Pigeonpea composite collection and identification of germplasm for use in crop improvement programmes

H. D. Upadhyaya*, K. N. Reddy, Shivali Sharma, R. K. Varshney, R. Bhattacharjee, Sube Singh and C. L. L. Gowda

Gene Bank, International Crops Research Institute for the Semi Arid Tropics (ICRISAT), Patancheru, Hyderabad, PO 502324, Andhra Pradesh, India

Abstract

Pigeonpea (*Cajanus cajan* (L.) Millsp. is one of the most important legume crops as major source for proteins, minerals and vitamins, in addition to its multiple uses as food, feed, fuel, soil enricher, or soil binder, and in fencing, roofing and basket making. ICRISAT's genebank conserves 13,632 accessions of pigeonpea. The extensive use of few parents in crop improvement is contrary to the purpose of collecting a large number of germplasm accessions and has resulted in a narrow base of cultivars. ICRISAT, in collaboration with the Generation Challenge Program, has developed a composite collection of pigeonpea consisting of 1000 accessions representing the diversity of the entire germplasm collection. This included 146 accessions of mini core collection and other materials. Genotyping of the composite collection using 20 microsatellite or simple sequence repeat (SSR) markers separated wild and cultivated types in two broad groups. A reference set comprising 300 most diverse accessions has been selected based on SSR genotyping data. Phenotyping of the composite collection for 16 quantitative and 16 qualitative traits resulted in the identification of promising diverse accessions for the four important agronomic traits: early flowering (96 accessions), high number of pods (28), high 100-seed weight (88) and high seed yield/plant (49). These accessions hold potential for their utilization in pigeonpea breeding programmes to develop improved cultivars with a broad genetic base. Pigeonpea germplasm has provided sources of resistance to abiotic and biotic stresses and cytoplasmic-male sterility for utilization in breeding programmes.

Keywords: characterization; composite set; cytoplasmic-male sterility; diversity; mini core collection; pigeonpea; reference set

Introduction

Pigeonpea (*Cajanus cajan* (L.) Millsp.), with its origination in India, is an important grain legume crop of the tropics and subtropics. It is an excellent source of protein, minerals and vitamins and has multiple uses as food, feed, fuel, soil enricher, or soil binder, and is also used in fencing, roofing and basket making. It also has wide applications in traditional medicine

(van der Maesen, 2006). Pigeonpea has wide adaptability to diverse climate conditions and is grown as a field and/or backyard crop in at least 82 countries (Nene and Sheila, 1990). FAO statistics are available for only 22 countries and, in these, pigeonpea as a field crop was grown on 4.86 million ha with a production of 4.1 million ton during 2008. India has the largest area under pigeonpea (3.73 m ha) followed by Myanmar (0.54 m ha), Kenya (0.20 m ha), Malawi (0.17 m ha), Uganda (0.09 m ha), Tanzania (0.07 m ha), Nepal (0.02 m ha) and Dominican Republic (0.02 m ha) (Food and Agriculture Organization, 2008). The productivity of pigeonpea (844 kg/ha) is low. The major constraints

to the productivity are biotic stresses such as pod borer, pod fly, fusarium wilt, sterility mosaic disease and abiotic stresses such as drought, salinity and frost/cold.

Germplasm assembly

ICRISAT, having the global responsibility as world repository for the germplasm of its mandate crops, made concerted efforts to acquire pigeonpea germplasm that was assembled at different national and international institutes, universities, National Agricultural Research System (NARS), etc. This was followed by systematic collecting missions in priority areas, which resulted in assembly of 13,632 accessions from 74 countries in the ICRISAT genebank. This is the world's largest single collection for pigeonpea. The biological status of the collection indicates the presence of landraces (8,215 accessions), breeding materials (4,795), advanced cultivars (67) and wild relatives (555 accessions, 67 species) in the collection (Upadhyaya and Gowda, 2009).

Characterization and evaluation

Any germplasm collection is of little value unless it is characterized, evaluated and documented properly for its enhanced utilization in crop improvement. All the cultivated accessions of pigeonpea have therefore been characterized and evaluated at the ICRISAT research farm, Patancheru (17.53°N, 78.27°E, 545 m.a.s.l.), India, for 18 qualitative and 16 quantitative characters following the 'Descriptors for Pigeonpea' (IBPGR and ICRISAT, 1993). A multi-disciplinary approach is followed at ICRISAT's genebank for characterization, including screening for diseases and pests by other disciplines, and a pigeonpea germplasm characterization database is maintained for use.

Patterns of diversity

The assembled germplasm represents a wide range of diversity for different morpho-agronomic characters. In an earlier report, Upadhyaya *et al.* (2005) studied the geographical pattern of diversity among 11,402 accessions from 54 countries for 14 qualitative and 12 quantitative traits. The accessions were grouped based on geographical proximity and similarity of the climate (Upadhyaya *et al.*, 2005). The range of variation for quantitative traits was maximum for the ASIA 4 region (south India, Maldives and Sri Lanka) and minimum for germplasm accessions from Europe and Oceania. The region ASIA 4 encompasses the primary centre of diversity of pigeonpea and, therefore, the high variation observed

in the germplasm from that region was not surprising. The accessions from Africa were of longer duration, tall and produced large seeds. Accessions from Oceania were conspicuous by short growth duration, short height, few branches, small seeds and low seed yield. Upadhyaya *et al.* (2005) used Shannon–Weaver diversity index (H′) (Shannon and Weaver, 1949) as a measure of diversity. A low H′ indicates an extremely unbalanced frequency classes for an individual trait and a lack of genetic diversity. The accessions from ASIA 6 (Indonesia, Philippines and Thailand) had the highest pooled H′ for qualitative traits (0.349 ± 0.059) and accessions from Africa for quantitative traits (0.613 ± 0.006). The accessions from Africa showed highest pooled (qualitative and quantitative traits) H′ (0.464 ± 0.039), whereas those from Oceania had lowest pooled H′ (0.337 ± 0.037). The H′ values across the regions were highest for primary seed colour (0.657 ± 0.050) followed by flower streak pattern, seed protein content and shelling per cent, whereas it was lowest for flowering pattern (0.087 ± 0.026). A hierarchical cluster analysis conducted on the first three principal component (PC) scores (92.3% variation) resulted in three clusters (Fig. 1).

Germplasm utilization

Low use of germplasm

The increase in accession numbers in genebanks and lack of corresponding increase in their utilization by the crop improvement scientists indicate that the collections are not being utilized to their full potential (Marshall, 1989; Ferreira, 2005). With the global responsibility, ICRISAT's genebank has supplied over 69,000 samples of pigeonpea accessions to scientists in 110 countries, besides 83,941 samples to researchers within ICRISAT in last 20 years. However, pattern of demand and
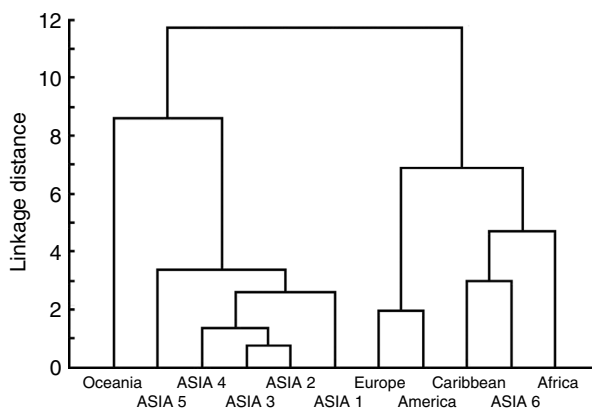

Fig. 1. Dendrogram of 11 regions in the entire pigeonpea germplasm based on the first three principal components.

consequent supply of pigeonpea germplasm has shown a greater demand for few specific germplasm accessions. India has a robust pigeonpea improvement programme for the last four decades and received 48,159 samples of 9917 accessions from ICRISAT. But the pedigree analyses of released cultivars indicate that T-1 and T-190 were the most frequently used parents (Kumar *et al.*, 2004). In a period of 35 years, one accession, ICP 7035, was supplied on demand for 309 times, seven accessions for more than 200 times and 29 accessions for more than 100 times. At the same time, 2959 accessions have not been requested at all and 7093 accessions have been supplied less than five times till now. This represents repeated demand and perhaps uses of a limited number of germplasm accessions in crop improvement programmes. At ICRISAT, a total of 1256 parents were involved in developing 527 advanced breeding lines between 1996 and 2000. Of these, only 54 were germplasm lines (50 landraces and 4 were selections from wild relatives), representing 0.4% of 13,362 of total accessions available in the genebank. The remaining 1202 parental lines were breeding lines (ICRISAT Pigeonpea Line (ICPL)), of which 997 lines were developed at ICRISAT and 205 were selections from breeding lines of Indian and Australian NARS. Of the parents, the top ten breeding lines were used 484 times (an average of 48.4 times) compared to only 38 times the top ten germplasm lines (average 3.8). This is comparably favourable with chickpea in which 91 accessions (0.4% of total 20,267 accessions) and unfavourable with groundnut where 171 (1.1% of total 15,445 accessions) accessions have been used in developing breeding lines (Upadhyaya *et al.*, 2003, 2006a, 2010).

The extensive use of few genotypes in crop improvement is contrary to the purpose of collecting large number of germplasm accessions and could result in vulnerability of cultivars to new pests and diseases due to their narrow genetic base. In addition, the use of working collections mostly consisting of elite breeding lines and some improved breeding lines by most of the breeders results in re-circulation of the same germplasm, thus narrowing the genetic base of the developed cultivars. This is true both in the International Agricultural Research Centres (e.g. Consultative Group on International Agricultural Research (CGIAR) institutes) and in the NARS. Hence, the fears of epidemics similar to the southern corn leaf blight in the USA in 1970 and late blight of potato causing Irish famine in Europe in 1845–46 on account of large-scale cultivation of genetically uniform high-yielding cultivars with narrow genetic base loom large. Therefore, there is an immediate need to diversify the genetic base of high-yielding cultivars for sustainable production and productivity of crop plants. To achieve this goal, one of the approaches is to utilize diverse plant genetic resources conserved in various genebanks.

Reasons for low use of diverse germplasm

The restricted use of diverse pigeonpea germplasm in crop improvement programmes is mainly due to the following reasons (Upadhyaya and Gowda, 2009):

(1) Difficulty in evaluating large collections for traits of economic importance such as resistance to biotic and abiotic stresses, agronomic and quality traits, which often display high genotype × environment interactions and require multilocation and replicated evaluation. Non-availability of such information discourages breeders who seek accessions for targeted traits with reliable information.

(2) Restricted access to the germplasm collections due to limited seed quantities particularly of wild relatives and unadapted landraces that are difficult to regenerate.

(3) Inadequate linkage between genebanks and the users.

(4) Lack of robust, cost-effective tools and limited capacity of breeding programmes to facilitate the efficient utilization of exotic germplasm in crop improvement programmes.

(5) Relatively low emphasis on pigeonpea research and lack of resources compared with other crops.

(6) Role of non-additive genetic variation when diverse exotic germplasm is used.

(7) Limited exposure to available germplasm and re-circulation of the same genotypes already available with the researchers (Duvick, 1995).

Enhancing the use of germplasm

With the establishment of genebank at ICRISAT and pigeonpea as a mandate crop, the availability of over 13,000 diverse pigeonpea germplasm accessions provide a vast scope to the researchers. However, many agronomic traits of pigeonpea show considerable genotype × environment interaction, necessitating a replicated multilocation evaluation to identify the germplasm with beneficial traits for use in crop improvement programmes. However, replicated multilocation evaluation of such a large collection of germplasm is resources demanding and labour intensive. Therefore, there is a need to find ways to diversify the genetic base of cultivars and enhance the utilization of assembled germplasm by following approaches:

(1) Developing core (10% of the entire collection) (Frankel, 1984; Brown, 1989) and mini core (10% of the core or 1% of the entire collection) (Upadhyaya and Ortiz, 2001) collections.

(2) Multilocation evaluation of core/mini core collections.
(3) Ensuring the availability of seed of all the accessions.
(4) Providing access to characterization and evaluation data for important traits.
(5) Information on diversity of germplasm.
(6) Identifying gaps in collections and exploration to fill in the gaps in collection before this diversity is lost.
(7) Developing genepools for important economic traits.
(8) Organizing field days facilitating the direct selection and access to material.

Generation challenge programme

The Generation Challenge Programme (GCP, www. generationcp.org) is a programme of the Consultative Group on International Agricultural Research (CGIAR), and is an international, multisectoral and interdisciplinary collaboration in the plant sciences which links the basic science with applied research. GCP aims to create a public platform that will utilize genomics and comparative biology to explore and exploit genetic diversity assembled in the germplasm collections, with special focus on drought tolerance. One of the important objectives of GCP is extensive genetic characterization of germplasm collections held by the participating institutions using molecular markers. The phenotyping of the germplasm collections for biotic and abiotic stresses would facilitate identifying the genes that can be utilized to develop cultivars tolerant to these stresses and thus would increase the efficiency, speed and scope of crop improvement programmes.

Developing a pigeonpea composite collection

To enhance utilization of germplasm in crop improvement programmes, scientists at ICRISAT have developed core (1290 accessions) and mini core collection (146 accessions) of pigeonpea representing diversity of the entire collection (Upadhyaya *et al.*, 2006b). Due to its greatly reduced size, a mini core collection provides an easy access to the scientists in crop improvement, which helps them evaluate it across multiple locations easily and economically for target traits to identify promising germplasm accessions for further use.

As part of GCP, to create a public platform that use molecular methods to unlock genetic diversity and put it to use in better crops for the world's poorest farmers, a global composite collection comprising 1000 accessions was developed (Table 1). The rationale for developing this composite collection was to capture the global diversity available in the ICRISAT genebank and other

materials such as released cultivars, reported sources of resistance to various biotic and abiotic stresses and wild species and relatives. The composite collection includes accessions from the mini core collection (146), mini core comparator (146), additional representative accessions from 79 clusters of core collection (236), control cultivars (4), 63 accessions of 7 wild species, diverse sources for biotic (77) and abiotic (16) stresses, promising germplasm accessions (59), released cultivars (16) and accessions with distinct morpho-agronomic traits (237) (Upadhyaya *et al.*, 2008) (Table 1) for genotypic characterization using microsatellites or simple sequence repeat (SSR). The biological status of the composite collection indicated 54.9% landraces, 35.2% breeding materials, 3.2% advanced cultivars, 6.3% wild relatives and 0.4% others. In total, 94% cultivated and 6% wild accessions were present in the composite collection. The composite collection captured accessions showing diversity for morpho-agronomic traits; for example, the flowering pattern indicated a maximum of 82.2% indeterminate (NDT), 9.1% determinate (DT) and 2.4% semi-determinate accessions (SDT) (6.3% accessions have no information). Geographically, 58.1% accessions were from Asian countries, 12.9% from African countries, 8.7% from America, 3.4% from Oceania, 0.9% from Europe, 15.6% from ICRISAT and 0.4% accessions had no information on country of origin. Overall, the composite collection included accessions from 54 countries and those developed/identified at ICRISAT. Four accessions had no information on center of origin. In terms of representation from different countries, 502 accessions were from India, 37 from Kenya, 33 from Australia, 23 from Nigeria, 22 from Tanzania and 20 from Venezuela. All other countries contributed < 20 accessions. One hundred fifty-six accessions were from ICRISAT. All the accessions in the composite collection are FAO designated and are held in trust, capturing a wide spectrum of genetic diversity present in the entire pigeonpea collection.

Molecular characterization of the composite collection

The composite collection was planted in Alfisols during 2005 rainy season at ICRISAT. Accessions were planted in one row of 9 m length with a spacing of 75 cm between rows and 25 cm between plants within a row. In recognition of the often-cross pollinated nature of the crop (Saxena *et al.*, 1990), leaf samples were collected from 12 representative plants/accession to extract DNA for molecular characterization. DNA was extracted from these randomly selected 12 plants/accession following a high-throughput procedure (Heckenberger *et al.*, 2002; Mace *et al.*, 2003) and pooled together with an objective

Table 1. Composition of pigeonpea composite collection developed at ICRISAT genebank

Stress/trait	Trait wise accessions	Total accessions
Cultivated accessions		
Mini core and its comparator	292	292
Core cluster representatives	236	236
Control cultivars	4	4
Biotic and abiotic stresses		
Pod borer and pod fly	Tolerant: 29; susceptible: 2	31
Sterility mosaic disease	Resistant: 16	16
Wilt	Resistant: 6; susceptible: 1	7
Alternaria blight	Resistant: 7	7
Phytophthora blight	Resistant: 5	5
Stem canker and root rot	Resistant: 5	5
Nematodes	Promising: 6	6
Drought tolerant	Tolerant: 7	7
Water-logging tolerant	Tolerant: 3	3
Salinity tolerance	Tolerant: 4; susceptible: 2	6
Promising germplasm accessions		
Nodulation	High nodulating: 2; non-nodulating: 3	5
Photoperiod response	Insensitive: 4; sensitive: 2	6
Suitable for agroforestry	7	7
Suitable for forage	6	6
Vegetable type	7	7
Seed protein content ($>$25%)	High ($>$25%): 20; low ($<$16): 8	28
Released cultivars	16	16
Growth habit	NDT: 6; DT: 5; SDT: 4	15
Plant height	Tall ($>$280 cm): 7; short ($<$75 cm): 8	15
Stem variants	10	10
Leaf variants	16	16
Selections for branches	32	32
Early flowering and maturity	14	14
Late flowering and maturity	6	6
Pod bearing length (cm)	High ($>$100 cm): 10; low ($<$5 cm): 6	16
Number of recemes/plant	High ($>$600): 7; low ($<$10): 6	13
Number of pods/plant	High ($>$1200): 6; low ($<$15): 3	9
Pod length ($>$11 cm)	Large ($>$11 cm): 8; small ($<$3.2 cm): 6	14
Seeds per pod (no.)	More ($>$6): 9; few $<$3): 6	15
Seed size (g)	Large ($>$21 g): 10; small ($<$4 g): 5	15
Shelling (%)	High (0.80): 8; low ($<$14): 8	16
Harvest index (%)	High $>$50: 9; low ($<$1.2): 3	12
Seed yield (g)	High ($>$2700 g): 10; low ($<$2 g): 5	15
Seed shape and colours	Oval shape: 2; black colour: 1; ivory colour: 1	4
Wild relatives	63 Species of 7 genera	63

to capture within-accession variation. This composite collection was genotyped using 20 SSR markers (Table 2) to study genetic diversity and population structure. Selected plants from each accession were harvested separately and equal quantity of seeds from each plant was bulked to reconstitute the accession.

Selection of SSR markers

At the time of initiation of this study in 2006, a limited number (164) of SSR markers were available in pigeonpea ($2n = 2x = 22$). Moreover, at that time, no genetic mapping populations or genetic maps were available. Therefore, in a preliminary experiment, 15 highly diverse accessions (based on morpho-agronomic traits) were screened using the then available 164 SSR markers at ICRISAT. A total of 33 SSR markers detected good quality (single peak) polymorphism between at least two of the examined genotypes. Since 12 plants from each of 1000 accessions were planned to genotype as we expect heterogeneity in the pigeonpea samples, it is also important to select the SSR markers, which can detect the interpretable heterogeneity in the sample. Therefore, a series of the artificial pools having different proportions of two genotypes, which showed polymorphism with a given SSR marker, were screened with the corresponding polymorphic SSR markers. As a result, a given SSR marker yielded two peaks (alleles) in different pools according to the compositions of the corresponding individual

Table 2. Features of SSR markers and molecular diversity in the pigeonpea composite collection (952 accessions)

SSR marker name	Repeat unit	Quality index	Allele size range (bp)	Major allele (size in bp)	Major allele frequency	Total number of alleles	Rare alleles (<0.01)	Common alleles	Gene diversity	Hetero- zygosity	Polymorphic Information Content (PIC) value
CCB1[a]	(CA)10	0.26	192–250	200	0.66	19	14	5	0.5	0.21	0.45
CCB10[a]	(CA)15	0.32	218–246	238	0.47	14	8	6	0.65	0.34	0.59
CCB7[a]	(CT)16	0.24	135–193	157	0.67	14	9	5	0.51	0.22	0.46
CCB8[a]	(CT)30	0.27	109–161	127	0.72	24	16	8	0.48	0.22	0.47
CCB9[a]	(CT)22	0.46	140–174	162	0.63	17	12	5	0.56	0.15	0.52
CCttc008[b]	(AGA)5	0.22	226–240	236	0.96	6	3	3	0.08	0.04	0.08
CCac035[b]	(AC)7	0.14	244–314	248	0.65	7	4	3	0.49	0.18	0.41
CCac036[b]	(CATA)3TA(TG)6	0.41	200–230	218	0.39	14	8	6	0.73	0.21	0.68
CCttc031[b]	(AAG)13	0.31	175–208	178	0.81	11	7	4	0.34	0.31	0.32
CCttc033[b]	(CTT)8	0.18	201–213	207	0.64	5	1	4	0.5	0.16	0.43
CCtc007[b]	(TC)8	0.26	229–247	245	0.92	7	4	3	0.15	0.06	0.15
CCttc003[b]	(GAA)5G(GAA)5	0.42	156–186	177	0.58	10	5	5	0.55	0.18	0.48
CCac009[b]	(TG)(TC)2(TG)7	0.2	241–253	249	0.9	5	3	2	0.19	0.07	0.17
CCttc005[b]	(GAA)6	0.23	192–210	204	0.94	8	5	3	0.11	0.02	0.11
CCac021[b]	(AC)6AAG(CTAA)3	0.3	197–217	215	0.82	6	1	5	0.32	0.15	0.31
CCggt001[c]	(GGT)4	0.11	208–211	208	1	2	1	1	0	0	0
CCtc002[c]	(CT)6TT(CT)2	0.23	141–156	150	0.94	6	3	3	0.12	0.05	0.12
CCttat001[c]	(TTAT)4	0.4	212–248	220	0.67	10	4	6	0.53	0.59	0.51
CCtta007[b]	(TTA)4	0.17	201–222	213	0.96	6	4	2	0.07	0	0.07
ICPM133[d]	(AT)6G(TAA)3	0.17	219–245	221	0.97	6	3	3	0.06	0.04	0.06
Mean					0.76	10	6	4	0.35	0.16	0.32
Total						197	115	82			
Minimum					0.39	2	1	1	0	0	0
Maximum					1	24	16	8	0.73	0.59	0.68

Source of primer sequences: [a]*Burns et al., 2001;* [b]*Odeny et al., 2009;* [c]*Odeny et al., 2007;* [d](ICPM133 F-CCAATCCTGGCAGTTTCT R-GCGGGCTTCATGACAACTT).

DNAs. On the basis of these results, 20 SSR markers (Table 2) were selected which had highly significant coefficient of correlations ($r^2 > 0.9$) for genotyping the pigeonpea composite collection. As no genetic map was available at the time of undertaking this study, SSR markers could not be selected based on their distribution on genetic map. Even today, when we have developed a framework genetic map, because of very limited level of polymorphism ($<5\%$) in parents of the mapping population, only 3 of 20 selected SSR markers have been integrated into genetic map.

SSR genotyping

Polymerase chain reaction (PCR) conditions for all 20 SSR markers were optimized following Taguchi method (Taguchi, 1986) as described in Cobb and Clarkson (1994). Fluorescent-based multiplex genotyping system was used to generate five multiplexes of four markers each. Capillary electrophoresis with an automated system (ABI 3700) was used to separate the amplified PCR products. Genotyper 3.7 software was used to determine the initial called allele sizes. The raw dataset was then analyzed using the Allelobin program (http://www.icrisat.org/gt-bt/biometrics.htm) developed at ICRISAT, which is based on the least squares algorithm of Idury and Cardon (1997). All the markers produced allele sizes expected on the basis of the repeat motifs for each of the SSR markers.

Data analysis

All raw allele calls were converted into best binned allele size based on the repeat units of the SSRs. A total of 20,000 data points (20 SSRs × 1000 accessions) were checked for quality and 48 accessions with high missing values were excluded for final data analysis. A total of 19,040 data points using 20 SSRs and 952 accessions showing less than 3% missing value were used from statistical analysis using Power Marker V3.0 (Liu and Muse, 2005; http://www.powermarker.net) for estimating basic statistics (Table 2). A neighbour-joining tree was constructed based on distance matrix using DARwin 5.0 (Perrier et al., 2003) for depicting the genetic structure of the composite collection.

Allelic diversity in composite collection

Analysis of 20 SSR markers data on 952 accessions detected 197 alleles, of which 115 were rare and 82 were common alleles (Table 2). Gene diversity varied

from 0.002 to 0.726. The group-specific 60 unique alleles were detected in 45 wild accessions and 64 unique alleles in 907 cultivated accessions. In pigeonpea, growth habit has been classified into DT, NDT and SDT. NDT type had 37 unique alleles, whereas DT had only one and SDT had no unique allele. Geographically, 32 unique alleles were found in ASIA 4 (Southern Indian provinces, Maldives and Sri Lanka), 7 in ASIA 6 (Indonesia, Philippines and Thailand), 5 in ASIA 3 (Central Indian provinces) and 4 in ASIA 1 (North western Indian provinces, Iran and Pakistan). Only two alleles in Africa differentiated them from other regions. Wild and cultivated types shared 73, DT and NDT shared 10, DT and wild shared 4 and the NDT and wild shared 20 alleles. ASIA 1 shared 4 alleles with ASIA 3 and 3 alleles with ASIA 4. ASIA 4 shared 6 alleles with ASIA 3 and 5 with ASIA 6. Wild types as a group were genetically more diverse than cultivated types. NDT types were more diverse than the other two groups in flowering pattern (DT and SDT) (Upadhyaya et al., 2008). A neighbour joining tree, based on the distance matrix using SSR data, grouped the cultivated and wild accessions into separate clusters (Fig. 2).

Identification of reference set

On the basis of genotypic data of composite collection, a reference set of 300 diverse accessions was chosen from the composite collection using 'max length sub tree' option of DARwin5.0, which creates the subset of units minimizing the redundancy between units and limiting the loss of diversity. This reference set captured 95% (187 alleles) of the 197 alleles found in the composite collection. Another reference set based on 16 qualitative traits captured 87% alleles of the composite collection (Upadhyaya et al., 2008).

Phenotyping the composite collection

The pigeonpea composite set was phenotyped by evaluating in an augmented design with four control cultivars (ICP 6971, ICP 7221, ICP 8863 and ICP 11 543) during 2006 rainy season at ICRISAT farm, Patancheru, AP, India (17.53°N latitude, 78.27°E longitude and 545 m.a.s.l.) in vertisols. The pooled seed lot harvested from 12 selected plants/accession from which leaf samples were collected for DNA extraction was used for planting. Accessions were sown in one row of 9 m long, with a spacing of 50 cm between plants and 75 cm between rows. The crop was fertilized with 20 kg N and 40 kg P_2O_5 per hectare as basal dose, and managed by recommended cultural and plant protection practices,
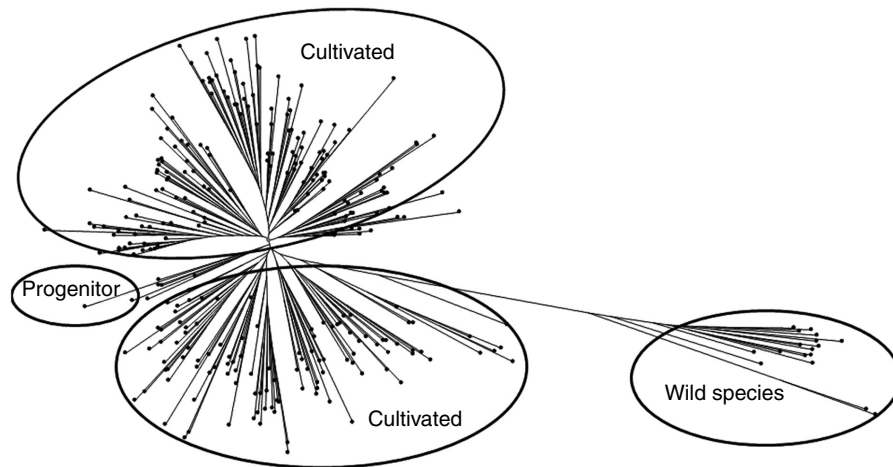
Fig. 2. Tree diagram of pigeonpea composite collection using 20 SSR markers and 952 accessions.

including supplementary irrigation whenever required. Observations on 16 qualitative and 3 quantitative traits (days to 50% flowering, days to 75% maturity and 100-seed weight) were recorded on plot basis. For the reaming 13 quantitative traits, data were recorded on 3 representative plants following descriptors for pigeonpea (IBPGR and ICRISAT, 1993). Analysis of morpho-agronomic data revealed a wide range of diversity for important traits (Table 3), revealing the importance of composite collection as a new source of diversity for important traits in pigeonpea improvement programmes.

Identification of promising germplasm lines

Promising germplasm lines for four important economic traits, days to 50% flowering, number of pods, 100-seed weight and seed yield/plant have been identified for use by the breeders in pigeonpea improvement. For days to 50% flowering, none of the accessions were found earlier than the earliest flowering check, ICP 11 543 (80 d); however, 96 accessions were at par with this check. Twenty-eight accessions produced higher number of pods (>264 pods) than the best check, out of which five accessions, ICP 9450, ICP 4167, ICP 14 225, ICP 11 970 and ICP 9558 produced significantly higher number of pods (356–393 pods) in comparison to the best check, ICP 8863 (264 pods). The 100-seed weight of 88 accessions was significantly higher (12.6–20.7 g) than the best check, ICP 7221 (9.8 g/100 seed), whereas none of the accessions produced significantly higher seed yield/plant compared with the best check, ICP 8863 (127.3 g). Based upon the *per se* performance, 49 accessions produced higher seed yield/plant (128.2–176.2 g) in comparison to the best check, ICP 8863.

Table 3. Range of variation for important traits in the pigeonpea composite collection developed and evaluated at ICRISAT, Patancheru, India

Character	Minimum	Maximum	Mean	CV%
Leaf area (cm^2)	3.4	98.1	22.3 ± 1.12	4.74
Days to 50% flowering	84.9	162.4	129.2 ± 11.78	13.14
Plant height (cm)	81.2	207.0	152.7 ± 13.77	10.98
Primary branches/plant (no.)	6.5	24.0	12.8 ± 2.44	26.05
Secondary branches/plant (no.)	3.0	35.1	8.7 ± 4.16	65.67
Tertiary branches/plant (no.)	0.8	5.0	2.0 ± 1.11	205.62
Days to 75% maturity	161.8	207.1	187.5 ± 11.03	13.07
No. of racemes/plant	16.0	195.8	70.1 ± 14.87	23.20
Pod bearing length (cm)	40.1	84.0	62.3 ± 7.27	16.68
No. of pods/plant	26.2	392.7	143.2 ± 32.33	24.72
Pod length (cm)	3.0	8.8	5.1 ± 0.45	9.78
Seeds/pod (no.)	3.3	4.6	3.8 ± 0.25	12.61
100-Seed weight (g)	3.9	20.7	9.0 ± 0.97	10.93
Seed yield/plant (g)	35.0	176.2	87.1 ± 21.42	33.54
Harvest index (%)	24.3	26.5	25.1 ± 1.34	31.11
Shelling percentage (%)	53.2	60.7	57.7 ± 3.25	18.55

The most diverse pairs of accessions have been identified among these promising accessions for the four important traits based upon the mean phenotypic diversity following Gower (1971) and SSR diversity following simple matching distance (Table 4). For early flowering, the accession ICP 15 391 from Nigeria (Africa) in combination with two accessions ICP 14 459 and ICP 11 737 both from ICRISAT (ASIA 4 region) showed high phenotypic diversity. Similarly, diverse accessions have been identified for higher number of pods, higher 100-seed weight, and high seed yield/plant (Table 4) for use in pigeonpea improvement programmes to develop improved cultivars with a broad genetic base. Among early flowering accessions, ICP 11 605 from ICRISAT (ASIA 4) exhibited maximum SSR diversity in combination with five accessions, ICP 14 486 and ICP 11 609, both from ICRISAT (ASIA 4), ICP 7629, ICP 6973 and ICP 6974, all three from India (ASIA 1). Diverse accessions were also identified for higher number of pods, higher 100-seed weight and high seed yield/plant (Table 4). There was no correspondence between the highly diverse pair of identified accessions using phenotypic and genotypic diversity in any of the four traits. This was not surprising as the correlation between the two measures of diversity (phenotypic and genotypic) was very low and non-significant in the four types of materials, early flowering ($r = -0.022$), high number of pods (0.181), high 100-seed weight (-0.014) and high seed yield/plant (0.009), which could be due to the fact that the diversity detected by these SSRs does not reflect the diversity associated with these important traits.

Use of germplasm in pigeonpea improvement

For pigeonpea, enormous genetic variability has been conserved in ICRISAT genebank in the form of landraces, breeding lines, advanced breeding cultivars and wild relatives. These are the reservoir of many useful genes for various agronomic traits and provide new sources of resistance to emerging insect-pests and diseases. Many traditional landraces have been released directly as cultivars through selection in several countries and contributed significantly to the increased production and productivity. ICP 8863, a wilt resistant selection from germplasm line ICP 7626, was released as cultivar 'Maruti' in India. This cultivar compared to the local variety has resulted in 57% gain for grain yield, 45% for fodder by-product and 27% for stalk yield, which yielded about 42% gain in the total value of output in comparison with the best cultivar. Further, the use of ICP 8863 reduced unit cost by 42% or Rs. 3820.47 (US$ 123)/ton of the grain, which resulted in the significant impact and large-scale cultivation of this cultivar in South India (Bantilan and Joshi, 1996).

Table 4. Promising diverse accessions in the pigeonpea composite collection

Traits	Pair of accessions showing maximum diversity			
	Phenotypic diversity based on morpho-agronomic traits		Genotypic diversity based on SSR markers	
Early flowering	ICP 14 459	ICP 15 391	ICP 14 486	ICP 11 605
	ICP 11 737	ICP 15 391	ICP 7629	ICP 11 605
	ICP 15 068	ICP 14 770	ICP 6973	ICP 11 605
	ICP 11 737	ICP 15 597	ICP 11 605	ICP 6974
	ICP 14 853	ICP 14 770	ICP 11 605	ICP 11 609
High number of pods	ICP 11 737	ICP 11 947	ICP 7426	ICP 14 209
	ICP 11 737	ICP 16 440	ICP 7194	ICP 14 209
	ICP 3230	ICP 16 440	ICP 7426	ICP 16 674
	ICP 3230	ICP 4167	ICP 7194	ICP 16 674
	ICP 11 737	ICP 228	ICP 8211	ICP 13 295
High 100-seed weight	ICP 13 170	ICP 8539	ICP 13 744	ICP 14 163
	ICP 13 170	ICP 7452	ICP 13 744	ICP 8539
	ICP 13 170	ICP 8003	ICP 13 744	ICP 13 033
	ICP 15 472	ICP 12 789	ICP 14 163	ICP 13 824
	ICP 13 170	ICP 15 394	ICP 13 744	ICP 15 180
High seed yield/plant	ICP 7949	ICP 13 483	ICP 12 773	ICP 16 674
	ICP 14 178	ICP 13 483	ICP 12 773	ICP 8424
	ICP 7170	ICP 13 483	ICP 12 773	ICP 13 216
	ICP 3230	ICP 13 483	ICP 12 773	ICP 13 203
	ICP 8424	ICP 7952	ICP 12 773	ICP 14 770

Similarly, ICP 14770, a pod borer tolerant selection from ICP 1903, was released as cultivar 'Abhaya' in India. Several other selections from germplasm lines have been released in USA, Fiji, India, Venezuela, Nepal and Malawi. In addition, several landraces have been used in hybridization programmes as sources for specific traits such as short duration, other important agronomic traits, resistance to biotic and abiotic stresses and nutrition quality traits to develop cultivars in India, Australia and Indonesia. Resistance sources have been identified in wild relatives, *C. acutifolius* against *Helicoverpa armigera* (Mallikarjuna and Saxena, 2002), and in six species, *C. albicans*, *C. platycarpus*, *C. cajanifolius*, *C. lineatus*, *C. scarabaeoides* and *C. sericeus* against three isolates of pigeonpea sterility mosaic viruses prevalent in peninsular India (Kumar *et al.*, 2005). For nutritional quality traits such as high protein (28–30%), *C. mollis*, *C. scarabaeoides* and *C. albicans* have shown promise as donors. ICPL 87162 with high seed protein content ($>27\%$) and good seed size has been developed by crossing with *C. scarabaeoides* (Reddy *et al.*, 1997). Higher levels of tolerance to salinity in *C. albicans* and *C. platycarpus* have been reported than in cultivated pigeonpea (Subbarao *et al.*, 1991). Except *C. platycarpus* and *C. mollis*, all other species are cross-compatible with cultivated pigeonpea and hold a great potential for their improvement. Further, cytoplasmic-male sterility (CMS) systems have been developed using wild *Cajanus* species such as *Cajanus sericeus*, denoted as A_1 CMS system (Ariyanayagam *et al.*, 1995), *C. scarabaeoides*, A_2 CMS system (Reddy and Faris, 1981; Tikka *et al.*, 1997; Saxena and Kumar, 2003), *C. cajanifolius*, A_4 CMS system (Saxena *et al.*, 2005) and *C. acutifolius*, A_5 CMS system (Mallikarjuna and Saxena, 2005). Though four male-sterility systems are available (Saxena *et al.*, 2006), only A_4 CMS system is now being utilized by ICRISAT and its public–private partners to develop the new generation of pigeonpea hybrids with good seed yield potential (Dalvi *et al.*, 2010). Fertility restorer and male-sterility maintainers have been identified among advanced breeding and germplasm lines based on their pollen fertility. Stable CMS system and restorers having high specific combining ability and resistance to important stresses are needed to develop high heterotic hybrids with wide adaptation for cultivation.

Conclusions

Germplasm subsets such as a mini core collection or a reference set capturing species diversity in a limited number of lines provide an excellent opportunity for the isolation of allelic variants of candidate genes for traits of economic importance, including functional genomic analysis (Upadhyaya *et al.*, 2006b; Glaszmann *et al.*, 2010). These subsets may be profiled with additional markers and extensively phenotyped for traits of economic importance to identify accessions for beneficial traits for utilization in pigeonpea breeding and genomics (Upadhyaya *et al.*, 2008). The promising diverse germplasm accessions identified in this study would play an important role in diversifying the genetic base of the working collection of plant breeders, for use in developing pigeonpea cultivars with a broad genetic base. Stable CMS systems have been developed following interspecific hybridization, using wild species. Identification of heterotic combination with resistance to diseases and insect-pests would revolutionize the pigeonpea production. The seeds of promising germplasm accessions, mini core collection and reference set are available upon request to pigeonpea researchers through Standard Material Transfer Agreement from ICRISAT genebank (Upadhyaya and Gowda, 2009).

Acknowledgements

References

Ariyanayagam RP, Rao AN and Zaveri PP (1995) Cytoplasmic-genic male-sterility in interspecific matings of *Cajanus*. *Crop Science* 35: 981–985.

Bantilan MCS and Joshi PK (1996) Adoption and impact of pigeonpea ICP 8863. In: Baidu-Forson J, Bantilan MCS, Debrah SK and Rohrback DD (eds) *Partners in Impact Assessment. Summary Proceedings of an ICRISAT/NARS Workshop on Methods and Joint Impact Targets in Western and Central Africa, 3–5 May 1995, Sadore, Niger*. Patancheru, AP: ICRISAT, pp. 36–39.

Brown AHD (1989) Core collections: a practical approach to genetic resources management. *Genome* 31: 818–824.

Burns MJ, Edwards KJ, Newbury HJ, Ford-Lloyd BV and Baggott CD (2001) Development of simple sequence repeat (SSR) markers for the assessment of gene flow and genetic diversity in pigeonpea (*Cajanus cajan*). *Molecular Ecology Notes* 1: 283–285.

Cobb BD and Clarkson JM (1994) A simple procedure for optimizing the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucleic Acids Research* 22: 3801–3805.

Dalvi VA, Saxena KB, Luo RH and Li YR (2010) An overview of male-sterility systems in pigeonpea [*Cajanus cajan* (L.) Millsp.]. *Euphytica* 173: 397–407.

Duvick DN (1995) Security and long-term prospects for conservation of genetic resources. *Research in Domestic and International Agribusiness Management* 11: 33–45.

Ferreira ME (2005) Molecular analysis of genebanks for sustainable conservation and increased use of plant genetic resources. Paper presented in International Workshop on The Role of Biotechnology for the Characterization and Conservation of Crop, Forestry, Animal and Fishery Genetic Resources 5–7 March 2005, Turin, Italy, pp. 78–82

Food and Agriculture Organization (2008) http://www/FAO. ORG/FAOSTAT database

Frankel OH (1984) Genetic perspective of germplasm conservation. In: Arber W, Limensee K, Peacock WJ and Stralinger P (eds) *Genetic Manipulations: Impact on Man and Society*. Cambridge: Cambridge University Press, pp. 161–170.

Glaszmann JC, Kilian B, Upadhyaya HD and Varshney RK (2010) Accessing genetic diversity for crop improvement. *Current Opinion in Plant Biology* 13: 167–173.

Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics* 27: 857–874.

Heckenberger M, Bohn M, Ziegle JS, Joe LK, Hauser JD, Hutton M and Melchinger AE (2002) Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data. *Molecular Breeding* 10: 181–191.

IBPGR and ICRISAT (1993) *Descriptors for Pigeonpea*. Patancheru/AP: IBPGR/ICRISAT.

Idury RM and Cardon LR (1997) A simple method for automated allele binning in microsatellite markers. *Genome Research* 7: 1104–1109.

Kumar S, Gupta S, Chandra S and Singh BB (2004) How wide is genetic base of pulse crops? In: Ali M, Singh BB, Kumar Shiv and Vishwa Dhar (eds) *Pulses in New Perspective. Proceedings of the National Symposium on Crop Diversification and Natural Resources Management, 20–22 December 2003*. Kanpur: Indian Society of Pulses Research and Development, Indian Institute of Pulses Research, pp. 211–221.

Kumar PL, Latha TK, Kulkarni NK, Raghavendra K, Saxena KB, Waliyar F, Rangaswamy KT, Muniyappa V, Doriswamy S and Jones AT (2005) Broad-based resistance to pigeonpea sterility mosaic disease in wild relatives of pigeonpea (*Cajanus*: Phaseoleae). *Annals of Applied Biology* 146: 371–379.

Liu K and Muse SV (2005) Power marker: integrated analysis environment for genetic marker data. *Bioinformatics* 21: 2128–2129.

Mace ES, Buhariwalla HK and Crouch JH (2003) A high-throughput DNA extraction protocol for tropical molecular breeding programs. *Plant Molecular Biology Reporter* 21: 459a–459h.

Mallikarjuna N and Saxena KB (2002) Production of hybrids between *Cajanus acutifolius* and *C. cajan*. *Euphytica* 124: 107–110.

Mallikarjuna N and Saxena KB (2005) A new cytoplasmic nuclear male sterility derived from cultivated pigeonpea. *Euphytica* 142: 143–148.

Marshall DR (1989) Limitations to the use of germplasm collections. In: Brown AHD, Frankel OH, Marshall DR and Williams JT (eds) *The Use of Plant Genetic Resources*. New York: Cambridge University Press, pp. 105–120.

Nene YL and Sheila VK (1990) Pigeonpea: geography and importance. In: Nene YL, Hall SD and Sheila VK (eds) *The Pigeonpea*. Wallingford, Oxon: CAB International, pp. 1–14.

Odeny DA, Jayashree B, Ferguson M, Hoisington D, Crouch J and Gebhardt C (2007) Development, characterization and utilization of microsatellite markers in pigeonpea. *Plant Breeding* 126: 130–137.

Odeny DA, Jayashree B, Gebhardt C and Crouch J (2009) New microsatellite markers for pigeonpea (*Cajanus cajan* (L.) Millsp.). *BMC Research Notes* 2: 35.

Perrier X, Flori A and Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X and Glazmann JC (eds) *Genetic Diversity of Cultivated Tropical Plants*. Montpellier: Enfield, Science Publishers, pp. 43–76.

Reddy LJ and Faris DJ (1981) A cytoplasmic-genetic male sterile line in pigeonpea. *International Pigeonpea Newsletter* 1: 16–17.

Reddy LJ, Saxena KB, Jain KC, Singh U, Green JM, Sharma D and Faris DJ (1997) Registration of high protein elite germplasm ICPL 87162. *Crop Science* 37: 94.

Saxena KB and Kumar RV (2003) Development of a cytoplasmic nuclear male-sterility system in pigeonpea using *C. scarabaeoides* (L.) Thouars. *Indian Journal of Genetics* 63: 225–229.

Saxena KB, Singh L and Gupta MD (1990) Variation for natural out-crossing in pigeonpea. *Euphytica* 46: 143–146.

Saxena KB, Kumar RV, Srivastava N and Shiying B (2005) A cytoplasmic–nuclear male-sterility system derived from a cross between *Cajanus cajanifolius* and *C. cajan*. *Euphytica* 145: 291–296.

Saxena KB, Kumar RV, Madhavi Latha K and Dalvi VA (2006) Commercial pigeonpea hybrids are just a few steps away. *Indian Journal of Pulses Research* 19: 7–16.

Shannon CE and Weaver W (1949) *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Subbarao GV, Johansen C, Jana MK and Kumar Rao JVDK (1991) Comparative salinity responses among pigeonpea genotypes and their wild relatives. *Crop Science* 31: 415–418.

Taguchi G (1986) *Introduction to Quality Engineering, Asian Productivity Organization*. Dearborn, Michigan: American Supplier Institute Inc..

Tikka SBS, Parmer LD and Chauhan RM (1997) First record of cytoplasmic genetic male sterility in pigeonpea (*Cajanus cajan* (L.) Millsp.) through wide hybridization. *Gujarat Agricultural University Research Journal* 22: 60–162.

Upadhyaya HD and Gowda CLL (2009) *Managing and Enhancing the Use of Germplasm – Strategies and Methodologies*. Technical Manual no. 10. Patancheru, Andhra Pradesh: International Crops Research Institute for the Semi-Arid Tropics, p. 236.

Upadhyaya HD and Ortiz R (2001) A mini core collection for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theoretical and Applied Genetics* 102: 1292–1298.

Upadhyaya HD, Ortiz R, Bramel PJ and Sube Singh (2003) Development of a groundnut core collection using taxanomical, geographical and morphological descriptors. *Genetic Resources and Crop Evolution* 50: 139–148.

Upadhyaya HD, Pundir RPS, Gowda CLL, Reddy KN and Sube Singh (2005) Geographical pattern of diversity for qualitative and quantitative traits in the pigeonpea germplasm collection. *Plant Genetic Resources: Characterization and Utilization* 3: 332–352.

Upadhyaya HD, Gowda CLL, Buhariwalla HK and Crouch JH (2006a) Efficient use of crop germplasm resources: identifying useful germplasm for crop improvement through core and min-core collections and molecular marker approaches. *Plant Genetic Resources* 4: 25–35.

Upadhyaya HD, Reddy LJ, Gowda CLL, Reddy KN and Chandra S (2006b) Development of a mini core subset for enhanced and diversified utilization of pigeonpea germplasm resources. *Crop Science* 46: 2127–2132.

Upadhyaya HD, Bhattacharjee R, Varshney RK, Hoisington DA, Reddy KN and Singh S (2008) Assessment of genetic diversity in pigeonpea using SSR markers. In: 2008 Joint Annual Meeting, 5–9 October 2008, Houston, Texas

Upadhyaya HD, Yadav Devvart, Dronavalli Naresh, Gowda CLL and Singh Sube (2010) Mini core germplasm collections for infusing genetic diversity in plant breeding programs. *Electronic Journal of Plant Breeding* 1: 1294–1309.

van der Maesen LJG (2006) *Cajanus cajan* (L.) Mill sp. In: Brink M and Belay G (eds) *Plant Resources of Africa 1. Cereals and Pulses*. Wageningen: Backhuys Publishers. http://database.prota.org/search.htm