Plant Breeding

doi:10.1111/j.1439-0523.2012.02009.x

Whole-genome scanning for mapping determinacy in Pigeonpea (*Cajanus* spp.)

Reyazul Rouf Mir[1,2], Rachit K Saxena[1], Kulbhushan Saxena[1], Hari D Upadhyaya[1], Andrzej Kilian[3], Douglas R Cook[4] and Rajeev K Varshney[1,5,6]

[1]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Hyderabad, India; [2]Division of Plant Breeding & Genetics, Shere-Kashmir University of Agricultural Sciences & Technology of Jammu (SKUAST-J), Chatha 180 009, Jammu, India; [3]Diversity Arrays Technology Pty Ltd, 1 Wilf Crane Crescent, Yarralumla, ACT 2600, Australia; [4]University of California, Davis, CA 95616-8680, USA; [5]CGIAR-Generation Challenge Programme (GCP), c/o CIMMYT, Int APDO Postal 6-641, 06600 Mexico, DF, Mexico; [6]Corresponding author, E-mail: r.k.varshney@cgiar.org

With 2 figures and 1 table

Received March 16, 2012/Accepted June 29, 2012
Communicated by H. Balyan

Abstract

Determinacy is an agronomically important trait in several crop species including pigeonpea. With an objective to investigate determinacy in pigeonpea, a set of 94 pigeonpea lines including 11 determinate (DT) and 83 indeterminate (IDT) lines were used for genotyping with DArT arrays (with 6144 features) and 768 SNP markers using GoldenGate assay. The polymorphism information content (PIC) for these markers varied from 0.02 to 0.50. Association analysis on marker genotyping and phenotyping data showed a significant association ($P \leq 0.01$) of determinacy with 19 SNP and 6 DArT markers explaining 8.05–8.58% and 7.26–14.53% phenotypic variation, respectively. Clustering based on entire DArT and SNP markers could not discriminate DT lines from IDT lines; however, analysis with associated markers discriminated DT lines from the IDT lines. Marker–trait associations after validation may prove useful in marker-assisted selection (MAS) involving the development of ideal DT genotypes for environments with moderate growth, tolerance to drought and water logging. This is the first report on mapping of determinacy trait as well as the first report on association mapping for any trait in pigeonpea.

Key words: Pigeonpea — determinacy — whole-genome scanning — diversity array technology arrays — GoldenGate SNP assays

The pattern and time of flowering are the two major adaptive traits in flowering plants. After a period of vegetative growth, plants undergo floral transition. The switch from vegetative to reproductive growth stages is controlled by physiological signals and interaction between positive and negative regulators that integrate environmental (photoperiod and temperature) and endogenous (stages of the plant) conditions (Colasanti and Sundaresan 2000, Foucher et al. 2003).

Pigeonpea [*Cajanus cajan* (L.) Millsp.] is one of the important legume crops for arid and semi-arid tropics of the world. Traditionally, pigeonpea lines have indeterminate (IDT) flowering pattern, where inflorescence develops as axillary racemes from all over the branches and flowering proceeds acropetally from base to apex, both within the raceme and on the branches. However, some genotypes have been reported, where apical buds of the main shoots develop into inflorescences. These genotypes are called determinate (DT) genotypes. The wild relatives and most of the cultivated pigeonpea are having indeterminate growth habit, and therefore, it is believed that determinacy trait has been selected during pigeonpea domestication. Occasionally, flowering pattern of some line has been found intermediate between IDT and DT, and the lines with this type of flowering pattern are called semi-determinate (SDT) lines (Craufurd et al. 2001).

The pigeonpea lines with DT flowering pattern has several advantages over lines with IDT/SDT flowering pattern. For instance, some of the advantages of DT lines include (i) much shorter heights and therefore increased lodging resistance, (ii) more main stem branches/plant, (iii) less lowest-pod heights, (iv) shorter flowering and reproductive periods and (v) flowering earlier and requiring a single-pass harvest of pods of nearly the same age, thus allowing mechanized harvesting (Robinson and Wilcox 1998, Kilgore-Norquest and Sneller 2000). Owing to significant importance of the trait, genetic dissection of determinacy/flowering time has been undertaken in many plant species including *Arabidopsis thaliana* (Ehrenreich et al. 2009), soybean (Tian et al. 2010), pea (Foucher et al. 2003), common bean (Kwak et al. 2008, Repinski et al. 2012). Majority of these studies have employed candidate gene sequencing approach; genome-scanning-based association mapping has also been used recently (Li et al. 2010). Extensive genetic and molecular studies have been conducted in long-day model plant species *Arabidopsis thaliana* that have led to the isolation and characterization of flowering time genes (Mouradov et al. 2002). Several of these genes were used to identify their functional orthologues in some warm-season legume crop species like soybean (Tian et al. 2010), common bean (Repinski et al. 2012). In the case of pigeonpea, determinacy trait, however, has not been studied yet. At the time of start of the study, very limited sequence data were available in pigeonpea. In recent years, Diversity Array Technology (DArT) platform comprising 6144 features (Yang et al. 2006, 2011) and 768-SNPs GoldenGate assay (Kassa et al. 2012) have been developed and became available for pigeonpea genotyping. These genotyping platforms have the capacity to generate genome-wide marker profile data in cost-effective and faster manner. The present study was therefore undertaken to map the DT trait using genome-scanning association mapping approach by deploying the DArT and SNP markers on a set of 94 pigeonpea lines segregating for determinacy trait.

Materials and Methods

Plant material and DNA extraction: A panel of 94 pigeonpea lines including 38 parents of 22 mapping populations and 56 germplasm lines were used (Table S1). These lines represent two distinct flowering patterns: 11 belong to DT flowering pattern and the remaining 83 showed IDT flowering pattern.

Genomic DNA of above-mentioned 94 pigeonpea lines was isolated and purified from the leaves of 2- to 3-week-old single plants from each line following the protocol as mentioned in the study by Cuc et al. (2008). The DNA quantity for each sample was assessed on 0.8% agarose gel, and DNA concentrations were normalized at 50 ng/μl.

DArT and SNP genotyping: For DArT genotyping, the genomic representations of samples were generated using the same complexity reduction method (*Pst*I/*Hae*III), which was developed for the construction of the DArT diversity library and hybridized with the array comprising 6144 clones (Yang et al. 2011). TIF images generated in the experiment were analysed using DARTSOFT 7 software. The hybridization signal in target channel (genomic representation) was divided by hybridization signal in reference channel (polylinker) to calculate the relative hybridization intensity of each clone on each slide. Fuzzy k-means was used to convert relative hybridization intensities of clones with variable relative hybridization intensity across slides into binary scores (presence vs. absence) and was subjected to clustering and classified polymorphic clones as being present ('1') or absent ('0') in the representation hybridized to a slide.

For SNP genotyping, GoldenGate assay comprising 728 tentative orthologous gene sequence (TOGs)-based SNPs (Kassa et al. 2012) was used to genotype the diversity panel using the Illumina Bead Station 500 G (Illumina, San Diego, CA, USA) at the University of California, Davis (USA). SNP genotyping data were analysed using the Illumina BeadStudio genotyping software that allows the visualization of the data directly for further analysis.

Analysis of polymorphism, genetic relationship and genetic structure: The PIC value of markers was calculated using the formula given by Anderson et al. (1993). To evaluate the relationship between 94 pigeonpea lines, SNP and DArT, allele call data obtained for polymorphic markers or selected associated markers were used to prepare a dissimilarity matrix and to construct a dendrogram using DARWIN V5.0.128 software (darwin.cirad.fr/darwin/Home.php, Perrier et al. 2003). The dissimilarity matrix thus obtained was subjected to cluster analysis using the unweighted neighbour–joining method, followed by bootstrap analysis with 1000 iterations to obtain dendrogram. Mantel test was conducted to test for the DArT and SNP markers dissimilarity matrices correspondence using software package GENALEX 6.41 (Peakall and Smouse 2006).

Model-based cluster analysis was performed to infer genetic structure and define the number of clusters (gene pools) in the data set using the software STRUCTURE V2.3.3 (Pritchard et al. 2000, Falush et al. 2003). The number of presumed populations (K) was set from 1 to 10, and each was repeated three times for each run burn-in, with iterations set to 100 000 and 200 000, respectively, and K model without admixture and correlated allele frequencies was used. The run with maximum likelihood was used to place individual genotypes into groups (subpopulations). Within a group, genotypes with affiliation probabilities (inferred ancestry) ≥80% were assigned to a distinct group, and those with < 80% were treated as 'admixture', that is, these genotypes seem to have a mixed ancestry from parents belonging to different gene pools or geographical origins. This assignment obtained through maximum-likelihood approach was further confirmed by a modified delta K (ΔK) method, which provides real number of clusters (Evanno et al. 2005). The information obtained through model-based clustering was used for working out marker–trait associations, thereby avoiding spurious associations for determinacy. STRUCTURE program was also used for calculating gene diversity (expected heterozygosity) of individuals within a subpopulation obtained through model-based clustering.

Association analysis: The Trait Analysis by Association, Evolution and Linkage (TASSEL version 2.1) was used for establishing association between markers and determinacy trait by using mixed linear model (MLM) approach. Subsequently, significant associations were also confirmed by general linear model (GLM). For MLM analysis, the population structure (Q-matrix) was inferred by program STRUCTURE 2.2, and kinship matrix (K-matrix) was inferred by TASSEL 2.1. Significant marker–trait associations were considered at P-value ($P \leq 0.01$). For GLM analysis, genotypic data, phenotypic data and Q-matrix were used. Significantly associated SNP markers were further analysed for their putative functions by blasting their original TOG sequences against the soybean nucleotide database (BLASTN) and further annotated by blasting against non-redundant protein database at NCBI using BLASTX.

Results

Marker analysis

Genotyping of the diversity panel with DArT array (6144 features) and GoldenGate assay (768 SNPs) on 94 lines showed a total of 3978 polymorphic markers including 3262 DArTs and 716 SNPs. The polymorphism information content (PIC), which indicates informativeness of a marker locus or marker system, varied from 0.02 to 0.50 for both DArT and SNP markers. However, the average PIC value was slightly more (0.18) for DArT markers than for SNP markers (0.15).

Genetic relationships among pigeonpea lines

DArT and SNP marker data sets generated above were used to prepare dendrograms (Figs. S1 and S2). Pairwise comparisons of genetic distance matrices for the DArT and SNP marker data sets through Mantel test were found correlated ($r = 0.49$), however, at relatively low level of significance ($P = 0.01$) (Fig. S3). Therefore, to obtain more accurate genetic distance estimate, combined analysis was carried out using both DArT and SNP markers together (Fig. 1). This dendrogram classified all the genotypes analysed into two main clusters (cluster I and cluster II) (Fig. 1). The cluster I includes a solitary IDT line (ICP 2376), while the cluster II contains all other lines and was divided into subclusters IIa and IIb. The subcluster IIa contains three wild species accessions (ICPW 69, ICPW 68 and ICP 15665), while subcluster IIb was further subdivided into IIb-1 and IIb-2. The IIb-1 contains 12 IDT lines including five wild species accessions, while IIb-2 contains all other lines including all DT lines (11) and remaining three wild species accessions. While comparing this combined dendrogram with the individual dendrograms constructed using DArT and SNP markers separately, it was found that clustering pattern of different pigeonpea lines in combined dendrogram is largely similar to dendrogram obtained by DArT markers separately (Fig. S1). However, the clustering in dendrogram obtained with SNP markers was divided into two main clusters (I and II), and cluster I was divided into two subclusters (IIa and IIb). The cluster I contains four IDT lines including three wild species accessions, while cluster IIa contains eight IDT lines including four wild species accessions. The subcluster IIb contains all other lines including all 11 DT lines and three wild species accessions (Fig. S2). It is important to note that, in all the three dendrograms, it was observed that wild species (*C. cajanifolius*) accessions, namely ICPW 29, source of the cytoplasmic male sterility (CMS) in cultivated pigeonpea, and

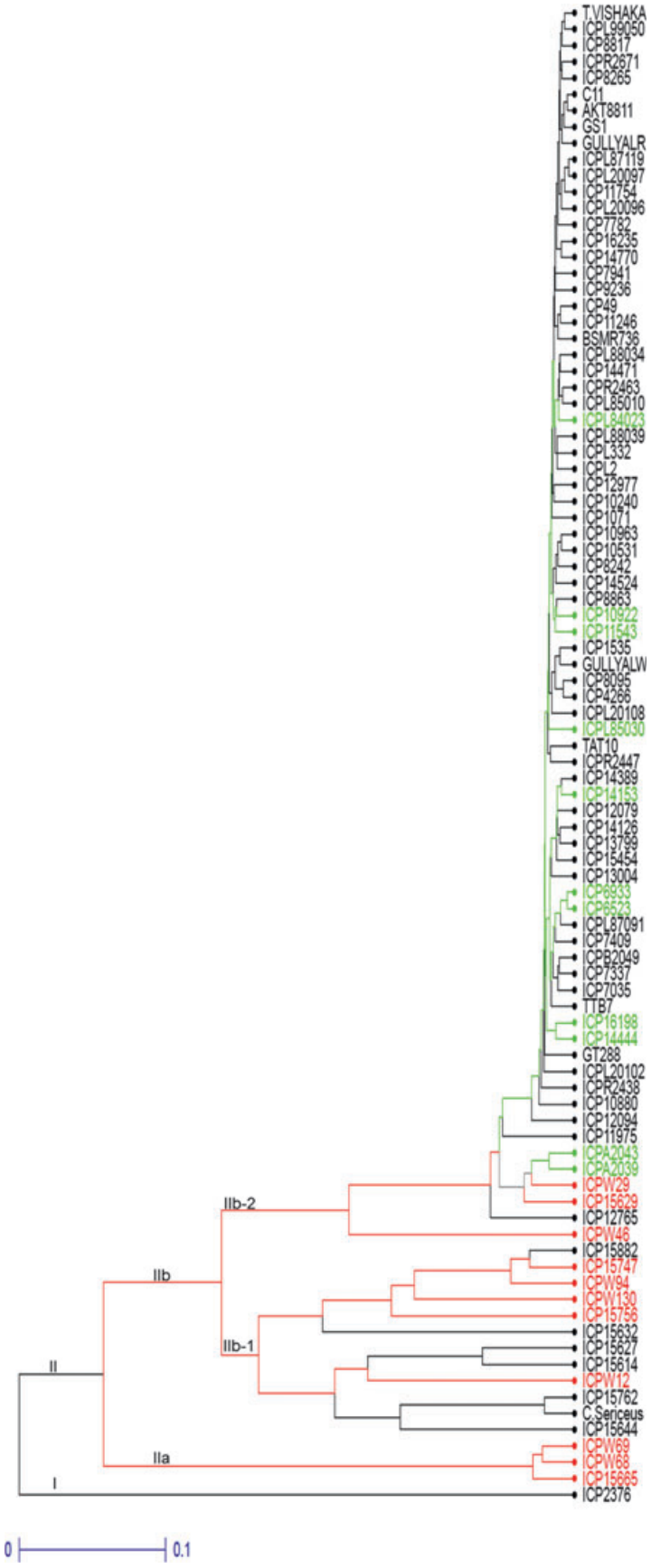

Fig. 1: Diversity analysis among 94 pigeonpea lines based on combined data set of SNP and DArT markers. The figure shows wild species accessions, determinate and indeterminate lines in red, green and black colour, respectively

ICP 15629, were clustered along with cultivated lines. In summary, the distance-based clustering pattern on the basis of all the DArT and SNP markers separately and together does not show clear distinctive clustering of pigeonpea accessions into DT/IDT. It was also found that wild pigeonpea species accessions, which are indeterminate in flowering pattern, were largely clustered together.

Genetic structure analysis and gene diversity

Model-based cluster analysis grouped 94 pigeonpea lines into three genetically distinct populations ($K = 3$) based on maximum natural log probability (LnPD) of data as well as ΔK method proposed by Evanno et al. (2005) using SNP (Fig. 2) and DArT (Fig. S4) markers separately. In the case of structural plot obtained by SNP markers, among the three populations, the *K1* contained only IDT lines including eight of eleven wild species accessions, while the other subpopulations (*K2* and *K3*) contained both DT/IDT lines. On the other hand, structural plot obtained by DArT markers contained DT/IDT lines in *K1* and *K2* subpopulations and the remaining subpopulation 3 (*K3*) contained only IDT lines including nine of eleven wild species accessions (Fig. S4). The population *K2* in case of DArT and *K3* in case of SNP markers also contained two lines as admixtures (membership probability < 0.8). It is interesting to note that in case of structural plots of both types of markers, among eleven wild species accessions, two accessions from *C. cajanifolius* species, namely ICPW 29 and ICP 15629, were clustered together (*K1* in DArT and *K2* in SNP plots) with cultivated lines rather than with other wild species accessions in structural plots.

Gene diversity (expected heterozygosity) values also indicated higher gene diversity (0.350) in subpopulation *K1* (containing wild species accessions) than in the other two subpopulations (0.008 for *K2* and 0.007 for *K3*) using SNP markers. Similarly, higher gene diversity (0.361) was observed in subpopulation *K3* (containing wild species accessions) than in other two subpopulations (0.044 for *K1* and 0.021 for *K2*) using DArT markers.

Marker–trait association

Detailed analysis using mixed linear model (MLM) and general linear model (GLM) approaches on the DArT and SNP marker genotyping data and the phenotyping data on the set of 94 lines yielded several significant marker–trait associations for determinacy trait. For instance, MLM approaches showed an association of determinacy with a total of 25 significant markers including six DArT and 19 SNPs. The 19 SNP markers are not linked with each other

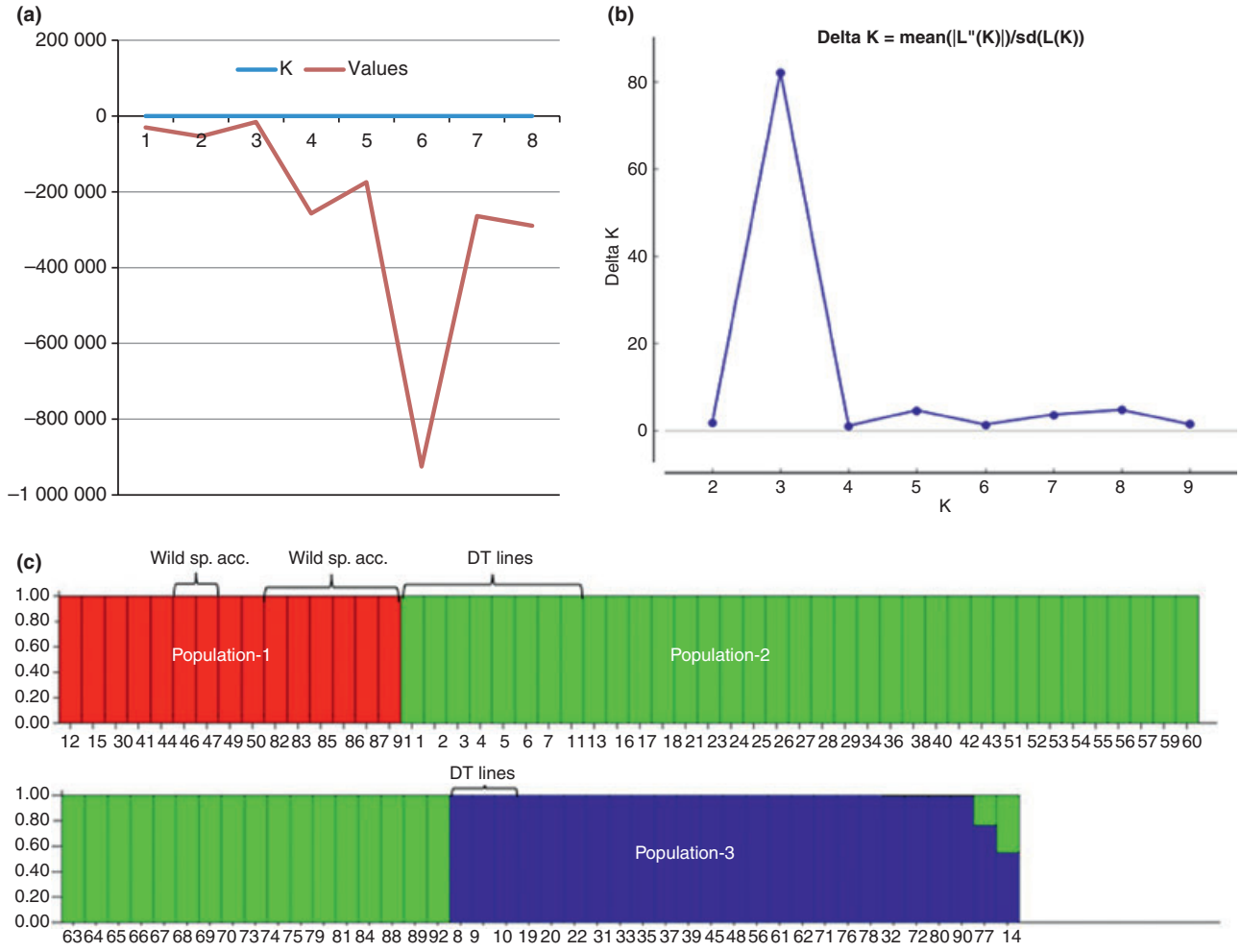

Fig. 2: Structural analysis of 94 pigeonpea lines using SNP markers. The figure has three panels: (a) shows the highest probability for having three subpopulations by using LnPD curve; (b) confirms the presence of three subpopulations by using delta *K* curve; (c) shows the structural plot of 94 pigeonpea lines. Wild species accessions are present in $K = 1$, and DT lines are present in $K = 2$ and $K = 3$

Table 1: Summary of marker–trait associations for determinacy trait in pigeonpea using (A) DArT and (B) SNP markers

	MLM	GLM	
Marker	P-value (P ≤ 0.01)	P-value (P ≤ 0.01)	R^2
(A) DArT marker			
Cc-693448	2.80E-04	2.70E-04	14.53
Cc-695833	7.74E-04	0.0023	10.38
Cc-693401	0.0011	0.0102	7.26
Cc-698210	0.0022	0.0022	10.09
Cc-693999	0.0022	0.0022	10.09
Cc-692675	0.0057	0.0042	9.79
(B) SNP marker			
TOG895724_478	0.0044	0.0044	8.57
TOG894880_514	0.0045	0.0045	8.53
TOG898078_1260	0.0045	0.0045	8.53
TOG910212_493	0.0045	0.0045	8.53
TOG923519_736	0.0045	0.0045	8.53
TOG913186_428	0.0045	0.0045	8.53
TOG897062_495	0.0045	0.0045	8.53
TOG905669_206	0.0045	0.0045	8.53
TOG895816_869	0.0045	0.0045	8.53
TOG910212_506	0.0045	0.0045	8.53
TOG896522_1889	0.0046	0.0046	8.58
TOG896850_41	0.0047	0.0047	8.54
TOG922448_261	0.0049	0.0049	8.49
TOG894864_284	0.005	0.005	8.6
TOG901166_1050	0.0051	0.0051	8.5
TOG899452_1243	0.0053	0.0053	8.44
TOG902834_914	0.0058	0.0058	8.05
TOG896976_486	0.0063	0.0063	8.3
TOG910323_779	0.007	0.007	8.21

but are distributed on six different linkage groups (LGs) with LG5 having one marker and LG10 with seven markers (D.R. Cook and R.V. Penmetsa).The P-value of these markers varied from 2.80E-04 to 0.0057 (for DArTs) and 0.0044 to 0.007 (for SNPs). All these significant marker–trait associations for determinacy were confirmed through GLM approach. Using GLM, the range of P-values varied from 2.70E-04 to 0.0102 (for DArT) and from 0.0044 to 0.007 (for SNPs). The phenotypic variation explained (R^2) by each marker varied from 7.26 to 14.53% for DArT and 8.05–8.60% for SNP markers (Table 1). It is important to mention here that analysis of allelic data for identified (19 SNP and six DArT) markers associated with determinacy grouped DT lines together as compared to the analysis based on the entire marker data set (Fig. S5).

Functional annotation of associated SNP markers

With an objective to have an idea about the genes associated with the determinacy trait, the context sequences for 19 significantly associated SNP markers were analysed against soybean genome. This provided a significant hit with the corresponding genes for 17 context sequences, source of the significantly associated SNP markers (see Table S2). BLASTX analysis (see Table S3), however, revealed that most of these sequences code for unknown proteins (10; ~53%), predicted proteins (4; ~21%), hypothetical proteins (2; ~11%) and phosphorylases (1; ~5% each).

Discussion

Domestication has resulted in changes in some floral and seed morphology traits in crop plants. These traits included shattering, free-threshing, seed dormancy, plant architecture, seed coat colour in cereals (Salamini et al. 2002, Dubcovsky and Dvorak 2007, Izawa et al. 2009), fruit size and shape in Solanaceae crops (Tanksley 2004) and determinacy/growth habit in legumes like soybean (Tian et al. 2010), pea (Foucher et al. 2003) and common bean (Kwak et al. 2008, Repinski et al. 2012). Molecular genetics studies including QTL mapping and candidate gene sequencing for flowering time/growth habit/determinacy have identified a large number of responsible genes/QTLs in several plant species including Arabidopsis, rice, barley, wheat, maize, pea, common bean, soybean (see Alonso-Blanco et al. 2009). Identification of multiple candidate genes including terminal flower 1 (*TFL1*), terminal flower 2 (*TFL2*) and polymorphisms has also allowed comparative analyses among some of these species. For instance, in the case of pea, '*PsTFL1a*', a homologue of *TFL1* of Arabidopsis, has been found responsible for determinacy trait (Foucher et al. 2003). In the case of common bean, a QTL was identified earlier for growth habit (Koinange et al. 1996, Poncet et al. 2004), and within this QTL region, a candidate gene homologous to *TFL1* '*PvTFL1y*' has been identified (Kwak et al. 2008). The function of this candidate gene has been recently validated for determinacy in common bean through candidate gene sequencing and transformation studies (Repinski et al. 2012). Similarly, in the case of soybean, a combination of genetic linkage analysis, candidate gene association analysis and heterologous transformation of Arabidopsis determinate (*tfl1*/*tfl*) mutants led to the isolation of a homologue of Arabidopsis *TFL1* in soybean for *Dt1* (Tian *et al.* 2010).

The study of determinacy in pigeonpea, despite being an important grain legume crop of the tropics and subtropics, has received little attention of breeders and molecular biologists. Only a few studies were undertaken long back on inheritance of determinacy, and there is hardly any report available where any gene has been identified for this important adaptive trait in pigeonpea (Waldia and Singh 1987, Gumber and Singh 1997) . For instance, inheritance of DT, SDT and IDT in pigeonpea was studied by using 15 different crossing combinations of DT, SDT and IDT lines (Gupta and Kapoor 1991). The segregation pattern in this study indicated that IDT is governed by a single dominant gene. The two flowering patterns in pigeonpea, DT and IDT, are domestication traits that can distinguish pigeonpea cultivars from one another. DT lines have been selected by breeders over years and are considered advantageous over IDT lines. The mechanism responsible for this transition of IDT to DT is not known in pigeonpea. This is probably the first study where an effort has been made to map determinacy trait in pigeonpea. At the time of undertaking this study, the crop did not have genome resources or genetic populations that could have been used for mapping or isolation of homologues of known DT/flowering time–related genes in pigeonpea. Availability of high-throughput marker genotyping platforms such as DArT arrays and SNP-GoldenGate assays (Yang et al. 2011, Kassa et al. 2012) has made it possible to generate genome-wide marker profile data for undertaking association mapping of determinacy in pigeonpea.

DArT and SNP markers that were used in the present study for genome scanning are considered two important marker types for genotyping natural or breeding populations, because both the marker systems used in the present study were predominantly bi-allelic markers and therefore, as expected, two alleles per DArT/SNP locus were observed for all the markers. Comparison of average PIC values showed slightly

higher average PIC values for DArT markers (0.18) over SNP markers (0.15). However, this may be just a function of larger number of polymorphic DArT markers (3262) as compared to SNP markers (716) used in the study.

Detailed analysis on genetic relationships between different pigeonpea lines on the basis of all SNP and DArT markers indicated genetic homogeneity of DT/IDT lines except some wild species accessions that tend to cluster separately. However, DT lines could be clearly discriminated from the IDT lines based on the dendrogram obtained from the marker data of only significantly associated SNP and DArT markers with determinacy. A close clustering of most of the lines in distance-based dendrogram and formation of only few clusters indicated moderate levels of genetic diversity/differentiation in the pigeonpea lines analysed. These findings obtain support from some of our earlier studies in pigeonpea, indicating moderate/less diversity in elite pigeonpea gene pools (Saxena et al. 2010a,b). Clustering of pigeonpea wild species accessions ICPW 29 and ICP 15629 with cultivated gene pool was expected and again confirmed that *C. cajanifolius* species is the most closely related species (progenitor) of pigeonpea (van der Maesen 1990). Similar results were also obtained through model-based structural analysis. This analysis also indicated moderate genetic differentiation and the presence of only three subpopulations in the lines analysed. Clustering of majority of the wild species accessions (except ICPW 29 and ICP 15629) in same subpopulation indicated that these lines are genetically distant/diverse from other cultivated lines. This has been also confirmed by gene diversity (expected heterozygosity) values, indicating high values for the subpopulations containing wild species (Mariette et al. 2010).

The study of marker–trait associations for determinacy trait in pigeonpea strongly indicates the involvement of several genomic regions/genes/markers responsible for this important trait in pigeonpea. The identification of four highly significantly associated DArT markers (Cc-693448, Cc-695833, Cc-698210 and Cc-693999) explaining >10% phenotypic variation for determinacy may prove useful in molecular breeding programmes for pigeonpea improvement. However, these marker–trait associations shall be validated on a large germplasm set, and these DArT markers shall be sequenced followed by their conversion into user-friendly PCR markers. The validated markers may prove useful in marker-assisted selection (MAS) in breeding programmes for selecting the lines carrying allele for DT trait. Such markers, eventually, may allow the development of ideal DT genotypes for environments with moderate growth (5–6 t/ha) having 30–35% harvest index, initial vigour and tolerance to drought, and water logging (Singh and Oswalt 1992).

Acknowledgement

The authors are thankful to Theme Leader Discretionary Grant of CGIAR Generation Challenge Programme (GCP) and Centre of Excellence (CoE) grant of Department of Biotechnology (DBT) of Government of India for providing funding in carrying out this research work.

References

Alonso-Blanco, C., M. G. M. Aarts, L. Bentsink, J. J. B. Keurentjes, M. Reymond, D. Vreugdenhil, and M. Koornneef, 2009: What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell **21,** 1877—1896.

Anderson, J. A., G. A. Churchill, J. E. Autrique, S. Tanksley, and M. E. Sorrells, 1993: Optimizing parental selection for genetic linkage maps. Genome **36,** 181—186.

Colasanti, J., and V. Sundaresan, 2000: 'Florigen' enters the molecular age: long-distance signals that cause plants to flower. Trends Biochem. Sci. **25,** 236—240.

Craufurd, P. Q., H. S. Soko, J. K. Jones, and R. J. Summerfield, 2001: Inheritance of duration from sowing to first flowering in pigeonpea. Euphytica **119,** 323—333.

Cuc, L. M., E. Mace, J. Crouch, V. D. Quang, T. D. Long, and R. K. Varshney, 2008: Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). BMC Plant Biol. **8,** 55.

Dubcovsky, J., and J. Dvorak, 2007: Genome plasticity a key factor in the success of polyploid wheat under domestication. Science **316,** 1862—1866.

Ehrenreich, I. M., Y. Hanzawa, L. Chou, J. L. Roe, P. X. Kover, and M. D. Purugganan, 2009: Candidate gene association mapping of Arabidopsis flowering time. Genetics **183,** 325—335.

Evanno, G., S. Regnaut, and J. Goudet, 2005: Detecting the number of clusters of individuals using the softwares STRUCTURE: a simulation study. Mol. Ecol. **14,** 2611—2620.

Falush, D., M. Stephens, and J. K. Pritchard, 2003: Inference of population structure using multi-locus genotype data: linked loci and correlated allele frequencies. Genetics **164,** 1567—1587.

Foucher, F., J. Morin, J. Courtiade, S. Cadioux, N. Ellis, M. J. Banfield, and C. Rameau, 2003: Determinate and late flowering are two terminal flower1/centroradialis homologs that control two distinct phases of flowering initiation and development in pea. Plant Cell **15,** 2742—2754.

Gumber, R. K., and S. Singh, 1997: Genetics of flowering patterns in pigeonpea: further evidence for two gene control. Euphytica **96,** 233—236.

Gupta, S. C., and R. K. Kapoor, 1991: Inheritance of growth habit in pigeonpea. Crop Sci. **31,** 1456—14.

Izawa, T., S. Konishi, A. Shomura, and M. Yano, 2009: DNA changes tell us about rice domestication. Curr. Opin. Plant Biol. **12,** 185—192.

Kassa, M. T., R. V. Penmetsa, N. Carrasquilla-Garcia, B. K. Sarma, S. Datta, H. D. Upadhyaya, R. K. Varshney, E. J. B. Wettberg, and D. Cook, 2012: Genetic patterns of domestication in pigeonpea (*Cajanus cajan* (L.) Millsp.) and wild *Cajanus* relatives. PLoS One **7,** e39563.

Kilgore-Norquest, L., and C. H. Sneller, 2000: Effect of stem termination on soybean traits in southern U.S production systems. Crop Sci. **40,** 83—90.

Koinange, E. M. K., S. P. S. Singh, and P. Gepts, 1996: Genetic control of the domestication syndrome in common bean. Crop Sci. **36,** 1037—1045.

Kwak, M., D. M. Velasco, and P. Gepts, 2008: Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (*Phaseolus vulgaris*). J. Hered. **99,** 283—291.

Li, Y., Y. Huang, J. Bergelson, M. Nordborg, and J. O. Borevitza, 2010: Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. Proc. Natl Acad. Sci. USA **107,** 21199—21204.

van der Maesen, L. J. G., 1990: Pigeonpea: origin, history, evolution and taxonomy. In: Y. L. Nene, S. D. Hall, and V. K. Sheila (eds), Pigeonpea, 15—46. CAB International, Wallingford.

Mariette, S., M. Tavaud, U. Arunyawat, G. Capdeville, M. Millan, and F. Salin, 2010: Population structure and genetic bottleneck in sweet cherry estimated with SSRs and the gametophytic self-incompatibility locus. BMC Genet. **11,** 77.

Mouradov, A., F. Cremer, and G. Coupland, 2002: Control of flowering time: interacting pathways as a basis for diversity. Plant Cell **14,** S111—S130.

Peakall, R., and P. E. Smouse, 2006: GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes **6,** 288—295.

Perrier, X., A. Flori, and F. Bonnot, 2003: Data analysis methods. In: P. Hamon, M. Seguin, X. Perrier, and J. C. Glaszmann (eds). Genetic diversity of cultivated tropical plants, 43—76. Enfield, Science Publishers, Montpellier, France.

Poncet, V., T. Robert, A. Sarr, and P. Gepts, 2004: Quantitative trait loci analyses of the domestication syndrome and domestication process. In: Goodman. R(ed.), Encyclopedia of plant and crop-science, 1069—1073. Marcel Dekker, New York, NY.

Pritchard, J. K., M. Stephens, and P. Donneily, 2000: Inference of population structure using multi-locus genotype data. Genetics **155,** 945—959.

Repinski, S., M. Kwak, and P. Gepts, 2012: The common bean growth habit gene *PvTFL1y* is a functional homolog of *Arabidopsis TFL1*. Theor. Appl. Genet. **124,** 1539—1547.

Robinson, S. L., and J. R. Wilcox, 1998: Comparison of determinate and indeterminate soybean near-isolines and their response to row spacing and planting date. Crop Sci. **38,** 1554—1557.

Salamini, F., H. Ozkan, A. Brandolini, R. Schafer-Pregl, and W. Martin, 2002: Genetics and geography of wild cereal domestication in the near east. Nat. Rev. Genet. **3,** 429—441.

Saxena, R. K., C. Prathima, K. B. Saxena, D. A. Hoisington, N. K. Singh, and R. K. Varshney, 2010a: Novel SSR markers for polymorphism detection in pigeonpea (*Cajanus spp.*). Plant Breeding **129,** 142—148.

Saxena, R. K., K. B. Saxena, R. V. Kumar, D. A. Hoisington, and R. K. Varshney, 2010b: Simple sequence repeat-based diversity in elite pigeonpea genotypes for developing mapping populations to map resistance to *Fusarium wilt* and sterility mosaic disease. Plant Breeding **129,** 135—141.

Singh, F., and D. L. Oswalt, 1992: Genetics and breeding of pigeonpea. Skill development series no. 10, F. Singh, and D. L. Oswalt, ICRISAT, Hyderabad, India.

Tanksley, S. D., 2004: The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. Plant Cell **16,** S181—S189.

Tian, Z., X. Wang, R. Lee, Y. Li, J. E. Specht, R. L. Nelson, P. E. McClean, L. Qiu, and J. Ma, 2010: Artificial selection for determinate growth habit in soybean. Proc. Natl Acad. Sci. USA **19,** 8563—8568.

Waldia, R. S., and V. P. Singh, 1987: Inheritance of stem termination in pigeonpea (*Cajanuscajan* (L.) Mill sp.). Euphytica **36,** 525—527.

Yang, S., W. Pang, G. Ash, J. Harper, J. Carling, P. Wenzl, E. Huttner, X. Zong, and A. Kilian, 2006: Low level of genetic diversity in cultivated pigeonpea compared to its wild relatives is revealed by diversity arrays technology. Theor. Appl. Genet. **113,** 585—595.

Yang, S. Y., R. K. Saxena, P. L. Kulwal, G. J. Ash, A. Dubey, J. D. I. Harper, H. D. Upadhyaya, R. Gothalwal, A. Kilian, and R. K. Varshney, 2011: The first genetic map of pigeonpea based on diversity arrays technology (DArT) markers. J. Genet. **90,** 103—109.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Distance based dendrogram based on DArT markers showing clustering of 94 pigeonpea lines.

Figure S2. Distance based dendrogram based on SNP markers showing clustering of 94 pigeonpea lines.

Figure S3. Mantel test between dissimilarity matrices obtained by DArT and SNP markers. The graph indicates a positive correlation between two marker systems.

Figure S4. The results of structural analysis of 94 pigeonpea lines using DArt markers.

Figure S5. Diversity analysis among 92 pigeonpea lines based on associated 19 SNP and 6 DArT markers.

Table S1. A list of 94 pigeonpea lines and their growth habit used during the present study.

Table S2. Nucleotide similarity of associated SNP marker sequences with soybean genome sequences.

Table S3. BlastX results of associated SNP marker sequences against nr-data base.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.