Single feature polymorphisms (SFPs) for drought tolerance in pigeonpea (Cajanus spp.)

Rachit K. Saxena, Xinping Cui, Vivek Thakur, Barbara Walter, Timothy J. Close & Rajeev K. Varshney

🐴 Springer

Springer

SHORT COMMUNICATION

Single feature polymorphisms (SFPs) for drought tolerance in pigeonpea (*Cajanus* spp.)

**Rachit K. Saxena · Xinping Cui · Vivek Thakur ·
Barbara Walter · Timothy J. Close ·
Rajeev K. Varshney**

Abstract Single feature polymorphisms (SFPs) are microarray-based molecular markers that are detected by hybridization of DNA or cRNA to oligonucleotide probes. With an objective to identify the potential polymorphic markers for drought tolerance in pigeonpea [*Cajanus cajan* (L.) Millspaugh], an important legume crop for the semi-arid tropics but deficient in genomic resources, Affymetrix Genome Arrays of soybean (*Glycine max*), a closely related species of pigeonpea were used on cRNA of six parental genotypes of three mapping populations of pigeonpea segregating for agronomic traits like drought tolerance and pod borer (*Helicoverpa armigiera*) resistance. By using robustified projection pursuit method on 15 pair-wise comparisons for the six parental genotypes, 5,692 SFPs were identified. Number of SFPs varied from 780 (ICPL 8755×ICPL 227) to 854 (ICPL 151×ICPL 87) per parental combination of the mapping populations. Randomly selected 179 SFPs were used for validation by Sanger sequencing and good quality sequence data were obtained for 99 genes of which 75 genes showed sequence polymorphisms. While associating the sequence polymorphisms with SFPs detected, true positives were observed for 52.6% SFPs detected. In terms of parental combinations of the mapping populations, occurrence of true positives was 34.48% for ICPL 151×ICPL 87, 41.86% for ICPL 8755×ICPL 227, and 81.58% for ICP 28×ICPW 94. In addition, a set of 139 candidate genes that may be associated with drought tolerance has been identified based on gene ontology analysis of the homologous pigeonpea genes to the soybean genes that detected SFPs between the parents of the mapping populations segregating for drought tolerance.

Keywords Single feature polymorphism · Microarray · Robustified projection pursuit · Molecular markers · Legumes

R. K. Saxena · V. Thakur · R. K. Varshney (✉)
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
Patancheru 502324, India
e-mail: r.k.varshney@cgiar.org

X. Cui · B. Walter · T. J. Close
University of California,
Riverside, CA 92521, USA

R. K. Varshney
CGIAR- Generation Challenge Programme (GCP),
c/o CIMMYT,
06600, Mexico DF, Mexico

Introduction

Pigeonpea [*Cajanus cajan* (L.) Millspaugh], with a genome size 858 Mbp and 11 pairs of chromosomes, is a major grain legume crop in the tropical and subtropical regions of the world. However, the crop productivity of pigeonpea has remained stagnant less than 1 ton per hectares for last 40 years as the crop is exposed to several biotic (e.g., *Fusarium* wilt, sterility mosaic disease, and pod borer) and abiotic (drought, salinity, and water logging) stresses. As genomics-assisted breeding has been very successful in several temperate cereals (Varshney et al. 2006) and some legume species like soybean (Varshney et al. 2010), the pigeonpea crop has remained untouched with genomics research. For instance, until recently a few hundred simple sequence repeat (SSR) markers were available (Burns et al. 2001; Odeny et al. 2007; Saxena et al. 2010a,b), some

efforts are being made to develop large-scale SSR markers after mining the end sequences of bacterial artificial chromosome (BAC) clones (Bohra et al. 2011). In addition to non-availability of appropriate genomic resources, occurrence of a very narrow genetic diversity pose another serious constraint for developing the genetic map and quantitative trait loci (QTL) analysis for traits of interest to breeders in pigeonpea (Odeny et al. 2007; Saxena et al. 2010a; Bohra et al. 2011; Yang et al. 2011). It is, therefore, evident that there is a need to develop large-scale genomic resources in pigeonpea that can be used not only for enhancing the basic genome research but also in crop improvement program.

Single feature polymorphisms (SFPs) in the context of oligonucleotide arrays, including single nucleotide polymorphism (SNPs) and insertions and deletions (INDELs), are particularly amenable to microarray-based genotyping (Shiu and Borevitz 2008). In case of food legume crops, microarrays (Affymetrix Genome Arrays) have been designed only in soybean (*Glycine max*). Because of occurrence of orthologous genes in closely related species that share high sequence similarity, microarrays developed for one species can be used in the other closely related species (Bar-or et al. 2006; Nuzhdin et al. 2004). In terms of phylogenetic relationships among legumes, pigeonpea, cowpea (*Vigna ungiculata*), and common bean (*Phaseolous vulgaris*) are grouped together with soybean under the *Phaseoloid* clade. This indicates that Affymetrix soybean genome arrays can be used to identify SFPs in genotypes of interest in aforementioned closely related *Phaseoloid* species (Das et al. 2008).

With an objective of extending repertoire of genomic resources in pigeonpea, the present study has deployed soybean genome arrays to identify the SFPs between the parental genotypes of three mapping populations of pigeonpea segregating for agronomic traits like drought tolerance and resistance to *Helicoverpa armigiera*, the pod borer insect. Subsequently, a subset of genes detecting SFPs was used for Sanger sequencing on the parental lines used in the study to measure the precision of SFP prediction. Furthermore, a set of candidate genes for drought tolerance has been identified by gene ontology analysis of the pigeonpea genes, homologues to the soybean genes predicting SFPs in the parental genotypes of mapping populations segregating for drought tolerance.

Materials and methods

Plant materials and soybean genome array

A total of six accessions namely ICP 28, ICPL 8755, ICPL 151, ICPL 87, and ICPL 227 from the cultivated genepool (*C. cajan*) and ICPW 94 a wild relative of pigeonpea (*Cajanus scarabaeoides*) were used in this study. These six genotypes are the parents of three different mapping populations segregating for important agronomic traits such as drought tolerance and pod borer resistance.

The Affymetrix Soybean Genome Arrays used in this study contained 37,500 probe sets derived from soybean unigenes. This represents 61% of the total probe sets on the chip, with the remainder targeting two pathogens important for soybean genetics research, of which 15,800 (26%) probe sets target *Phytophthora sojae* (a water mold) and 7,500 (12%) probe sets target *Heterodera glycines* (soybean cyst nematode). The genome array used probe sets composed of 11 probe pairs to measure the expression of each gene. Each probe pair consists of a perfect match (PM) probe and a mismatch (MM) probe.

RNA isolation and microarray hybridization

Root tissue samples were collected from all the six pigeonpea genotyped mentioned above after 15 days of sowing. RNA was isolated following the protocol of Schmitt et al. (1990). RNA quality was assessed using formamide gel electrophoresis and Agilent 2100 Bioannalyzer (Agilent technologies, Palo Alto, CA, USA).

Expression data were generated by hybridizing cRNA of pigeonpea genoypes to the soybean genome arrays. R version 2.10.1 and packages "Affy" and "gcrma" within BioConductor were used for data pre-processing. For correction of background and non-specific binding, GeneChip-Robust Multichip Average (GC-RMA; Wu et al. 2004) was used. Quantile normalization was used for probe-level normalization (Irizarry et al. 2003). To eliminate probe sets with absent transcripts when pigeonpea cRNAs were hybridized to the soybean genome array, we adopted the filtering procedure suggested by Schuster et al. (2007). Briefly, the MM probe values were replaced with the mean PM value (after GC-RMA transformation) of probe sets that were very likely to have absent target transcripts. By using the Micro array Suite version 5.0 (MAS 5.0) of Affyemetrix Inc., the present/absent calls were calculated based on the transformed PM and MM probe intensities. The probe sets that were "present" in both conditions under comparison were used for SFP detection (Das et al. 2008).

SFP prediction

We used robustified projection pursuit (RPP) method for SFP detection (Cui et al. 2005), for which only the GC-RMA adjusted PM probe values from "present" probe sets were utilized. A probe set was called "present" if it had present calls in all biological replicates of the two genotypes under comparison. Separate pair-wise comparisons were made among ICP 28, ICPW 94, ICPL 151,

ICPL 87, ICPL 8755, and ICPL 227. For each comparison, we used the top 15% outlying score as cutoff for calling SFP-containing probe sets, within which a probe will be identified as a SFP probe if it accounts for more than 40% of overall outlying score of its residing probe set.

Primer designing and sequencing

For validating the predicted SFPs, primer pairs were designed for the selected SFPs and used for amplification and sequencing of the genomic fragments. In this context, polymerase chain reaction (PCR) primer pairs were designed to bind at least 100 bases upstream or downstream of the probes predicted to contain SFPs. Sequence data for the soybean genes corresponding to the selected SFPs (available at www.Affymetrix.com) were used for BLASTN analysis with the pigeonpea transcriptome assembly (Dubey et al. 2011) and corresponding 200–500 bp sequence region from the pigeonpea genes was used to design the primer pair with help of Primer3 (Rozen and Skaletsky 2000).

The primer pairs were used to generate the amplicons using the same PCR conditions as given in Nayak et al. (2009). Subsequently, the amplified products were sequenced on an ABI 3730 DNA Analyzer using BigDye Terminator V1.1 (Macrogen, Seoul, Korea). Good quality sequence data were used to form contigs by using DNA Baser software (http://www.dnabaser.com). Further, contigs were aligned and viewed with the Bioedit software (http://www.mbio.ncsu.edu/BioEdit/). In-house developed software "Divest" (Jayashree et al. 2009) was used to detect the presence of SNPs and INDELs in the sequence data.

Gene ontology analysis

Functional assignment of pigeonpea tentative unique sequences (TUSs), homologues to the soybean genes that detected SFPs in the parental genotypes segregating for drought tolerance, was accomplished by finding significant hits in the UniProt database (e value$\leq 10^{-5}$; Jain et al. 2009). The gene ontology IDs were retrieved from the UniProt database using keywords obtained in the BLASTX descriptions of the most significant hits. Based on the Gene Ontology ID, unique sequences were categorized into three principal categories: biological processes, cellular localizations, and molecular functions.

Results and discussion

Single feature polymorphism discovery

The Affymetrix Soybean Genome Arrays were used with cRNA of six genotypes of pigeonpea and differentially expression data were analyzed using RPP method to identify SFPs in pigeonpea (Das et al. 2008). In the genome arrays, only 37,376 transcripts were represented by probe sets consisting of 11 PM. Therefore, the microarray data obtained for these 37,376 transcripts were analyzed for generating "present", "marginal", and "absent" calls for these transcripts in the genotypes for SFP analysis. Scatter plots of 411,136 PM probes for all pair-wise combinations revealed much less variation between two biological replicates of each accession (except replicates of ICP 28) compared to that between any two accessions, suggesting the feasibility of detecting SFP probes between accessions. The number of "present" calls varied, ranging from 4,882 to 5,810 probes for the 15 pair-wise comparisons. It is important to mention here that the number of probes showing differential hybridization and thus qualifying for potential SFPs in the combination ICPL 8755 and ICPL 227 (4,989) were less than those in ICP 28 and ICPW 94 (5,405) or ICPL 151 and ICPL 87 (5,455) pairs (Table 1).

By using stringent criteria and RPP analysis, a total of 5,692 potential SFPs were discovered across the six genotypes (Table 1). As an example, Fig. 1 shows detection of SFP between ICPL 8755 and ICPL 227 genotypes for the probe #5 of the SFP probe set GmaAffx.5953.1.A1_at. As the six genotypes analyzed in this study represents the parents of three mapping populations, an effort was made to identify the SFPs for different parental combinations. In this context, 850 SFPs were identified from 5,405 "present" probe sets in the ICP 28×ICPW 94 (cross 1), 854 SFPs out of 5,455 "present" probe sets in the ICPL 151×ICPL 87 (cross 2), and 780 SFPs out of 4,989 in the ICPL 8755×ICPL 227 (cross 3). However, many SFPs that were detected were unique to one parental combination. The number of SFPs in common between cross 1 and 2 was 17, between cross 1 and 3 it was 14, and between cross 2 and 3 a total of 19 SFPs were in common. A total of 10 SFPs were in common for all the three crosses used. These results reconfirm earlier observations on occurrence of low level of genetic diversity based on other marker systems such as amplified fragment length polymorphisms (Panguluri et al. 2006), Diversity Array Technology (Yang et al. 2011) and SSRs (Saxena et al. 2010a). Nevertheless, the present study adds a set of about 1,000 novel markers (SFPs) for genetics and breeding analysis in pigeonpea.

Validation of SFPs

With an objective to validate the SFPs at sequence level, a subset of SFPs was selected for allele-specific resequencing. In this context, homologues sequences in pigeonpea for the soybean probes detecting SFPs were identified by sequence analysis of the corresponding genes of soybean with the transcriptome assembly of pigeonpea (Dubey et al. 2011) comprising 127,754 TUSs defined based on cluster analysis

Table 1 Summary on identification and validation of SFP in pigeonpea

Parental genotype combination	SFPs identification		SFPs validation		
	No. of probes		No. of selected probes	No. of probes with sequence variation in	
	Present	Containing SFPs		Full length amplicon sequence	Probe sequence
ICP 28×ICPW 94	5,405	850	38	34	31
ICPL 151×ICPL 87	5,455	854	58	37	20
ICPL 8755×ICPL 227	4,989	780	43	34	18
Total	5,810	5,692	99	75	67

of 454/FLX transcript reads and Sanger ESTs of pigeonpea (Raju et al. 2010). By applying e-value$\leq10^{-5}$ and sequence similarity$\geq80\%$, 2,745 (48.2%) out of 5,692 SFP containing probes identified homologues in pigeonpea. These pigeonpea sequences were further examined for the presence of interrogation position for the corresponding soybean SFP probe. As a result, 1,815 pigeonpea TUSs were found positive and used for primer designing with an expected amplicon size of 200–500 bp. In summary, primer pairs could be designed for 1,131 TUSs containing SFP probe target regions (Electronic Supplementary Material Table 1). In other cases, primers could not be designed, this was mainly due to either the probe target region being very near (≤20 bp) the end of sequence or inability to fit the default parameters of the Primer3 software (Nayak et al. 2010).

In order to investigate whether the identified SFPs were related to sequence variations, 179 SFPs were randomly selected for validation. PCR with these 179 primer pairs on the same set of six genotypes used for SFP discovery provided strong and prominent amplicons in 102 (56.98%) cases. In the remaining cases, either no amplification or nonspecific amplification was observed. Amplicons generated for the 102 primers were sequenced using Sanger sequencing methodology. As a result, good quality sequence data were obtained for 99 primer pairs, which were further analyzed to identify the SNPs and INDELs. Most sequences were 250–650 bp, but some were as short as 207 bp or as long as 1,336 bp. Analysis of the sequence data with "Divest" tool (Jayashree et al. 2009) showed a total of 7,535 sequence polymorphisms (including SNPs

Fig. 1 A snapshot of SFP analysis in pigeonpea for the probe set GmaAffx.5953.1.A1_at of Affymetrix Soybean Genome Array. Two replicates of genotype ICPL 8755 and genotype ICPL 227 has been shown in *dark* and *dotted lines*, respectively. The *top panel* shows log intensities (*left side*) and hybridization affinities (*right side*) for all probes of the probe set for two genotypes. The *bottom panel* shows the difference in hybridization affinities (*left side*) and the individual outlying scores (*right side*) for each probe between two genotypes
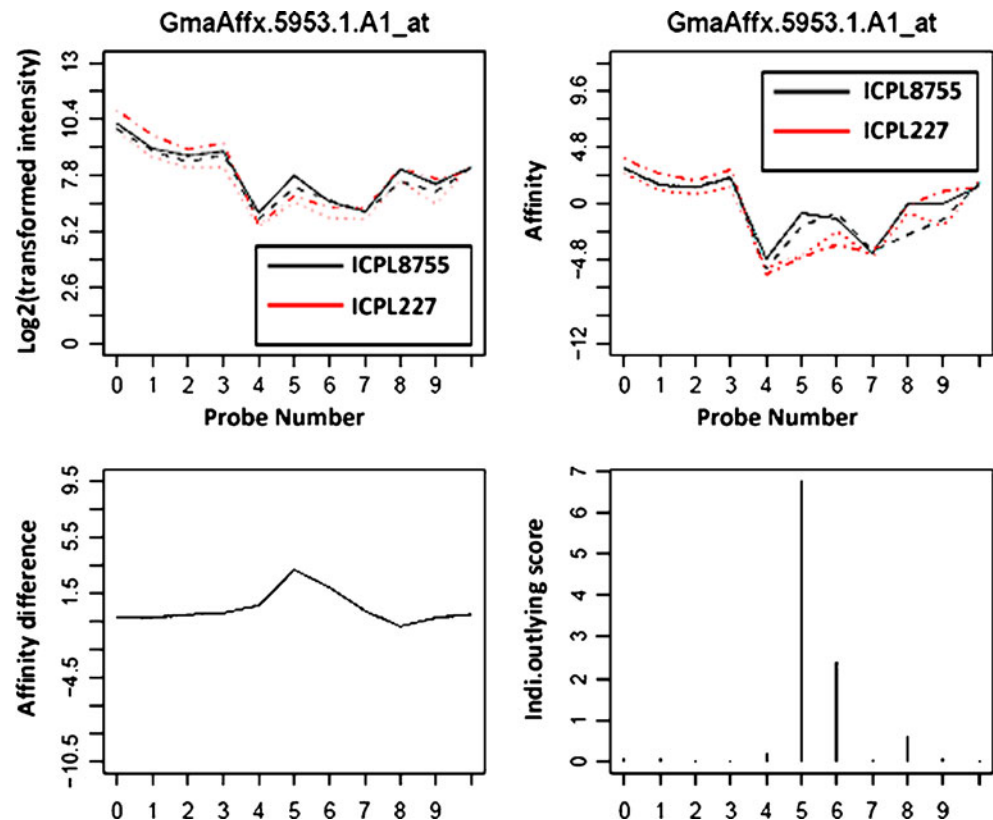
and INDELs) for 75 (75.7%) out of 99 primer pairs (Electronic Supplementary Material Table 2). Among all the sequence variations identified within 99 probe regions, 363 were SNPs and 44 INDELs. A representative alignment of genomic amplicon sequences examined for a putative SFP between ICP 28 and ICPW 94 genotypes for probe #3 of probe set Gma.12798.1.S1_at, revealed occurrence of four SNPs in the genotypes (Electronic Supplementary Material Figure 1). In addition, as compared to the pigeonpea TUS (TUS ID127906_2368_0221), there was a single base deletion in the two genotypes investigated. On the other hand, as compared to soybean gene (Gma.12798.1.S1_at), an insertion of 2 bp was observed in the both pigeonpea genotypes.

Based on SFP prediction, in aforementioned sequence data for 99 genes, the parental genotypes for cross 1 should have sequence polymorphisms for 38 genes, the parental genotypes for cross 2 should have sequence polymorphisms for 58 genes and 43 genes should have sequence polymorphisms for the parents for the cross 3. Sequence analysis for 75 polymorphic genes, however, confirmed SNPs or INDELs for 31 (81.58%), 20 (34.48%), and 18 (41.86%) genes for the parental genotypes of the cross 1, 2, and 3 respectively (Table 1). Across all the six genotypes, the accuracy of the array to predict the presence of a sequence variant based on SFP was 52.6%.

In summary, 52.6% predicted SFPs were found true, while the remaining 47.4% SFPs predicted were found false. The false discovery rate (FDR) of SFPs predicted in this study is relatively higher than FDRs reported in other studies using cRNA for SFP detection (Rostoks et al. 2005; Cui et al. 2005; Das et al. 2008; Yang et al. 2009). For instance, by using barley Affymetrix genome arrays to detect SFPs in barley genotypes, the FDR was 10–20% (Cui et al. 2005) and 40% (Rostoks et al. 2005). It is, however, noted that the present study deployed soybean genome arrays for SFP discovery in pigeonpea genotypes. Previously also, the soybean genome arrays were used to detect SFPs in cowpea (Das et al. 2008) and the barrel medic (*Medicago truncatula)* genome arrays were used for SFP discovery in alfalfa (*Medicago sativa*) (Yang et al. 2009). In these species, however, lower FDR (32% in cowpea and 17% in alfalfa) was observed. Higher (47.4%) FDR observed in the present study can be attributed to several factors. The phylogenetic distance between soybean and pigeonpea is higher than the phylogenetic distances between soybean–cowpea and Medicago–alfalfa. Owing to the relatively low level of sequence similarity between probes and transcripts in cross-species hybridization, more probes might be considered to cross-hybridize in cross-species hybridization as compared to self species hybridization (Bar-or et al. 2006). Some false positive cases may occur due to paralogs being sequenced in the tested

genotypes which do not exhibit polymorphism. Variation in post-transcriptional modification such as alternative splicing in some genes may also have contributed to the FDR observed (Rostoks et al. 2005; Das et al. 2008).

In addition to checking the SFPs specific to the given cross, analysis was also done to validate the SFPs that were common to parents of cross 1 and 2 (17), cross 1 and 3 (14), cross 2 and 3 (19), and cross 1, 2, and 3 (10). While a higher proportion of SFPs were confirmed for the parents of cross 1 and 2 (14, 82.35%), the least proportion of SFPs were confirmed for the parents of cross 2 and 3 (9, 47.36%). While 50% of SFPs (5) that were common to all the parents were confirmed at the sequence level, 64.28% (9) SFPs were confirmed for the parents of cross 1 and 3. Such polymorphisms generally represent haplotypes, which will be very useful in mapping of these genes especially for construction of consensus genetic maps. Moreover, these sequence variations represent expressed portion of the genomes and hence can directly be associated with some important phenotypic traits.

Gene ontology descriptions

Functional analysis was done for the 922 pigeonpea TUSs for the corresponding soybean transcripts that detected SFPs (soybean transcripts) between the parents of the mapping populations (ICPL 8755×ICPL 227 and ICPL 151×ICPL 87) segregating for drought tolerance. Analysis of these sequences against UniProt database (Uniref50) showed that 752 TUSs (81.56%) had similarity to the proteins available in the database at a stringent criterion of e-value$\leq 10^{-5}$. Subsequently, these TUSs were analyzed for gene ontology descriptions based on their BLASTX functions. As a result, 724 (78.52%) TUSs could be assigned into three principal categories: molecular function (599), biological process (559), and cellular component (570). While distributing these TUSs into various subcategories of three main categories, the highest number of TUSs fell into cell part (555) followed by cellular process (453), nucleotide binding (426), and metabolic process (409) subcategories (Electronic Supplementary Material Figure 2). Gene ontology analysis was also extended to identify the genes related to stress responses. As a result, a total of 139 TUSs were found under "response to stimulus" subcategory which includes both abiotic and biotic responses. This provides an indication of the involvement of these genes for drought tolerance. Some functional genomics approaches like qRT-PCR may validate the function of these genes (Hu et al. 2009). Furthermore, the linkage mapping of these SFPs using aforementioned mapping populations may provide association of these genes for QTLs for drought tolerance that can be used as "functional markers" in marker-assisted selection approaches in pigeonpea breeding.

In summary, the present study identified 5,692 unique candidate SFPs extending the marker repertoire with functional marker systems in pigeonpea. Allele-specific sequencing of a set of selected genes detecting SFPs showed association of 52.6% SFPs analyzed with actual sequence polymorphisms in the probe sets. Gene ontology analysis of the genes detecting SFPs provided a set of candidate genes that may have association with drought tolerance. These candidate genes are useful resource for undertaking the gene expression analysis as well development of functional markers for both basic and applied research, especially for drought tolerance in pigeonpea improvement.

References

Bar-Or C, Bar-Eyal M, Gal TZ, Kapulnik Y, Czosnek H, Koltai H (2006) Derivation of species-specific hybridization-like knowledge out of cross-species hybridization results. BMC Genomics 7:110

Bohra A, Dubey A, Saxena RK, Penmetsa RV, Poornima KN, Kumar N, Farmer AD, Srivani G, Upadhyaya HD, Gothalwal R, Ramesh R, Singh D, Saxena KB, PB Kavi kishor, Town CD, May GD, Cook DR, Varshney RK (2011) Analysis of BAC-end sequences (BESs) and development of BES-SSR markers for genetic mapping and hybrid purity assessment in pigeonpea (*Cajanus* spp.). BMC Plant Biol 11:56

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

Cui X, Xu J, Asghar R, Condamine P, Svensson JT, Wanamaker S, Stein N, Roose M, Close TJ (2005) Detecting single-feature polymorphisms using oligonucleotide arrays and robustified projection pursuit. Bioinformatics 21:3852–3858

Das S, Bhat PR, Sudhakar C, Ehlers JD, Wanamaker S, Roberts PA, Cui X, Close TJ (2008) Detection and validation of single feature polymorphisms in cowpea (*Vigna unguiculata* L. Walp.) using a soybean genome array. BMC Genomics 9:107

Dubey A, Farmer A, Schlueter J, Cannon S, Abernathy B, Tuteja R, Woodward J, Shah T, Mulasmanovic B, Kudapa H, Raju NL, Gothalwal R, Pande S, Xiao Y, Town CJ, Singh N, May GD, Jackson S, Varshney RK (2011) Defining the transcriptome assembly and its use for genome dynamics and transcriptome profiling studies in pigeonpea (*Cajanus cajan* L.). DNA Research (in press)

Hu Y, Li WC, Xu YQ, Li GJ, Liao Y, Fu FL (2009) Differential expression of candidate genes for lignin biosynthesis under drought stress in maize leaves. J Appl Genet 50:213–223

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. Nuc Acid Res 31:e15

Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P, Gasteiger E (2009) Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinform 10:136

Jayashree B, Bhanuprakash A, Jami A, Reddy SP, Nayak S, Varshney RK (2009) Perl module and PISE wrappers for the integrated analysis of sequence data and SNP features. BMC Res Notes 2:92

Nayak SN, Balaji J, Upadhyaya HD, Hash CT, Kavi Kishor PB, Chattopadhyay D, María Rodriguez L, Blair M, Baum M, McNally K, This D, Hoisington DA, Varshney RK (2009) Isolation and sequence analysis of DREB2A homologues in three cereal and two legume species. Plant Sci 177:460–467

Nayak SN, Zhu H, Varghese N, Datta S, Choi H-K, Horres R, Jüngling R, Singh J, Kavi Kishor PB, Sivaramakrihnan S, Hoisington DA, Kahl G, Winter P, Cook DR, Varshney RK (2010) Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with Medicago truncatula genome. Theor Appl Genet 120:1415–1441

Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM (2004) Common pattern of evolution of gene expression level and protein sequence in Drosophila. Mol Biol Evol 21:1308–1317

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

Panguluri SK, Janaiah K, Govil JN, Kumar PA, Sharma PC (2006) AFLP fingerprinting in pigeonpea (*Cajanus cajan* (L.) Millsp.) and its wild relatives. Genet Resour Crop Evol 53:523–531

Raju NL, Gnanesh BN, Lekha PT, Balaji J, Pande S, Byregowda M, Singh NK, Varshney RK (2010) The first set of EST resource for gene discovery and marker development in pigeonpea (*Cajanus cajan* L.). BMC Plant Biol 10:45

Rostoks N, Borevitz J, Hedley P, Russell J, Mudie S, Morris J, Cardle L, Marshall D, Waugh R (2005) Single-feature polymorphism discovery in the barley transcriptome. Genome Biol 6:R54

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana, Totowa, pp 365–386

Saxena RK, Prathima C, Saxena KB, Hoisington DA, Singh NK, Varshney RK (2010a) Novel SSR markers for polymorphism detection in pigeonpea (*Cajanus* spp.). Plant Breed 129:142–148

Saxena RK, Saxena KB, Varshney RK (2010b) Application of SSR markers for molecular characterization of hybrid parents and purity assessment of ICPH 2438 hybrid of pigeonpea [*Cajanus cajan* (L.) Millspaugh]. Mol Breed 26:371–380

Schmitt ME, Brown TA, Trumpower BL (1990) A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. Nucl Acid Res 18:3091–3092

Schuster EF, Blanc E, Partridge L, Thornton JM (2007) Correcting for sequence biases in present/absent calls. Genome Biol 8:R125

Shiu S-H, Borevitz JO (2008) The next generation of microarray research: applications in evolutionary and ecological genomics. Heredity 100:141–149

Varshney RK, Hoisington D, Tyagi AK (2006) Advances in cereals genomics and applications in crop breeding. Trends Biotechnol 24:490–499

Varshney RK, Glaszmann J-C, Leung H, Ribaut JM (2010) More genomic resources for less-studied crops. Trends Biotechnol 28:452–460

Wu Z, Irizarry RA, Gentleman R, Murillo FM, Spencer F (2004) A model based background adjustment for oligonucleotide expression data. J Am Stat Assoc 99:909–917

Yang SY, Saxena RK , Kulwal PL, Ash GJ, Dubey A, Harpe JDI, Upadhyaya HD, Gothalwal R, Kilian A, Varshney RK (2011) The first genetic map of pigeonpea based on Diversity Arrays Technology (DArT) markers. Journal of Genetics 90: (in press)

Yang SS, Xu WW, Tesfaye M, Lamb JAFS, Jung HJG, Samac DA, Vance CP, Gronwald JW (2009) Single-feature polymorphism discovery in the transcriptome of tetraploid alfalfa. Plant Genome 2:224–232