RESEARCH ARTICLE

Analysis of genetic diversity inpigeonpeagermplasm using retrotransposon-based molecular markers

Maneesha[1]and Kailash C. Upadhyaya[1,2,*]

[1]School of Life Sciences, Jawaharlal Nehru University, New Delhi – 110067, India

[2]Amity Institute of Molecular Biology and Genomics, Amity University, Noida-201313, India

Keywords: Molecular markers, Pigeonpea, Molecular diversity,Retrotransposon, *Panzee*, SSAP, REMAP

Abstract

Pigeonpea (*Cajanuscajan*), an important legume crop is predominantly cultivated in tropical and subtropical regions of Asia and Africa. It is normally considered to have a low degree of genetic diversity, an impediment in undertaking crop improvement programs. We have analyzed genetic polymorphism of domesticated pigeonpeagermplasm (47 accessions) across the world using earlier characterized *panzee*retrotransposon-based molecular markers. It was conjectured that since retrotransposons are interspersed throughout the genome, retroelements–based markers would be able to uncover polymorphism possibly inherent in the diversity of retroelement sequences. Two PCR-based techniques, SSAP (Sequence-SpecificAmplified Polymorphism) and REMAP (REtrotransposonMicrosatellite Amplified Polymorphism) were utilized for the analysis. We show that a considerable degree of polymorphism could be detected using these techniques. Three primer combinations in SSAP generated 297 amplified products across 47 accessions with an average of 99 amplicons per assay. Degree of polymorphism varied from 84-95%. In the REMAP assays, the number of amplicons was much less but up to 73% polymorphism could be detected. On the basis of similarity coefficients, dendrograms were constructed. The results demonstrate that the retrotransposon-based markers could serve as a better alternative for the assessment of genetic diversity in crops with apparent low genetic base.

[*]Corresponding author: Prof. K. C. Upadhyaya, Amity Institute of Molecular Biology and Genomics, Amity University, Sector 125, Noida 201313, India.
E-mail:kcupadhyaya@gmail.com; kcupadhyaya@amity.edu

Introduction

Grain legumes constitute an important component of protein intake in human diet especially in the context of India. Pigeonpea(*Cajanuscajan*) is an important major pulse crop predominantly cultivated in semi-arid tropical and subtropical regions of Asia and Africa. It is an important grain that ranks among the five edible legumes that is used in intercropping systems, for food, fodder and firewood. The pigeon pea seeds are rich in sulfur containing amino acids, methionine and cysteine and rich source of proteins and vitamins and minerals.*C.cajan* is normally self-pollinated species with great adaptability in wide range of environmental conditions. India is the "centre of origin" of pigeonpea and Africa could be considered second major centre of origin for this crop (van der Maesen 1989).

Pigeonpea is diploid (2n=22) with a genome size of ~833 Mb and contains over 48,600 genes (Varshney*et al*. 2012). Its genome contains over 51% sequences as transposable elements with majority of them as retrotransposons (Singh *et al*. 2012; Varshney*et al*. 2012). It is known to contain a low level of genetic diversity a definite impediment in breeding programs (Yang *etal*. 2006). DNA based molecular markers are increasingly becoming important tools for selection in breeding programs and for analyzing genetic polymorphism. Several techniques have been developed such as RFLP (restriction fragment length polymorphism), RAPD (random amplified polymorphic DNA), SSR (simple sequence repeats), AFLP (amplified fragment length polymorphism), SNPs (single nucleotide polymorphisms), DArT (diversity array technology), genic-SSR (simple sequence repeats) etc.(see review by Varshney*etal.* 2013). Since retrotransposons are distributed across the entire genome,it was conjectured that due to their large proportion and dispersed localization within the genome, the retrotransposon-based markers could be utilized to uncover polymorphism possibly inherent in diversity of retroelement sequences. Therefore, theretrotransposon-based markerscould be considered a desirable alternative to analyze molecular polymorphism.

The retrotransposon-based molecular markers techniques have recently been developed as efficient and informative. The plant genomes including those of legumes contain a large proportion of retroelementsnormally dispersed throughout the genome. They exist as highly heterogeneous families, and due to their "copy-and-paste" retrotranspositional mechanism, their

original insertion site(s) remain undisturbed. Thus, retroelements appear to be useful candidates as molecular markers for assessment of genetic diversity and DNA fingerprinting in crop plants especially for those with low genetic base (Yang et al. 2006; Varshney*et al*. 2013).

We isolated and characterized a*copia*family of retrotransposonfrom pigeonpea, named *panzee* (Lall*et al*. 2002). The element is present in very high copy number in the pigeonpeagenome. The element contains similar 372-bp of 5' LTR and 383-bp of 3' LTR sequences except 11-bp additional sequences in the 3' LTR. We have utilized the *panzee*retroelement-based SSAP (Sequence-SpecificAmplified Polymorphism) and REMAP (REtrotransposonMicrosatelliteAmplified Polymorphism) techniques for analyzing the genetic diversity of pigeonpeagermplasm. A high degree of polymorphism could be shown using these marker techniques in spite of lowgenetic polymorphism in the pigeonpea. We show that the techniques are useful and could be a better alternative to other marker techniques in analyzing the genetic diversity of various pigeonpeaaccessions and could also be utilized for fingerprinting cultivated varieties.

Materials and methods

Plant material and DNA preparations

Forty-sixpigeonpea (*C.cajan*) accessions belonging to different eco-geographical regions of the world were obtained from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad (Table 1). PUSA84 was obtained from Indian Agricultural Research Institute (IARI), New Delhi. Three-week-old freeze-dried seedlings were used for DNA extraction. DNA was isolated using the original protocol provided with the Wizard genomic DNA kit (Promega).

SSAP (Sequence-Specific Amplified Polymorphism)

SSAP is a modified AFLP, where one primer comes from the conserved termini of the LTR retrotransposon sequence and the other is based on the presence of a nearby restriction site (Mse I) It detects variation in the DNA flanking the retrotransposon insertion site and displays individual retrotransposon insertions as bands.

The selective amplification employs one adaptor-homologous primer, together with primers designed facing outward from the long terminal repeats (LTRs) of the *panzee*retrotransposon. Samples were amplified adding 3 selective nucleotides to the MseI primer to reduce the number of bands to a level that is easily scored for their presence /absence.The SSAP method as described by Waugh *et al.* (1997) was used with modifications. Genomic DNA (1µg) was digested with enzymes EcoRI (5 U) and MseI (5 U). Enzyme-specific adapter and primer sequences for EcoRI and MseI primer sequences for EcoRI and MseI primer sequences were as described by Vos*et al.* (1995). Pre-amplifications were carried out in 20 µl reaction mixture containing ten-fold diluted (digested and ligated) DNA with EcoRI and MseI primers having A and C respectively, as selective nucleotides. The cycling conditions were: one cycle of 45 s at 94°C, 30 s at 65°C, 1 min at 72°C, followed by a touch-down profile for the annealing step (13 cycles in which the annealing temperature was decreased at a rate of 0.7°C/cycle), followed by 18 cycles at a constant annealing temperature of 55.9°C, and a final extension step at 72°C for 5 min. Selective amplification was performed using an extended MseI + 3 primer (the selective bases used were CTG, CTA, CAC) in combination with *panzee*3' LTR-based primer K2 (5'-CTCTTGTTTTAGCTGTGAG-3') labeled at the 5' end with gamma ^{32}P-ATP (5000Ci/mmole) and 2 units T4 kinase. Amplifications were carried out using the amplification profile described above. After PCR, 8µl of loading buffer (98% formamide, 2% dextran blue, 0.25 mM EDTA) was added to each sample. Samples were denatured at 90°C for 5 min and then immediately placed on ice. An aliquot (6µl) of each sample was loaded onto a 6% denaturing polyacrylamide gel, which had been pre-run for 2 h at 80 W. Gels were dried and autoradiographed.

REMAP (Retrotransposon Microsatellite Amplified Polymorphism)

REMAP uses one outward facing primer based on a LTR target sequence and the other primer based on a simple sequence motive (microsatellite) with a repeat length of 7 to 10. It detects polymorphism among amplicons produced between retrotransposonsand microsatellites generated by PCR.

The reverse primer corresponded to either 3' end or 5' end of the *Panzee* LTR was used for REMAP. The following forward primers based on two dinucleotides repeats, (GA)n and (CT)n; and threetrinucleotide repeats, (CAC)n and (GTG)n, were used. The primers are anchored at the 3' ends

of the microsatellite repeat; one is anchored at 5' end. Following is a list of the LTR and microsatellite PCR primers:

LTR primers

K1 (5' LTR) 5'-TGTTGGCCGAAAGGAAAA -3'

K2 (3' LTR) 5'-CTCTTGTTTTAGCTGTGAG-3'

K3 (3'-LTR) 5'-GCCTACTCTGGACCTAACA-3'

Microsatellite primers

MS1 $(GA)_9C$

MS4 $(GTG)_7C$

MS5 $(CA)_{10}G$

MS7 $GT(CAC)_7$

The PCR reactions were performed in a 20 µl reaction mixture containing 100 ng DNA, 1x PCR buffer, 2 mM $MgCl_2$, 10 pmol each primer, 200 nMdNTP, 1U Taq polymerase (Promega). Amplifications were performed in a PTC-200 (MJ Research) in 0.2-ml tubes (Perkin Elmer). The PCR reaction program consisted of: 1 cycle at 94°C, 2 min; 1cycle at 94°C, 30 sec; annealing temperature (T_a) as specified; ramp + 0.5°Csec^{-1} to 72°C; 30 cycles of 72°C, 2 min + 3 sec; 1 cycle at 72°C, 10 min; 4°C. Products wereelectrophoresed on 5% polyacrylamide gel and analyzed following ethidium bromide staining.

Data analyses

SSAP and REMAP amplification products were scored for their presence (1) and absence (0) across all accessions for all primer combinations employed to generate a binary matrix. The data were analyzed with NTSYS-pc software version 1.8 (Rohlf 1989) and he genetic distance was calculated for all possible pair wise comparisons according to Nei and Li (1979). Genetic similarity (S) was calculated by making a pair-wise comparison among the accessions using Jaccard's coefficient. The formula is given as follows: S_{ij} = 2x (number of common bands)/total bands in accession 'i' + total bands in accession 'j'. The similarity matrix thus generated was used to construct a dendrogram using unweighted pair group method of averages (UPGMA). Principal correspondence analysis was carried out on the similarity matrix, with the scores on the principal axes plotted pair-wise.

Results

SSAP analysis based on panzeeLTR

The techniques of SSAP and REMAP based on LTR sequences of panzeeretrotransposon was utilized for analyzing the genetic diversity in various pigeonpea accessions from different geographical regions. Since the*panzee*retroelement is present in high copy number across the pigeonpea genome, it was expected that the amplified sequences would cover a large part of the genome at least in SSAP. The different MseI primers with the LTR primer resulted 297 amplified fragments from 39 accessions. Figure 1 shows a representative SSAP profile (LTR x M – CTG). The details of amplicons with three different MseI primers are given in Table 2. The M-CTG produced the maximum number of amplicons (124) followed by M-CTA (114). The least number (59) were generated in M-CTA.

On an average 92% of polymorphic bands were observed. A separate analysis of 19 Indian accessions with three-primer combination is presented in Table 2. These accessions produced on an average 254 bands. The amplicon, labeled B1 is rare polymorphic while B2 and B3 are polymorphic with frequent occurrence across the accessions. The amplicons, B4 and B5 are accession-specific, where as B4 is specific to Indian accession (105) from Orissa and B5 is specific to Australian accession (60). B6 is a band absent in accession 60 but present in rest of the accessions. B7 is a monomorphic band present across all the accessions. These polymorphic bands reveal heterozygosity in pigeonpea. Fingerprints of accessions 60 from Australia and 105 from India exhibit a different pattern. Table 2 summarizes the total number of bands amplified with different primer combinations and percentage polymorphism (calculated as polymorphic bands divided by total number) detected across all the pigeonpea accessions. Analysis of three SSAP (MseI) primer combinations generated 297 amplified products with an average of 99 amplifications per assay across 39 accessions ranging from 129-59 bands in primer combinations K2 (LTR) + M-CTG and K2 (LTR) + M-CTA, respectively.

Analysis revealed an average percentage polymorphism of 91% that ranged from a high of 95% in primer combination K2 (LTR) + M-CTG to a low of 84% in the primer combination K2 (LTR) + M-CTA.A separate analysis for the 254 amplification products for the 19 Indian accessions with the three primer combinations is given in Table 2. The average amplification products per

primer combinations were 85, which ranged from 101 bands in primer combination, K2 (LTR) + M-CTG to 55 bands in primer combination K2 (LTR) + M-CAC. The overall percentage polymorphism across the Indian accessions was observed to be 77%. It ranged from 70% in the primer combination, K2 (LTR) + M-CTG to 87% for K2 (LTR) + M-CTA. Further analysis showed that 4% polymorphism was contributed by accession 60 and 3% by 105. PUSA84 contributed 2% in the Indian data set. Together 60 and 105 contributed 4% in the overall polymorphism and 2% in the Indian data.

REMAP analysis

Three outward-facing LTR primers were utilized in combination with four anchored microsatellite primers (Table 2). The LTR primers used were same as the ones used in SSAP. Different combinations were tried with three LTR primers and seven microsatellite primers. The best band-yielding combinations were considered for the present studies. A total of 45accessions from various eco-geographical regions of the world were compared by REMAP. Figure 2 shows a typical REMAP profile obtained employing primer combination K3 (LTR) + $(CA)_{10}G$. For each primer combination three gels were run. The band sizes range from 200 to 1200bp. The total number of amplified products with this primer combination was 21. The percentage polymorphism detected was 95% with this primer combination.

Table 2 gives the total number of fragments amplified per primer combination, percentage polymorphism detected across the 45 pigeonpea accessions. A total of 113 REMAP fragments were amplified with a mean of 16 fragments per assay with maximum number of bands (21) in primer combinations K3(LTR) + $(GA)_9C$ and K3 (LTR) + $(CA)_{10}G$ to primer combination K1(LTR) + $GT(CAC)_7$ with 11 bands. Average percentage polymorphism with 7 primer combinations was observed to be 85% across 45 accessions that ranged from a high of 100% to 54% in the primer combinations K1(LTR) + MS7[$GT(CAC)_7$] and K3(LTR) + MS4$(GA)_9C$, respectively. In a separate analysis for 20 Indian accessions, the average number of bands amplified per primer combination was found to be 16. The average percentage polymorphism in Indian accessions was 73%. The percentage polymorphism ranged from a high of 100 in primer combination K2(LTR) + MS1 $(GA)_9C$ to a low of 40 in primer combination K1(LTR) + MS7 [$GT(CAC)_7$]. A separate set of analysis was also done to see the impact of the accessions exhibiting a different banding pattern on

the overall REMAP analysis. In order to explore the possibility of generating some PCR bands from the SSR primer alone in REMAP reactions, in control experiments for REMAP, the retrotransposon primers were left out of the reaction. It was observed that the REMAP bands were not generated by amplification between the microsatellite repeats (ISSR) as none of the bands were shared with ISSR pattern, indicating that all the bands span intervening domains between LTRs and microsatellites in the genome (Fig. 3).

The marker efficiency in terms of the average number of fragments amplified per assay revealed that SSAP generated more fragments. Comparison of SSAP and REMAP showed that, SSAP detects more polymorphic fragments per assay. The data analysis revealed that the overall average percentage polymorphism within pigeonpea accessions was 91 and 81% in SSAP and REMAP, respectively. The average percentage polymorphism within Indian accessions as detected by SSAP was higher (77%) as compared to REMAP (73%). Thus SSAP is able to detect higher degree of polymorphism as compared to REMAP among various pigeonpea varieties.

Genetic similarity matrix and cluster analysis

The accessions were clustered based on genetic similarities in a pheneticdendrogram using UPGMA algorithm for SSAP and REMAP (Fig. 4-5). The individual phenograms generated from SSAP and REMAP revealed consistent pattern of groupings. The major difference being the similarity values at which the pigeonpea accessions clustered together depicting the same underlying pattern of grouping. Two major clusters could be observed in the phenograms generated with all the accessions, where the Australian accession (60) formed a separate cluster and rest of the accessions formed a second large cluster, which separated at 0.34 similarity value in SSAP phenograms, whereas in REMAP analysis, the same separated at 0.54 similarity value. This observation reiterates that SSAP markers detect more polymorphic loci.

In SSAP the 39 genotypes were grouped into two main clusters (Fig. 4), cluster I has the Australian accession 60 and cluster II having the rest 37 accessions, which are not tightly grouped. Cluster II was divided into many sub-clusters with Indian accessions scattered all over. Sub-cluster IIa has 16 Indian accessions which is further consists of many groups like accessions from contiguous states or same state grouped together (i.e. two accessions (57, 100) from Andhra Pradesh (AP), accessions (49, 50) from Andhra Pradesh and Maharashtra are placed together. Indian

accessions from Orissa (105) and AP (15) do not group with any Indian accessions. Accessions from Uganda (41) and Germany (106) are placed together. A separate dendrogram of 19 Indian accessionsrevealed that accessions from Orissa (105), AP(15) and PUSA-84 separate out from the rest of the accessions while within the large cluster there are several accessions belonging to same state or adjoining states are placed together (e.g. two accessions (57, 100) from Andhra Pradesh (AP), accessions (49, 50) from Andhra Pradesh and Maharashtra.

The dendrogramderived from REMAP data (Fig. 5) reveals that the 45 accessions are grouped into two clusters where Australian accession (60) stands out singularly while cluster II having the rest of the accessions. Cluster II was divided into two main sub-clusters with Indian accession from Orissa (15) forming a solitary sub-cluster and rest of the accessions forming a separate sub-cluster. The sub-cluster II was further divided into many groups (i.e. Indian accessions formed two major groups one consisting of 11 Indian accessions was present where PUSA-84 forms a separate subgroup with accession from Myanmar, and second group with four accessions. Likewise African accessions also formed two major groups, one with 7 accessions at a high GS value of 0.75, which contains2 accessions from Nigeria (33 and 80) and 3 from Kenya (103, 23 and 104) and another group with 3 accessions (37, 88 and 101). Accessions from Barbados (67), Jamaica (98), and Venezuela (32 and 46) were grouped together, which are geographically very close. Accessions from Uganda and Germany were together in REMAP as they were in SSAP. The dendrogram of 20 Indian accessionsshowed that accessions 105 from AP and 15 from Orissa stand out of the main cluster. Many sub-clusters were present in the main cluster, forming groups of accessions from either the same states or geographically contiguous states.

Discussion

Molecular marker technology has evolved into a very powerful tool in plant biology for marker-assisted breeding, DNA finger printing, derivation of linkage maps and analysis of genetic relationship and molecular diversity. Almost all the crop plants contain a high proportion of retroelements even in those with known low genetic base. Since their retrotranspositional movement is mediated viaan RNA intermediate by “copy-and-paste mechanism”, the elementsinvariablyremain

integrated at the original locations; making retrotransposon-based markers as stable markers.A variety of molecular marker techniques have evolved over the years, however, retotransposon-based markers offer an advantage due to their presence across the genome, thereby able to cover broad spread of the genome. Different retrotransposon-based markers and their relative advantages have been well discussed (Kalendar et al. 2011). The retroelements markers have been utilized in several plant species such as barley (Kalendaret al. 2003; 2011), sweet potato (Berenyi2002), rice (Brancoet al.2007), cucurbits (Lon and Chen 2007),*Vicia* species (Sanzet al. 2007), pea (Pearce et al. 2000; Smykalet al. 2006, Jing et al. 2012), sunflower (Vukichet al.2009), flex (Smykalet al.2011), Japanese pear (Kim et al 2012), wheat (Nasriet al, 2013), potato (Sharma and Nandineni2014), and grape vines (Villanoet al. 2014).

The SSAP technique is similar to AFLP but uses only one restriction site adaptor-primer and the other primer is derived from retroelement LTR sequence (Vosset al.1995; Waghet al., 1997). The variations in length of bands obtained depend on the distance between restriction sites and the LTRs. REMAP on the other hands requires no enzyme digestion but relies on the amplified polymorphism between an LTR and a microsatellite. The appearance of new bands would arise from the insertion of new *Panzee* copies into the genome within amplifiable distance of a microsatellite, or from an increase in the repeat number of a microsatellite to a point sufficient for amplification.

In the present work, comparison between SSAP and REMAP revealed that SSAP markers exhibit higher levels of genetic polymorphism between accessions as compared to REMAP. Cluster analysis indicated certain congruency in the basic pattern of grouping between REMAP and SSAP markers based on their phenetic relationships. However, some differences in the groupings were observed in the placement of certain accessions in both the phenograms as well as principal correspondence plots. This was attributed to the fact that various markers target different regions of the genome and explores genetic variations differently resulting in estimates of genetic diversity differing considerably.A number of marker systems such asAFLP (Panguluriet al. 2006), DArT (Yang et al. 2006; 2011), SSR (Bohraet al. 2011; Duttaet al. 2011), and SNPs(Saxenaet al. 2012; 2014) have been used for analysis of genetic diversity in pigeonpea. A comparative account of these marker techniques has been discussed by Saxenaet al.(2014). Using SNP profiling over 75% polymorphism among pigeonpeaaccessions including both wild relatives and domesticated ones could be detected (Kassaet al.2012;Saxenaet al. 2014). We show that retrotransposon-based markers, especially SSAP,

are able to detect much higher degree of polymorphism in pigeonpea. Only disadvantage of the technique is that it utilizes radioactivity. However, it is possible to use nonradioactive labels in future and technique is also amenable to automation.

Both the SSAP and REMAP markers are efficient in identifying species but SSAP is more useful in distinguishing species and accessions with close genetic background.Accessions belonging to different geographical regions of the world are clustered together indicating that there is no relationship between accession and geographic origin at a worldwide level, which is in accordance with earlier studies. Our results are in by and large in conformity with the SNP polymorphism data (Saxena et al. 2014) except two accessions (one from Australia and the other from India), which show divergence from the other accessions.This possibly indicatesto the fact that the present worldwide distribution of pigeonpea is due to its relatively recent dispersal from its site of origin, which is a common feature in legumes.

These techniques did reveal an association between Indian accessions collected from same geographical region. It can be conjectured that the accessions growing in these areas have undergone some common evolution at the local level. This analysis substantiates the fact that there is high degree of polymorphism present in the existent pigeonpeagermplasm. The high level of genetic variability may have arisen due to the retrotransposonsof pigeonpea that might have been mobilized during adaptation of the species to new environment, as they are get activated in various biotic and abiotic stresses (Grandbastien1998;Bui and Grandbastien 2012). Phylogenetic resolution between closely related germplasm accessions benefits from increasing the number of polymorphic markers scored, and several retrotransposons can be used together in SSAP and REMAP to achieve this purpose (Kalendar et al. 2003, 2011; Kalendar and Schulman, 2014) as each retrotransposons chosen for analysis may be distinct in their pattern of activity over evolutionary time. An active retrotranspson may bepolymorphic between individuals within species, while ancient and inactive elements will be more informative at the inter-specific higher taxonomic levels.

Acknowledgements

Technical help provided by Dr. MalathiLakshmikumaran, then at The Energy Research Institute, New Delhi, is gratefully acknowledged. Maneesha thanks University Grants Commission, New Delhi for a fellowship.

References

Bohra, A., Dubey, A., Saxena, R.K., Varma,P.R., Poornima, K.N.*et al*. 2011Analysis of BAC-end sequences (BESs) and development of BES-SSR markersfor genetic mapping and hybrid purity assessment in pigeonpea (*Cajanus*spp).*BMC Plant Biol*.**11**, 56.

Branco, C.J.S., Vieira, E.A., Malone, G., Kopp, M.M., Malone, E.*et al*. 2007 IRAP and REMAP assessments of genetic diversity in rice. *J. Appl. Genet*.**48**, 107-113.

Berenyi, M., Gichuki, T., Schmidt, J. and Burg, K. 2002 Ty1-copia retrotransposon-based S-SAP (sequence-specific amplified polymorphism) for genetic analysis of sweet potato.*Theor. Appl. Genet*.**105**, 862-869.

Bui,Q.T.and Grandbastien, M-A. 2012 LTR retrotransposons as controlling elementsof genome responseto stress?In M-A.Grandbastien and J.M. Casacuberta (eds.), Plant Transposable Elements,*Topics Current Genetics***24**, DOI 10.1007/978-3-642-31842-9_14.

Dutta, S., Kumawat, G.,Singh, B.P., Gupta, D.K., Singh, S., Dogra, V.*et al*. 2011 Development of genic-SSR markers by deeptranscriptome sequencing in pigeonpea[*Cajanuscajan* (L.) Millspaugh].*BMC Plant Biol*.**11**, 17.

Grandbastien, M-A. 1998 Activation of plant retrotransposons under stress conditions.*Trends Plant Sci*.**3**, 181-187.

Jing, R., Ambrose, M.A., Knox,M.R., Smykal,P., Hybl, M., Ramos, A., Caminero, C.*et al.*2012Genetic diversity in European *Pisum*germplasm collections. Theor. Appl. Genet.**125**,367–380

Kalender, R.,Flavell, A.J., Ellis,T.H.W., Sjakste, T., Moisy, C.and Schulman, A.H. 2011 Analysis of plant diversity with retrotransposon–based molecular markers. *Heredity***106**,520-530.

Kalendar,R.and Schulman, A.H. 2014 Transposon-based tagging:IRAP, REMAP and IPBS. In "Molecular Plant Taxonomy: Methods and Protocols" (Pascale Besse, ed.), *Methods in Molecular Biology*. Vol. **1115**, 233-255.

Kassa,M.T., Varma,P.R., Carrasquilla-Garcia,N., Sarma, B.K., Datta, S.*et al*. 2012Genetic patterns of domestication in pigeonpea (*Cajanuscajan* (L)Millsp) and wild *Cajanus* relatives. *PloSOne***7**, 6.

Kim, H., Terakami, S., Nishitani, C., Kurita, K., Kanamuri, H.*et al.* 2012Development of cultivar-specific DNA markers based on retrotransposon-based insertional polymorphism in Japanese pear. *Breeding Science***62**, 53-62.

Lall, I.P., Maneesha and Upadhyaya, K.C. 2002*Panzee*, a copia-like retrotransposon from the grain legume, pigeonpea (*Cajanuscajan* L.). *Mol. Genet. Genomics***267**, 271-280.

Lon, Q. and Chen, J. 2007 Ty1-copia retrotransposon-based SSAP marker development and its potential in the genetic study of cucurbits. Genome **50**,802-810.

van der Maesen, L.J.G. (1989) *Cajanuscajan* (L) Millsp. In van der Maesen, l.J.G. andSomaatmadja, S. (eds.) Plant Resources of South-East Asia No. 1 pulses. Pudoe/Prosea, Wageningen, Netherlands, pp. 39-42.

Nasri, S., Mandoulakani, B.A., Darvishzadeh, R. andBernousi, I. 2013Rerotransposoninsertional polymorphism in Iranian bread wheat cultivars and breeding lines revealed by IRAP and REMAP markers. *Biochem. Genet*.**51**, 927-943.

Nei, M. and Li, W.H. 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci.USA***76**, 5269-5273.

Panguluri, S.K., Janaiah, K., Govil, J.N., Kumar, P.A., and Sharma, P.C. 2006 AFLPfingerprinting in pigeonpea (*Cajanuscajan* (L) Millsp.) and its wild relatives.*Genetic Resources Crop Evolution***53**, 523–531.

Pearce, S.R., Knox, M., Ellis, T.H., Flavell, A.J. and Kumar, A. 2000 Pea *Ty1-copia* group retrotransposons: transpositional activity and use as markers to study genetic diversity in *Pisum*. *Mol. Gen. Genet.***263**, 898-907.

Rohlf,F.J. 1989 NTSYS-pc. Numerical Taxonomy and Multivariate Analysis System, Version 1.80, Exeter Publications, New York.

Sanz, A.M., Gonzalez, S.G., Syed, N.H., Suso, M.J., Saldan, C.C.and Flavell, A.J. 2007 Genetic diversity analysis in *Vicia* species using retrotransposon-based SSAP markers. *Mol. Genet.Genomics***278**, 433-441.

Saxena, R.K., Varma,P.R., Upadhyaya, H.D., Kumar, A, Carrasquilla-Garcia, N., *et al*. 2012 Large-scale development of cost-effective singlenucleotidepolymorphism marker assays for genetic mapping in pigeonpea andcomparative mapping in legumes. *DNA Res*.**19**, 449–461.

Sharma,V.and Nandinemi, M.R. 2014 Assessment of genetic diversity among Indian potato (*Solanumtuberosum* L) collection using microsatellite and retrotransposon based marker system. *Mol.Phylogenet.Evol*.**73**, 10-17.

Singh, N.K., Gupta, D.K., Jayaswal, P.K., Mahto, A.K. *et al*. 2012 The first draft of pigeonpea genome sequence. *J. Plant Biochem. Biotech.* **21**, 98-112.

Smykal, P. 2006 Development of an efficient retrotransposon–based finger printing method for rapid pea variety identification.*J.Appl.Genet*.**47**,221-230.

Smykal, P., Bacova-Korteszova, N., Kalendar, R., Corander, J., Schulman, A.H.and Pavelek, M. 2011 Genetic diversity of cultivated flax (*Linumusitatissimum* L.) germplasm assessed by retrotransposon-based markers. *Theor. Appl. Genet*.**122**, 1385-1392.

Varshney, R.K., Chen, W., Li, Y., Bharti, A.K., Saxena, R.K.*et al*.2012 Draft genome sequence of pigeonpea (*Cajanuscajan*), an orphan legume crop of resource-poor farmers. *Nature Biotech*.**30**,83-89.

Varshney, R.K., Mohan, S.M., Gaur, P.M., Gangarao, N.V.P.R., Pandey, M.K.*et al*. 2013 Achievements and prospects of genomics–assisted breeding in three legume crops of the semi-acid tropics. *Biotech.Adv*.**31**,1120-1134.

Villano, C., Carputo, D., Frusciante, L., Santoro,X.and Aversano, R. 2014 Use of SSR and retrotransposon-based markers to interpret the population structure of native grapevine from Southern Italy. *Mol.Biotechnol*.**56**,1011-1020.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van der Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., *et al*. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*. **23**, 4407-4414.

Vukich, M., Schulman, A.H., Giordani, T., Natali, L., Kalendar, R. andCavallini, A. 2009 Genetic variability in sunflower (*Helanthusannus* L.) and in the *Helianthus* genus as assessed by retrotransposon-based molecular markers. *Theor. Appl. Genet*.**119**, 1027-1038.

Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, B.B. and Powell, W. 1997 Genetic distribution of Bare-1-like retrotransposable elements in the barley Genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet*.**253**, 687-694.

Yang, S., Pang, W., Ash, G., Harps, J., Caring, J., Wenze, P., Hutter, E., Zong, X. and Killian, A. 2006 Low level of genetic diversity in cultivated pigeonpea compared to its wild relatives as revealed by diversity array technology.*Theor. Appl. Genet*.**113**,585-595.

Yang, S., Saxena, R.K., Kulwal, P.L., Ash, G.J., Dubey, A., *et al*. 2011 First geneticmap of pigeonpea based on diversity array technology (DArT) markers.*J. Genetics*.**90**,103–109.

Received 4 June 2016, in revised form 13 September 2016; accepted 21 November 2016
Unedited version published online: 25 November 2016

Table 1: Pigeonpea accessions used

Source Country	Number of Accessions	ICRISAT Accessions*
India	15	ICP11767**(7)**, ICP9826**(8)**, ICP938**(12)**,ICP4659**(15)**, ICP9665**(16)**, ICP234**(18)**, ICP893**(19)**, ICP7629**(20)**, ICP7817**(48)**,ICP8896**(49)**, ICP8904**(50)**, ICP8754**(56)**, ICP11825**(100)**, ICP10385**(105)**, PUSA84
ICRISAT	3	ICP11593**(2),** ICP9881**(3),** ICP15193**(57)**
Nigeria	3	ICP13573 **(33),** ICP15336 **(75)**, ICP15420**(80)**
Kenya	3	ICP13217**(23),**ICP9156**(103),**ICP13059**(104)**
Zambia	2	ICP11458**(37)**, ICP14268**(38)**
Malawi	1	ICP13396**(101)**
Tanzania	2	ICP12111**(39),** ICP12141**(40)**
Zaire	1	ICP15138**(77)**
Australia	3	ICP10897 **(21),** ICP10913**(38)**, ICP15763 (wild *Cajanussericeus*) **(60)**
Puerto Rico	1	ICP13884**(76)**
Philippines	1	ICP10880**(22)**
Trinidad & Tobago	1	ICP13804**(99)**
Thailand	1	ICP14567**(108)**
Italy	1	IPC14301**(66)**
Barbados	1	ICP134847**(67)**
Taiwan	1	ICP10875**(25)**
Jamaica	1	ICP14132**(98)**
Venezuela	2	ICP14084**(32)**, ICP14101**(46)**
UK	1	ICP15134**(78)**
Myanmar	1	ICP15188**(6)**
Uganda	1	ICP15510**(41)**
Germany	1	ICP13578 **(106)**

*****Number in the brackets indicate the ones used in the Figures.

Table 2: Amplified bands in SSAP and REMAP

Primer combination	Total bands	% Monomorphic bands	% Polymorphic bands	Total accessions
SSAP (All accessions)				
K2 (LTR)xM-CTG	124	4.84	95.16	39
K2 (LTR)xM-CTA	59	15.25	84.75	39
K2 (LTR)xM-CAC	114	7.89	92.11	39
Average	99	8.08	91.92	39
Total	**297**	**8.08**	**91.92**	**39**
SSAP (Indian Accessions)				
K2 (LTR)xM-CTG	101	12.87	87.13	19
K2 (LTR)xM-CTA	55	29.09	70.91	19
K2 (LTR)xM-CAC	98	29.59	70.41	19
Average	85	22.83	77.17	19
Total	**254**	**22.83**	**77.17**	**19**
REMAP (All accessions)				
K1MS1 $(GA)_9C$	15	26.67	73.33	39
K1MS7 $GT(CAC)_7$	11	45.45	54.55	39
K2MS1 $(GA)_9C$	18	11.11	88.89	39
K2MS4 $(GTG)_7C$	12	0	100	39
K3MS1 $(GA)_9C$	21	14.29	85.71	39
K3MS4 $(GTG)_7C$	15	13.33	86.67	39
K3MS5 $(CA)_{10}G$	21	4.76	95.24	39
Average	16	15.04	84.96	39
Total	**113**	**15.04**	**84.96**	**39**
REMAP (Indian accessions)				
K1MS1 $(GA)_9C$	15	33.33	66.67	19
K1MS7 $GT(CAC)_7$	10	60	40	19
K2MS1 $(GA)_9C$	17	23.53	76.47	19
K2MS4$(GTG)_7C$	12	0	100	19
K3MS1 $(GA)_9C$	21	28.57	71.43	19
K3MS4 $(GTG)_7C$	15	26.67	73.33	19
K3MS5 $(CA)_{10}G$	21	19.05	80.95	19
Average	16	26.13	73.87	19
Total	**111**	**26.13**	**73.87**	**19**

Figure legends

Figure 1. A typical representation of S-SAP fingerprint profile obtained with the outward-going LTR (K2) primer together with M-CTG primer showing polymorphic and monomorphic bands The accession 60 from Australia, a different species of *Cajanus* shows entirely different profile than the other accessions. Different pigeonpea accessions are: I- India, Ic- ICRISAT, N-Nigeria, K-Kenya, Za- Zambia, M-Malawi, T- Tanzania, Z- Zaire, A- Australia, Pr- Puerto Rico, P-Philippines, Tt- Trinidad & Tobago, Th- Thailand, It- Italy, B- Barbados, Ta- Taiwan, J-Jamaica, V- Venezuela, UK- United kingdom, My- Myanmar, U- Uganda, and G- Germany.

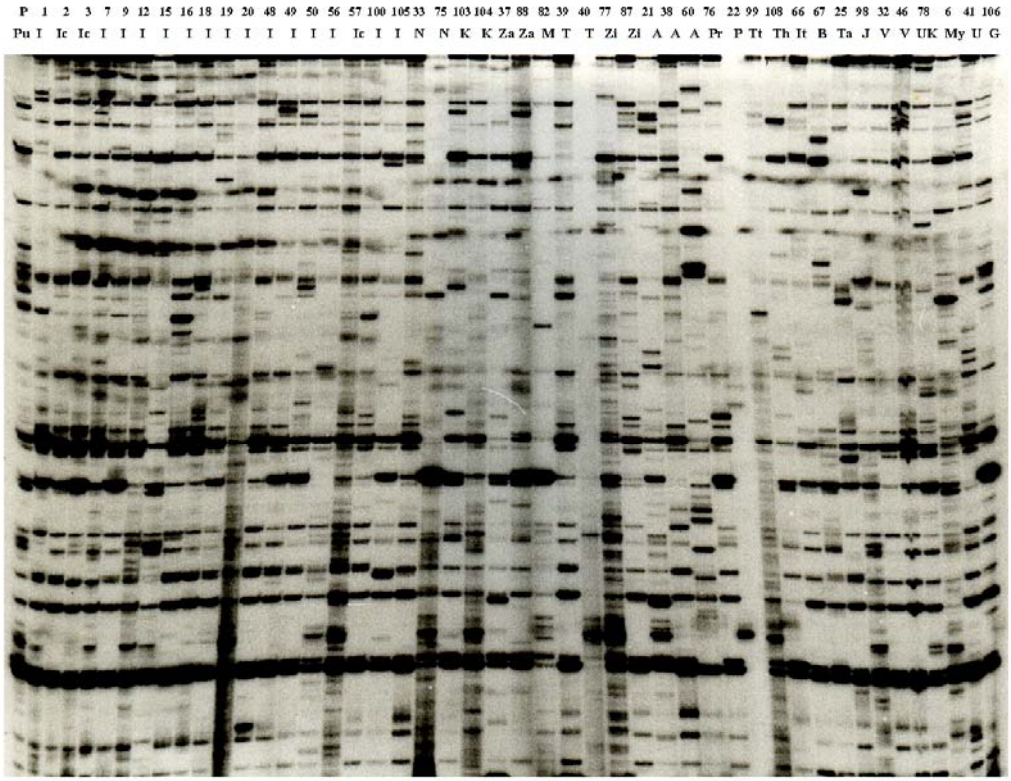


Figure 2. A typical representation of REMAP banding profile of different pigeonpea accessions generated by REMAP with K3(LTR) + $(CA)_{10}G$, outward LTR primer along with microsatellite primer. The PCR amplification products were separated on 5% polyacrylamide gel and stained with ethidium bromide. The gel is shown as a negative image. M- 100-bp DNA marker.Different pigeonpea accessions are: I- India, Ic- ICRISAT, N-Nigeria, K-Kenya, Za- Zambia, M-Malawi, T- Tanzania, Z- Zaire, A- Australia.

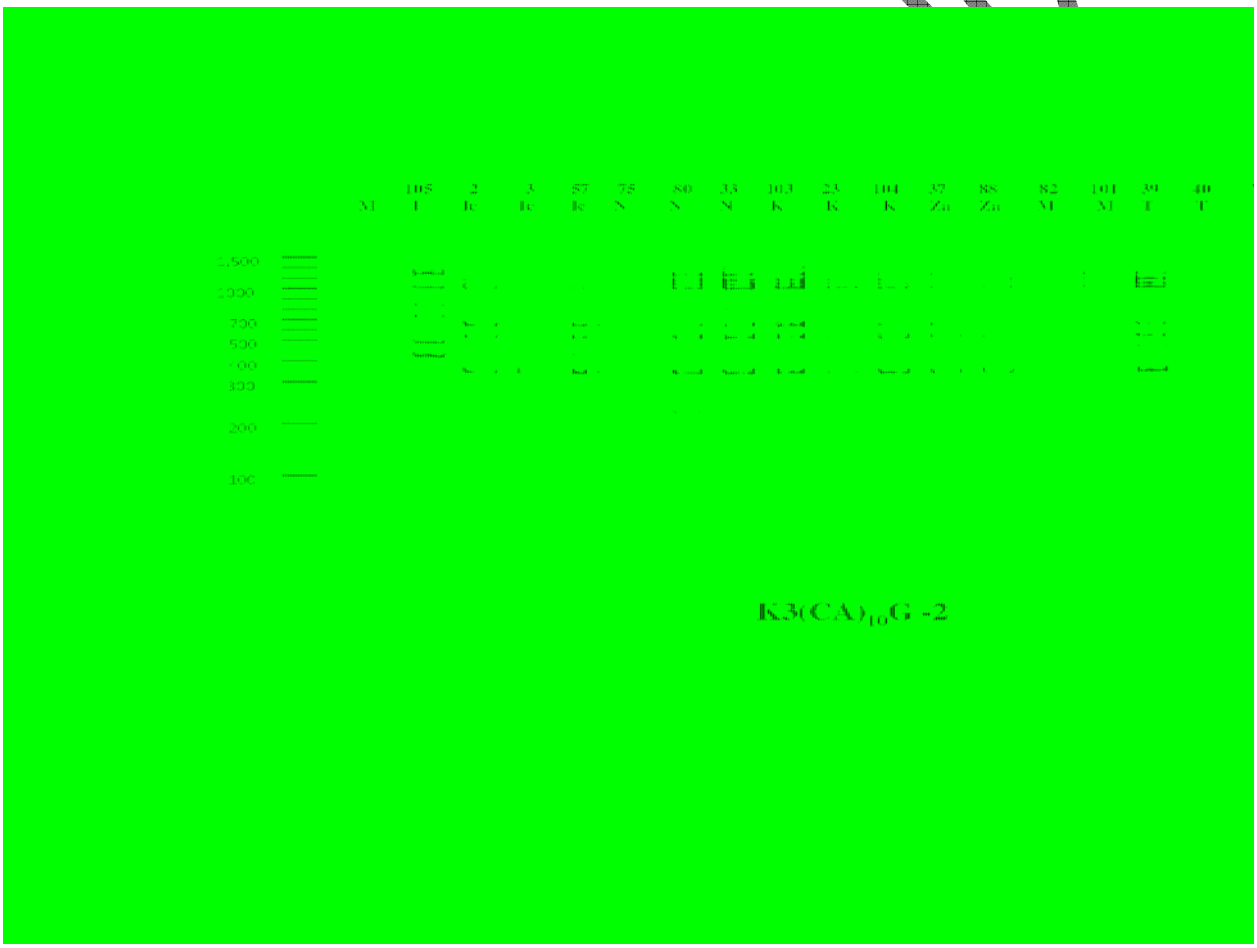


Figure 3. Inter-simple sequence repeat (ISSR) with primer $(GA)_9C$

(a) Showing basic strategy of ISSR with unanchored or anchored primes.

(b) ISSR pattern to show that no bands are shared with REMAP, indicating that all the bands span intervening domains between LTRs and microsatellites in the genome. The gel is shown in a negative image.

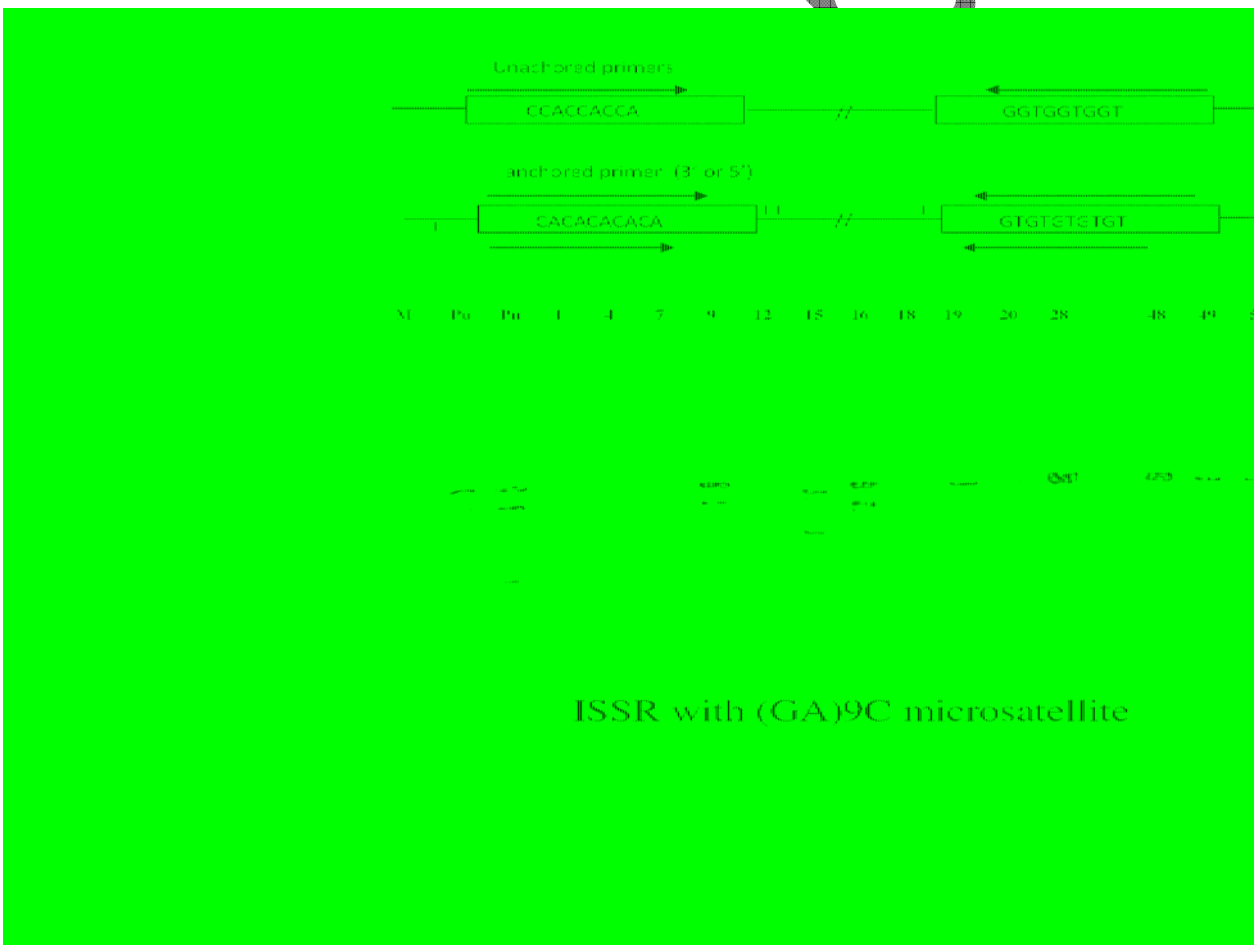


Figure 4.UPGMA pheneticdendrogram based on Jaccard's coefficient among pigeonpea accessions derived from S-SAP data. The scale represents the similarity coefficient values.

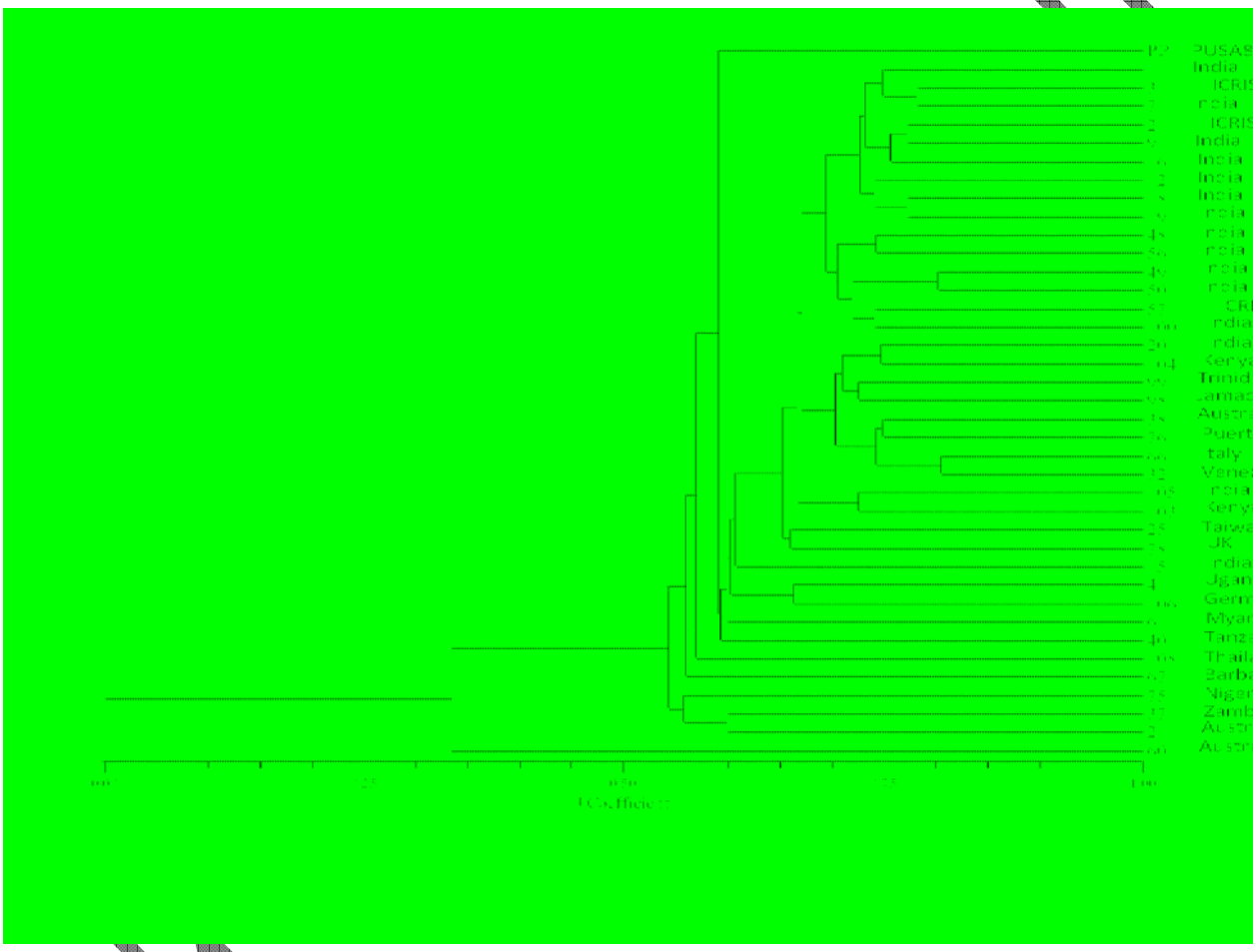


Figure 5. UPGMA pheneticdendrogram based on Jaccard's coefficient derived from REMAP. The scale represents the similarity coefficient values.

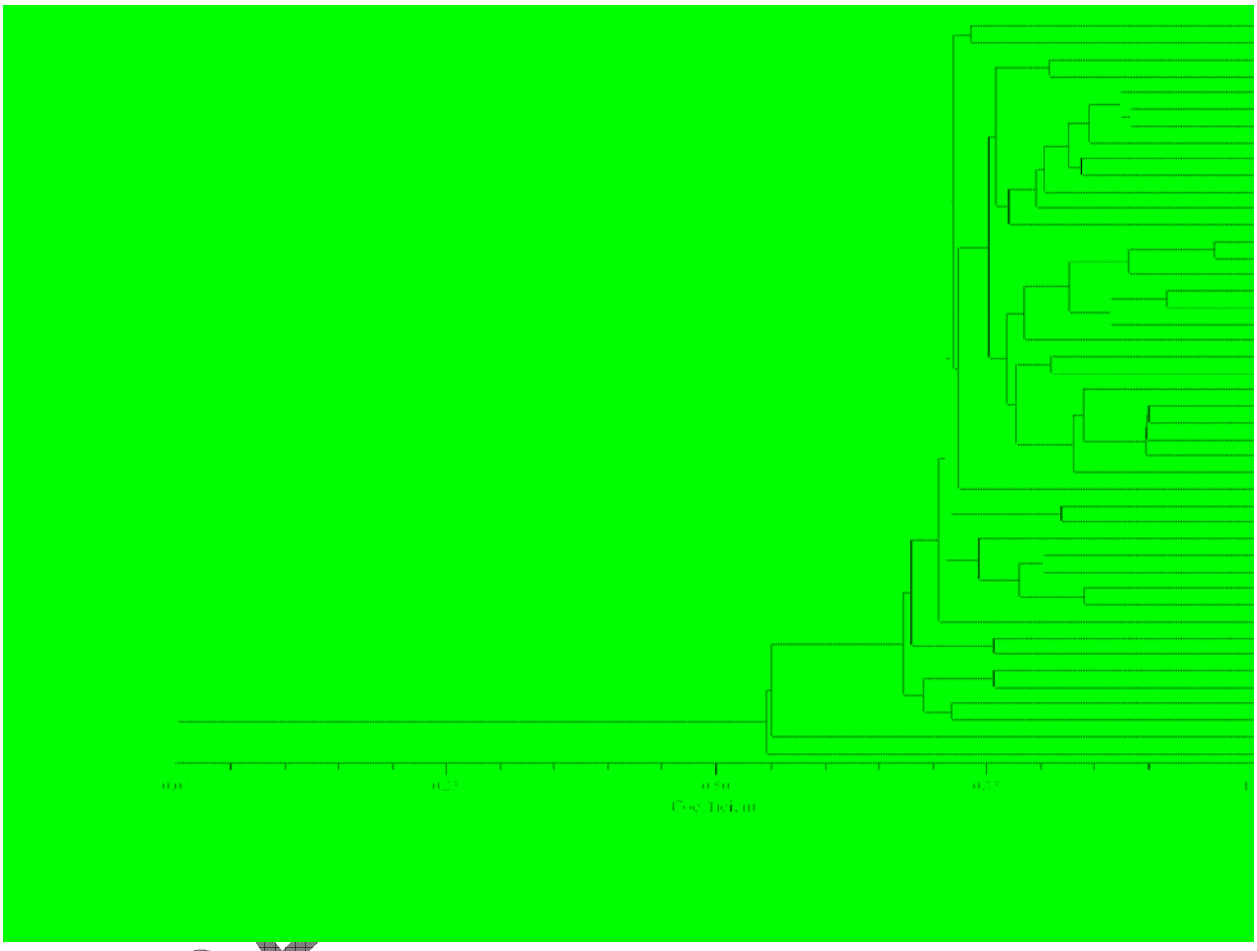
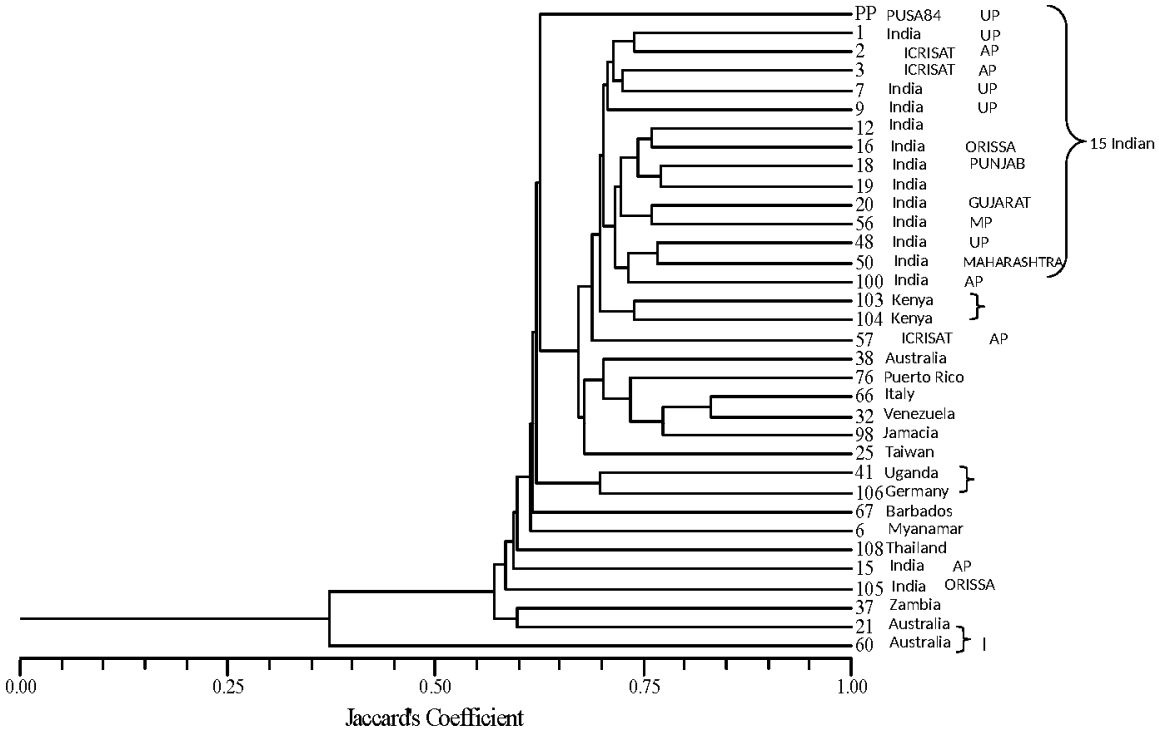


Figure 6: UPGMA pheneticdendrogram based on Jaccard's coefficient among various pigeonpea accessions derived from SSAP and REMAP pooled data. The scale represents the similarity coefficient values.


PP PUSA84 UP
1 India UP
2 ICRISAT AP
3 ICRISAT AP
7 India UP
9 India UP
12 India
16 India ORISSA
18 India PUNJAB
19 India
20 India GUJARAT
56 India MP
48 India UP
50 India MAHARASHTRA
100 India AP
15 Indian
103 Kenya
104 Kenya
57 ICRISAT AP
38 Australia
76 Puerto Rico
66 Italy
32 Venezuela
98 Jamacia
25 Taiwan
41 Uganda
106 Germany
67 Barbados
6 Myanamar
108 Thailand
15 India AP
105 India ORISSA
37 Zambia
21 Australia
60 Australia
I
0.00
0.25
0.50
0.75
1.00
Jaccard's Coefficient