resources.

**Citation:** Gumede, M.T.; Gerrano, A.S.; Amelework, A.B.; Modi, A.T. Analysis of genetic diversity and population structure of cowpea [*Vigna unguiculata* (L.) Walp] genotypes using single nucleotide polymorphism markers. *Plants* **2022**, volume number, x. <https://doi.org/10.3390/xxxxx>

Academic Editor(s):

Received: date

Accepted: date

Published: date

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Cowpea; genetic diversity; population structure; single nucleotide polymorphism

## 1. Introduction

Cowpea [*Vigna unguiculata* (L.) Walp, 2n=2x=22] is an important legume crop that belongs to the genus *Vigna*, family Fabaceae and order Fabales, with a genome size of 620 million base pairs [1]. The crop is an annual herbaceous plant widely grown in tropical and subtropical regions of Sub-Saharan Africa, East Asia and other developing countries [2-3]. The crop plays a major role in both human and animal nutrition and food as well as feed security and income generation for farmers and agro-traders [3]. In addition to the crop's importance in sustaining food security, cowpea possess high tolerance to heat and drought stresses which is extremely importance in improving the resilience of the crop to the current climate change [4-5].

Moreover, cowpea has a significant importance to cropping systems on account of its ability to grow in low fertility soils, as a complementary crop in rotation with cereals

to break the life cycle of pathogens of cereals infested in the soil and consequently improve the fertility of the soil [6]. Globally, the cowpea average yield ranges from 0.1 t ha<sup>-1</sup> to 0.59 t ha<sup>-1</sup> is lower than its expected potential yield of 1.5 t ha<sup>-1</sup> to 3 t ha<sup>-1</sup> under suitable environmental conditions [7]. This is due to the narrow genetic base of improved varieties and their susceptibility to abiotic and biotic stresses. It is of high importance to develop improved varieties of cowpea to increase productivity in order to help alleviate poverty in Sub-Saharan Africa and also to meet the market demand. The production of improved cowpea varieties for traits such as high yield and nutritional status will largely benefit both subsistence and commercial farmers. Therefore, characterizing genetic diversity in any crop species is important for optimal germplasm utilization, conservation and crop improvement programmes. In support, Kondwakwenda [8]) also stated that the availability of appropriate genetic diversity is imperative for the success and sustainability of any crop breeding programme.

The assessment of genetic diversity of a particular crop is achieved using morphological, biochemical and molecular markers [3,9]. Although biochemical markers are reliable than morphological markers, they are both reported to be influenced by environmental factors. These markers provide genetic diversity information on the basis of genotype performances using agronomic traits and may differ at different stages of growth and development as well as growing environment. This may hinder the real genetic variation among genotypes [10-11] and reduce the accuracy of the results obtained. Hence, there was an evolution of the development of DNA molecular markers [12-15]. Molecular markers are neutral to environmental effects and the genetic diversity is reviewed at genomic DNA level, therefore, it is able to visualize the accurate genetic diversity among genotypes [3,16]. However, they may not be associated to any agronomic traits and needs to be supplemented with morphological markers in order to infer a meaningful conclusion.

There are several DNA markers that have been developed to determine the genetic diversity of cowpea. These include restriction fragment length polymorphism (RFLP) [16-17], simple sequence repeats (SSR) [18], single nucleotide polymorphism (SNP) [19], amplified fragment length polymorphic (AFLP) [20] and random amplified polymorphic DNA (RAPD) (Ba et al., 2004). In recent advances in molecular genetics and molecular biology, the use of SNP marker has emerged to be the most preferred molecular markers because of its high genomic abundance, cost-effectiveness, reliability and ease of application compared to other polymerase chain reaction (PCR)-based molecular markers [21-22]. Hence, the SNP markers were used in the current study. The study by Desalegne [23] compared the efficiency of SNP and SSR marker based analysis of genetic diversity in 95 cowpea accessions collected from East Africa and International Institute of Tropical Agriculture (IITA) inbred lines. Their study revealed that SNP markers were found to be more effective than SSR marker to determine the relationship between cowpea varieties, hence, the study suggested the utilization of SNP markers in future analysis of genetic diversity and population structure in cowpea. Similarly, the recent studies by Nkhoma [19] evaluated the genetic diversity in 90 cowpea genotypes using SNP markers and phenotypic traits and the study showed that SNP markers were more efficient in discriminating the diversity among and within the cowpea genotypes evaluated.

There are sequencing technology-based tools that has emerged through SNP markers, which includes genotyping-by-sequencing (GBS) and next generation sequencing (NGS) which has been reported to be efficient, inexpensive and fast developing in sequencing plant genomes [24-25]. In addition, a new type of marker known as diversity arrays technology (DArT) has been recently developed for genotyping and genome sequencing needlessly of sequence information [3]. In cowpea, DArT marker has been recently used by Gbedevi [3] to study the genetic diversity and population structure of 498 cowpea accessions collected from Republic of Togo and their study revealed the presence of four major clusters among the accessions studied and the accessions were not clustered according to the regions where they were collected suggesting that the clustering did not closely resemble the geographical areas of the collections. The DArT marker has developed a GBS method called DArTseq sequencing technology for genotyping in legumes

including cowpea [3, 26], chickpea [27], common beans [28] and Bambara groundnut [29]. Hence, the objective of this study was to assess the magnitude of the genetic diversity and population structure among cowpea genotypes using single nucleotide polymorphisms (SNP) markers.

## 2. Results

### *Allele polymorphism*

SNP distribution per chromosome and the gene diversity parameters measured from 90 cowpea accessions is presented in Table 1. The genetic diversity parameters analysis was conducted using 5864 (49%) SNPs that remained after filtering out monomorphic and minor allele frequency of less than 2%. The number of polymorphic SNPs per chromosome ranged from 345 on chromosome 1 to 668 on chromosome 3 with an overall mean of 488 per chromosome. The proportion of polymorphic SNPs per chromosome varied from 31.82% on SNPs with unknown chromosome origin to 57.35% SNPs on chromosome 9, with an overall mean of value of 49.11% per chromosome. The mean number of effective allele ( $N_e$ ) per chromosome was the highest on chromosome 7 ( $1.48 \pm 0.013$ ), followed by chromosome 10 ( $1.47 \pm 0.013$ ) and chromosome 4 ( $1.47 \pm 0.013$ ) whilst the lowest values were observed on chromosome 9 ( $1.34 \pm 0.014$ ) and chromosome 8 ( $1.37 \pm 0.013$ ). The observed heterozygosity ranged from 7.6% to 9.6% with the mean value of 8.4%. The unbiased gene diversity ( $uH_e$ ) values ranging from 0.221 to 0.291 per chromosome with an overall mean of 0.267. The fixation index ( $F_{IS}$ ) values ranged from 64% on chromosome 11 to 72% on chromosome 7 with the mean value of 67%. The mean polymorphic information content (PIC) value was 0.27, in which the PIC values per chromosome ranged from 0.22 to 0.32.

**Table 1.** Genetic diversity within and among 90 cowpea accessions genotypes based on 5864 SNPs markers.

| Chromosome   | NSU   | NPS  | %P    | $N_e$        | $H_o$         | $H_e$         | $F_{IS}$     | PIC          |
|--------------|-------|------|-------|--------------|---------------|---------------|--------------|--------------|
| 1            | 670   | 345  | 51.49 | 1.43 (0.017) | 0.088 (0.004) | 0.269 (0.008) | 0.66 (0.012) | 0.27 (0.008) |
| 2            | 725   | 377  | 52.00 | 1.41 (0.016) | 0.077 (0.003) | 0.258 (0.008) | 0.68 (0.011) | 0.26 (0.008) |
| 3            | 1222  | 668  | 54.66 | 1.46 (0.012) | 0.082 (0.003) | 0.288 (0.006) | 0.70 (0.008) | 0.29 (0.005) |
| 4            | 1123  | 533  | 47.46 | 1.47 (0.013) | 0.094 (0.004) | 0.290 (0.006) | 0.67 (0.010) | 0.29 (0.006) |
| 5            | 965   | 453  | 46.94 | 1.42 (0.015) | 0.081 (0.004) | 0.264 (0.007) | 0.68 (0.011) | 0.26 (0.007) |
| 6            | 915   | 503  | 54.97 | 1.42 (0.015) | 0.075 (0.003) | 0.259 (0.007) | 0.68 (0.009) | 0.26 (0.007) |
| 7            | 1066  | 580  | 54.41 | 1.48 (0.013) | 0.079 (0.003) | 0.295 (0.006) | 0.72 (0.009) | 0.29 (0.006) |
| 8            | 823   | 459  | 55.77 | 1.38 (0.014) | 0.079 (0.004) | 0.246 (0.007) | 0.67 (0.011) | 0.25 (0.007) |
| 9            | 762   | 437  | 57.35 | 1.34 (0.014) | 0.076 (0.004) | 0.221 (0.007) | 0.66 (0.013) | 0.22 (0.007) |
| 10           | 1110  | 538  | 48.47 | 1.47 (0.014) | 0.088 (0.003) | 0.289 (0.007) | 0.69 (0.009) | 0.29 (0.007) |
| 11           | 1110  | 510  | 45.95 | 1.42 (0.014) | 0.096 (0.004) | 0.265 (0.007) | 0.64 (0.011) | 0.26 (0.007) |
| UN           | 1449  | 461  | 31.82 | 1.37 (0.015) | 0.086 (0.005) | 0.234 (0.008) | 0.59 (0.016) | 0.32 (0.008) |
| Overall mean | 11940 | 5864 | 49.11 | 1.43 (0.004) | 0.084 (0.001) | 0.267 (0.002) | 0.67 (0.003) | 0.27 (0.002) |

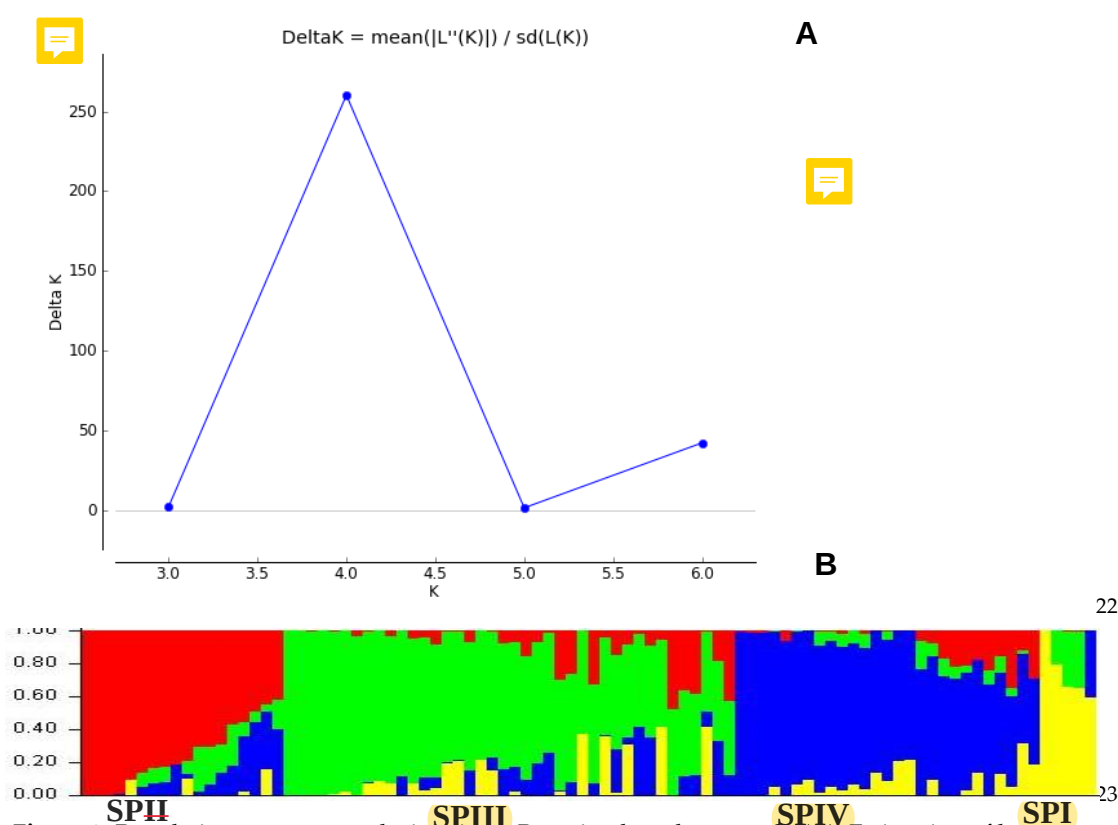
Note: %P = Polymorphism, NPS = Number of polymorphic SNPs, NSU = number of SNPs used,  $N_e$  = number of effective alleles per locus,  $H_o$  = observed heterozygosity,  $uH_e$  = unbiased gene diversity,  $F_{IS}$  = inbreeding coefficient, PIC = polymorphic information content, UN- unknown, the values within the brackets are standard error.

### *Population structure and clustering*

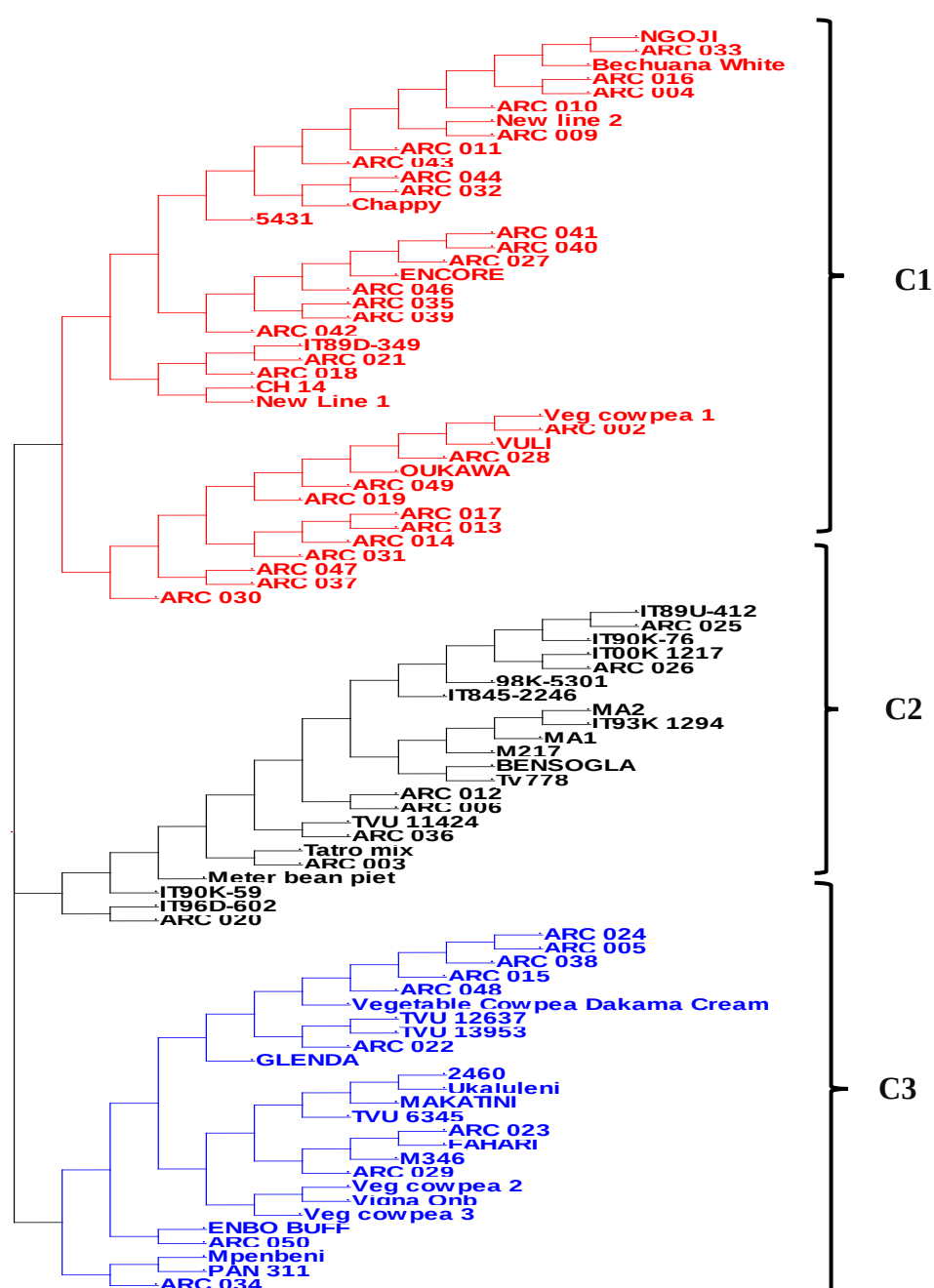
The population structure of the 90 accessions was examined using a model- and distance-based structure analyses. The STRUCTURE analysis estimated the most suitable number of subpopulations was at  $K = 4$  (Fig. 2A), indicating that the 90 accessions could be grouped into four subpopulations (SP1 – SP4) based on differences in their genetic makeup (Figure 2B). SP1 (yellow) contained only five accessions with admixtures from SP2 and SP3. SP2 (red) contained 18 accessions that share more admixture membership

with the SP3 and SP4. P3 (green) contained 40 accessions, which share admixture membership with the other three subpopulations. SP4 (blue) contained 27 accessions with admixture from the other subpopulations. The admixture level in the four subpopulations ranged from 50% to 63%, which clearly indicated that these subpopulations shared more admixture memberships.

However, the distance-based cluster analysis generated using Nei's genetic distance using a neighbor-joining algorithm revealed the presence of three distinct clusters in the population represented by the 90 accessions (Figure 3). The clustering patterns of the two approaches were similar and constituted similar sets of accessions. For example, C1 (red) contained 41 accessions, 5 accessions from SP1, 27 from SP4 and 9 from SP3 of the STRUCTURE generated clusters. C2 (black) consisted of 24 accessions, of which 18 accessions from SP2 and 5 were from SP3. C3 (blue) contained 26 accessions all from SP3. The discrepancy between the two clustering approaches could result from admixtures since only 43% of the tested genotypes were considered pure.



**Figure 2.** Population structure analysis using a Bayesian-based approach. (A). Estimation of hypothetical subpopulations using K-values showing the highest Delta k value was observed at the number of populations (K) = 4. (B). Population structure analysis of 90 cowpea accession at K = 4 based on inferred ancestry (Q matrix).



**Figure 3.** Neighbour joining analysis showing the genetic relationships among 90 cowpea accession tested using 5864 SNP markers. The different colours indicate different clusters.

#### Genetic diversity among subpopulation

The population genetic diversity estimates on 90 accessions were analyzed based on the four subpopulations generated by STRUCTURE and three populations generated by DARwin (Table 3). SP3 revealed the highest values for most of the genetic parameters and displayed the highest level of genetic diversity ( $H_e = 0.247$  and  $I = 0.381$ ). SP1 displayed the lowest level of genetic diversity ( $H_e = 0.162$  and  $I = 0.189$ ) but 68.4% of the loci were fixed (Table 2). SP3 had the highest number of private alleles (508) and the highest percentage of polymorphic loci (91%). Based on the three subpopulations generated by cluster analysis, C2 revealed the highest genetic diversity for most of the genetic parameters except for fixation index. C1, on the other hand, revealed the lowest for all the studied genetic parameters. In C1 68% of the alleles were fixed and 87% of loci were polymorphic.

**Table 2.** Genetic diversity within and among the 90-cowpea accession classified by growth habit.

| Pop.                                         | Na | Ne                   | I                    | Ho                   | He                   | F <sub>IS</sub>      | %P           | P <sub>A</sub> |
|----------------------------------------------|----|----------------------|----------------------|----------------------|----------------------|----------------------|--------------|----------------|
| Model based population structure analysis    |    |                      |                      |                      |                      |                      |              |                |
| SP1                                          | 5  | 1.189 (0.006)        | 0.189 (0.004)        | 0.035 (0.002)        | 0.162 (0.003)        | <b>0.684</b> (0.007) | 30.68        | 3              |
| SP2                                          | 18 | 1.368 (0.004)        | 0.352 (0.003)        | <b>0.116</b> (0.002) | 0.233 (0.002)        | 0.398 (0.005)        | 80.64        | 229            |
| SP3                                          | 40 | <b>1.397</b> (0.004) | <b>0.381</b> (0.003) | 0.083 (0.001)        | <b>0.247</b> (0.002) | 0.617 (0.004)        | <b>90.96</b> | <b>508</b>     |
| SP4                                          | 27 | 1.258 (0.004)        | 0.257 (0.003)        | 0.068 (0.001)        | 0.165 (0.002)        | 0.437 (0.005)        | 68.66        | 42             |
| Overall                                      | 90 | 1.303 (0.002)        | 0.295 (0.002)        | 0.075 (0.001)        | 0.202 (0.001)        | 0.514 (0.003)        | 67.74        | -              |
| Distance-based population structure analysis |    |                      |                      |                      |                      |                      |              |                |
| C1                                           | 41 | 1.342 (0.004)        | 0.340 (0.003)        | 0.059 (0.001)        | 0.218 (0.002)        | 0.682 (0.004)        | 86.77        | 65             |
| C2                                           | 24 | <b>1.404</b> (0.004) | <b>0.387</b> (0.003) | <b>0.108</b> (0.001) | <b>0.254</b> (0.002) | 0.512 (0.005)        | <b>88.93</b> | <b>234</b>     |
| C3                                           | 25 | 1.385 (0.004)        | 0.366 (0.003)        | 0.097 (0.001)        | 0.240 (0.002)        | 0.526 (0.005)        | 85.44        | 177            |
| Overall                                      | 90 | 1.377 (0.003)        | 0.364 (0.002)        | 0.088 (0.001)        | 0.237 (0.001)        | 0.573 (0.003)        | 87.05        | -              |

Note: Ne = average number of effective alleles per locus per subpopulation, I = Shannon information index, Ho = observed heterozygosity per subpopulation, He = expected heterozygosity per subpopulation, FIS = inbreeding coefficient, PA = Private alleles, %P = percentage of polymorphic loci, the values within the brackets are standard error.

#### Analysis of Molecular Variance (AMOVA)

AMOVA was performed in both model-based and distance-based populations (Table 3). The results of AMOVA in model-based populations indicated that the majority of the variance occurred among individuals within populations and accounted for 59.4% of the total variation. However, 18.4% and 22.1% of the total variation was attributed to differences within individuals and among populations, respectively. Similarly, in the distance-based population, the majority of the variance was observed among individuals within the population and accounted for 68.9% of the total genetic variance. The mean fixation index within individuals was significantly high and positive in all classes of populations suggesting that outcrossing among the tested cowpea populations was low. The variation existed among populations was positive and significant suggesting that these populations were highly differentiated. Similarly, the relatively low level of variation observed within individuals was attributed to the high fixation index value.

**Table 3.** Analysis of molecular variance (AMOVA) among 90-cowpea accessions classified based on SNP markers.

| Source                            | DF  | SS       | MS      | Est. Var. | Per. Var | F-statistics                        |
|-----------------------------------|-----|----------|---------|-----------|----------|-------------------------------------|
| Model-based structure analysis    |     |          |         |           |          |                                     |
| Among Population                  | 3   | 37575.3  | 12525.1 | 270.2     | 22.14    | F <sub>ST</sub> = 0.221 (P < 0.001) |
| Among Individual                  | 86  | 144053.5 | 1675.0  | 725.2     | 59.44    | F <sub>IS</sub> = 0.763 (P < 0.001) |
| Within Individual                 | 90  | 20217.5  | 224.6   | 224.6     | 18.42    | F <sub>IT</sub> = 0.816 (P < 0.001) |
| Total                             | 179 | 201846.2 | -       | 1220.0    | 100.00   | -                                   |
| Distance based structure analysis |     |          |         |           |          |                                     |
| Among Population                  | 2   | 20380.1  | 10190.1 | 143.8     | 12.16    | F <sub>ST</sub> = 0.122 (P < 0.001) |
| Among Individual                  | 87  | 161248.8 | 1853.4  | 814.4     | 68.85    | F <sub>IS</sub> = 0.784 (P < 0.001) |
| Within Individual                 | 90  | 20217.5  | 224.64  | 224.6     | 18.99    | F <sub>IT</sub> = 0.810 (P < 0.001) |
| Total                             | 179 | 201846.4 | -       | 1182.8    | 100.0    | -                                   |

Note: DF =degrees of freedom, SS=sum of squares, MS= mean sum of squares, Est. Var= estimated variance, Per. Var = Percentage variation.

Genetic differentiation (F<sub>ST</sub>) estimates among the subpopulations ranged from 0.103 between SP3 and SP4 to 0.239 between SP1 and SP2. The gene flow ranged from 0.8 (between SP1 and SP2) to 2.2 (between SP3 and SP4). The genetic distance among populations ranged from 0.089 between SP3 and SP4 to 0.214 between SP1 and SP2. The genetic identity (GI) ranged from 0.81 between SP1 and SP2 to 0.92 between SP3 and SP4 (Table 4).

According to Wright (1984) standard guidelines for the interpretation of genetic differentiation, all pairs of subpopulations showed a moderate level of population differentiation. However, SP1 showed a relatively higher degree of differentiation (0.239) from the rest of the subpopulations. Gene flow among the subpopulation was relatively high between SP2, SP3, and SP4 (Table 4). The observed high genetic differentiation among the subpopulations could be explained by the low gene flow among subpopulations.

**Table 4.** Pair-wise estimates of gene flow (above diagonal, within the brackets), genetic differentiation ( $F_{ST}$ ) (above diagonal off brackets); genetic distance (GD) (lower diagonal off brackets) and genetic identity (GI) (lower diagonal within the brackets).

|     | SP1            | SP2           | SP3           | SP4           |
|-----|----------------|---------------|---------------|---------------|
| SP1 | -              | 0.239 (0.796) | 0.142 (1.511) | 0.192 (1.052) |
| SP2 | 0.214 ( 0.807) | -             | 0.106 (2.108) | 0.149 (1.428) |
| SP3 | 0.112 (0.894)  | 0.103 (0.902) | -             | 0.103 (2.172) |
| SP4 | 0.134 (0.874)  | 0.135 (0.874) | 0.089 (0.915) | -             |

## Discussion

The analysis of genetic diversity in crops is a prerequisite for the success of any plant breeding programs [39]. Therefore, assessing the population structure and genetic diversity of crops is fundamental to implement efficient genetic resources management and conservation strategies. Application of high-throughput molecular markers provides a better understanding of genomic diversity and population structure of germplasm and can speed up the identification of superior groups for further hybrid development [40]. The current study used 5864 SNP markers to assess the pattern and level of genetic variation and genetic structure among 90 cowpea accessions collected from various geographic origins.

The quality and the discriminatory power of a given marker system is assessed by its PIC values [41]. It is importance to note that SNP markers are bi-allelic in nature, hence their PIC values are restricted to 0.5, which is considered to be low or moderate informative as compared to SSR markers [42]. The mean PIC value of 0.27 reported in this study was in agreement with Gbedevi [3] who reported PIC value of 0.25 but relatively higher than the one reported by Sodedji [43] (PIC = 0.22). The results suggests that the SNP markers that were used in this study showed a moderate level of polymorphism and revealed the existence of genetic diversity among the tested genotypes. The number of SNP markers used and the number of accessions studied might explain the observed differences among the reported PIC values and allelic polymorphism in this study and other previously reported studies. Kondwakwenda [8] also indicated that the observed variation in the quality and performance of SNP markers in different studies depends on the number of accessions studied, type of markers used and the type of germplasm studied. Nonetheless, the SNP markers used in this study were relatively informative and reliable in assessing the diversity of cowpea accessions studied.

The mean number of effective alleles per locus reported in this study was 1.43, which was comparative to the 1.41 reported by Fatokun [44]. The gene diversity was further expressed using the probability of observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity, which were the true indicator for the degree of genetic variation within and among the population assessed. The average  $H_o$  and  $H_e$  in the present study were 0.075 and 0.202, respectively for the 90 accessions which were was comparable to the  $H_o$  of 0.05 and  $H_e$  of 0.31 reported by Gbedevi [3] for 70 cowpea accessions. Xiong [45], on the other hand, assessed 768 worldwide cowpea germplasm collections maintained at USDA GRIN and reported the  $H_o$  value of 0.06 and the  $H_e$  of 0.35. Nonetheless, the  $H_e$  was moderately low in this study but generally higher than the  $H_o$  for all subpopulations. Fatokun [44] also reported similar trend in cowpea. Govindaraj [46] alluded that the low observed heterozygosity suggests the high level of inbreeding within the subpopulations.

A moderate fixation index ( $F_{IS} = 67\%$ ) was observed in this study indicating 67% of the SNP loci used were fixed in the studied accessions. However, relatively high ( $F_{IS} = 0.83$ ) was reported by Gbedevi [3]. The low observed heterozygosity and the relatively high rate of fixation index exhibited by the populations was explained by the fact that cowpea is a self-pollinated crop possessing a low out-crossing rate and low within-accession variability. Although the outcrossing rate in cowpea is low that ranges from less than 0.15 up to 1.58% depending on the genotypes involved and environment, where its grown [47], further purification (self-pollination) of the accession is needed. The self-pollination nature and low outcrossing rate of cowpea has been reported to be the major contributor for the observed low genetic variation among cowpea accessions [6,48].

Population structure analysis is the key to assess the genetic structure of a given population and the basis for complex marker-trait association analysis [49]. The model based population structure analysis grouped the 90 accessions into four subpopulations based on the peak of delta K ( $\Delta K$ ) at  $K = 4$ . The admixture level ranges from 50% in SP4 to 60% SP1. The high proportion of admixture detected indicates either these subpopulation share the same ancestral progenitor or there was gene flow between the subpopulations. The distance-based cluster analysis using neighbor-joining method showed the presence of three distinct clusters, which was not consistent with the results of structure analysis. However, the pattern and the number of accessions maintained in each clustering approached were similar. The differences observed between the model based on STRUCTURE analysis and distance-based cluster analysis in the size and number of subgroups can be explained by the presence of admixtures within the subpopulations. In both model-based and distance-based clustering approaches, the grouping patterns were inconsistent with the growth habit and geographic origins of the studied accessions. Similar phenomena whereby subpopulations were grouped irrespective of the grouping criteria has also been reported by other researcher's studying the genetic diversity in cowpea accessions [3,43,50]. In contrast, Ravelombola [51] reported clustering of genotypes based on growth habit resulted in two subpopulations that were highly differentiated.

The seed size and seed coat color preference can highly influence the genetic diversity in cowpeas [5]. Classification based on growth habit and other agronomic traits such as seed shape and seed coat color are significantly influenced by the breeding programs because these traits have been used to classify genotypes. The qualitative traits such as growth habit, seed size, seed shape and seed coat color are also important traits for farmers and consumers preferences [52]. Therefore, it is important to incorporate farmers- and consumers-preferred traits in future selections to enhance varietal adoption among farmers. In the UPGM clustering, C1 (red) was the highest group comprising of 41 (46%) genotypes with brown (22%), red (12%) and cream (24%) seed coat color. Similarly, 23 accessions were grouped in C2 (black) comprises of 35% of brown seed coat color and 26 accessions grouped in C3 (blue) comprised of 35% black seed coat and 19% brown seed coat. In addition, the majority of accessions clustered in C1 and C2 were kidney shaped while C3 was dominated with Rhomboid shaped accessions. In terms of phenotypical variation, the erect and prostrate type were not different with respect to seed shape. The studies by Hellens [53] in peas has reported the lighter seed coat color as human preference during domestication.

Similarly, C1 was dominated by erect (32%), while C2 was dominated by accessions with unknown growth habits (44%). C3 was mainly dominated by the prostate (54%) growth habit type. Regarding the wide distribution of cowpea, accession was studied based on growth habit, the prostrate and erect types were more dominant than the semi-erect and climbing types. Growth habit is morphological important qualitative traits in cowpea production that highly affects crop yield and tillage method and further defines the shape of the plants and dictate how the plant should be harvested [54]. Plant growth habit has been a major breeding target for crop improvement. Therefore, determining the genetic mechanisms that controls the plant type will assist in cowpea growth development improvements. However, the results revealed that growth habit could not be used

as an indices for evaluating genetic diversity and for genotypes classifications. The study by Khan [55] on Bambara groundnut has also reported the same findings.

The AMOVA results revealed that the majority of the total molecular variation (59% and 69%) were due to differences among individuals within a population, 22% and 12% of the variation were attributed to the difference among population and 18% and 19% was due to variation within individuals. In the model-based approach the among population variation was higher than the within individuals variation. The magnitude of variations among and within populations were further quantified by genetic differentiation observed among the populations ( $F_{ST} = 0.221$ ). The results indicate high genetic differentiation between four subpopulations based on the standard guideline of Wright [56]. The studies by Gbedevi [3] also reported a high genetic differentiation ( $F_{ST}$ ) value of 0.423 between two major reported subpopulations of 498 cowpea accessions assessed. However, Sarr [50] also reported a low genetic differentiation ranging from 0.018 to 0.100. The differences reported in genetic differentiations could be attributed to the diversity and number of accessions used and the number of markers involved to assess the genetic diversity.

The reported level of genetic differentiation reported in this study could be explained by the gene flow among subpopulations [57]. The gene flow among the studied populations ranged from 0.796 to 2.172 and according to Wright [58], gene flow  $< 1$  indicates limited gene exchange among population. This result suggested that a moderate gene flow occurred in this study and led to high genetic differentiation between the populations. Furthermore, genetic distance was used to measure the relatedness between individuals in a population. In this study, the low pairwise genetic distance were observed ranging from 0.089 to 0.214 revealing a wide genetic variations among the tested cowpea accessions. This result suggests that the accessions studied are unique and have greater potential to contribute to new varieties to the breeding programmes in South Africa. The understanding of genetic diversity among cowpea populations studied in this study will improve the subsequent planning in future cowpea breeding and contribute useful information in conservation and managing genetic diversity required for vigorous and successful breeding program

#### 4. Materials and Methods

##### *Plant materials*

The study evaluated the genetic diversity of 90 cowpea accessions sourced from Agricultural Research Council - Vegetables, Industrial and Medicinal Plants (ARC-VIMP) gene bank, Pretoria, South Africa. These accessions were collected from different parts of the world including South Africa, Kenya and Nigeria. The geographic origin and growth habits of the accessions is presented in Supplementary Table 1.

##### *DNA extraction and sequencing*

The cowpea genotypes were grown in a seed germination chamber at the Biosciences eastern and central Africa -International Livestock Research Institute in Nairobi (BecA-ILRI) for genotyping. Ten-day-old leaf materials were sampled from the seedlings and the leaf samples were frozen in a liquid nitrogen and stored at  $-80^{\circ}\text{C}$  for genotyping. DNA extraction was done using NucleoMag Plant DNA extraction kit (Takara Bio USA, Inc). The genomic DNA extracted was in the range of 50-100ng/ul. DNA quality and quantity were checked on 0.8% agarose gel.

Libraries were constructed according to Killian [30] using DArTSeq complexity reduction method through digestion of genomic DNA using a combination of PstI and MseI enzymes and ligation of barcoded adapters and common adapter followed by PCR amplification of adapter-ligated fragments. Libraries were sequenced using Single Read sequencing runs for 77 bases. Next generation sequencing was carried out using Hiseq2500. DArTseq markers scoring was achieved using DArTsoft14 which is an in house marker

scoring pipeline based on algorithms. Two types of DArTseq markers were scored, SilicoDArT markers and SNP markers which were both scored as binary for presence /absence of the restriction fragment with the marker sequence in genomic representation of the sample. Both SilicoDArT markers and SNP markers were aligned to cowpea reference genome, Vunguiculata\_469\_v1.0 to identify chromosome positions.

#### Data analysis

Ninety-cowpea **accession** were genotyped with 11940 SNP markers. Monomorphic and SNPs with minor allele frequency of less than 2% were filtered out and 5864 (49%) SNPs were retained for further analysis. Genotypic data were subjected to analyses of molecular variance (AMOVA) and various measures of genetic diversity within and among inferred subpopulations using GenAlex software version 6.5 [31]. Genetic diversity parameters such as number of effective alleles per locus ( $N_e$ ), Shannon's Information Index ( $I$ ), gene diversity ( $H_e$ ) and polymorphic information content (PIC) were determined using the protocol of Nei and Li [32]. The genotypic data were used to obtain a dissimilarity matrix using the Jaccard index as described by Debener [33]. The matrix was then used to run a cluster analysis based on neighbor-joining algorithm using the unweighted pair group method with arithmetic average (UPGMA) in DARwin 6.0 software [34]. Bootstrap analysis was performed for node construction using 10,000 bootstrap values.

The Bayesian genotypic clustering approach of STRUCTURE 2.3.4 [35] was used to determine the population structure. An admixture model with independent allele frequencies, without prior population information was used to simulate the population. The STRUCTURE program was set as follows: a burn-in period length of 100,000, and after burn-in, 100,000 Markov Chain Monte Carlo (MCMC) were used. This model assumes that the genome of each individual is a mixture of genes originating from  $K$  unknown ancestral populations. For joint inference of the population substructure,  $K$  ranging from 1 - 6 was set up, with ten independent runs for each  $K$ . The most probable value of  $K$  for each test was detected by  $\Delta K$  [36] using the STRUCTURE HARVESTER [37]. Each individual genotype was grouped into a given cluster using 'membership coefficient' for each cluster interpreted as a probability of membership. The genotype membership was determined by the computer program CLUMPP [38].

#### 5. Conclusions

The current study revealed the existence of genetic diversity among and within the cowpea accessions ~~assessed~~ and showed the effectiveness and reliability of SNP markers. The study revealed that 49% of the selected SNP markers were highly polymorphic and efficiently discriminate the tested cowpea accessions. The low heterozygosity and the high inbreeding coefficients observed among cowpea varieties indicates that the accessions reached an acceptable level of homozygosity. The model based (structure analysis) and distance-based (UPGM) clustering approaches were used in this study. The model-based analysis revealed the presence of four subpopulations at  $K = 4$  whereas the distance-based cluster analysis classified the cowpea accessions into three distinct clusters. The subpopulations identified exhibited high level of genetic diversity and moderately differentiated. These subpopulations could serve as heterotic groups and a relevant source of genes for future breeding and selection of diverse cultivars with different traits. Therefore, the results obtained from this study will be highly valuable in development of new varieties adapted to diverse environments. Consequently, this study will substantially contribute to utilization, conservation and broadening of the use of genetic resources of cowpeas for future improvement.

**Supplementary Materials:** The following supporting information can be downloaded at: [www.mdpi.com/xxx/s1](http://www.mdpi.com/xxx/s1), Supplementary Table 1: title; Video S1: title.

**Author Contributions:** Conceptualization: Abe Shegro Gerrano, Mbali Gumede, Albert Thembinkosi Modi; Formal analysis: Mbali Gumede, Amelework Assefa; Data validation: Mbali Gumede, Amelework Assefa; Abe Shegro Gerrano, Albert Thembinkosi Modi; Funding acquisition: Abe Shegro Gerrano; Investigation: Mbali Gumede, Abe Shegro Gerrano, Albert Thembinkosi Modi; Methodology: Abe Gerrano; Project administration: Abe Shegro Gerrano; Resources: Abe Shegro Gerrano; Supervision: Abe Shegro Gerrano, Albert Thembinkosi Modi; Writing – original draft, Mbali Gumede; Writing – review & editing, Mbali Gumede, Abe Shegro Gerrano, Amelework Assefa, Albert Thembinkosi Modi.

**Funding:** This research was funded by the National Research Foundation, South Africa, unique grant no: 116344 and Agricultural Research Council of South Africa.

**Data Availability Statement:** Not available

**Acknowledgments:** The University of KwaZulu-Natal for the support of the project and the Agricultural Research Council for providing plant material, research fund and research support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Badiane, F.A., Diouf, M., Diouf, D. Cowpea. In Broadening the genetic base of grain legumes. Singh M., Bisht I.S., Dutta M.; Springer. **2014**. 79, pp. 95-114.
- Coulibaly, S., Pasquet, R.S., Papa, R., Gepts, P. AFLP analysis of the phenetic organization and genetic diversity of *Vigna unguiculata* L. Walp. reveals extensive gene flow between wild and domesticated types. *Theor. Appl. Genet.* **2002**, *104* (2), 358-366.
- Gbedevi, K.M., Boukar, O., Ishikawa, H., Abe, A., Ongom, P.O., Unachukwu, N., Rabbi, I., Fatokun, C. Genetic diversity and population structure of cowpea [*Vigna unguiculata* (L.) Walp.] germplasm collected from Togo based on DArT markers. *Gen.* **2021**, *12* (9), 1451. DOI.org/10.3390/genes12091451.
- Ehlers, J.D., Hall, A.E. Cowpea (*Vigna unguiculata* L. Walp.). *Field crops res.* **1997**, *53* (3), 187-204.
- Seo, E., Kim, K., Jun, T., Choi, J., Kim, S., Muñoz-Amatriaín, M., Sun, H., Ha, B. Population structure and genetic diversity in Korean cowpea germplasm based on SNP markers. *Plants.* **2020**, *9* (9), 1190.
- Carvalho, M., Muñoz-Amatriaín, M., Castro, I., Lino-Neto, T., Matos, M., Egea-Cortines, M., Rosa, E., Close, T., Carnide, V. Genetic diversity and structure of Iberian Peninsula cowpeas compared to world-wide cowpea accessions using high density SNP markers. *BMC Genom.* **2017**, *18* (1), 1-9.
- Gbaye, O.A., Holloway, G.J. Varietal effects of cowpea, *Vigna unguiculata*, on tolerance to malathion in *Callosobruchus maculatus* (Coleoptera: Bruchidae). *J. Stored Prod. Res.* **2011**, *47*, 365-371.
- Kondwakwenda, A., Sibiya, J., Amelework, A.B., Zengeni, R. Diversity analysis of provitamin A maize inbred lines using single nucleotide polymorphism markers. *Acta Agric Scand B Soil Plant Sci.* **2020**, *70* (4), 265-271.
- Bhandari, H.R., Bhanu, A.N., Srivastava, K., Singh, M.N., Shreya, H.A. Assessment of genetic diversity in crop plants-an overview. *Adv. Plants Agric. Res.* **2017**, *7*(3): 279-286.
- Campos, A., Tedesco, S., Vasconcelos, V., Cristobal, S. Proteomic research in bivalves: Towards the identification of molecular markers of aquatic pollution. *J. Proteom.* **2012**, *75* (14), 4346-4359.
- Horn, L.N., Shimelis, H. Production constraints and breeding approaches for cowpea improvement for drought prone agro-ecologies in Sub-Saharan Africa. *Ann. Agric. Sci.* **2020**, *66*, 93-91.
- De Vicente, M.C., Lopez, C., Fulton, T. Genetic diversity analysis with molecular marker data: learning module. International Plant Genetic Resources Institute (IPGRI). **2004**. DOI: 10.2135/cropsci2005.0015br.
- Bucheyeki, T.L., Gwanama, C., Mgonja, M., Chisi, M., Folkertsma, R., Mutegi, R. Genetic variability characterisation of Tanzania sorghum landraces based on simple sequence repeats (SSRs) molecular and morphological markers. *African Crop Sci. J.* **2009**, *17* (2), 71-86. DOI: 10.4314/acsj.v17i2.54201.
- Sami, R.A., Yeye, M.Y., Ishiyaku, M.F., Usman, I.S. Heritability studies in some sweet sorghum (*Sorghum Bicolor*. L. Moench) genotypes. *J. Biol. Agric. Health.* **2013**, *3* (17), 49-51.
- Sejake, T., Shargie, N., Christian, R., Amelework, A.B., Tsilo, T.J. Genetic diversity in sorghum ('*Sorghum bicolor*'L. Moench) accessions using SNP based Kompetitive allele-specific (KASP) markers. *Aust. J. Crop. Sci.* **2021**, *15* (6), 890-898.
- Cui, C., Mei, H., Liu, Y., Zhang, H., Zheng, Y. Genetic diversity, population structure, and linkage disequilibrium of an association-mapping panel revealed by genome-wide SNP markers in sesame. *Front in plant sci.* **2017**, *8*, 1189.
- Malviya, N., Sarangi, B.K., Yadav, M.K., Yadav, D. Analysis of genetic diversity in cowpea (*Vigna unguiculata* L. Walp.) cultivars with random amplified polymorphic DNA markers. *Plant Syst. Evol.* **2012**, *298* (2), 523-526.
- Mafakheri, K., Bihanta, M.R., Abbasi, A.R. Assessment of genetic diversity in cowpea (*Vigna unguiculata* L.) germplasm using morphological and molecular characterisation. *Cogent Food Agric.* **2017**, *3* (1), 1327092.
- Nkhoma, N., Shimelis, H., Laing, M., Shayanowako, A., Mathew, I. Assessing the genetic diversity of cowpea [*Vigna unguiculata* (L.) Walp.] germplasm collections using phenotypic traits and SNP markers. *BMC Genom.* **2020**, *21*(1): 1-16.

20. Fang, J., Chao, C., Roberts, P.A., Ehlers, J.D. Genetic diversity of cowpea [*Vigna unguiculata* (L.) Walp.] in four West African and USA breeding programs as determined by AFLP analysis. *Genet. Resour. Crop Evol.* **2007**, *54* (6), 1197-1209.
21. Rafalski, A. Applications of single nucleotide polymorphisms in crop genetics. *Plant Biol.* **2002**, *5* (2), 94-100.
22. Kujur, A., Bajaj, D., Upadhyaya, H.D., Das, S., Ranjan, R., Shree, T., Saxena, M.S., Badoni, S., Kumar, V., Tripathi, S. Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Front plant sci.* **2015**, *6*:162.
23. Desalegne, B.A., Dagne, K., Melaku, G., Boukar, O., Fatokun, C.A. Efficiency of SNP and SSR-based analysis of genetic diversity, population structure, and relationships among cowpea (*Vigna unguiculata* (L.) Walp.) germplasm from East Africa and IITA inbred lines. *J. Crop Sci. Biotech.* **2017**, *20* (2), 107-128.
24. Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E. A robust, simple sequence genotyping-by-sequence (GBS) approach for high diversity species. *PLoS one.* **2011**, *6*, e19379.
25. Bastien, M., Sonah, H., Belzile, F. Genome wide association mapping of resistance in soybean with a genotyping-by-sequencing approach. *Plant Genom.* **2014**, *7*, 1-13.
26. Ketema, S., Tesfaye, B., Keneni, G., Fenta, B.A., Assefa, E., Greliche, N., Machuka, E., Yao, N. DArTSeq SNP-based markers revealed high genetic diversity and structured population in Ethiopian cowpea [*Vigna unguiculata* (L.) Walp] germplasms. *PLoS one.* **2020**, *15* (10), e0239122.
27. Farahani, S., Maleki, M., Mehrabi, R., Kanouni, H., Scheben, A., Batley, J., Talebi, R. Whole genome diversity, population structure, and linkage disequilibrium analysis of chickpea (*Cicer arietinum* L.) genotypes using genome-wide DArTseq-based SNP markers. *Genes* **2019**, *10* (9), 676.
28. Valdisser, P.A., Pereira, W.J., Filho, J.E., Müller, B.S., Coelho, G.R., de Menezes, I.P., Vianna, J.P., Zucchi, M.I., Lanna, A.C., Coelho, A.S. In-depth genome characterization of a Brazilian common bean core collection using DArTseq high-density SNP genotyping. *BMC Genom.* **2017**, *18* (1), 1-19.
29. Uba, C.U., Oselebe, H.O., Tesfay, A.A., Abtey, W.G. Genetic diversity and population structure analysis of Bambara groundnut (*Vigna subteranea* L) landraces using DArT SNP markers. *PLoS one.* **2021**, *16*(7), e0253600.
30. Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., Heller-Uszynska, K., Jaccoud, D., Hopper C. Diversity arrays technology: a generic genome profiling technology on open platforms, in: Data production and analysis in population genomics. *Methods Mol. Biol.* **2012**, *888*, 67-89. DOI: 10.1007/978-1-61779-870-2\_5.
31. Peakall, R., Smouse, P.E. Genalex 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. *Bioinformatics* **2012**, *28*, 2537-2539.
32. Nei, M., Li, W. Mathematical model for studying genetic variation in terms of restriction endonucleases. *PNAS.* **1979**, *76* (10), 5269-5273.
33. Debener, T., Salamini, F., Gebhardt, C. Phylogeny of wild and cultivated *Solanum* species based on nuclear restriction fragment length polymorphisms (RFLPs). *Theor. Appl. Genet.* **1990**, *79* (3): 360-368.
34. Perrier, X., Jacquemoud-Collet, J.P. DARwin software. Dissimilarity analysis and representation for windows. <http://www.darwin.cirad.fr/darwin.html>. **2006**.
35. Pritchard, J.K., Stephens, M., Rosenberg, N.A., Donnelly P. Association mapping in structured populations. *American J. Human Gen.* **2000**, *67* (1), 170-181.
36. Evanno, G., Regnaut, S., Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **2005**, *14* (8), 2611-2620.
37. Earl, D.A., VonHoldt, B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Cons. Gene Res.* **2012**, *4* (2), 359-361.
38. Jakobsson, M., Rosenberg, N. ACLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23* (14), 1801-1806.
39. Ebert, A.W., Engels, J.M.M. Plant biodiversity and genetic resources matter. *Plants.* **2020**, *9*(12), 1706.
40. Serba, D.D., Muleta, K.T., Amand, P., Bernado, A., Bai, G., Perumal, R., Bashir, E. Genetic diversity, population structure and linkage disequilibrium of pearl millet. *Plant Gen.* **2019**, *12*, 180091.
41. Kumar, D., Chhokar, V., Sheoran, S., Singh, R., Sharma, P., Jaiswal, S., Iqbal, M.A., Jaiswar, A., Jaisri, J., Angadi, U.B., Rai, A., Singh, G.P., Kumar, D., Tiwari, R. Characterization of genetic diversity and population structure in wheat using array based SNP markers. *Mol Biol.* **2020**, *47*, 293-306.
42. Eltaher, S., Sallam, A., Belamkar, V. Genetic diversity and population structure of F3:6 Nebraska winter wheat genotypes using genotyping-by-sequence. *Front Genet.* **2018**, *9*, 11-9.
43. Sodedji, F.A., Agbahoungba, S., Agoyi, E.E., Kafoutchoni, K.M., Kim, H., Nguetta, S.A., Assogbadjo, A.E. Diversity, population structure, and linkage disequilibrium among cowpea accessions. *Plant Genom.* **2021**, *14*(3), e20113.
44. Fatokun, C., Girma, G., Abberton, M., Gedil, M., Unachukwu, N., Oyatomi, O., Yusuf, M., Rabbi, I., Boukar, O. Genetic diversity and population structure of a mini-core subset from the world cowpea (*Vigna unguiculata* (L.) Walp.) germplasm collection. *Sci Rep.* **2018**, *8*, 1-10.
45. Xiong, H., Shi, A., Mou, B., Qin, J., Motes, D., Lu, W., Ma, J., Weng, Y., Yang, W., Wu, D. Genetic diversity and population structure of cowpea (*Vigna unguiculata* L. Walp). *PLoS One.* **2016**, *11* (8), e0160941.
46. Govindaraj, M., Vetriventha, M., Srinivasan, M. Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genet Res Int.* **2015**, *15*, 1-14.

- 
47. Fatokun, C.A., Ng, Q. Outcrossing in cowpea. *J. Food, Agric. Environ.* **2007**, *5*, 334–338. 385
48. Wamalwa, E.N., Muoma, J., Wekesa, C. Genetic diversity of cowpea (*Vigna unguiculata* (L.) Walp.) accession in Kenya gene bank based on simple sequence repeat markers. *Int. J. Genom.* **2016**, 8956412. 386
49. Wu, X., Li, Y., Li, X., Li, C., Shi, Y., Song, Y., Zheng, Z., Li, Y., Wang, T. Analysis of genetic differentiation and genomic variation to reveal potential regions of importance during maize improvement. *BMC Plant Biol.* **2015**, *15* (1), 1–13. 387
50. Sarr, A., Bodian, A., Gbedevi, K.M., Ndir, K.N., Ajewole, O.O., Gueye, B., Foncéka, D., Diop, E.A., Diop, B.M., Cissé, N. Genetic diversity and population structure analyses of wild relatives and cultivated cowpea (*Vigna unguiculata* (L.) Walp.) from Senegal using simple sequence repeat markers. *Plant Mol. Biol. Rep.* **2020**, *39*, 112–124. 388
51. Ravelombola, W., Qin, J., Shi, A., Weng, Y., Bhattarai, G., Dong, L., Morris, J.B. A SNP-based association analysis for plant growth habit in worldwide cowpea (*Vigna unguiculata* (L.) Walp) germplasm. *Plant Genet. Resour.* **2017**, *213*, 284. 389
52. Nkhata, W., Shimelis, H., Melis, R., Chirwa, R., Mzengeza, T., Mathew, I., Shayanowako, A. Population structure and genetic diversity analyses of common beans germplasm collections of East and Southern Africa using morphological traits and high-density SNP markers. *Plos one.* **2010**, *12*, e0243238. 390
53. Hellens, R.P., Moreau, C., Lin-Wang, K., Schwinn, K.E., Thomson, S.J., Fiers, M., Frew, T.J., Murray, S.R., Hofer, J.M., Jacobs, J.M. Identification of Mendel's white flower character. *Plos One.* **2010**, *5*, e13230. 391
54. Li, L., Cui, S., Dang, P. GWAS and bulked segregant analysis reveal the Loci controlling growth habit-related traits in cultivated Peanut (*Arachis hypogaea* L.). *BMC Genom.* **2022**, *23*, 403. 392
55. Khan, M., Rafii, M., Ramlee, S., Jusoh, M., Mamun, M. Genetic analysis and selection of Bambara groundnut (*Vigna subterranea* [L.] Verdc.) landraces for high yield revealed by qualitative and quantitative traits. *Sci. Rep.* **2021**, *11* (1), 1–21. 393
56. Wright, S. *Evolution and the genetics of populations: variability within and among natural populations* (Vol 4). Chicago (IL).University of Chicago press. **1984**. 394
57. Gepts, P., Papa, R. Possible effects of (trans) gene flow from crops on the genetic diversity from landraces and wild relatives. *Env. Bios Res.* **2003**, *2*, 89–103. 395
58. Wright, S. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evol.* **1965**, *19*, 395–420. 396