Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Cross-species transferability of soybean SSR markers to yam bean (Pachyrhizus erosus L.): an underutilized crop for diversity analysis

Biswajit Jena
CTCRI: Central Tuber Crops Research Institute
Kalidas Pati (kalidas.pati@icar.gov.in)
CTCRI: Central Tuber Crops Research Institute https://orcid.org/0000-0002-9472-1874
Ravindra Donde
CTCRI: Central Tuber Crops Research Institute
Varsha Acharya
CTCRI: Central Tuber Crops Research Institute
Goutam Kumar Dash
CTCRI: Central Tuber Crops Research Institute
Alok Kumar Giri
CTCRI: Central Tuber Crops Research Institute
Vijay Bahadur Singh Chauhan
CTCRI: Central Tuber Crops Research Institute
Maniyam Nedunchezhiyan
CTCRI: Central Tuber Crops Research Institute

Research Article

Keywords: SSR, yam bean, soybean, transferability, genetic diversity

Posted Date: June 2nd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2995547/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at Genetic Resources and Crop Evolution on August 3rd, 2023. See the published version at https://doi.org/10.1007/s10722-023-01692-8.

Abstract

Yam bean (Pachyrhizus erosus L.) is an underutilized tuber crop that has recently gained popularity because of its enormous advantages in terms of nutrition and health. To date, there are no such potential molecular and genomic tools are available for the improvement of the yam bean crop that could help in studying genetic diversity or to identify and selecting elite cultivars for their conservation. The present study has used 67 SSR markers from soybean crops and checked their transferability in the 140 accessions of the yam bean, which identified 15 polymorphic novel SSR markers that showed cross-species transferability. These SSR markers show diversity dissimilarity percentage of 0.42 among the all accession and polymorphic information content (PIC) ranging from 0.03 to 0.84 with a mean of 0.36. Among the 140 accessions of yam bean, the SSR markers showed major allele frequency varying from 0.22 to 0.99 with an average of 0.68. The highest allele frequency, 0.99, was observed in SSR marker GMES0011. The population structure K-3 was identified based on the values of delta K and maximum likelihood (K) among the population of the yam bean. In the AMOVA analysis, identified genetic variance obtained among the population and within the population was 21% and 79%, respectively. These 15 soybean SSR markers successfully differentiated the yam bean accessions in to three groups. The present finding could be used to design a breeding strategy for the yam bean, to discover donor parent with untapped genetic potential for improvement and promote crop preservation.

Introduction

The human population relies on a variety of domesticated plant species that are utilized for a variety of reasons. Few plant species are used for the majority of the food supply; nevertheless, several are still underutilized crops and has not gained much popularity as food crops. Due to their adaptability and the restricted growing and geographical circumstances, some crops might not be consumed worldwide. However, in a diet with a well-balanced nutritional intake, staple crops are for a specific region and provide a substantial contribution to the food supply. Despite being nutritionally rich, staple crops also play an essential role in sustaining and strengthening our country's economic growth. To date, little information is available on underutilized crops that hinder their development and sustainability.

The Yam bean (*Pachyrhizus erosus* L.) Urban is an underutilised tuber crop that has lately risen to prominence due to its significant nutritional and health benefits. The name *Pachyrhizus* is derived from the Greek terms Pachys, which means thick-ended, and Rhiza, signifies root (Sørensen 1996). The crop is a member of the *Leguminosae* family and the Fabaceae (*Papilionaceae*)subfamily, closely related to the Glycininae and Phaseolinae subtribes (Tay Fernandez et al. 2021). The genome size of yam bean vary from 572–597 Mbp and diploid in nature (Pati et al. 2019). In India, it is grown in Bihar, West Bengal, Odisha, Assam, eastern Uttar Pradesh and North-eastern states. In different parts of India, it is known as Misrikand, Sankesh, Shankalu, or Kesaru (Pati et al. 2021). Yam bean has several health benefits due to its nutritional and medicinal properties, as the fresh tuber contains carbohydrates (14.9%), crude fiber (1.4%), protein (1.2%), and lipid (0.1%) (Noman et al. 2007). *Pachyrhizus erosus* dried seeds contain an approximate composition of protein (29.9%) and oil (24%) (Gruneberg et al. 1988). Moreover, Rotenone, a potentially toxic isoflavonoid, is found in the seeds. Therefore, yam bean seeds are not used as a food source (Lautié et al. 2012). However, Rotenone is also used as an insecticide to control thrips, aphids, and whitefly larvae which are fed on crops. The extract of rotenoids

can be used in insecticide products and for industrial purposes (Alavez-Solano et al. 1998; Noman et al. 2007; Lautié et al. 2012).

Yam beans are grown for their starchy roots which is only edible part and they can only be propagated via seeds (Pati et al. 2021). In order to establish the conservation strategies, it is necessary to comprehend the dynamics of the crop's genetic diversity, but this requires molecular tools that can collect data on vital elements like heterozygosity and allelic frequencies, which are needed for computing the majority of the population genetic statistics (Tay Fernandez et al. 2021). In addition, to develop and improve yam bean variety using molecular breeding approaches, it is essential to know the genetic background of the parents. In contrast to this, the conventional breeding technique is time-consuming and laborious. So it cannot address the complicated genetic behaviour of desirable features since they are affected by the yam bean's growth phases and have multiple environmental issues (Pati et al. 2020). However, to document this, unexplored genetic riches have been hampered by a lack of molecular data for yam bean. *Pachyrhizus* species' phylogenetic relationships are still mainly unclear. Being a crop of immense potential lags behind its competitive legume crop, Soybean (*Glycine max*), in terms of genetic study. *Pachyrhizus* species' phylogenetic relationships are still mainly unclear. Presently, no such potential genomic markers are available for this genus or species, which can help in its diversity study, assisting in selecting elite cultivars and proper identification and conservation. However, recently Tay Fernandez et al, 2021 studied the genetic diversity of yam bean through Draft Genome Assembly(Tay Fernandez et al. 2021).

Every crop breeding and conservation of germplasm effort starts with assessing genetic variety, and SSR markers are quick and reliable tools for choosing diverse parents over other molecular markers. Due to their abundance, co-dominance, hypervariability, and transportability between species, microsatellites or Simple Sequence Repeats (SSRs) are the genetic markers of choice for plant and animal systems (Roeder 1995; Powell et al. 1996; Gupta and Varshney 2000). SSRs are effective markers for genotyping, diversity analysis, and genetic mapping in forest tree species (Dayanandan et al. 1997; Brondani et al. 1998; Butcher et al. 2000). Developing SSR markers for an unexplored species is a tiresome job. The Expressed Sequence Tag (EST) collections that are publicly accessible have recently been utilised to create SSR markers (Kantety et al. 2002). But it is only possible for investigated species with abundant sequencing data. An alternate approach is being followed to curb this conflict, sourcing SSR markers developed for other species and utilising them to study the species of interest. It is less cumbersome if the source plant belongs to the same family, subfamily, subtribe or genus. According to the study, comparative genome mapping has revealed that plants have retained various levels of genome conservation during evolution depending on their phylogenetic distance (Choi et al. 2004; Zhu et al. 2005). When compared either genetically or physically, these conserved sections, also known as syntenic or orthologous regions, exhibit collinear gene contents (Phan et al. 2007).

Soybean (*Glycine max*), is an East Asian native species of legume that is commonly cultivated for its edible bean, which has a variety of uses. Both yam bean and soybean belong to same Family i.e., *Fabaceae*, Subfamily: Faboideae and subtribe *Glycininae*. The soybean (*G. max*) crop has been well studied and explored. The soybean genomic size of around 1115 Mb and is a diploidized tetraploid with 2n = 40 chromosomes. A well-developed composite genetic map, a sizable number of ESTs, and a sizable number of BAC libraries are among the accessible well-developed genomic resources. As WGS scaffolds and chromosome pseudomolecules, the soybean genome sequence was uploaded to NCBI in 2010. The NCBI created the RefSeq

chromosomal accessions in 2011. About 2120 SSR markers have been developed for soybean (Hisano et al. 2007; Grant et al. 2010). Therefore, in the current investigation, we are examining the transferability of these SSR markers from soyabean (*Glycine max*) to yam bean through cross-species amplification. As a result, 140 Yam bean accessions were studied using 67 SSR primers for soybean that were chosen from earlier investigations.

Material And Methods

Selection of Plant Materials and DNA extraction

A total of 140 germplasm lines of Yam bean (*P. erosus*) were acquired from the Central Tuber Crop Research Institute-Regional Centre, ICAR, Bhubaneswar germplasm repository. Young, fresh leaves were collected and placed in sample bags that were clearly labelled for each germplasm for DNA isolation. The genomic DNA was isolated from the leaves of each germplasm using the DNA isolation method described by Doyle and Doyle (Doyle and Doyle 1987). The purified DNA was quantified using a NanoDrop spectrophotometer (ND-1000) (Thermo Fisher Scientific, USA). The DNA samples were diluted to a concentration of 25 ng/µl required for PCR amplification.

Selection of Soybean SSR markers

A total of 67 SSR primers of soybean (*G. max*) were selected for the present study to check their transferability in yam beans. Forty one primers were taken from earlier reports that showed amplification in soybean (Hipparagi et al. 2017; Hwang et al. 2020). The total 67 SSR markers were chosen for preliminary screening. For each primer, gradient PCR was carried out using a few randomly chosen yam bean DNA samples to standardise the annealing temperature. Out of 67 SSR markers, 15 demonstrated good and consistent amplification and were selected for further investigation. Taking these 15 markers (Table 1), PCR was performed in 140 germplasm of yam beans.

Table 1
The list of SSR markers were selected for the present study to access the genetic diversity among yam bean accessions

Primer Name	Forward (5−3)	Reverse (3−5)	Annealing Temp(^{0}C)
GMES0007	CCACGTACGTTGTACGCAGT	AAAAGTCAGTTACCCACACCAAA	55
GMES0010	CTGTTCGACACCTCCGATTT	GGTCATCGTAAAGGGACCAA	55
GMES0011	TTTTCCCCTTTGTTTCATGC	CGGAGAGGAAAGAGCAGAGA	55
GMES0015	ATGAGCACCGAAAAGGAGAA	GATACACCTCGGTGGAGGAA	55
GMES0027	AGATGTCGCCATGTGTCTCA	CTCCATTTGGTGGGAGAGAA	52
GMES0041	ACCTTCGCAAATGCTTCTGT	GCGCAATCCACTAGGTGTTAAT	55
SAT179	ATCTACAACGACGCCGTCTC	GCGGCTGGTGTATGGTGAAATCA	51
SAT180	GCCCCTTGCACACTAAGAAG	GCGCCAAAACCAAAAGTTATATC	54
SAT277	GCAGATGCGGTTACGAATTT	TGGCCATGTCAAACCATAACAACA	53
SAT411	AAACCGTCAACCTGGTCAAG	GCGTTGAAGCCGCCTACAAATATAAT	55
SAT415	CTGTTCGACACCTCCGATTT	GCGATATGCTTTGTAAGAAAATTA	56
SAT577	GGTCATCGTAAAGGGACCAA	GCGCAAGCCCAAATATTACAAATT	51
SAT385	AATGGGTACTTTGGTGCTGC	GCGACATATTGCATTAAAAACATACTT	52
GMES0042	TTTGGTTCATGAGCATTGGA	GCGGACTAATTCTATTTTACACCAACAAC	55
SAT183	TCGTTCCTAACATTCATTTTTGT	TAGGTCCCAGAATTTCATTG	55

Note : From the above list of 15 polymorphic SSR markers were used for further molecular genetic diversity analysis

PCR amplification

The PCR reaction was set up in a total volume of 20 µl of reaction mixture containing 1 µl of genomic DNA (25 ng/µl), 2.5 µl of 10X buffer with 15 mM MgCl2, 0.2 µl of 10 mM dNTPs, 1 µl of both the forward and reverse primers (10 nmol), 0.17 µl of Taq DNA polymerase (5 units/µl) (DreamTaq DNA Polymerase, Applied Biosystems, Thermo Fischer Scientific, USA) and 14.03 µl nuclease free water. PCR amplification was performed in a thermocycler (GeneAmp 9700 & Veriti 96 well, Applied Biosystems, USA) with initial denaturation at 96°C for 5 min, followed by 35 cycles of denaturation at 94°C for 1 min, annealing (different for a different set of primers) for 1 min, extension at 72°C for 2 min, and a final extension at 72°C for 7 min. The amplified products were analysed on 3% Low EEO agarose gel using an electrophoresis unit (Bio-Rad Laboratories, PowerPac HV High-Voltage Power Supply, USA), and the photographs of the agarose gels were captured using a gel documentation system (Gel-Doc, Bio-Rad Laboratories, USA). The size of the amplicons on the agarose gels was determined by their migration relative to the standard 50 bp DNA ladder (ThermoScientific, USA.)

Genetic Diversity Analysis and Population STRUCTURE

Visual analysis was used to confirm the gel pictures from the Gel Documentation system for reproducible bands. For each primer, all of the SSR fragments with polymorphic banding patterns were scored in two different methods i.e., based on their base pairs using the Image Lab software (Version 6.0.1 build 34, Bio-Rad Laboratories, Inc.) and presence or absence of amplicons, with 1 denoting the presence of amplicons and 0 denoting their absence to produce binary matrix data. Power Marker V3.25 (Liu and Muse 2005) was used to calculate the polymorphic information content (PIC) value and allele frequency. The polymorphic information content (PIC) of each marker was calculated as per Botstein et al. (Botstein et al. 1980). Using Brookfield's estimate, Micro-Checker 2.2.1 (Van et al. 2004) was used to check for probable genotyping errors, including allelic dropouts, stuttering, or null alleles (Brookfield, 1996). Based on the Ruchi et al. formula, the percentage amplification of each soybean SSR primer was determined (Vir et al. 2009). Jaccard's coefficient of similarity (Jaccard 1908) was measured, and a dendrogram was constructed based on the similarity coefficient generated by Unweighted Pair Group Method using Arithmetic average (UPGMA) using statistical package NTSYS-pc (Sneath and Sokal 1973). We employed a model-based clustering approach implemented in the STRUCTURE software to investigate population structure. The Monte Carlo Markov chain approach was used to estimate allele frequencies in each K populations and the level of mixture for each individual plant. We performed 100,000 iterations with a 100,000 iteration burn-in. For each specified number of populations(K), the results of ten or more separate runs were evaluated (Falush et al. 2003).

Result And Discussion

Genetic Diversity Analysis

After initial screening, 15 soybean primers showed reproducible amplification for 140 yam bean samples. Out of the 15 primers that showed transferability, 6 were EST-SSR markers while the rest belonged to genomic origin. Eight primers showed an amplification percentage of 83.49%, whereas the rest settled with a score of less than 70%. SSRs have been used to study diversification and have been reported to be transferable across species. The transferability of molecular markers across species is crucial for species without sequence data. This study is the first that we are aware of to transfer SSR markers from soybean to yam bean. In all the 15 primers studied, the number of alleles per marker ranged from 1 to 2 with a total of 3928 bands. SAT277 and GMES0007 provided the most polymorphic bands. The most common and uncommon alleles for each SSR locus were the same across all cultivars. The mean PIC value for the SSR markers employed was 0.36 ranging from 0.03 to 0.84, based on the statistical analysis of the binary data (Table 2). The average main allele frequency was determined to be 0.68, ranging from 0.22 to 0.99. GMES0011 had the highest allele frequency of 0.99. Even though the average PIC value calculated in this experiment was significantly lower, it was nevertheless adequate for characterising the yam bean (Fig. 1).

Table 2
Characteristics of amplified fragments in 140 yam bean cultivars using 15 soybean SSR markers

Marker	Major.Allele.Frquency	GenotypeNo	AlleleNo	GeneDiversity	Heterozygosity	PIC
GMES0007	0.60	6	4	0.52	0.68	0.43
GMES0010	0.22	13	11	0.86	0.16	0.84
GMES0011	0.99	2	2	0.03	0.00	0.03
GMES0015	0.57	2	2	0.49	0.86	0.37
GMES0027	0.76	4	3	0.38	0.09	0.33
GMES0041	0.87	2	2	0.22	0.00	0.20
SAT179	0.71	3	3	0.43	0.00	0.37
SAT180	0.51	17	17	0.65	0.00	0.59
SAT277	0.79	2	2	0.34	0.43	0.28
SAT411	0.99	3	3	0.03	0.01	0.03
SAT415	0.72	3	3	0.44	0.00	0.40
SAT577	0.79	2	2	0.34	0.43	0.28
SAT385	0.45	4	3	0.61	0.06	0.53
GMES0042	0.69	2	2	0.43	0.00	0.34
SAT183	0.60	2	2	0.48	0.00	0.36
Mean	**0.68**	**4.46**	**4.06**	**0.42**	**0.18**	**0.36**

Multiple alleles generated from a single locus have been produced due to the significant variance in ploidy levels among the many yam beans (Table 2) (Schmutz et al. 2010). According to PIC, a marker's informativeness or discriminatory power is a relative measure that depends on the population's relative frequency of an allele and its number of alleles (Serrote et al. 2020).

This, implies that a few markers can distinguish between the various yam bean species, which is essential for the preservation of genetic resources containing rare alleles. There was no proof of substantial allele drop-out or stuttering for any of the SSR sites. One hundred forty yam bean varieties were sorted into two main clusters and two subclusters based on Jaccard's similarity and a dendrogram produced using the UPGMA clustering technique from SSR markers (Fig. 2). The subclusters were agained divided to further subclusters based on the distance from the rest accessions. YBBL-13 formed a separate cluster from the rest followed by YB-564 which remained different from the rest in second level cluster. Similarly in the unrooted tree it can be clearly seen that there are three different groups accumulating denoting the three population (Fig. 3).

Population STRUCTURE Analysis

In this work, we found transferability rates for genomic SSR and EST-SSR of 62% and 38%, respectively. The result seems near to the study conducted to check the transferability of cowpea SSR primers to the African Yam bean (*Sphenostylis stenocarpa*). (Shitta et al. 2016). According to their source reports, the size of the amplicon produced by transferable SSRs in yam bean was different from the size of the amplicon produced in soybean. This suggests that several of the soybean primers had amplifications different from the anticipated amplicon size. Datta et al. made comparable observations and took another plant species(Datta et al. 2010). Using STRUCTURE 2.3.4 software and the 15 soybean SSR markers, the population structure of the yam germplasm was deduced. Based on the greatest likelihood and delta K (K) values, the number of subpopulations (K) was determined. The germplasm was split into three major populations at K = 3 (Fig. 4). Similarly, when checking the bar plot the three colour regions claerlt mentioned the three populations derived after structure analysis (Fig. 5).

Transferability studies are being undertaken at a greater pace compared to developing species-specific SSR markers. SSRs have been used to study diversification and have been reported to be transferable across species, such as transferability from *A. officinalis* to *A. racemosus* (Geetha and Siril 2022), *Cryptomeria japonica* var. *sinensis* to *Glyptostrobus pensilis*, *Taiwania cryptomerioides*, *C. japonica*, and *Cunninghamin lanceolata* (Yan et al. 2019), *Glycine max* to *Arachis hypogaea* (He et al. 2006); V. *angularis* to *V. mungo* (Chaitieng et al. 2006)d *unguiculata* to *Vigna sesquipedialis* (Gupta and Gopalakrishna 2010), etc. These markers have been proven to be very useful and effective tools for genetic diversity investigations of various yam bean species. The abundance of the unusual alleles in the chosen yam bean species shows that this broad range of genetic resources has not been widely exploited, even though the importance of the rare alleles cannot be inferred from the current results. Our study can lead to broader area of diversity study and conservation of elite cultivars or species of yam bean. These 15 primers can be used for diversity study of yam bean species and can help in exploring the untouched genetic resource and helping in the conservation of the crop.

Analysis of Molecular Variants Analysis (AMOVA) and Principal Coordinate Analysis (PCoA) Analysis

GenAlEx 6.51 (Peakall and Smouse, 2006, 2012) was used to do molecular variance analysis (AMOVA), which estimates the components of variation related to differences in the yam bean population and within the population on the basis of the Euclidean Square distance. The associations between the members of various populations were evaluated using principal coordinate analysis (PCoA). Three populations as per the report from STRUCTURE analysis, were clustered to three different clusters. The molecular variance among populations was found to be 21% whereas within the population to be 79%(Table 3)(Fig. 6). From the PCoA graph three clear and distinct clusters of population were obtained(Fig. 7).

Table 3
AMOVA analysis table for 3 populations of yam bean

Source	df	SS	MS	Est. Var.	Mol. Var.%
Among Pops	6	177.106	29.518	0.676	21%
Within Pops	273	696.234	2.550	2.550	79%
Total	279	873.339	—	3.227	100%

Conclusion

This study is one among the first to be reported which showed the transferability of Soybean SSR markers to yam bean. These markers have been proven to be very useful and effective tools for genetic diversity investigations of a variety of yam bean species. The abundance of the unusual alleles in the chosen yam bean species shows that this broad range of genetic resources has not been widely exploited, even though the importance of the rare alleles cannot be inferred from the current results. Our study can lead to wider area of diversity study and conservation of elite cultivars or species of yam bean. These 15 primers of soybean can be used for diversity study of yam bean species and can help in conservation of wild species of yam bean and accession. The study could help in exploring the untouched genetic resource that could be used in future breeding program to develop climate smart and nutrition rich yam bean variety.

Declarations

Acknowledgement

The authors are thankful to the Director, ICAR, CTCRI, Trivandrum, Kerala, India for kind support to carry out the research work.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

All authors contributed to the study conception and design. Kalidas Pati and Goutam Kumar Dash designed the experiment. Biswajit Jena, Varsha Acharya and Alok Kumar Giri performed the experiments. Biswajit Jena and Ravindra Donde performed the data analysis. Biswajit Jena and Ravindra Donde prepared the manuscript and reviewed by Kalidas Pati, Goutam Kumar Dash, Vijay Bahadur Singh Chauhan and Maniyam Nedunchezhiyan. All authors read and approved the final manuscript.

References

1. Alavez-Solano D, Chilpa JRR, Jimenez-Estrada MA, et al (1998) Chemical study and insecticidal activities of seeds of *Pachyrhizus erosus* (L.) Urban. In: International Symposium on Tuberous Legumes, Celaya (Mexico), 5-8 Aug 1996. KVL
2. Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331
3. Brondani RPV, Brondani C, Tarchini R, Grattapaglia D (1998) Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. Theor Appl Genet 97:816–827

4. Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455. https://doi.org/https://doi.org/10.1046/j.1365-294X.1996.00098.x
5. Butcher PA, Decroocq S, Gray Y, Moran GF (2000) Development, inheritance and cross-species amplification of microsatellite markers from *Acacia mangium*. Theor Appl Genet 101:1282–1290
6. Chaitieng B, Kaga A, Tomooka N, et al (2006) Development of a black gram [*Vigna mungo* (L.) Hepper] linkage map and its comparison with an azuki bean [*Vigna angularis* (Willd.) Ohwi and Ohashi] linkage map. Theor Appl Genet 113:1261–1269. https://doi.org/10.1007/s00122-006-0380-5
7. Choi H-K, Kim D, Uhm T, et al (2004) A sequence-based genetic map of Medicago truncatula and comparison of marker colinearity with *M. sativa*. Genetics 166:1463–1502. https://doi.org/10.1534/genetics.166.3.1463
8. Datta S, Kaashyap M, Kumar S (2010) Amplification of chickpea-specific SSR primers in Cajanus species and their validity in diversity analysis. Plant Breed 129:334–340. https://doi.org/https://doi.org/10.1111/j.1439-0523.2009.01678.x
9. Dayanandan S, Bawa K, Kesseli R (1997) Conservation of microsatellites among tropical trees (Leguminosae). Am J Bot 84:1658
10. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull
11. Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics 164:1567–1587. https://doi.org/10.1093/genetics/164.4.1567
12. Geetha CM, Siril EA (2022) Cross-species transferability of genomic SSR markers and genetic diversity among Asparagus racemosus Willd. Accessions. Plant Gene 31:100361. https://doi.org/https://doi.org/10.1016/j.plgene.2022.100361
13. Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res 38:D843-6. https://doi.org/10.1093/nar/gkp798
14. Gruneberg WJ, Buttner G, Delgado-Licon E (1988) Eucarpia International Symposium on Breeding of Protein and Oil crops, Pontevedra, Spain. 95–97
15. Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–185. https://doi.org/10.1023/A:1003910819967
16. Gupta SK, Gopalakrishna T (2010) Development of unigene-derived SSR markers in cowpea (*Vigna unguiculata*) and their transferability to other Vigna species. Genome 53:508–523
17. He G, Woullard FE, Marong I, Guo BZ (2006) Transferability of Soybean SSR Markers in Peanut (*Arachis hypogaea* L.). Peanut Sci 33:22–28. https://doi.org/10.3146/0095-3679(2006)33[22:tossmi]2.0.co;2
18. Hipparagi Y, Singh R, Choudhury DR, Gupta V (2017) Genetic diversity and population structure analysis of Kala bhat (*Glycine max* (L.) Merrill) genotypes using SSR markers. Hereditas 154:9. https://doi.org/10.1186/s41065-017-0030-8
19. Hisano H, Sato S, Isobe S, et al (2007) Characterization of the soybean genome using EST-derived microsatellite markers. DNA Res an Int J rapid Publ reports genes genomes 14:271–281.

https://doi.org/10.1093/dnares/dsm025

20. Hwang TY, Gwak BS, Sung J, Kim HS (2020) Genetic diversity patterns and discrimination of 172 korean soybean (*Glycine max* (L.) merrill) varieties based on SSR analysis. Agric 10:. https://doi.org/10.3390/agriculture10030077
21. Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull la Société Vaudoise des Sci Nat 44:223. https://doi.org/10.5169/seals-268384
22. Kantety R V, La Rota M, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. Plant Mol Biol 48:501–510. https://doi.org/10.1023/A:1014875206165
23. Lautié E, Rozet E, Hubert P, Quetin Leclercq J (2012) Quantification of rotenone in seeds of different species of yam bean (*Pachyrhizus* sp.) by a SPE HPLC–UV method. Food Chem 131:1531–1538. https://doi.org/https://doi.org/10.1016/j.foodchem.2011.09.125
24. Liu K, Muse S V (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129. https://doi.org/10.1093/bioinformatics/bti282
25. Noman ASM, Hoque MA, Haque MM, et al (2007) Nutritional and anti-nutritional components in *Pachyrhizus erosus* L. tuber. Food Chem 102:1112–1118. https://doi.org/https://doi.org/10.1016/j.foodchem.2006.06.055
26. Pati K, Behera S, Ray DK, et al (2021) Element profiling of thirty genotypes of yam bean in eastern India by using proton induced X-ray emission (PIXE). J Food Compos Anal 102:104006. https://doi.org/https://doi.org/10.1016/j.jfca.2021.104006
27. Pati K, Kaliyappan R, Chauhan VBS, et al (2020) Phenological growth stages of underutilised crop yam bean (*Pachyrhizus erosus* L. Urban) according to the extended BBCH scale. Ann Appl Biol 177:417–423. https://doi.org/https://doi.org/10.1111/aab.12637
28. Pati K, Zhang F, Batley J (2019) First report of genome size and ploidy of the underutilized leguminous tuber crop Yam Bean (*Pachyrhizus erosus* and *P. tuberosus*) by flow cytometry. Plant Genet Resour Charact Util 17:456–459. https://doi.org/DOI: 10.1017/S1479262119000170
29. Peakall R, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. https://doi.org/https://doi.org/10.1111/j.1471-8286.2005.01155.x
30. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460
31. Phan HTT, Ellwood SR, Hane JK, et al (2007) Extensive macrosynteny between Medicago truncatula and Lens culinaris ssp. culinaris. Theor Appl Genet 114:549–558. https://doi.org/10.1007/s00122-006-0455-3
32. Powell WW, Koput KW, Smith-Doerr L (1996) Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. Adm Sci Q 41:116–145. https://doi.org/10.2307/2393988
33. Roeder E (1995) Medicinal plants in Europe containing pyrrolizidine alkaloids. Pharmazie 50:83–98

34. Schmutz J, Cannon SB, Schlueter J, et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183. https://doi.org/10.1038/nature08670
35. Serrote CML, Reiniger LRS, Silva KB, et al (2020) Determining the Polymorphism Information Content of a molecular marker. Gene 726:144175. https://doi.org/https://doi.org/10.1016/j.gene.2019.144175
36. Shitta NS, Abberton MT, Adesoye AI, et al (2016) Analysis of genetic diversity of African yam bean using SSR markers derived from cowpea. Plant Genet Resour Characterisation Util 14:50–56. https://doi.org/10.1017/S1479262115000064
37. Sneath PHA, Sokal RR (1973) Numerical Taxonomy: The Principles and Practice of Numerical Classification. W. H. Freeman and Co.
38. Sørensen M (1996) Yam Bean: Pachyrhizus DC. - Promoting the conservation and use of underutilized and neglected crops. 2.: IPGRI
39. Tay Fernandez CG, Pati K, Severn-Ellis AA, et al (2021) Studying the Genetic Diversity of Yam Bean Using a New Draft Genome Assembly. Agronomy 11:. https://doi.org/10.3390/agronomy11050953
40. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. https://doi.org/https://doi.org/10.1111/j.1471-8286.2004.00684.x
41. Vir R, Bhat K V, Lakhanpaul S (2009) Transferability of sequence tagged microsatellite sites (STMS) primers to pulse yielding taxa belonging to *Phaseolae*. IJIB 5:62–66
42. Yan Y-D, Li X-Y, Worth JRP, et al (2019) Development of chloroplast microsatellite markers for *Glyptostrobus pensilis* (Cupressaceae). Appl Plant Sci 7:e11277. https://doi.org/https://doi.org/10.1002/aps3.11277
43. Zhu H, Choi H-K, Cook DR, Shoemaker RC (2005) Bridging model and crop legumes through comparative genomics. Plant Physiol 137:1189–1196. https://doi.org/10.1104/pp.104.058891

Figures

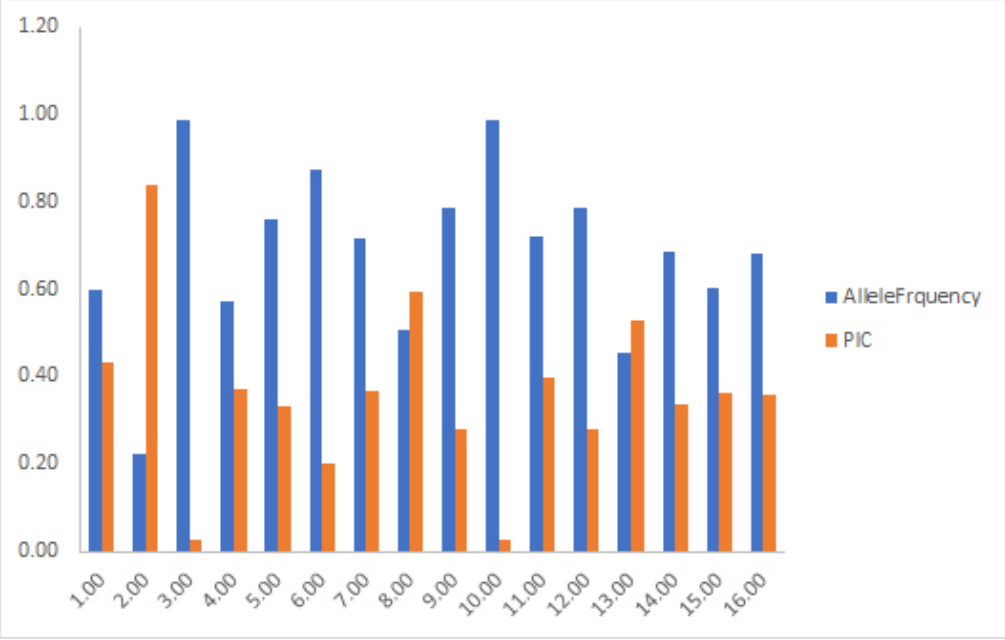


Figure 1

PIC and allele frequency produced in 140 yam bean germplasms by using 15 soybean SSR markers.

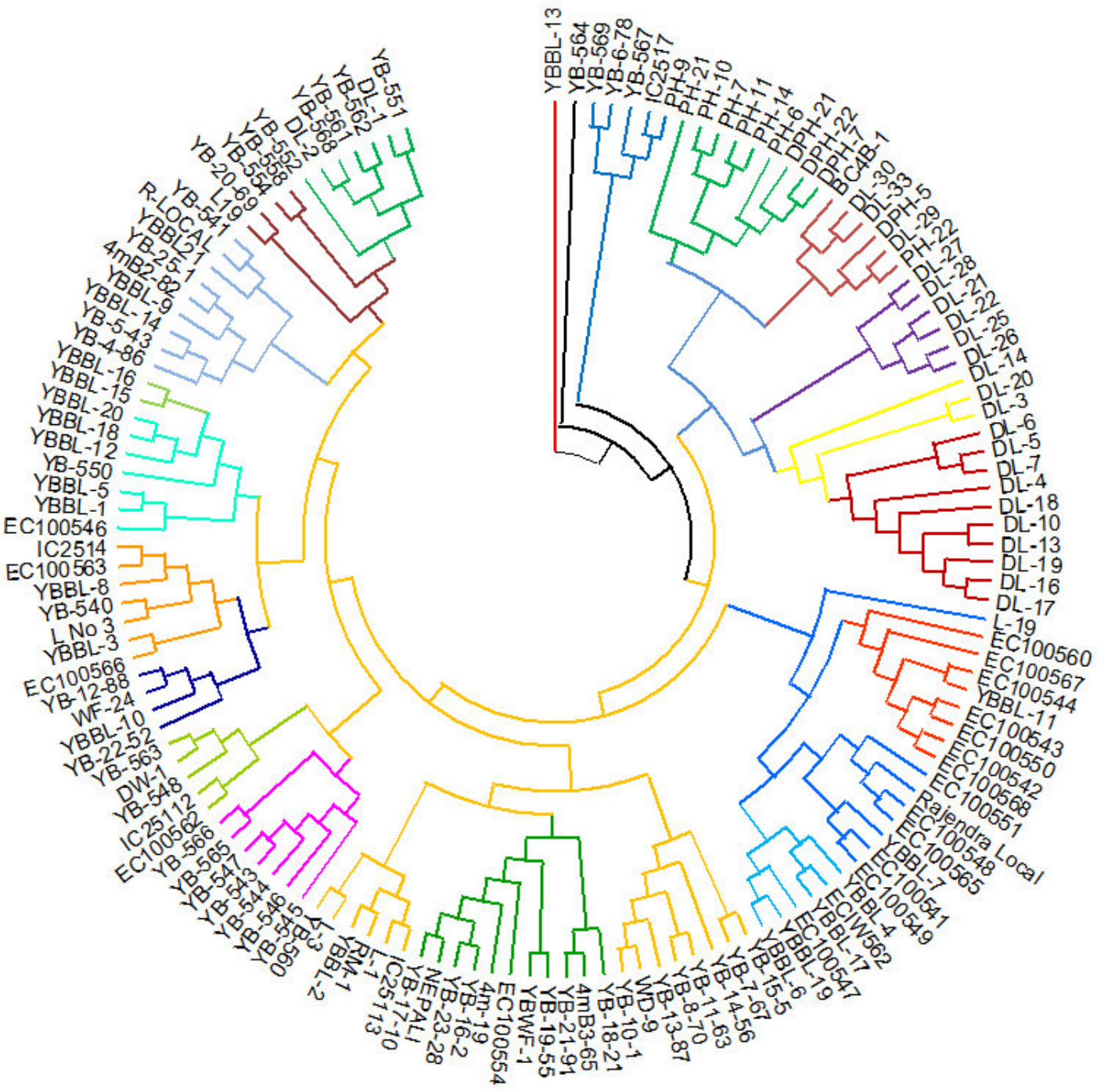


Figure 2

Dendrogram prepared taking the 140 cultivars of yam bean using UPGMA clustering technique.

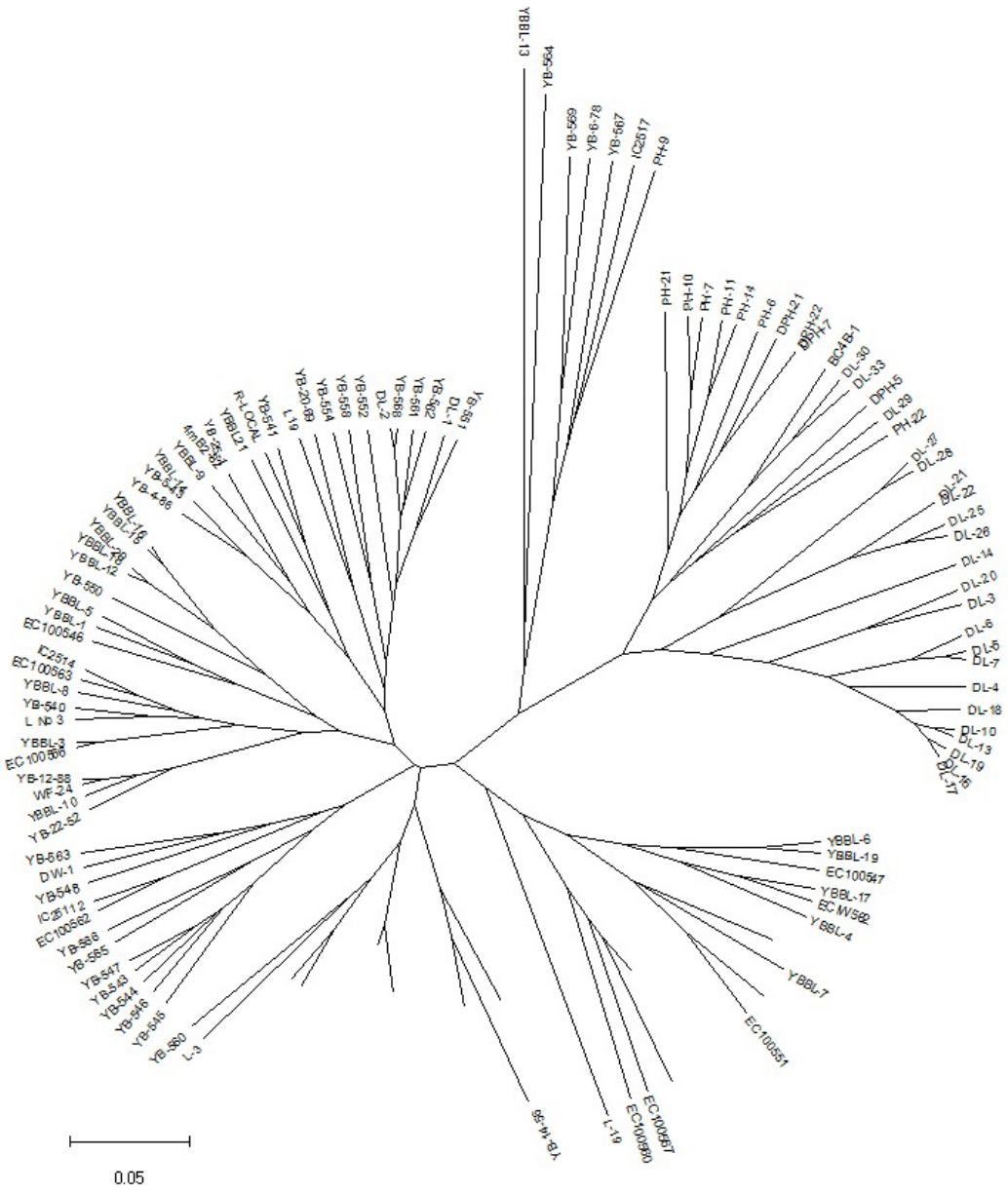


Figure 3

Unrooted tree based on UPGMA clustering method.

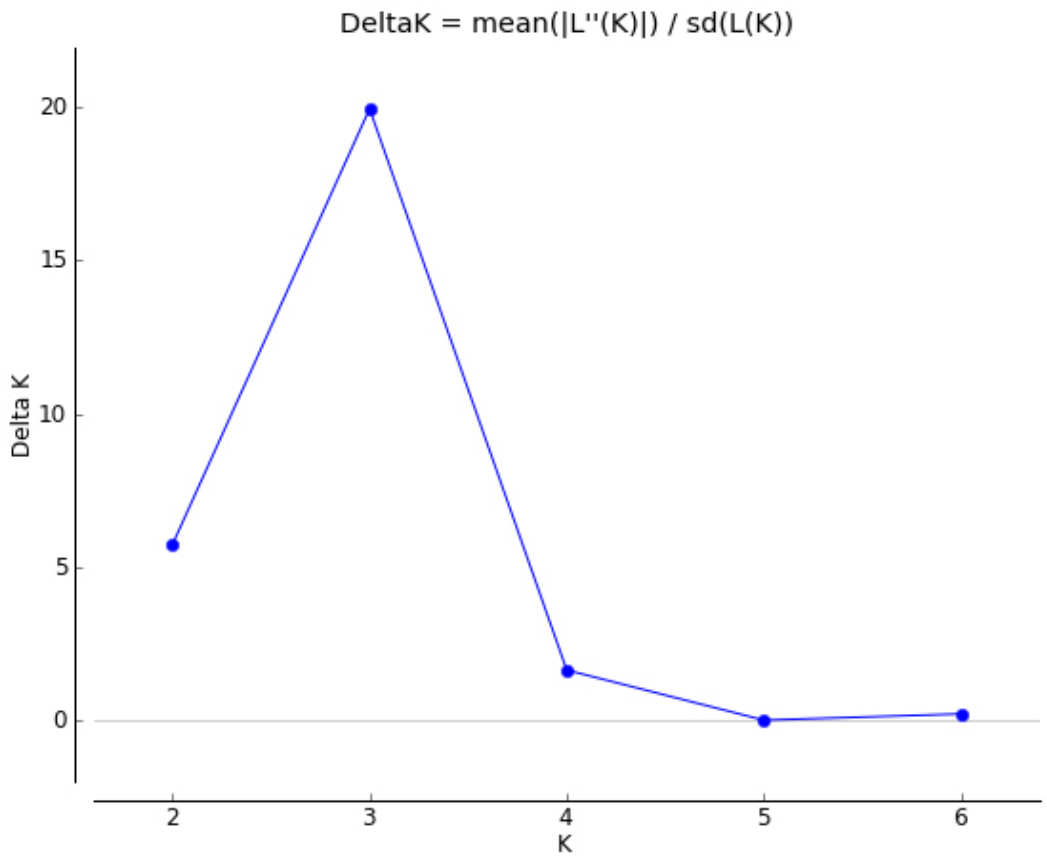


Figure 4

Delta K plot for germplasms analyzed using model-based clustering of STRUCTURE.

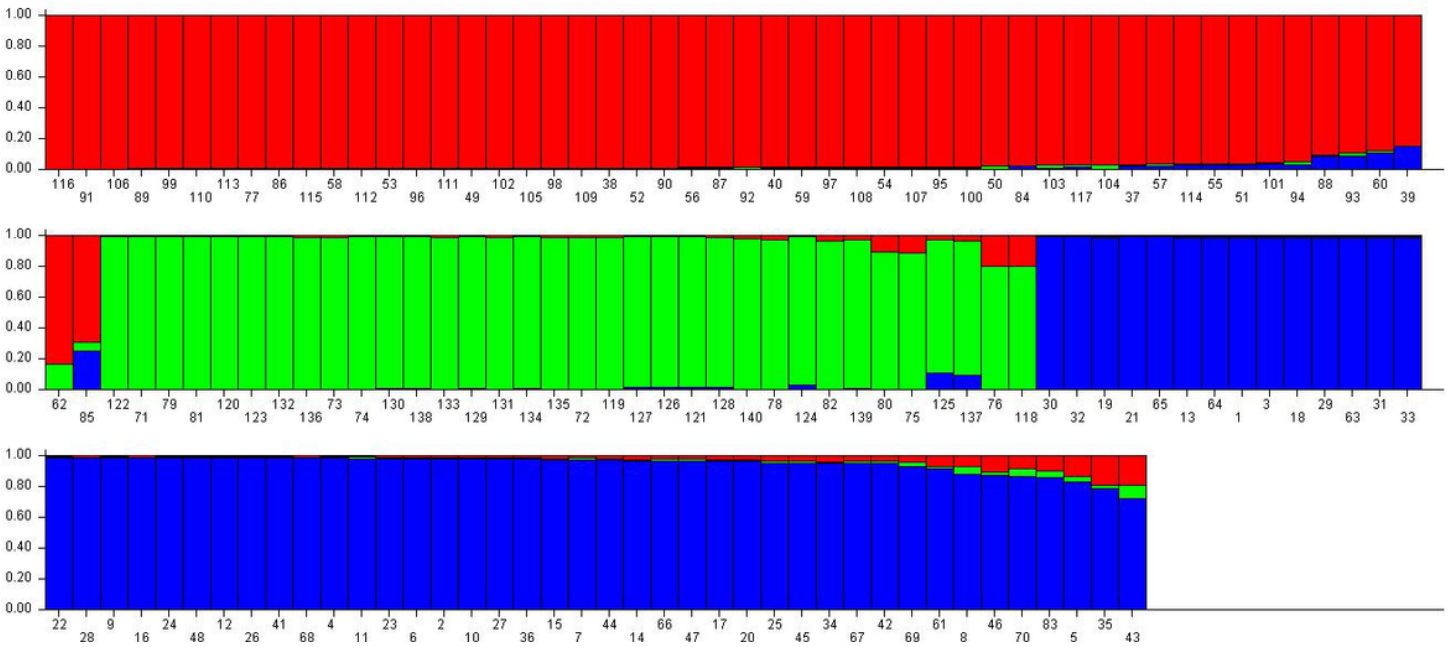


Figure 5

Bar Plot showing three different populations for 140Yam bean accessions.

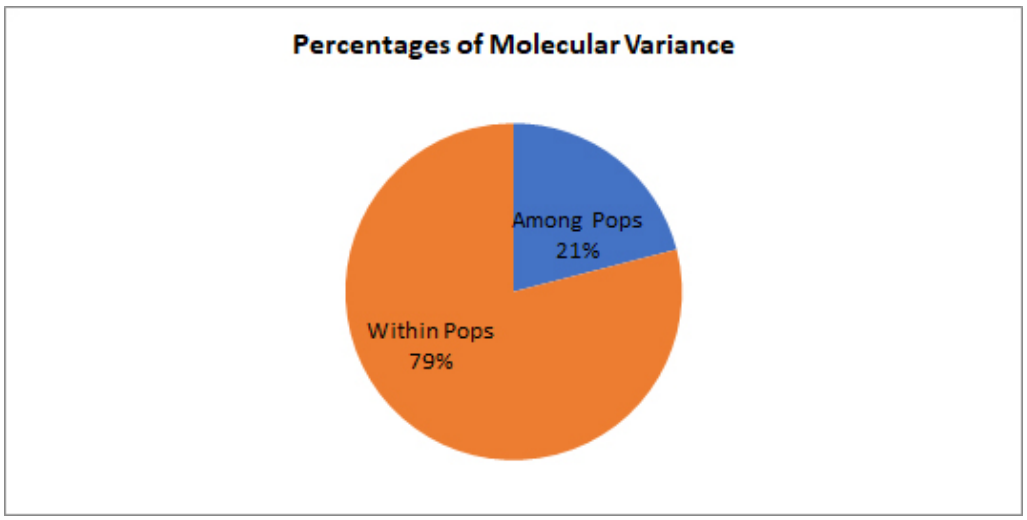


Figure 6

Percentages of Molecular Variance within and among populations of yam bean used for the present study.

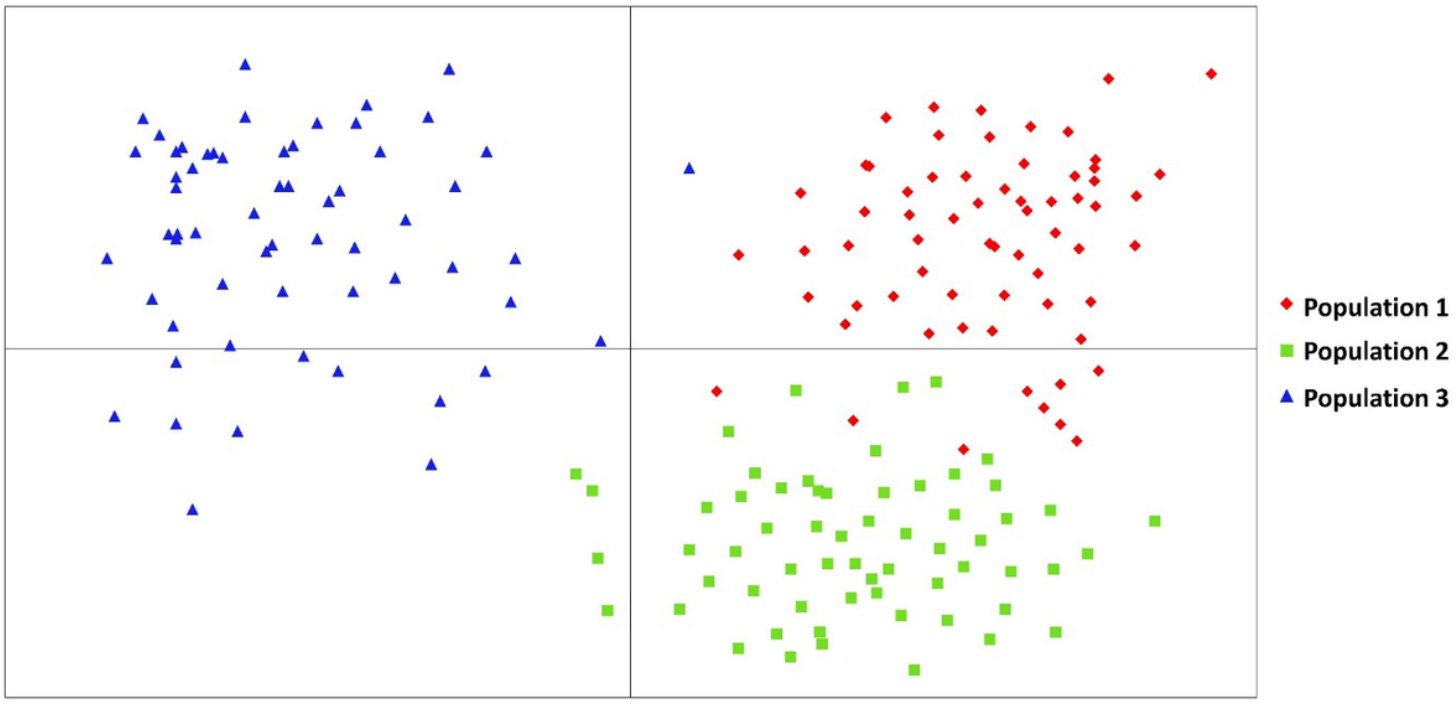


Figure 7

PCoA Analysis Graph.